

Федеральное государственное автономное образовательное учреждение высшего
образования

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТОМСКИЙ
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

На правах рукописи

ЕМЕЛЬЯНОВА ЕКАТЕРИНА ЮРЬЕВНА

**АГРЕГИРОВАНИЕ ПРЕДПОЧТЕНИЙ НА ОСНОВЕ ТОЧНОГО
РЕШЕНИЯ ЗАДАЧИ О РАНЖИРОВАНИИ КЕМЕНИ**

Специальность: 2.3.1 – Системный анализ, управление и обработка информации

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель – доктор технических наук,
профессор Муравьев Сергей Васильевич

Томск – 2022

Содержание

Введение	4
ГЛАВА 1 Алгоритм точного решения задачи о ранжировании Кемени	11
1.1 Агрегирование предпочтений	11
1.2 Ранжирование и профиль предпочтений	12
1.2.1 Слабый порядок.....	13
1.2.2. Профиль предпочтения.....	16
1.2.3 Взвешенная турнирная матрица.....	19
1.3 Правила агрегирования предпочтений	20
1.3.1 Правило простого большинства	20
1.3.2 Правило Борда.....	22
1.3.3 Правило Кондорсе.....	23
1.3.4 Правило Кемени.....	26
1.4 Проблемы практического применения правила Кемени	29
1.4.1 <i>NP</i> -трудность задачи о ранжировании Кемени.....	30
1.4.2 Множественность решений.....	31
1.5 Точный алгоритм нахождения всех ранжирований Кемени.....	33
Выводы к главе 1.....	41
ГЛАВА 2 Множественность решений задачи о ранжировании Кемени.....	43
2.1 Причины множественности	43
2.2 Предыдущие исследования множественности	44
2.3 Экспериментальные исследования транзитивности входного профиля.....	50
2.3.1. Критерий транзитивности входного профиля	50
2.3.2. План экспериментов.....	51
2.3.3. Транзитивность входного профиля при различных m и n	54
2.3.4. Оценка вероятности транзитивного входного профиля	59
2.4 Транзитивность выходного профиля.....	65
Выводы к главе 2.....	68

ГЛАВА 3 Преобразование выходного профиля в единственное оптимальное решение	70
3.1 Взвешенный турнир выходного профиля	70
3.1.1 Матрица взвешенного турнира	70
3.1.2 Параметры взвешенного турнира.....	71
3.2 Построение правила свёртки	77
3.2.1 Характеризация выходного профиля ЗРК.....	77
3.2.2 Желательные свойства правила свёртки	79
3.2.3 Свёртка на основе выигрышей z_i	81
3.2.4 Свёртка с использованием общих рангов r_i	82
3.2.5 Расстояние Кемени и итоговое ранжирование консенсуса.....	83
Выводы к главе 3.....	87
ГЛАВА 4 Практические аспекты агрегирования предпочтений на основе правила Кемени	88
4.1 Две основные парадигмы применения правила Кемени	88
4.1.1 Измерение в многомерной порядковой шкале.....	88
4.1.2 Комплексирование интервалов агрегированием предпочтений.	89
4.2 Метод комплексирования интервальных данных IF&PA.....	90
4.2.1 Этапы метода IF&PA.....	90
4.2.2 Инранжирования.....	94
4.3 Свойства инранжирований и связанных с ними пространств.....	97
4.3.1 Пространства инранжирований.....	97
4.3.2 Мощность классов A_k	102
4.3.3 Порождающие инранжирования множества	103
Выводы к главе 4.....	107
Заключение	108
Список сокращений и обозначений.....	110
Список используемой литературы.....	114
Приложение А. Акты внедрения диссертационной работы.....	124

Введение

Актуальность темы. *Агрегирование предпочтений* – это классическая задача определения ранжирования консенсуса для заданного исходного профиля предпочтений, состоящего из m ранжирований, возможно, включая толерантности, n альтернатив (или кандидатов). Эта задача с конца XVIII века исследуется в основном в рамках *теории голосования* (или *социального выбора*).

Ранжирование консенсуса должно удовлетворять определенным естественным требованиям (правилам), обычно формулируемым в аксиоматической форме. Различные *правила агрегирования предпочтений* отличаются друг от друга с учетом их соответствия определенным аксиомам. Одним из наиболее аксиоматически обоснованных правил считается правило Кондорсе, основанное на парном сравнении альтернатив. Однако это правило может приводить к парадоксу не-транзитивности Кондорсе. Разумным способом преодоления этой трудности является использование правила Кемени, которое позволяет найти такой линейный порядок альтернатив (ранжирование Кемени), что расстояние от него до исходного профиля предпочтений будет минимальным. Это *расстояние Кемени* определяется в терминах числа попарных несоответствий между ранжированиями.

Правило Кемени обладает положительными свойствами правила Кондорсе и не приводит к парадоксу, что позволяет рекомендовать его для применения в различных ситуациях агрегирования предпочтений, например, при многокритериальном (или групповом) принятии решений, в области искусственного интеллекта и, в частности, машинного обучения. Соответствующие методы используются для решения таких проблем, как краудсенсинг/лэйбелинг, анализ настроений, создание мета-поисковых систем и т.п.

Однако, для практического использования правила Кемени существуют два серьезных препятствия: (1) задача о ранжировании Кемени (ЗРК) является *NP-трудной* и (2) правило Кемени может приводить к множественным ранжированиям консенсуса.

NP-трудность означает, что вряд ли будет существовать алгоритм с поли-

номиальным временем (пропорциональным размерности n задачи) решения для правила Кемени. Этот факт требует разработки различных подходов для преодоления для нахождения оптимального решения за приемлемое время. Однако известны реализующие правило Кемени алгоритмы, обеспечивающие нахождение точного решения за время порядка нескольких секунд при размерности задачи до $n < 20$. Эта размерность вполне подходит для практических применений.

Множественность решений ЗРК означает, что ранжирований Кемени может быть слишком много. Хотя эти оптимальные решения, количество которых будем обозначать N , находятся на одном и том же минимальном расстоянии Кемени от m исходных ранжирований, они могут быть довольно далеки друг от друга. При определенных значениях m и n (особенно при малых и четных $m \geq 4$) значение N может достигать 10^7 . При этом возникает хаосо-подобная ситуация, когда $N \gg m$, и неопределенность после решения задачи значительно превышает исходную неопределенность.

Применение правила Кемени в агрегировании предпочтений требует проведения систематических исследований и разработки основанных на них рекомендаций по преодолению указанных выше трудностей.

Целью диссертационной работы является разработка и экспериментальные исследования процедур агрегирования предпочтений на основе точного решения задачи о ранжировании Кемени.

В связи с поставленной целью должны быть решены следующие **задачи**:

- анализ основных правил агрегирования предпочтений и исследование агрегирования предпочтений в форме правила Кемени и реализующего это правило точного алгоритма;
- проведение экспериментальных исследований транзитивности/нетранзитивности входного профиля предпочтений и его связи с явлением множественности решений ЗРК;
- разработка правил преобразования множества оптимальных решений ЗРК в единственное итоговое оптимальное решение;
- исследование пространства слабых порядков при применении правила

Кемени в методе комплексирования интервалов агрегированием предпочтений IF&PA.

Методы исследования. Использованы методы теории голосования (социального выбора), математической статистики, комбинаторики, теории измерений и агрегирования предпочтений. Экспериментальные численные исследования проводились с использованием метода Монте-Карло с генерацией наборов индивидуальных входных профилей, состоящих из случайных ранжирований, элементы которых распределены по равномерному закону.

Достоверность полученных результатов диссертационной работы подтверждается совпадением с достаточной точностью аналитических расчетов и результатов численных экспериментов.

Научная новизна

1. Предложена стратегия получения итогового ранжирования консенсуса при применении правила Кемени, заключающаяся в нахождении всех ранжирований консенсуса, являющихся точными оптимальными решениями, достижимыми при размерности задачи $n < 20$, и нахождении для них единственного подходящего итогового ранжирования консенсуса.
2. Экспериментальные исследования фактора множественности, связанного с проявлением транзитивности/нетранзитивности входного профиля предпочтения, показали, что вероятность транзитивности значительно выше при чётных значениях m , чем при нечётных, и плавно уменьшается с ростом n , тогда как вероятность множественности резко увеличивается при уменьшении чётных значений m , но плавно увеличивается при уменьшении нечётных значений m .
3. Предложено правило свёртки всех N оптимальных решений задачи о ранжировании Кемени в точное единственное итоговое ранжирование: в итоговом ранжировании консенсуса альтернативы располагаются в порядке убывания строковых сумм, вычисленных для N -взвешенной турнирной матрицы выходного профиля; две альтернативы толерантны друг другу, если они имеют одинаковые строковые суммы.
4. При применении правила Кемени в методе комплексирования интервалов агре-

гированием предпочтений IF&PA введено новое понятие инражирования как слабого порядка, наведенного входными интервалами, которое появляется в результате ограничений, накладываемых на обычные ранжирования интервальным характером исходных данных; показано, что мощность пространства инражирований определяется треугольным числом для заданного числа n дискретных элементов.

Практическая ценность работы. Результаты диссертационной работы могут быть использованы для реализации процедур агрегирования предпочтений и нахождения отношения консенсуса при принятии решений (например, при выборе цветовой модели для цветометрического анализа состава веществ), оценивании качества продукции и услуг, решении многокритериальных задач, и т.п. В задачах комплексирования интервалов агрегированием предпочтений результаты работы могут использоваться для обработки интервальных данных разной природы, в частности, для организации межлабораторных сличений, для установления значений фундаментальных констант, повышения точности оценок измеряемых величин и т.д.

Реализация и внедрение результатов работы. Результаты исследований использованы при выполнении НИР по гранту 18-19-00203 Российского научного фонда "Агрегирование предпочтений для решения задач обработки многомерных гетероскедастичных измерительных данных", 2018-2020 гг.

Результаты работы также используются: в научно-исследовательской лаборатории "Безопасность и электромагнитная совместимость радиоэлектронных средств" федерального государственного бюджетного образовательного учреждения высшего образования "Томский государственный университет систем управления и радиоэлектроники"; в лаборатории службы релейной защиты и испытаний электрооборудования ООО "Горсети" г. Томска и в учебном процессе в отделении автоматизации и робототехники Инженерной школы информационных технологий и робототехники ТПУ. Акты внедрения приложены к диссертационной работе.

Положения, выносимые на защиту

1. Предложенная стратегия получения итогового ранжирования консенсуса обеспечивает возможность применения правила Кемени для широкого круга практических задач агрегирования предпочтений с максимально возможной точностью.
2. Полученные оценки вероятности появления нетранзитивности входного профиля соответствуют утверждениям, что при нечетных m вероятность появления нетранзитивного профиля всегда больше или равна вероятности парадокса Кондорсе, а при четных m вероятность появления нетранзитивного профиля всегда меньше вероятности парадокса Кондорсе; эти оценки согласуются с известными аналитическими оценками вероятности альтернативы Кондорсе и оценками вероятности появления множественности решений ЗРК.
3. Предложенная процедура свёртки всех N оптимальных решений задачи о ранжировании Кемени в точное единственное итоговое ранжирование посредством вычисления строковых сумм для N -взвешенной турнирной матрицы выходного профиля, имеет естественное обоснование на основе правила Борда.
4. Применение правила Кемени в методе комплексирования интервалов агрегированием предпочтений IF&PA приводит к пространству инражирований, которое является подмножеством пространства всех слабых порядков с единственным символом строгого порядка.

Апробация результатов работы. Основные результаты диссертационной работы докладывались на следующих конференциях: XIV Международная научно-практическая конференция студентов, аспирантов и молодых ученых "Современные техника и технологии", г. Томск, 2008 г.; XV Международная научно-практическая конференция студентов, аспирантов и молодых учёных "Молодёжь и современные информационные технологии", г. Томск, 2017 г.; XVI Международная научно-практическая конференция студентов, аспирантов и молодых ученых "Молодежь и современные информационные технологии" г. Томск, 2018 г.; II всероссийская научно-практическая конференция с международным участием им. В.В. Губарева "Интеллектуальный анализ сигналов, данных и знаний: методы и

средства", г. Новосибирск, 2018 г.; XVII Международная научно-практическая конференция студентов, аспирантов и молодых учёных "Молодежь и современные информационные технологии", г. Томск, 2020 г.; XIV Всероссийская научная конференция молодых ученых "Наука. Технологии. Инновации", г. Новосибирск, 2020 г.; Международный симпозиум ТК1-ТК-7-ТК13-ТК18 ИМЕКО, г. Санкт-Петербург, 2019 г.

Публикации. Основные результаты исследований отражены в 11 публикациях: 5 статей в ведущих научных журналах и изданиях, рекомендуемых ВАК, все проиндексированы в базах данных Scopus и (или) Web of Science, из них две статьи в журналах квартиля Q1; 6 статей в сборниках трудов международных и российских конференций.

Структура и объем диссертации. Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы из 101 наименования. Работа содержит 127 страницы основного текста, включая 33 рисунка и 29 таблиц.

Во введении обоснована актуальность темы диссертации, сформулирована цель исследований, определены решаемые задачи, указаны научная новизна и практическая ценность результатов работы.

В первой главе введены необходимые термины и обозначения, в частности понятия ранжирования (слабого порядка) как формы представления предпочтений; проведен анализ основных правил выбора ранжирования консенсуса для заданных m ранжирований n альтернатив; подробно рассмотрена процедура агрегирования предпочтений в форме правила Кемени, а также свойства матрицы входного профиля предпочтения; дано описание рекурсивного алгоритма ветвей и границ для точного решения задачи о ранжировании Кемени.

Во второй главе рассматриваются причины явления множественности. Обсуждаются результаты экспериментальных численных исследований транзитивности / нетранзитивности входного профиля предпочтения $\Lambda(m, n)$ в зависимости от числа ранжирований (избирателей), числа альтернатив (кандидатов) и плотности толерантностей. Рассматриваются случаи нетранзитивности выходного профиля оптимальных решений $B(N, n)$.

В третьей главе детально рассмотрен выходной профиль ЗРК, исследуются свойства взвешенной турнирной матрицы, построенной для выходного профиля; а также соответствующие понятия выигрыша, проигрыша и ранга. Формулируются правила преобразования выходного профиля в единственное оптимальное решение, т.е. правила свёртки. Обсуждаются справедливость и обоснованность предлагаемых правил свёртки на основе счета Борда и понятия отношения "между", заданного на ранжированиях.

В четвертой главе рассмотрены две основные парадигмы применения правила Кемени: (1) измерение в многомерной порядковой шкале и (2) комплексирование интервалов агрегированием предпочтений. Основное внимание уделено рассмотрению процедуры комплексирования интервальных данных агрегированием предпочтений (IF&PA), которая оперирует "инранжированиями" – слабыми порядками, наведенными входными интервалами. Показано, что мощность пространства инранжирований определяется треугольным числом для заданного числа n дискретных элементов. Исследованы мощности смежных подпространств слабых порядков.

ГЛАВА 1

АЛГОРИТМ ТОЧНОГО РЕШЕНИЯ ЗАДАЧИ О РАНЖИРОВАНИИ КЕМЕНИ

В этой главе введены необходимые термины и обозначения, в частности понятия *ранжирования* (слабого порядка) как формы представления предпочтений; проведен анализ основных правил выбора ранжирования консенсуса для заданных m ранжирований n альтернатив; подробно рассмотрена процедура агрегирования предпочтений в форме правила Кемени, а также свойства матрицы входного профиля предпочтения; дано описание рекурсивного алгоритма ветвей и границ для точного решения задачи о ранжировании Кемени.

1.1 Агрегирование предпочтений

Определение 1.1. *Агрегирование предпочтений* (или рангов) – это классическая задача определения *ранжирования консенсуса* β для заданного профиля предпочтений, состоящего из m ранжирований, возможно включающих толерантности, n альтернатив (или кандидатов).

Другими словами, проблема агрегирования состоит в представлении индивидуальных предпочтений, определенных на множестве кандидатов (альтернатив), коллективным (групповым) ранжированием (или *ранжированием консенсуса*) этих кандидатов [35,46].

Проблему агрегирования предпочтений по праву можно назвать классической проблемой *теории социального выбора*, т.к. с конца XVIII века тщательным образом она исследуется в основном в рамках *теории голосования*. Каким образом может быть осуществлен наиболее справедливый выбор победителей из n кандидатов группой, состоящей из m избирателей, предложил Ж.-К. де Борда, Н. де Кондорсе, К.Л. Доджсон и многие другие ученые [49,59]. В XX веке для теории социального выбора К. Эрроу, Н.Р. Янг, Р.С. Фишберн и др. [30,47,59,99,100] заложили аксиоматические основы, формализующие требования к ранжированию

консенсуса.

Понятие величины, сформулированное аксиоматически, и понятийный аппарат бинарных отношений и отображений сделали возможным построение *репрезентационной теории измерений*, созданной в середине XX века усилиями П. Суппеса, Р.Д. Льюса, Д.Г. Кранца и др. [1,20,45,64,72,75,82,83,88]. Можно спорить о практических преимуществах этой теории [66], но, несомненно, ее идеи способствовали формированию более полной и целостной *науки об измерениях* (*measurement science*), охватывающей большинство классов измерений во многих прикладных областях – от физики и техники до экономики и психологии.

Определение 1.1 включает в себя многие понятия, определения которым и соответствующие обозначения будут даны ниже.

1.2 Ранжирование и профиль предпочтений

Понятие *ранжирования* можно интерпретировать как процедуру или процесс расстановки нескольких альтернатив (объектов, кандидатов, ...) по степени их важности в контексте поставленной задачи выбора.

В диссертационной работе под ранжированием будем понимать результат процесса расстановки n альтернатив. Для формального представления ранжирования воспользуемся аппаратом бинарных отношений.

Пусть дано некоторое множество $A = \{a_1, a_2, \dots, a_n\}$, $n \geq 2$, состоящее из элементов (альтернатив, объектов, ...) a_i , $i = 1, \dots, n$, подлежащих оцениванию по m свойствам (критериям, экспертам...).

Определение 1.2. *Ранжирование* – это бинарное отношение предпочтения λ на множестве A в виде цепочки

$$\lambda = a_2 \succ a_1 \succ \dots \succ a_s \sim a_t \succ \dots \succ a_p \sim a_q, \quad (1.1)$$

где символы $\succ$ и $\sim$ обозначают бинарные отношения строгого порядка и толерантности соответственно.

Заметим, что порядок предпочтения в ранжировании (1.1) всегда ориентирован слева направо, т.е. более предпочтительные альтернативы находятся слева. Благодаря этому, для удобства и краткости обозначений, наряду с представлением (1.1), будем использовать следующее очевидное компактное обозначение для ранжирования:

$$\lambda = a_2 \succ a_1 \succ \dots \succ a_s \sim a_t \succ \dots \succ a_p \sim a_q \Leftrightarrow 21 \dots (st) \dots (pq), \quad (1.2)$$

где элементы ранжирования обозначены их индексами, а толерантные элементы показаны путем заключения в скобки соответствующих им подмножеств индексов. Тогда в сокращенной записи символы $\succ$ и $\sim$ могут быть опущены.

Пример 1.1. Следуя выражению (1.2), в этих конкретных случаях краткие обозначения получаются следующим образом:

$$a_2 \succ a_5 \sim a_1 \succ a_3 \sim a_4 \succ a_6 \Leftrightarrow 2(51)(34)6 \text{ или}$$

$$a_3 \succ a_1 \succ a_5 \succ a_2 \succ a_6 \succ a_4 \Leftrightarrow 315264. \blacklozenge$$

1.2.1 Слабый порядок

Ранжирование можно определить в терминах бинарных отношений как *отношение слабого порядка*. Для того, чтобы определить отношение слабого порядка, воспользуемся известными свойствами бинарных отношений [16], приведенными в таблице 1.1:

Таблица 1.1 – Основные свойства бинарных отношений на множестве A

Обозначение	Свойство	Формальная запись для $i, j, k = 1, \dots, n$
P1	Рефлексивность	$(a_i, a_i) \in \lambda$, для всех i
P2	Антирефлексивность	$(a_i, a_i) \notin \lambda$, для всех i
P3	Симметричность	$(a_i, a_j) \in \lambda \Rightarrow (a_j, a_i) \in \lambda$, для всех i, j
P4	Антисимметричность	$(a_i, a_j) \in \lambda \wedge (a_j, a_i) \in \lambda \Rightarrow a_i = a_j$
P5	Асимметричность	$(a_i, a_j) \in \lambda \Rightarrow (a_j, a_i) \notin \lambda$
P6	Транзитивность	$(a_i, a_j) \in \lambda \wedge (a_j, a_k) \in \lambda \Rightarrow (a_i, a_k) \in \lambda$
P7	Полнота	$(a_i, a_j) \in \lambda \vee (a_j, a_i) \in \lambda, i \neq j$

Свойства, представленные в таблице 1.1, в определенной комбинации приписанные конкретному отношению λ , позволяют отнести его к одному из пяти хоро-

шо известных *типов* отношений (выделенных жирным шрифтом на рисунке 1.1). Основные типы отношений, порождаемые комбинациями свойств бинарных отношений, изображены вертикальными ориентированными влево и вправо дугами [69] на рисунке 1.1.

Пример 1.2. Из рисунка 1.1 видно, что симметричное и рефлексивное бинарное отношение является отношением *толерантности*. Сочетание таких свойств как симметричность, рефлексивность и транзитивность порождает *эквивалентность*, и т.п. ♦



Рисунок 1.1 – Основные типы бинарных отношений и их свойства

В свою очередь, используя некоторые из упомянутых типов бинарных отношений, можно моделировать различные практически полезные случаи типов отношений [39]. В частности, таким путем можно смоделировать ранжирование λ как объединение

$$\lambda = \rho \cup \tau \quad (1.3)$$

двух отношений: строгого порядка ρ (которое является асимметричным и транзитивным отношением), т.е. $a_i > a_j$, и толерантности τ (которое является симметричным и рефлексивным отношением, называемым также "безразличием" или "ничьей"), т.е. $a_i \sim a_j$.

Тогда отношение λ вида (1.3) задает бинарное *отношение слабого порядка* (или *отношение предпочтения*, или общий предпорядок) на множестве A . Легко показать, что $\rho = \overline{\lambda^{-1}}$ и $\tau = \lambda \cap \lambda^{-1} = \rho \cup \rho^{-1}$ [16].

Говорят, что толерантность τ образует *симметричную часть* слабого по-

рядка λ , а строгий порядок ρ составляет его *асимметричную часть*. Слабый порядок обычно рассматривается как *полное, рефлексивное и транзитивное отношение*.

С учетом выражения (1.3), слабый порядок λ можно представить $(n \times n)$ **матрицей отношения** $[l_{ij}] = L$, строки и столбцы которой соответствуют альтернативам из A , а элементы l_{ij} определяются как

$$l_{ij} = \begin{cases} 1, & \text{если } (a_i \succ a_j) \vee (a_i \sim a_j) \\ 0, & \text{если } a_i \prec a_j \end{cases} = \begin{cases} 1, & \text{если } (a_i, a_j) \in \lambda \\ 0, & \text{если } (a_i, a_j) \notin \lambda \end{cases}. \quad (1.4)$$

Пример 1.3. Для ранжирования $a_2 \succ a_5 \sim a_1 \succ a_3 \sim a_4$ матрица L имеет следующий вид:

$$[l_{ij}] = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}. \quad \blacklozenge \quad (1.5)$$

Если требуется различать все три возможных варианта проявления слабого порядка на паре альтернатив, а именно $\succ$, $\prec$ и $\sim$, то отношение предпочтения λ можно представлять $(n \times n)$ **матрицей отношения** $R = [r_{ij}]$, строки и столбцы которой соответствуют альтернативам из A , а элементы r_{ij} определяются как

$$r_{ij} = \begin{cases} 1, & \text{если } a_i \succ a_j \\ 0, & \text{если } a_i \sim a_j \\ -1, & \text{если } a_i \prec a_j \end{cases}. \quad (1.6)$$

Заметим, что матрица R является **антисимметрической**, т.е. $r_{ij} = -r_{ji}$ для всех i, j , и **транзитивной**, т.к., в силу транзитивности отношения λ , если $r_{ij} \geq 0$ и $r_{jk} \geq 0$, то $r_{ik} \geq 0$ [44].

Пример 1.4. Для ранжирования из Примера 1.3 матрица R имеет следующий вид:

$$[r_{ij}] = \begin{bmatrix} 0 & -1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & -1 \\ -1 & -1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 1 & 0 \end{bmatrix} . \blacklozenge \quad (1.7)$$

Часто бывает удобно или необходимо вместо представления (1.6) для слабого порядка λ использовать *кусочно-постоянную функцию* $\text{sgn}()$, которая отображает знак (или ориентацию) пары $(a_i, a_j) \in \lambda$:

$$\text{sgn}(a_i, a_j) = \begin{cases} 1, & a_i \succ a_j \\ 0, & a_i \sim a_j \\ -1, & a_i \prec a_j \end{cases} . \quad (1.8)$$

1.2.2. Профиль предпочтения

Пусть имеется множество $\Lambda(m, n) = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$, состоящее из m ранжирований n альтернатив, определенных на множестве $A = \{a_1, a_2, \dots, a_n\}$.

Определение 1.3. *Профилем предпочтения* называется множество $\Lambda(m, n)$ для заданных m ранжирований

$$\lambda_k = (a_1 \succ_k a_2 \succ_k \dots \sim_k a_s \sim_k a_t \succ_k \dots \sim_k a_n), k = 1, \dots, m. \quad (1.9)$$

В дальнейшем для краткости вместо обозначения профиля $\Lambda(m, n)$ часто будем использовать обозначение Λ .

Ранжирование $\lambda_k \in \Lambda$ может быть задано указанием *некоторого признака* или свойства k из списка $K = \{1, 2, \dots, m\}$, по которому каждая пара элементов из A находится друг с другом в отношении строгого предпочтения ρ , т.е. $a_i \succ_k a_j$, или в отношении толерантности τ , т.е. $a_i \sim_k a_j$.

В другой эквивалентной интерпретации под множеством K можно понимать *группу индивидуумов* (агентов, экспертов, избирателей и т.п.), выражающих свое мнение путем формирования отношения строгого предпочтения или толерантности для каждой пары $(a_i, a_j)_k$.

Рассмотрим упрощенный пример формирования профиля для одной из возможных практических задач агрегирования предпочтений.

Пример 1.5. Предположим, необходимо оценить публикационную активность трех кафедр a_1 , a_2 и a_3 , т.е. $n = 3$, одного из факультетов университета. Пусть многомерное свойство "публикационная активность" оценивается по $m = 3$ признакам из списка K : количество опубликованных журнальных статей (1), количество изданных монографий (2) и число цитирований публикаций (3). Каждая кафедра за определенный период времени достигла значений признаков, приведенных в таблице 1.2.

Таблица 1.2 – Список K признаков публикационной активности и их значений для каждой кафедры $a \in A$

Кафедра	Публикационная активность		
	Статьи, шт.	Монографии, шт.	Цитирования, шт.
	Номер признака k		
	1	2	3
a_1	35	2	28
a_2	30	2	40
a_3	35	4	40

Данные из таблицы 1.2 позволяют проранжировать три кафедры по каждому из трех признаков из списка $K = \{1, 2, 3\}$, и получить следующий профиль предпочтения:

$$\lambda_1: a_1 \sim a_3 \succ a_2$$

$$\lambda_2: a_3 \succ a_1 \sim a_2$$

$$\lambda_3: a_2 \sim a_3 \succ a_1.$$

Найденное для этого профиля ранжирование консенсуса β (см. Определение 1.1) покажет, какое место каждая кафедра занимает в общем рейтинге с точки зрения публикационной активности. ♦

Введем понятие ранга альтернативы в ранжировании.

Определение 1.4. Рангом r_i^k альтернативы a_i называется позиция (номер места) альтернативы a_i в ранжировании $\lambda_k \in \Lambda(m, n)$, $k = 1, \dots, m$.

Введем понятие общего ранга альтернативы во всем профиле предпочтения.

Определение 1.5. Общим рангом r_i альтернативы a_i в профиле предпочтения $\Lambda(m, n)$ называется сумма ее рангов во всех ранжированиях этого профиля, т.е.

$$r_i = \sum_{k=1}^m r_i^k. \quad (1.10)$$

Ясно, что любой профиль $\Lambda(m, n)$ может быть однозначно представлен *таблицей рангов*.

Пример 1.6. Пусть дан следующий профиль предпочтения $\Lambda(5, 6)$ для $m = 5$ и $n = 6$:

$$\begin{aligned} \lambda_1: a_1 \succ a_2 \sim a_5 \succ a_6 \succ a_4 \succ a_3 \\ \lambda_2: a_5 \succ a_6 \succ a_2 \succ a_3 \sim a_4 \succ a_1 \\ \lambda_3: a_3 \succ a_1 \succ a_6 \sim a_4 \succ a_2 \succ a_5. \\ \lambda_4: a_3 \succ a_6 \succ a_1 \succ a_2 \sim a_4 \succ a_5 \\ \lambda_5: a_6 \succ a_1 \succ a_2 \sim a_3 \succ a_4 \succ a_5 \end{aligned} \quad (1.11)$$

Короткая форма этого профиля имеет вид:

$$\begin{aligned} &1 \ (2 \ 5) \ 6 \ 4 \ 3 \\ &5 \ 6 \ 2 \ (3 \ 4) \ 1 \\ &3 \ 1 \ (6 \ 4) \ 2 \ 5. \\ &3 \ 6 \ 1 \ (2 \ 4) \ 5 \\ &6 \ 1 \ (2 \ 3) \ 4 \ 5 \end{aligned} \quad (1.12)$$

Ранги альтернатив для каждого ранжирования этого профиля, полученные в соответствии с определением 1.9, представлены в таблице 1.3. В нижней строке таблицы приведены общие ранги для каждой альтернативы.

Таблица 1.3 – Таблица рангов профиля $\Lambda(m, n) = \Lambda(5, 6)$ для примера 1.6

r_i^k	a_1	a_2	a_3	a_4	a_5	a_6
r_i^1	1	2	5	4	2	3
r_i^2	5	3	4	4	1	2
r_i^3	2	4	1	3	5	3
r_i^4	3	4	1	4	5	2
r_i^5	2	3	3	4	5	1
$r_i = \sum_{k=1}^m r_i^k$	13	16	14	19	18	11

1.2.3 Взвешенная турнирная матрица

Каждое ранжирование в профиле предпочтения $\Lambda(m, n)$ можно рассматривать как *круговой (round-robin) турнир*, который представляет собой игру (соревнование), в которой каждый игрок по очереди встречается с каждым другим игроком [35,46]. В контексте данной работы это процедура *парного сравнения альтернатив*, в которой каждая альтернатива сравнивается по очереди с каждой другой альтернативой.

Если игроки встречаются m раз, то турнир называется m -взвешенным или просто взвешенным. Тогда профиль $\Lambda(m, n)$ можно представить $(n \times n)$ взвешенной турнирной матрицей $S = [s_{ij}]$, где

$$s_{ij} = \sum_{k=1}^m \mathbf{I}_k(a_i, a_j), i \neq j = 1, \dots, n, \quad (1.13)$$

— элемент матрицы и

$$\mathbf{I}_k(a_i, a_j) = \begin{cases} 1 & \text{если } a_i \succ_k a_j \\ 0,5 & \text{если } a_i \sim_k a_j \\ 0 & \text{если } a_i \prec_k a_j \end{cases} \quad (1.14)$$

— индикаторная функция.

Очевидно, для турнирной матрицы S выполняется следующее свойство:

$$s_{ij} + s_{ji} = m \text{ для } i \neq j; i, j = 1, \dots, n, \quad (1.15)$$

где значение s_{ij} равно доле значения m , которую альтернатива a_i выигрывает (приобретает) в результате сравнения с альтернативой a_j . Таким образом, взвешенная турнирная матрица S содержит информацию о том, в какой степени альтернатива a доминирует любую другую альтернативу после проведения m турниров (сравнений) $\lambda_k, k = 1, \dots, m$ [39].

Пример 1.7. Турнирная матрица для профиля (1.7) из примера 1.6 приведена в таблице 1.4. ♦

Таблица 1.4 – Матрица турнира для профиля $\Lambda(5, 6)$ из примера 1.6.

	a_1	a_2	a_3	a_4	a_5	a_6	z_i
a_1	0	4	2	4	4	2	16
a_2	1	0	2,5	3,5	3,5	1	11,5
a_3	3	2,5	0	3,5	3	2	14
a_4	1	1,5	1,5	0	3	0,5	7,5
a_5	1	1,5	2	2	0	2	8,5
a_6	3	4	3	4,5	3	0	17,5

1.3 Правила агрегирования предпочтений

Цель агрегирования предпочтений состоит в определении *ранжирования консенсуса* β , которое является единственным отношением предпочтения, дающим интегральную характеристику профиля $\Lambda(m, n)$.

Математический подход к проблеме агрегирования предпочтений и нахождения отношения консенсуса был впервые исследован в рамках теории социального выбора в качестве проблемы голосования или группового решения, в которой роль множества A выполняет множество n кандидатов (альтернатив), ранжируемых группой из m избирателей [38,49,59,86]. Задача голосования состоит в определении такого кандидата из n кандидатов, которому m избирателей оказывают наибольшую степень поддержки. Существует много правил определения ранжирования консенсуса. Подробный формальный анализ многих из них дан в работах [15,30,44,57,95,100]. В данной работе рассмотрим только четыре из правил голосования, являющихся наиболее популярными и, по-видимому, наиболее изученными: правило простого большинства, правило Борда, правило Кондорсе и правило Кемени.

1.3.1 Правило простого большинства

Древний принцип простого большинства (*simple majority* или *plurality*) гласит, что победителем голосования является тот кандидат, который *большинством*

избирателей поставлен на *первое место*. Дадим формальное определение этому правилу в терминах рангов.

Пусть $w(a_i)$ обозначает число случаев, когда альтернатива a_i заняла первое место в k -ом ранжировании профиля $\Lambda(m, n)$, т.е.

$$w(a_i) = |\{k \mid r_i^k = 1, k = 1, \dots, m\}|. \quad (1.16)$$

Определение 1.6. Альтернатива a_i называется *победителем* $SM(a_i)$ по правилу простого большинства, который определяется выражением:

$$SM(a_i) = \{a_i \mid \max[w(a_i)], i = 1, \dots, n\}. \quad (1.17)$$

Заметим, что определение 1.6, по сути, описывает процедуру *относительного простого большинства*. Существует также вариант, который называется *абсолютным простым большинством*. Для него требуется, чтобы победитель набрал более 50 % голосов. Это значит, что вместо (1.17) используется выражение

$$SM(a_i) = \{a_i \mid w(a_i) > m/2, i = 1, \dots, n\}. \quad (1.18)$$

Применим правило простого большинства к профилю (1.7) из примера 1.6. Альтернативы получили следующие голоса за первое место: a_1 получает 1 голос, a_2 – 0 голосов, a_3 – 2 голоса, a_4 – 0 голосов, a_5 – 1 голос, a_6 – 1 голос. Следовательно, победителем по этому правилу является альтернатива a_3 .

Однако такое решение связано с противоречием, так как из профиля (1.7) видно, что три избирателя из пяти предпочли кандидата a_6 кандидату a_3 , следовательно, кандидат a_6 должен выиграть у кандидата a_3 хотя бы во втором туре выборов.

При таком способе голосования остается также неясным, кто должен занять *вторую позицию* в результатах голосования: кандидат a_1 занял один раз первое место и два раза – второе. Кандидат a_3 – два раза первое и ни одного второго, вроде бы кандидат a_1 должен занять второе место, но из кандидатов a_3 и a_1 тремя избирателями отдано предпочтение кандидату a_3 и лишь два избирателя предпочли кандидата a_1 .

Таким образом, правило простого большинства не является надежным, т.к. далеко не всегда приводит к корректным результатам, если число кандидатов

$n \geq 2$. Очевидно, это связано с недостаточным использованием информации о взаимном доминировании альтернатив.

1.3.2 Правило Борда

В конце XVIII века известный французский исследователь Жан-Шарль де Борда предложил систему голосования для выборов из двух и более кандидатов, получившую известность как "счет Борда". Это правило до сих пор остается одним из самых популярных по частоте применения в избирательных системах всего мира. Согласно этому правилу, результаты голосования выражаются в виде числа баллов, набранных каждым из кандидатов.

Каждой альтернативе присваивается количество баллов в зависимости от ее позиции в ранжировании: 0 баллов за последнее место, 1 балл за предпоследнее место и т.д. до $n - 1$ балл за первое место. Баллы каждой альтернативы суммируются по всем ранжированиям, и ее место в отношении консенсуса β определяется полученной суммой: чем больше сумма, тем лучше (левее) место альтернативы. Дадим формальное определение этому правилу в терминах турнирной матрицы.

Пусть сумма элементов i -й строки турнирной матрицы S профиля $\Lambda(m, n)$ называется *выигрышем* z_i альтернативы a_i для $i = 1, \dots, n$ и определяется как

$$z_i = \sum_{j=1}^n s_{ij}. \quad (1.19)$$

Выигрыш z_i показывает общий доход от всех выигрышей альтернативы a_i во всех $m(n - 1)$ сравнениях. Легко убедиться, что сумма баллов, полученных альтернативой a_i в каждом ранжировании λ_k по правилу Борда, равна сумме z_i элементов соответствующей строки турнирной матрицы S .

Определение 1.7. Альтернатива a_i называется *победителем* $Bo(a_i)$ по правилу Борда, который определяется выражением:

$$Bo(a_i) = \{a_i \mid \max(z_i), i = 1, \dots, n\}. \quad (1.20)$$

Применим правило Борда к профилю (1.7) из примера 1.6. В приведенной в таблице 1.4 строковые суммы z_i турнирной матрицы S этого профиля показаны в

крайнем правом столбце. Из порядка полученных сумм z_i следует, что отношение консенсуса, определяемое правилом Борда, имеет вид: $a_6 \succ a_1 \succ a_3 \succ a_2 \succ a_5 \succ a_4$.

Однако, полученное решение, как и решение по правилу простого большинства, *нельзя признать удовлетворительным*. Действительно, из турнирной матрицы S видно, что три из пяти (т.е. большинство) избирателей считают, что $a_3 \succ a_1$ и $a_4 \succ a_5$, а в полученном отношении консенсуса отражено противоположное мнение.

Правило Борда относится к семейству *позиционных методов голосования*, в которых баллы альтернатив рассчитываются на основе суммарного учета их позиций в предпочтениях отдельных избирателей. Это правило не всегда приводит к наилучшему отношению консенсуса [57,65], но часто используется в приложениях агрегирования предпочтений, благодаря простоте необходимых расчетов и отсутствию стратегии слишком явной поддержки большинства при голосовании. Однако известно, что это правило подвержено манипулятивным действиям [98]. Победитель по правилу Борда с высокой вероятностью может оказаться победителем по правилу Кондорсе, когда он существует [30]. Обсуждение правила Борда будет продолжено в разделе 3.2.

1.3.3 Правило Кондорсе

Пытаясь преодолеть недостатки правила Борда, Николя де Карита, маркиз де Кондорсе в 1785 г. предложил свое правило голосования: в каждом парном сравнении кандидатов предпочтительным кандидатом является тот, что предпочитается большинством избирателей, т.е. $a_i \succ a_j$, если и только если $s_{ij} > s_{ji}$ для всех $i \neq j$.

Иными словами, если альтернатива a_i получила большинство голосов в парном сравнении со всеми другими альтернативами, эта альтернатива называется

ся победителем (альтернативой) Кондорсе. Дадим формальное определение правилу Кондорсе в терминах турнирной матрицы.

Определение 1.8. Альтернатива a_i называется победителем Кондорсе $Co(a_i)$, если она определяется выражением:

$$Co(a_i) = \{a_j \mid s_{ij} > s_{ji}, i \neq j, i, j = 1, \dots, n\}. \quad (1.21)$$

Заметим, что, с учетом свойства (1.11), выражение (1.17) может быть переписано в виде

$$Co(a_i) = \{a_j \mid s_{ij} > m/2, i \neq j, i, j = 1, \dots, n\}. \quad (1.22)$$

Применительно к рассматриваемому профилю (1.6), правило Кондорсе дает следующее ранжирование консенсуса: $a_6 \succ a_3 \succ a_1 \succ a_2 \succ a_4 \succ a_5$.

Действительно, турнирная матрица S (см. таблицу 1.4) позволяет построить таблицу 1.5, в колонках которой, обозначенных символами $\succ$, $\prec$ и $\sim$, содержатся соответственно числа побед (т.е. когда $s_{ij} > m/2 = 2,5$ для фиксированного индекса i), поражений (т.е. когда $s_{ij} < 2,5$ для фиксированного индекса i) и ничьих (т.е. когда $s_{ij} = 2,5$ для фиксированного индекса i) для каждой из альтернатив a_i . Например, в строке a_1 матрицы S мы видим следующие значения s_{1j} : 0, 4, 2, 4, 4, 2, — откуда ясно, что число побед альтернативы a_1 во всех парных сравнениях равно 3 (это случаи, когда $s_{12} = s_{14} = s_{15} = 4 > 2,5$), число поражений равно 2 (это случаи, когда $s_{13} = s_{16} = 2 < 2,5$), а число ничьих равно 0. Поэтому строка a_1 таблицы 1.5 заполнена значениями: 3, 2, 0. Аналогичным образом заполнены остальные строки таблицы 1.5.

Таблица 1.5 – Числа итогов парных сравнений альтернатив из профиля (1.6)

a_i	Число итогов парных сравнений альтернативы a_i		
	Победы ($\succ$)	Поражения ($\prec$)	Ничьи ($\sim$)
a_1	3	2	0
a_2	2	2	1
a_3	3	1	1
a_4	1	4	0
a_5	0	5	0
a_6	5	0	0

Из таблицы 1.5 видно, что явным лидером по числу побед является альтернатива a_6 , на второе место претендуют две альтернативы a_1 и a_3 , как имеющие по 3 победы, но a_1 имеет два поражения, а a_3 только одно поражение и одну ничью. Поэтому второе место занимает альтернатива a_3 . Дальнейший анализ доказывает корректность приведенного выше ранжирования консенсуса.

Правило Кондорсе совершенно обоснованно могло бы претендовать на роль наилучшего правила формирования отношения консенсуса, т.к. оно опирается на полную информацию о парных сравнениях альтернатив, содержащуюся в турнирной матрице S .

Однако, к сожалению, правило Кондорсе страдает существенным недостатком: *победитель Кондорсе может не существовать*.

Сам Кондорсе в своей работе 1785 г. [55] указал на возможность *нетранзитивности* определяемого его правилом ранжирования консенсуса β . Пусть исходный профиль $\Lambda(m, n)$ содержит только строгие отношения предпочтения.

Определение 1.9. Ситуация, когда построение ранжирования консенсуса β в соответствии с выражением (1.21) приводит к одновременному выполнению трех отношений строгого предпочтения

$$(a_i \succ a_j) \in \beta, (a_j \succ a_k) \in \beta, (a_k \succ a_i) \in \beta, \quad (1.23)$$

называется *парадоксом Кондорсе*.

В этом случае победителя Кондорсе не существует [30,55,99,100]. Таким образом несмотря на то, что каждое из ранжирований, входящих в профиль предпочтения $\Lambda(m, n)$, всегда является транзитивным, нет никакой гарантии, что ранжирование консенсуса β будет также транзитивным.

Пример 1.8. Простейшим и стандартным примером парадокса Кондорсе является следующий профиль $\Lambda(3, 3) = \{1\ 2\ 3, 2\ 3\ 1, 3\ 1\ 2\}$. Его турнирная матрица имеет вид:

$$S = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \end{bmatrix}. \quad (1.24)$$

Пытаясь применить к профилю принцип Кондорсе, на основании данных матрицы S легко убедиться, что каждая из альтернатив может претендовать на роль победителя. В результате получается так называемая циклическая неопределенность (*цикл* Кондорсе), которую нередко рекомендуют разрешать, рассматривая её как своего рода ничью (отношение толерантности) между всеми тремя альтернативами a_1 , a_2 и a_3 , являющимися участниками этого парадокса. То есть ранжирование консенсуса β в этом случае может иметь вид: $a_1 \sim a_2 \sim a_3$. ♦

Обсуждение правила и парадокса Кондорсе будет продолжено в главе 2. Отметим, что одним из самых разумных способов избежать парадокса Кондорсе является использование правила голосования, предложенного Дж. Кемени.

1.3.4 Правило Кемени

Правило Кемени [14] относится к семейству кондорсетовских методов голосования, т.е. сохраняет положительные свойства правила Кондорсе. В то же время это правило не приводит к парадоксу Кондорсе. Поэтому оно считается одним из наиболее приемлемых правил агрегирования предпочтений [31,56].

Дж. Кемени в работе [14] ввел понятие расстояния между ранжированиями в терминах *числа парных несоответствий* (противоположных порядков альтернатив или инверсий), имеющих в этих ранжированиях.

Определение 1.10. *Расстоянием Кемени* называется расстояние между двумя ранжированиями λ_k и λ_l , определяемое выражением

$$d(\lambda_k, \lambda_l) = \sum_{i < j} |r_{ij}^k - r_{ij}^l| = \sum_{i < j} |\text{sgn}(a_i^k, a_j^k) - \text{sgn}(a_i^l, a_j^l)|, \quad (1.25)$$

где r_{ij}^k и r_{ij}^l – элементы матриц отношений вида (1.6) ранжирований λ_k и λ_l соответственно, а $\text{sgn}(a_i^k, a_j^k)$ и $\text{sgn}(a_i^l, a_j^l)$ – соответствующие знаковые функции вида (1.8). В формуле (1.25) суммируются только элементы верхней треугольной подматрицы r_{ij} , $i < j$, матрицы R .

Расстояние между произвольным ранжированием λ и профилем $\Lambda(m, n)$ можно теперь определить следующим образом:

$$D(\lambda, \Lambda) = \sum_{k=1}^m d(\lambda, \lambda_k) = \sum_{i < j} \sum_{k=1}^m |\operatorname{sgn}(a_i^k, a_j^k) - \operatorname{sgn}(a_i, a_j)|. \quad (1.26)$$

Расстояние Кемени удовлетворяет обычным метрическим аксиомам [55,56]:

$$\begin{aligned} (1) \quad & d(\lambda_k, \lambda_l) = d(\lambda_l, \lambda_k); \\ (2) \quad & d(\lambda_k, \lambda_l) \geq 0 \text{ и } d(\lambda_k, \lambda_l) = 0 \Leftrightarrow \lambda_k = \lambda_l; \\ (3) \quad & d(\lambda_k, \lambda_l) \leq d(\lambda_k, \lambda) + d(\lambda, \lambda_l). \end{aligned} \quad (1.27)$$

Пространство слабых порядков. Расстояние Кемени наводит определенную структуру на множество всех слабых порядков Ω_0 (см. п. 4.3.1), которая позволяет рассматривать *пространства слабых порядков* для различных n , где ранжирования являются точками (узлами) пространств. На рисунке 1.2 показаны все такие пространства для $n = 1, 2, 3$ и 4 . К сожалению, для $n \geq 5$ графические изображения этих пространств теряют свою наглядность.

Мощность $|\Omega_0|$ пространства слабых порядков (ранжирований) связана с числами Стирлинга 2-го рода $S_{n,q}$, которые определяют *количество неупорядоченных разбиений* n -элементного множества на q непустых подмножеств и определяется по формуле [2]:

$$|\Omega_0| = \sum_{q=0}^n q! S_{n,q}.$$

Наряду с турнирной таблицей профиль предпочтения $\Lambda(m, n)$ можно представлять $(n \times n)$ *матрицей профиля* $P = [p_{ij}]$, строки и столбцы которой соответствуют номерам (индексам) альтернатив из множества A , а элементы определяются как

$$p_{ij} = \sum_{k=1}^m [1 - \operatorname{sgn}(a_i^k, a_j^k)], \quad i, j = 1, \dots, n \quad (1.28)$$

Очевидно, для матрицы профиля P выполняется аналогичное свойству (1.11) турнирной матрицы S свойство, но только для удвоенного m :

$$p_{ij} + p_{ji} = 2m; \quad i, j = 1, \dots, n. \quad (1.29)$$

Пример 1.9. Матрица профиля для профиля предпочтения (1.7) выглядит следующим образом:

$$[p_{ij}] = \begin{bmatrix} 0 & 2 & 6 & 2 & 2 & 6 \\ 8 & 0 & 5 & 3 & 3 & 8 \\ 4 & 5 & 0 & 3 & 4 & 6 \\ 8 & 7 & 7 & 0 & 4 & 9 \\ 8 & 7 & 6 & 6 & 0 & 6 \\ 4 & 2 & 4 & 1 & 4 & 0 \end{bmatrix} \cdot \blacklozenge \quad (1.30)$$

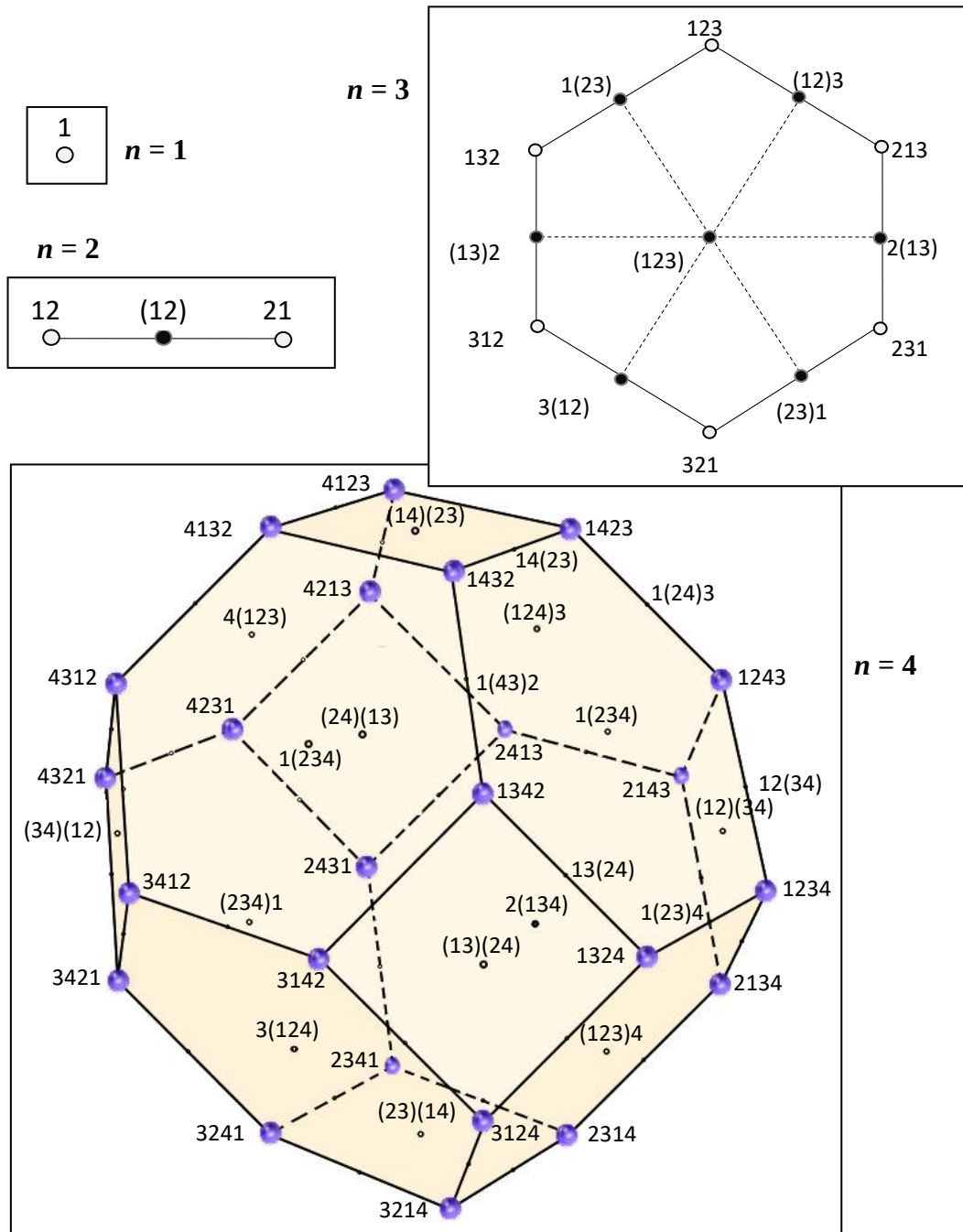


Рисунок 1.2 – Пространства всех слабых порядков для $n = 1, 2, 3$ и 4 .
 Дадим теперь определение правилу Кемени. Пусть Π_n является множеством

всех $n!$ отношений строгого порядка $\succ$ на множестве альтернатив A , тогда каждый строгий порядок $\rho \in \Pi_n$ находится во взаимно однозначном соответствии с перестановкой первых n натуральных чисел $\mathbb{N}_n = \{1, 2, \dots, n\}$; в этом смысле мы будем рассматривать понятия строгого (или линейного) порядка и перестановки как синонимы.

Определение 1.11. *Правило Кемени* заключается в нахождении такого линейного порядка (*ранжирования Кемени*) $\beta \in \Pi_n$ альтернатив из A , что расстояние $D(\beta, \Lambda)$ от β до ранжирований исходного профиля $\Lambda(m, n)$ минимально для всех возможных строгих порядков (перестановок) $\rho \in \Pi_n$, т.е.

$$\beta = \arg \min_{\rho \in \Pi_n} D(\rho, \Lambda) = \arg \min \sum_{i < j} p_{ij}. \quad (1.31)$$

В терминах матрицы профиля P правило Кемени заключается в нахождении такой перестановки строк и столбцов матрицы профиля, что сумма элементов её верхней треугольной подматрицы минимальна.

Правило Кемени (1.31) представляет собой оптимизационную задачу, которую будем называть *задачей о ранжировании Кемени* (ЗРК). Оно является одним из наиболее обоснованных и глубоко исследованных правил агрегирования предпочтений.

1.4 Проблемы практического применения правила Кемени

При всех достоинствах правила Кемени, на пути его широкого практического применения для агрегирования предпочтений существуют два серьезных препятствия:

- *во-первых*, доказано, что задача о ранжировании Кемени является *NP-трудной*, и
- *во-вторых*, для заданного исходного профиля предпочтения правило Кемени может приводить к множественным ранжированиям консенсуса вместо одного.

Рассмотрим эти проблемы более подробно.

1.4.1 NP-трудность задачи о ранжировании Кемени

Пространство решений ЗРК велико, и известно, что она является NP-трудной [38,46]. В ряде работ показано, что вряд ли будет существовать алгоритм с полиномиальным временем (пропорциональным размерности n задачи) решения для правила Кемени [31,35,46].

Точного решения ЗРК можно достичь только с помощью так называемого полного перебора. Так пространство решений совпадает со множеством строгих порядков мощность которого равна $n!$ перестановок альтернатив, то полный перебор потребует экспоненциального времени для поиска решения. Для примера в таблице 1.6 приведены оценки затрат времени некоторых полиномиальных и экспоненциальных алгоритмов на поиск решений задач размерности n .

Таблица 1.6 – Оценки временной сложности некоторых полиномиальных и экспоненциальных функций

Функция временной сложности	Размерность n задачи					
	10	20	30	40	50	60
n	0,00001 с	0,00002 с	0,00003 с	0,00004 с	0,00005 с	0,00006 с
n^2	0,0001 с	0,0004 с	0,0009 с	0,0016 с	0,0025 с	0,0036 с
n^3	0,001 с	0,008 с	0,027 с	0,064 с	0,125 с	0,216 с
n^5	0,1 с	3,2 с	243 с	1,7 мин	5,2 мин	13,0 мин
2^n	0,001 с	1,0 с	17,9 мин	12,7 дней	35,7 лет	366 столетий
3^n	0,059 с	58 мин	6,5 лет	3855 столетий	$2 \cdot 10^8$ столетий	$1,3 \cdot 10^{13}$ столетий

Ясно, что принадлежность ЗРК классу NP-трудных задач требует разработки различных подходов для нахождения решения за приемлемое время.

Многие исследователи предложили алгоритмы, реализующие *полиномиально-временную аппроксимацию* (*polynomialtime approximation*) решения задачи с определенной гарантией близости к оптимуму, см., например, [3,4,5,6,15,46]. Другой популярный подход – *разрешимость с фиксированными параметрами* (*fixed-parameter tractability*), когда параметризованный алгоритм пытается ограничить любой сверх-полиномиальный рост времени выполнения алгоритма предвари-

тельно выбранными параметрами задачи [37]. Наконец, существует обширная литература, описывающая различные *эвристики* и их комбинации для нахождения ранжирований Кемени [28,31,35,39,46].

В работе Бартелеми [31] продемонстрирована взаимосвязь между медианными линейными порядками и так называемой проблемой минимизации реверсирования дуг. На основе подхода Смита и Пэйна в работе [92] были предложены эвристические методы решения проблемы, связанной с минимизацией реверсирования дуг как для невзвешенных турниров, так и для общего случая (взвешенных турниров). Однако экспериментальным путем доказано, что предложенные алгоритмы вычисляют только часть медианных (средние) порядков, или даже не всегда позволяет найти хотя бы один медианный порядок [40].

Следует отметить, что известны реализующие правило Кемени алгоритмы, которые обеспечивают точное решение за разумное время порядка нескольких секунд при размерности задачи до $n < 30 \dots 50$ [40,46]. Хотя эта размерность считается небольшой, она вполне подходит для практических применений правила агрегирования Кемени (см. главу 4). В частности, во всех реализациях (и соответствующих экспериментах) метода IF&PA, рассматриваемого в п. 4.2.1, размерность задачи составляет $n < 20$ [43,69,71,76,79]. Этой размерности вполне достаточно для достижения целей соответствующих практических задач агрегирования предпочтений.

1.4.2 Множественность решений

Ранжирований Кемени может быть слишком много [28,31,35,71]. Хотя эти оптимальные решения, количество которых будем обозначать N , находятся на минимальном расстоянии от m исходных ранжирований, они могут быть довольно далеки друг от друга. Численные эксперименты показывают, что при определенных значениях m и n , особенно при малых и четных $m \geq 4$, значение N может достигать десятков миллионов [71]. При этом возникает хаосо-подобная ситуация,

когда $N \gg t$, и неопределенность после решения задачи значительно превышает исходную неопределенность.

При этом оптимальные решения могут ранжировать альтернативы по-разному, что приводит к гораздо большей неоднозначности, чем та, которая присуща исходному профилю Λ .

Пример 1.10. Пусть задан следующий профиль: $\Lambda(3,3) = \{(13)2; 3(12); (23)1\}$. Тогда для данной ЗРК будут существовать два решения: $\beta_1 = 312$ и $\beta_2 = 321$, оба решения обладают минимальным расстоянием до профиля $D(\beta_1, \Lambda) = D(\beta_2, \Lambda) = 5$. Рассматриваемый профиль $\Lambda(3,3)$, принадлежащий пространству слабых порядков для $n = 3$, показан на рисунке 1.3, где также приведены две матрицы профиля, строки и столбцы которых соответствуют двум перестановкам альтернатив, задаваемым каждым из оптимальных решений β_1 и β_2 . Суммы элементов верхних треугольных подматриц этих матриц равны 5. ♦

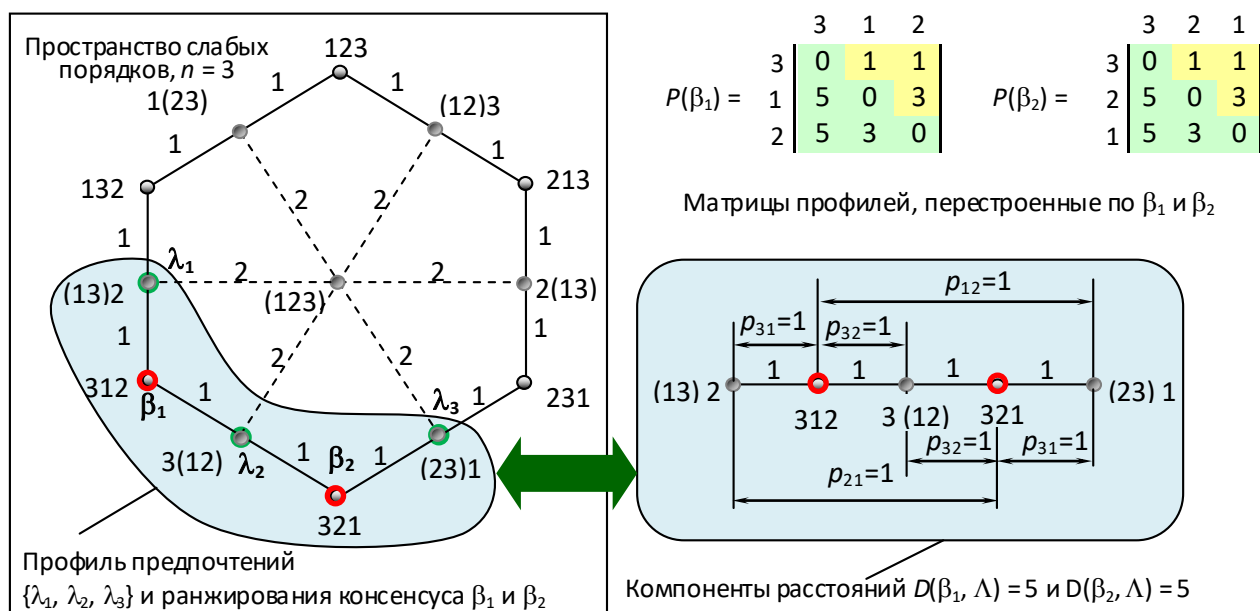


Рисунок 1.3 – Иллюстрация множественности решений ЗРК: профиль предпочтений $\{\lambda_1, \lambda_2, \lambda_3\}$ и отношение консенсуса β_1 и β_2 в пространстве слабого порядка при $n = 3$

Принимая во внимание две рассмотренные выше проблемы, будем в данном диссертационном исследовании придерживаться следующей **стратегии получения надежного отношения консенсуса при применении правила Кемени:**

- (1) всегда находить **полный список всех N ранжирований консенсуса**, которые являются точными оптимальными решениями, достижимыми при приемлемой размерности задачи $n < 20$, и
- (2) находить для этого списка **единственное подходящее итоговое ранжирование консенсуса**.

Заметим, что не каждый из подходов, перечисленных в разделе 1.4.1, позволяет перечислить (найти) *все* возможные ранжирования Кемени (отношения консенсуса). Это возможно сделать только теми методами, которые генерируют частичные решения при построении дерева поиска. Как раз таким методом является *метод ветвей и границ*, компактная рекурсивная версия которого, позволяющая перечислять все ранжирования Кемени, предложена в [77] (см. также следующий параграф) и используется во всех упомянутых выше приложениях правила Кемени.

В цитируемых выше источниках явление множественности ранжирований Кемени лишь обсуждается, но в них отсутствуют систематическим образом изложенные рекомендации о том, как бороться с этим явлением. Поэтому проблема сведения всех отношений консенсуса, определяемых правилом Кемени, к единственному, требует безотлагательного решения, которое будет рассмотрено в главе 3 диссертационной работы.

1.5 Точный алгоритм нахождения всех ранжирований Кемени

Задача о ранжировании Кемени в общем случае может быть решена методом полного перебора [73,80]. На практике метод полного перебора для решения ЗРК не применяется, так как требует слишком больших временных затрат.

Для нахождения ранжирования консенсуса по правилу Кемени можно использовать различные методы решения оптимизационных задач. Известны работы по применению к решению ЗРК методов динамического программирования [85], целочисленного линейного программирования [33], различных поисковых методов [23]. В качестве метода решения ЗРК достаточно хорошо зарекомендовал себя

один из широко применяемых поисковых методов – *метод ветвей и границ*, описанный в работах [17,77].

Метод ветвей и границ является общим алгоритмическим методом, благодаря которому находятся оптимальные решения различных задач. При этом применяются основные две процедуры для поиска решений: ветвление и нахождение оценок. Процедура ветвления состоит в разбиении области допустимых решений на подобласти меньших размеров. Процедура нахождения оценок заключается в поиске верхних и нижних границ для дальнейшего принятия решения. Основной идеей этого подхода является отсечение подмножеств допустимых решений, которые не содержат оптимальные решения.

В данной диссертационной работе для решения ЗРК используется разработанный на основе метода ветвей и границ в научной группе под руководством проф. Муравьева С.В. рекурсивный алгоритм RECURSALL [73]. Ниже представлена формальная запись алгоритма.

RECURSALL: Алгоритм нахождения ранжирования консенсуса по правилу
Кемени

Вход: $(n \times n)$ матрица $[p_{ji}]$ профиля предпочтения $\Lambda(m, n)$

Пусть:

$S = \{s_1, s_2, \dots, s_K\}$: частичное решение (лидер)

$T = \{t_1, t_2, \dots, t_{K-n}\}$: дополнение S до полного порядка

$K = 1, \dots, n - 1$: номер уровня дерева поиска

D_u : верхняя граница

D_{low} : нижняя граница

β_N : N -ое ранжирование консенсуса

1: $D_{least} \leftarrow \sum_{\substack{i=1, \dots, n \\ j=i+1, \dots, n}} \min(p_{ij}, p_{ji}); D_u \leftarrow \infty; S \leftarrow \emptyset; T \leftarrow N_n$ [инициализация]

2: **LEADER** (1, 0); [вызов рекурсивной процедуры]

3: **procedure LEADER** (K, D): [определение рекурсивной процедуры]

4: **for** $l = 1$ **to** $n - K + 1$ **do**

5: $s_K \leftarrow t_l; T \leftarrow T \setminus \{t_l\}$; [ветвление]

6: $D_e \leftarrow D + \sum_{i=1}^{n-K} p_{s_K t_i}$; [изменение расстояния из-за расширения лидера S]

```

7:    $D_{\text{low}} \leftarrow D_e + \sum_{\substack{i=1, \dots, n-K \\ j=i+1, \dots, n-K}} \min(p_{t_{t_j}}, p_{t_{t_i}});$  [изменение нижней границы]
8:   if  $D_{\text{low}} \leq D_u$  then [проверка условия ветвления/отсечения]
9:       if  $K < n - 1$  then  $LEADER(K + 1, D_e);$  [вызов рекурсивной процедуры]
10:      else
11:          if  $D_{\text{low}} = D_u$  then  $N \leftarrow N + 1;$  else  $N \leftarrow 1;$ 
12:           $\beta_N \leftarrow S; D_u \leftarrow D_{\text{low}};$  [запись полного  $N$ -го решения]
13:      end else
14:       $T \leftarrow T \cup \{s_K\}; S \leftarrow S \setminus \{s_K\};$  [отсечение]
      end if
    end if
  end for [завершение рекурсивной функции]

```

Выход: выходной профиль $B(N, n) = \{\beta_1, \beta_2, \dots, \beta_N\} \subset \Pi; D_u$ и D_{least}

Входом алгоритм служит матрица $[p_{ji}]$ профиля предпочтения $\Lambda(m, n)$, определяемая выражением (1.28).

Наименьшее возможное расстояние. Матрицу P можно характеризовать *наименьшим возможным расстоянием* D_{least} от описываемого ею профиля предпочтения Λ до некоторого линейного порядка [17]. В него включаются меньшие значения из каждой пары (p_{ij}, p_{ji}) , то есть

$$D_{\text{least}} = \sum_{i=1}^n \sum_{j=i+1}^n \min(p_{ij}, p_{ji}) = \sum_{i < j} \min(p_{ij}, p_{ji}). \quad (1.32)$$

В монографии Б.Г. Литвака [17] показано, что, если матрица P транзитивна, т.е. $p_{ik} \leq p_{ki}$, если $p_{ij} \leq p_{ji}$ и $p_{jk} \leq p_{kj}$, $i \neq j \neq k = 1, \dots, n$ (это значит, что все исходные ранжирования совместимы), то $D_{\text{least}} = D(\beta, \Lambda)$. Справедливо и обратное утверждение. Очевидно также, что значение D_{least} не зависит от перестановки строк и столбцов матрицы P .

В силу парадокса Кондорсе, равенство $D_{\text{least}} = D(\beta, \Lambda)$ выполняется не всегда, и его справедливость может быть установлена только после нахождения β . Этим фактом можно воспользоваться при организации поиска решения ЗРК методом ветвей и границ.

Дерево решений. Алгоритм исследует древовидную модель пространства решений (рисунок 1.4). Каждой вершине (узлу) дерева решений соответствует множество $S = \{s_1, s_2, \dots, s_K\}$, которое рассматривается как представитель (*лидер*) всех решений, содержащих его в качестве начальной части. Корень дерева является лидером абсолютно всех возможных решений и для него $S = \emptyset$. На следующем (первом) уровне дерева находятся n лидеров с числом элементов, равным 1. Каждый из этих лидеров имеет $n - 1$ последователей мощности 2 на втором уровне. Таким образом, каждый из лидеров K -го уровня имеет $n - K$ последователей мощности $|S| = K + 1$ при $K = 1, \dots, n - 1$. Если $|S| < n - 1$, то соответствующее решение S будем называть текущим *частичным решением*. При $|S| = n - 1$ множество S будет *полным решением*, так как в этом случае оно определяет порядок всех элементов множества A .

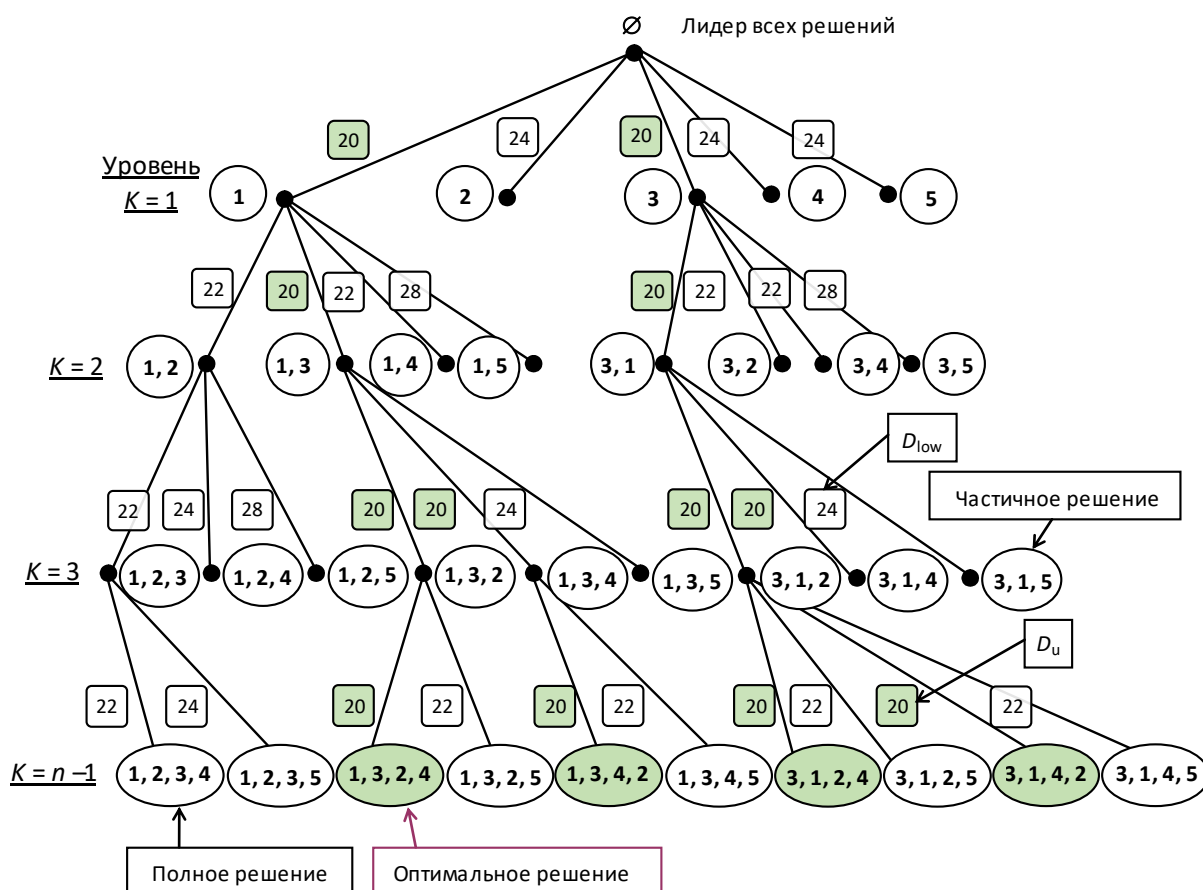


Рисунок 1.4 – Древовидное пространство решений алгоритма RECURSALL для профиля $\Lambda(m, n)$ на рисунке 1.7

Лидер формируется из элементов $\mathbb{N}_n = \{1, 2, \dots, n\}$, следовательно, для всяко-

го лидера S имеется его дополнение $\bar{S} = T = \{t_1, t_2, \dots, t_{K-n}\} = \mathbb{N}_n \setminus S$, содержащее элементы из $\mathbb{N}_n$, не вошедшие в S . Заметим, что элементы в T всегда расположены в порядке возрастания, а частичные решения S порождаются в лексикографическом порядке.

Каждый лидер S образуется путем конкатенации его предшественника с первым по порядку элементом t_l множества T , т.е. $S = \{s_1, \dots, s_{K-1}, s_K = t_l\}$, при этом t_l из T удаляется. Например, пусть $n = 6$, тогда если на уровне $K = 3$ дерева поиска лидер $S = \{2, 1, 3\}$ и $T = \{4, 5, 6\}$, то на следующем четвертом уровне $S = \{2, 1, 3, 4\}$ и $T = \{5, 6\}$.

Проверка лидера на перспективность. Каждому лидеру (частичному решению) соответствует значение оценки D_{low} функции расстояния от профиля Λ до оптимального линейного порядка β . Будем называть эту оценку D_{low} *нижней границей*. *Верхней границей* D_u будем называть наименьшее значение функции расстояния Кемени для полученных к данному моменту полных решений.

Лидер S жестко определяет положение строк и столбцов матрицы P с индексами $s_1, \dots, s_{K-1}$. Соответствующая сумма элементов верхней треугольной подматрицы матрицы P имеет вид

$$D = \sum_{\substack{i=1, \dots, K-1 \\ j=i+1, \dots, K-1}} p_{s_i s_j} + \sum_{i, j=1, \dots, n-K} p_{t_i t_j}. \quad (1.33)$$

Учет расширения лидера за счет конкатенации с элементом t_l дает

$$D_e = D + \sum_{i=1, \dots, n-K} p_{s_K t_i}. \quad (1.34)$$

Наконец, для оставшейся части матрицы, определяемой элементами T , воспользуемся тем же принципом, что и при определении наименьшего возможного расстояния (1.32). Тогда получаем

$$D_{\text{low}} = D_e + \sum_{\substack{i=1, \dots, n-K \\ j=i+1, \dots, n-K}} \min(p_{t_i t_j}, p_{t_j t_i}). \quad (1.35)$$

В приведенном выше примере $D = 51$, $D_e = 51 + 6 = 57$, $D_{\text{low}} = 57 + 4 = 61$ (см. рисунок 1.5).

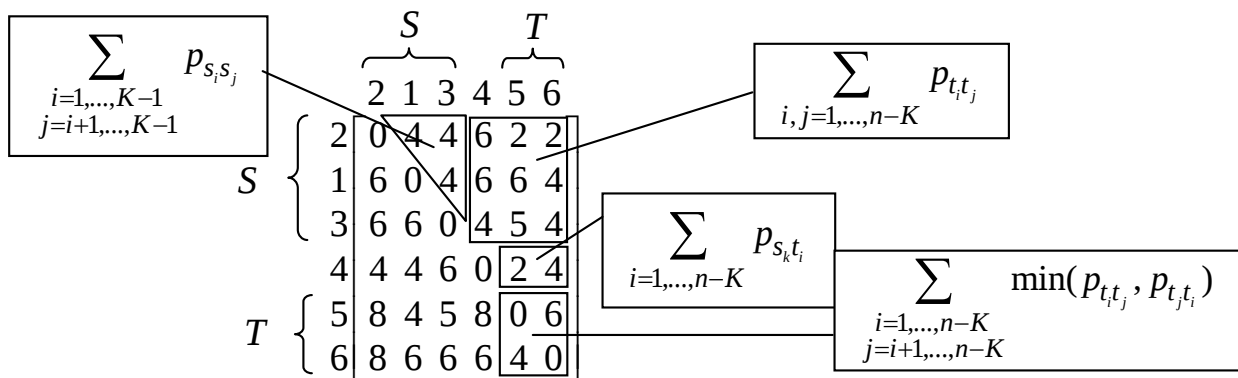


Рисунок 1.5 – К вычислению нижней границы D_{low}

Перспективность лидера определяется соотношением рассчитанной для него нижней границы D_{low} и принятой за верхнюю границу D_u наименьшей из достигнутых к данному моменту нижних границ. Если $D_{low} \geq D_u$ (а ведь, в силу (1.35), D_{low} – это наименьшее возможное значение для данного лидера S), то ясно, что все решения с этим лидером являются бесперспективными, т.е. не могут быть оптимальными.

Алгоритм. Поиск решения в методе ветвей и границ можно организовать двумя способами: итерационным и рекурсивным [21,22,24]. Нами была выбрана рекурсивная организация в силу ее способности очень компактного описания необходимых действий. При этом появления лишних ветвей дерева поиска удалось избежать в силу присущей ЗРК специфики.

Алгоритм RECURSALL содержит всего два этапа:

- *инициализацию* (шаг 1), при которой для данной матрицы P вычисляется D_{least} и все необходимые переменные приобретают исходные значения, и
- *вызов рекурсивной процедуры LEADER* (K, D) (шаг 2), которая принимает два параметра K и D , где K – номер уровня дерева поиска, а D – составляющая нижней границы, определяемая фиксированным порядком элементов лидера, как в формуле (1.27). Исходные значения параметров: $K = 1$ и $D = 0$.

Процедура $LEADER(K, D)$ (шаги 3 - 14) содержит основной цикл **for** (шаг 4) по l , где l – номер лидера на K -ом уровне. На каждом этапе цикла формируется текущее частичное решение в виде лидера S . Новый лидер S всякий раз задает но-

вое положение строк и столбцов матрицы P , и минимально возможное значение D_{low} суммы элементов верхней треугольной подматрицы как функции расстояния. В качестве D_u используется наименьшее значение D_{low} для полученных к данному моменту полных решений (шаг 12).

Если $D_{\text{low}} \leq D_u$ (шаг 8) и текущее решение неполное (т.е. $K < n - 1$), то вызывается процедура *LEADER* (шаг 9) с новыми параметрами, обеспечивающими ветвление дерева поиска. Если $D_{\text{low}} > D_u$, то все решения с текущим лидером не могут быть оптимальными и соответствующая часть дерева поиска не проверяется, т.е. отсекается (шаг 14). Если при $D_{\text{low}} \leq D_u$ текущее решение является полным, т.е. определяющим порядок всех элементов ранжирования, то оно запоминается как пара $\beta_N = S$ и $D_u = D_{\text{low}}$ (шаг 12). Поиск других неполных и полных решений продолжается, пока все неперспективные решения не будут отсечены.

После нахождения первого оптимального решения при $N = 1$ каждое новое найденное оптимальное решение запоминается при соответствующем увеличенном на единицу N (шаг 11). Таким образом, последнее значение N (после останова работы алгоритма) указывает на общее число оптимальных решений.

Результатом работы алгоритма *RECURSALL* является N оптимальных решений ЗРК (ранжирований консенсуса Кемени) $\{\beta_1, \beta_2, \dots, \beta_N\}$; соответствующее этим решениям значение верхней границы функции расстояния D_u и характеризующее матрицу профиля P наименьшее возможное расстояние D_{least} . Пример решений алгоритма *RECURSALL* приведен в таблице 1.7.

Как уже было сказано в параграфе 1.3.2, количество оптимальных решений ЗРК может быть весьма велико даже для небольших m и n .

Определение 1.12. Будем называть множество ранжирований Кемени $B(N, n) = \{\beta_1, \beta_2, \dots, \beta_N\} \subset \Pi$ *выходным профилем* ЗРК. Тогда профиль предпочтения $\Lambda(m, n) = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$ можно называть *входным профилем* ЗРК, см. рисунок 1.6.

Таблица 1.7 – Пример таблицы решений алгоритма RECURSALL для профиля $\Lambda(m, n)$ на рисунке 1.7

Шаг	K	S	T	D_{low}	$\geq D_u$	Решение и комментарии
		{}	1, 2, 3, 4, 5			$D_u = \infty, D_{least} = 20$
1	1	1	2, 3, 4, 5	20	<	ветвление
2	2	1, 2	3, 4, 5	22	<	ветвление
3	3	1, 2, 3	4, 5	22	<	ветвление
4	4	1, 2, 3, 4	5	22	<	полное решение {1, 2, 3, 4, 5}, $D_u = 22$
5	4	1, 2, 3, 5	4	24	<	полное решение {1, 2, 3, 5, 4}, $D_u = 24$
6	3	1, 2, 4	3, 5	24	>	отсечение
7	3	1, 2, 5	3, 4	28	>	отсечение
8	2	1, 3	2, 4, 5	20	=	ветвление
9	3	1, 3, 2	4, 5	20	=	ветвление
10	4	1, 3, 2, 4	5	20	=	полное решение {1, 3, 2, 4, 5}, $D_u = 20$
11	4	1, 3, 2, 5	4	22	>	полное решение {1, 3, 2, 5, 4}, $D_u = 22$
12	3	1, 3, 4	2, 5	20	=	ветвление
13	4	1, 3, 4, 2	5	20	=	полное решение {1, 3, 4, 2, 5}, $D_u = 20$
14	4	1, 3, 4, 5	2	22	>	полное решение {1, 3, 4, 5, 2}, $D_u = 22$
15	3	1, 3, 5	2, 4	24	>	отсечение
16	2	1, 4	2, 3, 5	22	>	отсечение
17	2	1, 5	2, 3, 4	28	>	отсечение
18	1	2	1, 3, 4, 5	24	>	отсечение
19	1	3	1, 2, 4, 5	20	=	ветвление
20	2	3, 1	2, 4, 5	20	=	ветвление
21	3	3, 1, 2	4, 5	20	=	ветвление
22	4	3, 1, 2, 4	5	20	=	полное решение {3, 1, 2, 4, 5}, $D_u = 20$
23	4	3, 1, 2, 5	4	22	>	полное решение {3, 1, 2, 5, 4}, $D_u = 22$
24	3	3, 1, 4	2, 5	20	=	ветвление
25	4	3, 1, 4, 2	5	20	=	полное решение {3, 1, 4, 2, 5}, $D_u = 20$
26	4	3, 1, 4, 5	2	22	>	полное решение {3, 1, 4, 5, 2}, $D_u = 22$
27	3	3, 1, 5	2, 4	24	>	отсечение
28	2	3, 2	1, 4, 5	22	>	отсечение
29	2	3, 4	1, 2, 5	22	>	отсечение
30	2	3, 5	1, 2, 4	28	>	отсечение
31	1	4	1, 2, 3, 5	24	>	отсечение
32	1	5	1, 2, 3, 4	24	>	отсечение; непроверенные решения не существуют, последние полные решения оптимальны.

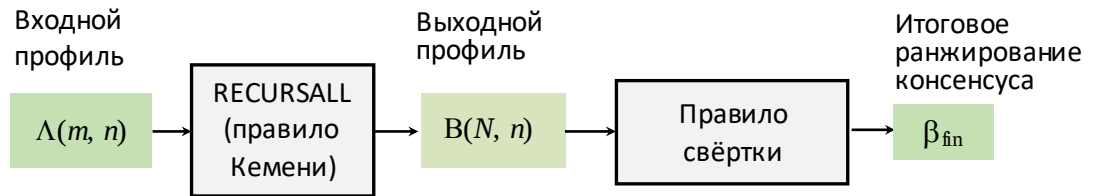


Рисунок 1.6 – Входной и выходной профили алгоритма RECURSALL для решения ЗПК

На рисунке 1.7. показаны примеры конкретных входного и выходного профилей ЗПК. Для входного профиля $\Lambda(m, n)$, приведенном на рисунке 1.7, дерево решений алгоритма RECURSALL показано на рисунке 1.4. Правило свертки для преодоления множественности решений ЗПК будет предметом рассмотрения главы 3.

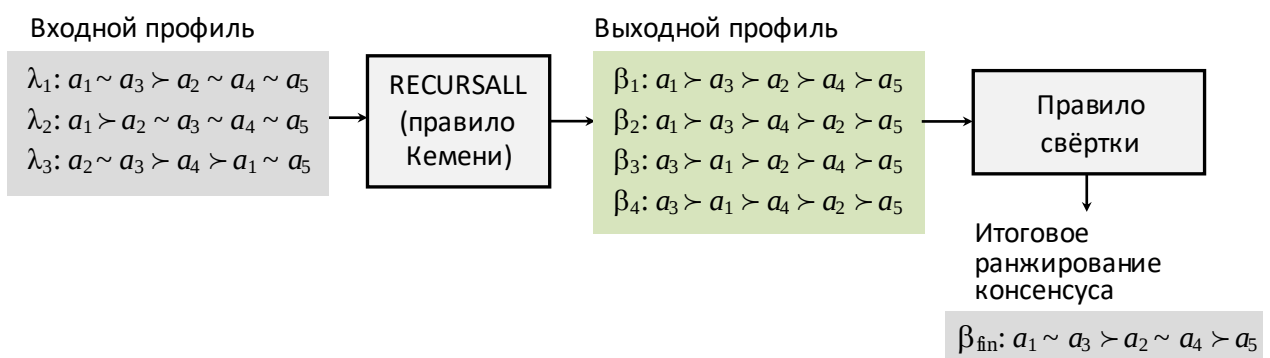


Рисунок 1.7 – Примеры входного и выходного профилей алгоритма RECURSALL для решения ЗПК

Выводы к главе 1

1. Проведен анализ основных правил выбора ранжирования консенсуса для заданных m ранжирований n альтернатив: правило простого большинства, правило Борда, правило Кондорсе и правило Кемени. Одним из наиболее справедливых и аксиоматически обоснованных правил является правило Кондорсе, основанное на полном парном сравнении альтернатив. Однако ранжирование консенсуса Кондорсе не обязательно является транзитивным, что приводит к парадоксу Кондорсе, когда ранжирование консенсуса не существует. Правило Кемени сохраняет

положительные свойства правила Кондорсе, не приводя к парадоксу. Поэтому оно считается одним из наиболее приемлемых правил агрегирования предпочтений.

2. Для применения правила Кемени при агрегировании предпочтений существуют два серьезных препятствия: (1) задача о ранжировании Кемени является *NP*-трудной, и (2) для заданного исходного профиля предпочтения правило Кемени может приводить к множественным ранжированиям консенсуса. Поэтому в данном диссертационном исследовании выбрана следующая стратегия: (1) всегда находить *полный список всех N ранжирований консенсуса*, которые являются точными оптимальными решениями, достижимыми при приемлемой размерности задачи $n < 20$, и (2) находить для этого списка *единственное* подходящее итоговое ранжирование консенсуса.

3. Для нахождения всех ранжирований консенсуса будем использовать реализующий правило Кемени рекурсивный алгоритм ветвей и границ RECURSALL.

ГЛАВА 2

МНОЖЕСТВЕННОСТЬ РЕШЕНИЙ

ЗАДАЧИ О РАНЖИРОВАНИИ КЕМЕНИ

В этой главе исследуется проблема множественности решений задачи о ранжировании Кемени (ЗРК). Рассматриваются причины явления множественности. Обсуждаются результаты экспериментальных численных исследований транзитивности/нетранзитивности входного профиля предпочтения $\Lambda(m, n)$ в зависимости от числа ранжирований (избирателей), числа альтернатив (кандидатов) и плотности толерантностей. Рассматриваются случаи нетранзитивности выходного профиля оптимальных решений $B(N, n)$.

2.1 Причины множественности

В главе 1 уже упоминалась существенная особенность правила Кемени, заключающаяся в неединственности оптимальных решений (ранжирований Кемени). Ситуации, когда ранжирований Кемени может быть слишком много, рассматривались лишь очень поверхностно в немногочисленных работах [28,31,35]. В целом можно утверждать, что этот вопрос почти не освещен в научной литературе. Возможно, это связано с тем, что далеко не каждый из возможных подходов к решению ЗРК, перечисленных в разделе 1.4.1, позволяет перечислить (найти) *все* возможные ранжирования Кемени. Эту возможность предоставляют только те методы, которые генерируют частичные решения при организации дерева поиска оптимального решения. Как раз таким методом является *метод ветвей и границ*, лежащий в основе алгоритма RECURSALL (см п. 1.4).

Так как множество Π_n всех отношений строгого предпочтения (перестановок элементов множества A) обладает конечной мощностью n , ранжирование Кемени всегда существует, однако оно далеко не всегда является единственным, т.к. значение оптимального расстояния $D(\beta, \Lambda)$ может соответствовать не единствен-

ной перестановке строк и столбцов матрицы профиля P , т.е.

$$D(\beta_1, \Lambda) = \sum_{i < j} p_{ij}^1 = D(\beta_2, \Lambda) = \sum_{i < j} p_{ij}^2 = \dots = D(\beta_N, \Lambda) = \sum_{i < j} p_{ij}^N, \quad (2.1)$$

где N – количество ранжирований Кемени (оптимальных перестановок). Оптимальные решения могут ранжировать альтернативы по-разному, что приводит к гораздо большей неоднозначности, чем та, которая изначально присуща исходному профилю $\Lambda(m, n)$.

Среди причин, приводящих к неединственности оптимальных решений ЗРК, следует выделить два основных фактора.

1. ***Несоответствие друг другу мощностей множества возможных решений и множества исходных ранжирований.*** Действительно, каждое оптимальное (точное) решение, по определению (п. 1.3.4), принадлежит пространству строгих (линейных) ранжирований, а исходные ранжирования принадлежат более широкому пространству слабых порядков. При этом велика вероятность существования одинаковых расстояний Кемени от исходного профиля до разных строгих ранжирований.
2. ***Взаимно некомпенсированная несогласованность исходных ранжирований между собой.*** Действительно, несогласованность исходных ранжирований проявляется как наличие в *разных* ранжированиях противоположных порядков (или *инверсий*) альтернатив и/или *нетранзитивных циклов*. В конкретных ситуациях эта несогласованность может быть взаимно скомпенсирована или не скомпенсирована. В случае отсутствия компенсации также возникает вероятность существования одинаковых расстояний Кемени от исходного профиля до разных строгих ранжирований.

2.2 Предыдущие исследования множественности

Значимые исследования множественности были впервые проведены Муравьевым С.В. в 2009-2012 гг. и опубликованы в работе [71]. В ходе исследований было показано, что правило Кемени приводит к числу N оптимальных решений (выходному профилю), которое может во много раз превышать количество ран-

жирований во входном профиле $\Lambda(m, n)$, т.е. при решении ЗРК нередки случаи, когда $N \gg m$.

Для исследования множественности правила Кемени был проведен вычислительный эксперимент с использованием алгоритма RECURSALL, реализующего рекурсивный метод ветвей и границ (см. п. 1.5).

В ходе эксперимента генерировались по 100 индивидуальных профилей (задач) с начальными точками от 1 до 100 для каждого набора параметров $m = 4, 5, 6, 7, 10, 15$; $n = 10, 15, 20$ и трех значений плотности толерантностей t (представленных в таблице 2.1), рассчитанных по формуле:

$$t = n_{\tau} / m(n - 1), \quad (2.2)$$

где n_{τ} – количество толерантностей в $\Lambda(m, n)$.

Таблица 2.1 – Три принятых значения плотности толерантностей

t_0	t_1	t_2
$t = 0$	$0,02 \leq t \leq 0,1$	$0,2 \leq t \leq 0,5$

Исходные матрицы профиля рассчитывались по ранжированиям, полученным путем объединения псевдослучайных строгих порядков p и толерантностей τ , сгенерированных по отдельности на основе библиотечной функции Microsoft Visual C++ `randomize()`. Строгие порядки, представленные N перестановками, генерировались на основе равномерного закона распределения целых чисел в заданном диапазоне от 1 до n . Толерантности также генерировались в соответствии с равномерным законом распределения.

Функция `randomize()` принимает произвольное целое число, так называемую *стартовую точку* S_p (starting point), $S_p = 1, 2, \dots, 100$, которая служит номером (идентификатором) генерируемого профиля (соответствующего *индивидуальной задаче* о ранжировании Кемени). Например, для генерации задачи с номером 35 используется целое число 35 в качестве стартовой точки. Таким образом, используя количество ранжирований m , количество альтернатив n и стартовую

точку, можно в любой момент восстановить (заново сгенерировать) конкретный индивидуальный профиль с этими параметрами.

Таким образом, в ходе экспериментов были сгенерированы 4600 индивидуальных профилей (задач), каждый из которых служил входными данными для алгоритма RECURSALL. Затем фиксировалось количество оптимальных решений N алгоритма для каждой индивидуальной задачи и для каждого прогона из 100 задач строился график зависимости $N = f(S_p)$.

Следует отметить важную особенность построения графика $N = f(S_p)$: для придания ему большей информативности и наглядности, значения S_p всегда следует *сортировать в порядке убывания значений N* . На рисунке 2.1. показан пример графика $N = f(S_p)$, построенного для неупорядоченных значений S_p , для следующего набора параметров $m = 49$, $n = 18$ и t_1 . По данным такого графика трудно судить о характере изменения множественности оптимальных решений ЗРК, поэтому в данной работе все графики вида $N = f(S_p)$ приводятся с использованием процедуры сортировки.

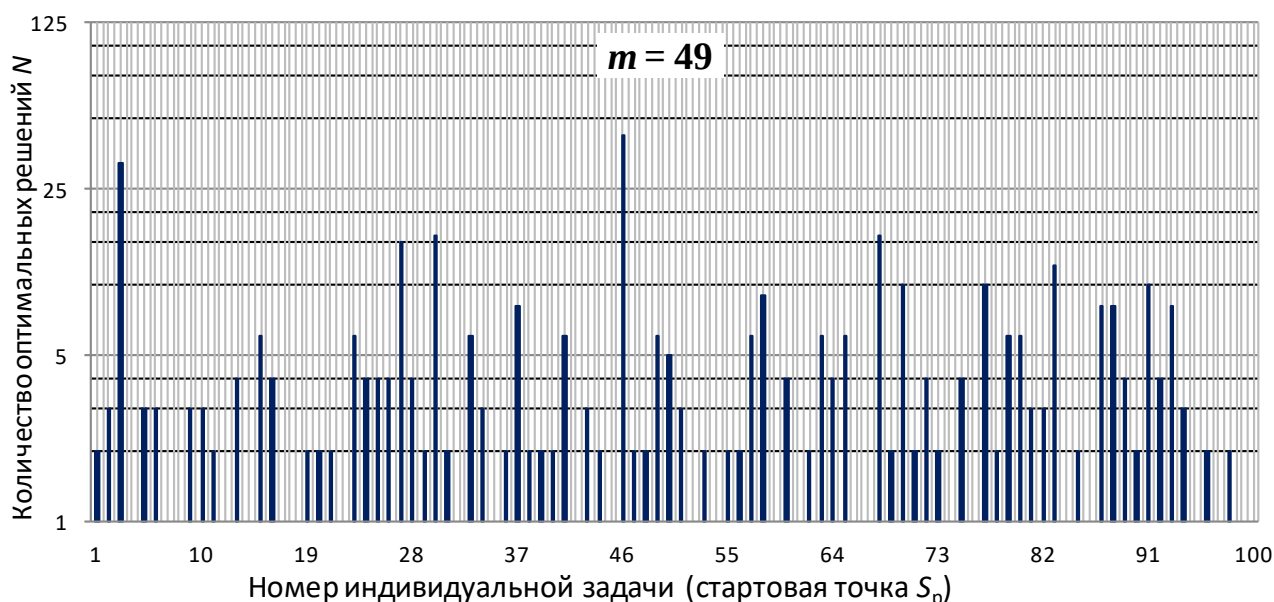


Рисунок 2.1 – Пример неупорядоченных чисел ранжирований Кемени при решении 100 индивидуальных задач для $m = 49$, $n = 18$ при t_1

Проведенные экспериментальные исследования показали, что хотя N оптимальных решений ЗРК находятся на минимальном расстоянии Кемени от m исходных ранжирований, они могут быть довольно сильно удалены друг от друга. При определенных значениях m и n , особенно при *малых и четных* $m \geq 4$, значение N может достигать десятков миллионов [71]. При этом возникает хаосоподобная ситуация, когда $N \gg m$, и неопределенность после решения задачи значительно превышает исходную неопределенность.

Для иллюстрации этого явления, на рисунках 2.2 - 2.5 приведены кривые, показывающие множественность оптимальных решений N (вертикальная ось) для каждой группы из 100 индивидуальных задач (горизонтальная ось) для $m = 4; 5; 49$ и 50 при различных значениях t (сплошная линия – t_0 , штрих-пунктирная t_1 ; пунктирная t_2 , см. таблица 2.1). Цветом выделены различные значения n : красный цвет для $n = 20$; синий цвет для $n = 15$; зелёный цвет для $n = 10$.

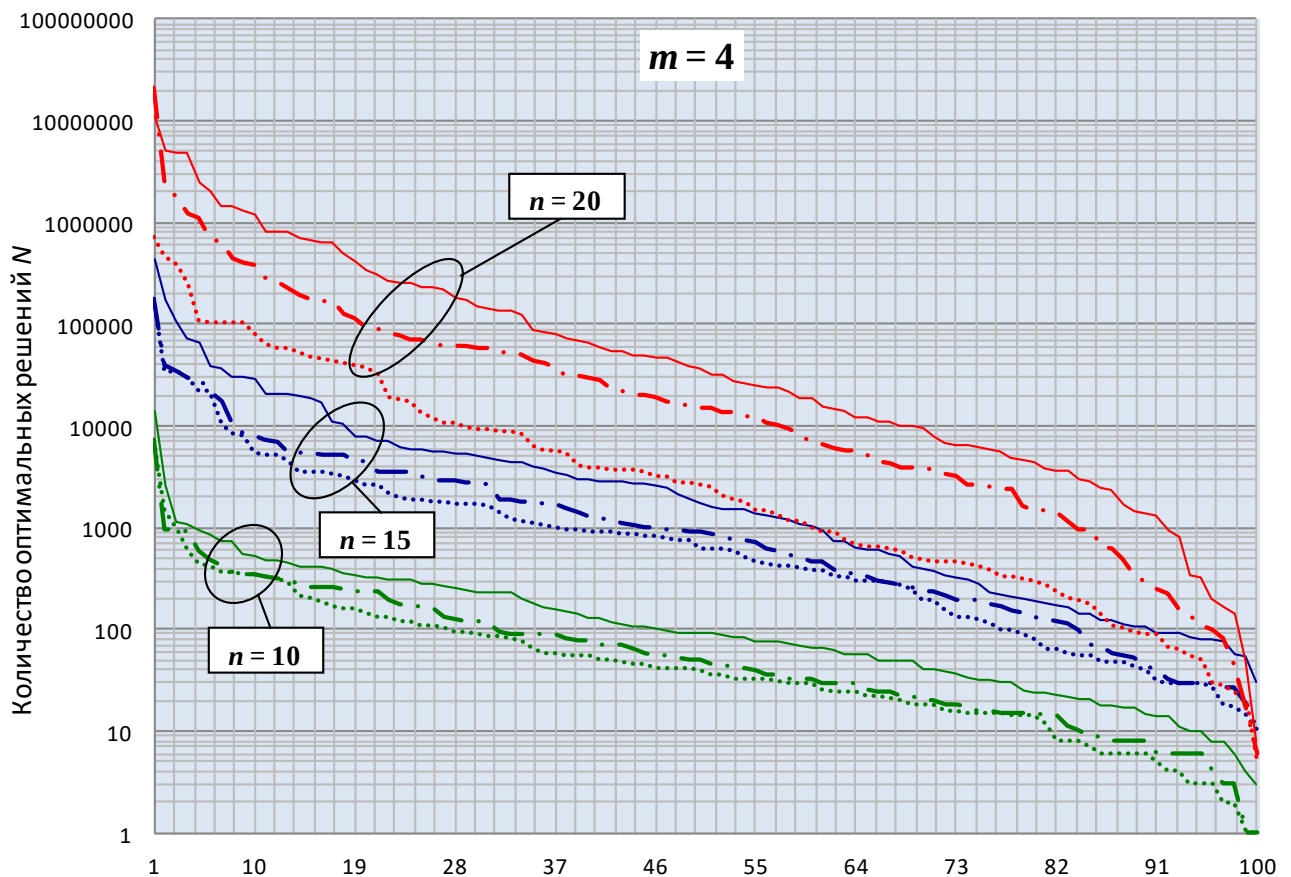


Рисунок 2.2 – Множественность оптимальных решений N для $m = 4$ при t_0, t_1, t_2

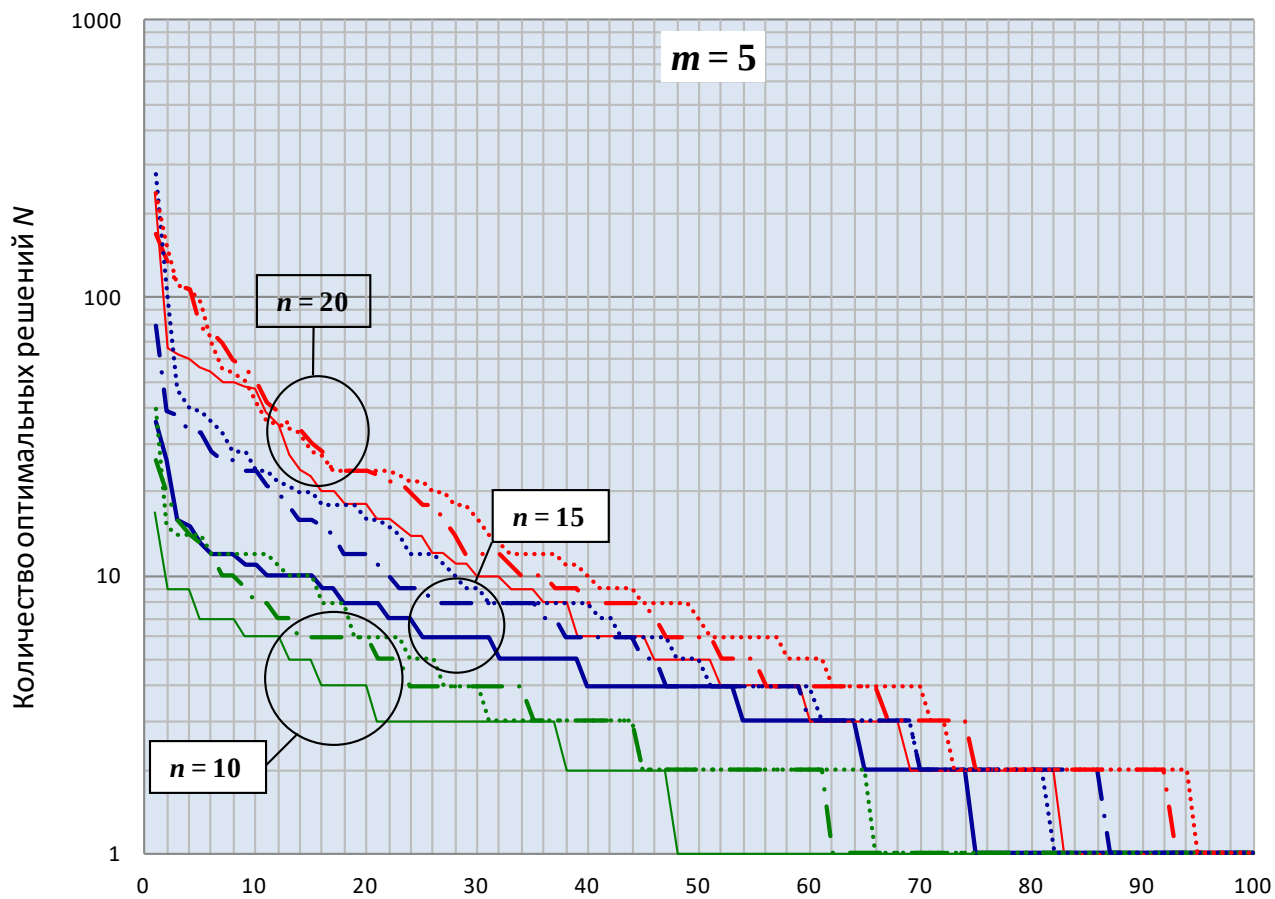


Рисунок 2.3 – Множественность оптимальных решений N для $m = 5$ при t_0, t_1, t_2

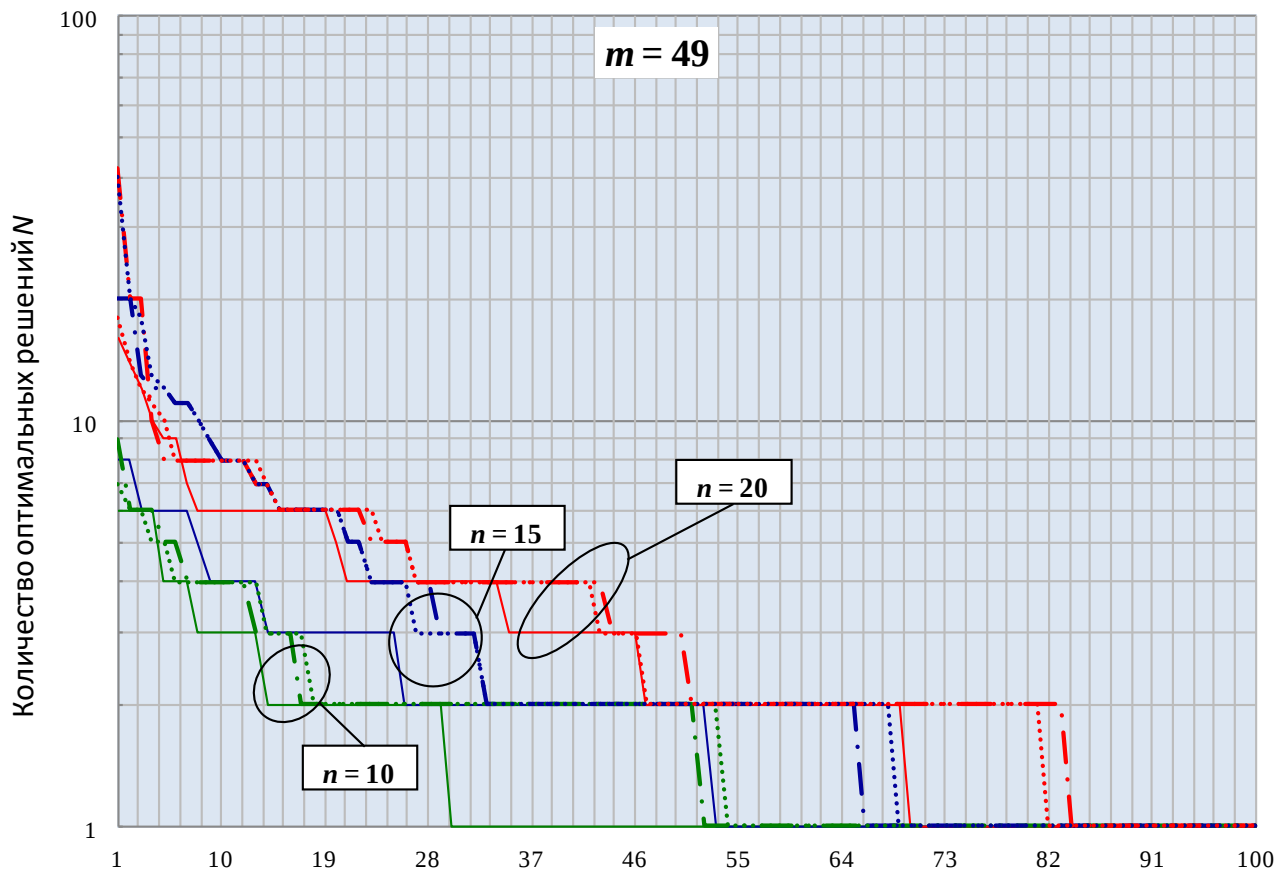


Рисунок 2.4 – Множественность оптимальных решений N для $m = 49$ при t_0, t_1, t_2

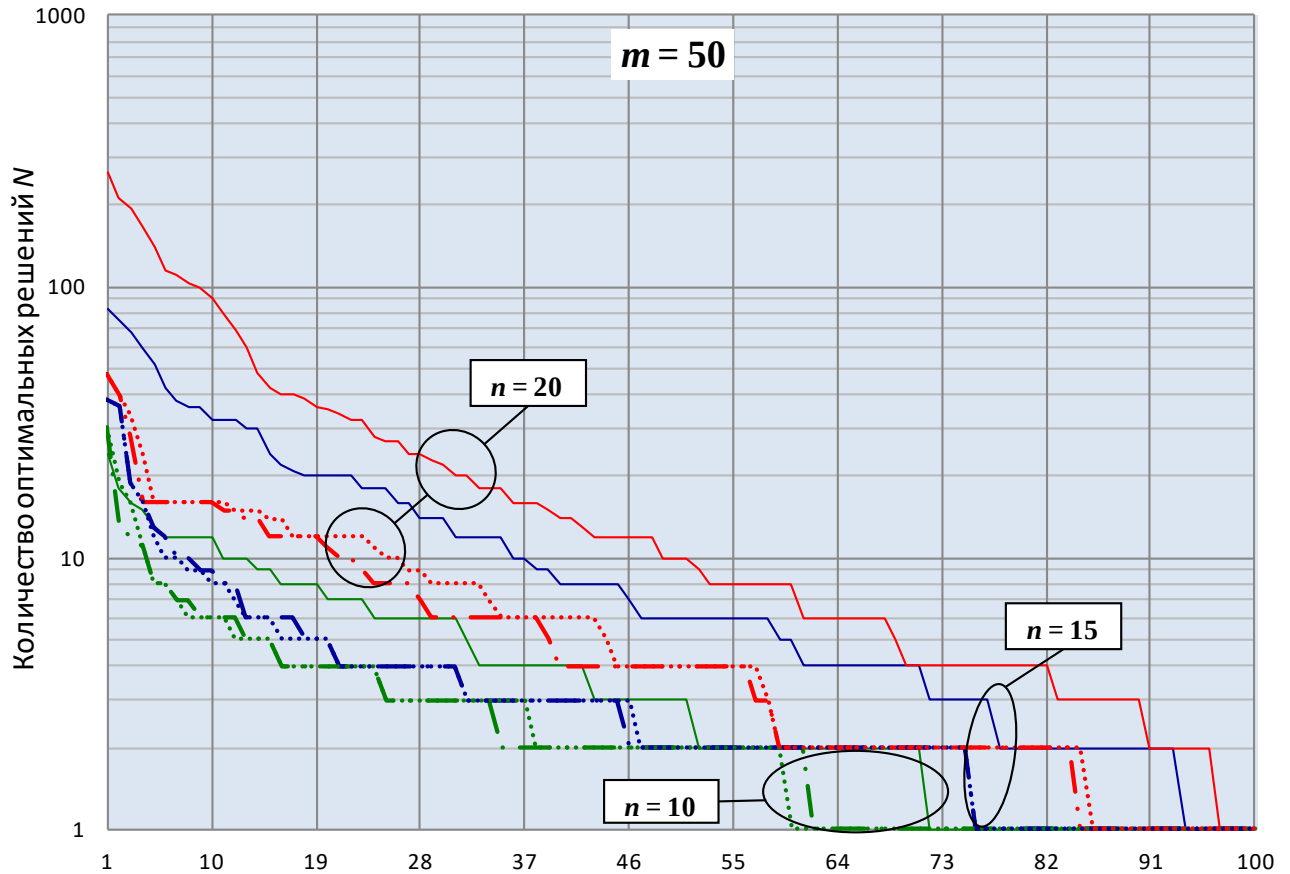


Рисунок 2.5 – Множественность оптимальных решений N для $m = 50$ при t_0, t_1, t_2

Анализ полученных результатов экспериментальных исследований позволяет сформулировать следующие утверждения.

1. Число N увеличивается с увеличением количества альтернатив n .
2. Множественность проявляется в значительно большей степени при чётных значениях m , чем при нечётных. Рекордное значение N во всех экспериментах для параметров $m = 4, n = 20, t_0$ составило $N = 11\,279\,826$ оптимальных решений.
3. Множественность резко увеличивается при уменьшении чётных значений m , но плавно увеличивается при уменьшении нечётных значений m .
4. Наличие толерантностей во входном профиле $\Lambda(m, n)$ влияет на множественность по-разному: увеличение плотности толерантностей t при нечётных m приводит к плавному увеличению N , в то время как для чётных m наоборот, N плавно уменьшается.

2.3 Экспериментальные исследования транзитивности входного профиля

Описанные в параграфе 2.1. исследования множественности решений ЗРК фактически связаны с первым фактором множественности, обусловленным несоответствием друг другу мощностей множества возможных решений и множества исходных ранжирований. Вторым фактором множественности, связанный с проявлением транзитивности/нетранзитивности входного профиля, является предметом проведенных в ходе диссертационной работы экспериментальных исследований и рассматривается в этом параграфе.

2.3.1. Критерий транзитивности входного профиля

Несмотря на то, что каждое k -е ранжирование профиля предпочтения $\Lambda(m, n)$ транзитивно, матрица профиля P является *транзитивной* только в том случае, если все исходные ранжирования *совместимы* между собой, т.е.

$$p_{ik} \leq p_{ki}, \text{ если } p_{ij} \leq p_{ji} \text{ и } p_{jk} \leq p_{kj}, i \neq j \neq k = 1, \dots, n. \quad (2.3)$$

Матрица P всякий раз будет оказываться *нетранзитивной*, если условия (2.3) нарушаются. В случае применения правила Кондорсе такое нарушение транзитивности привело бы к парадоксу Кондорсе, при котором победителя Кондорсе не существует (см. п. 1.3.3).

В монографии [17] Литвака Б.Г. показано, что если матрица P транзитивна, то *наименьшее возможное расстояние Кемени* $D_{\text{least}} = D(\beta, \Lambda)$ и D_{least} является достижимым значением, см. (1.32), определяемым по формуле:

$$D_{\text{least}} = \sum_{i < j} \min(p_{ij}, p_{ji}), i, j = 1, \dots, n. \quad (2.4)$$

Ясно, что транзитивность/нетранзитивность входного профиля $\Lambda(m, n)$ определяется транзитивностью/нетранзитивностью его матрицы. Таким образом, транзитивность любого входного профиля, после определения ранжирования консенсуса β по правилу Кемени (1.31), может быть установлена с использованием следующего критерия:

$$\text{Профиль } \Lambda(m, n) \begin{cases} \text{транзитивен,} & \text{если } D_{\text{least}} = D(\beta, \Lambda) \\ \text{нетранзитивен,} & \text{если } D_{\text{least}} < D(\beta, \Lambda) \end{cases} \quad (2.5)$$

Введем обозначение для признака транзитивности профиля:

$$\alpha = \begin{cases} 1, & \text{если профиль } \Lambda(m, n) \text{ транзитивен} \\ 0, & \text{если профиль } \Lambda(m, n) \text{ нетранзитивен} \end{cases} \quad (2.6)$$

2.3.2. План экспериментов

Для исследования транзитивности/нетранзитивности входного профиля ЗРК был проведен вычислительный эксперимент с использованием алгоритма RECURSALL, реализующего рекурсивный метод ветвей и границ (см. п. 1.4).

Как и в случае экспериментальных исследований множественности ЗРК (п. 2.2), в ходе эксперимента генерировались по 100 индивидуальных профилей (задач) со стартовыми точками $S_p = 1, \dots, 100$ для каждого набора параметров $m = 3, 4, 5, 6, 7, 10, 15, 49, 50$; $n = 7, 10, 15, 18, 20$ и трех значений плотности толерантностей t , представленных в таблице 2.1 [8].

Малые чётные и нечётные значения m ($m = 3, 4, \dots, 15$) исследовались достаточно детально, то есть с малым шагом изменения m , с учетом результатов исследований множественности ЗРК (п. 2.2). Значения $m = 49, 50$ были выбраны с целью проведения анализа транзитивности при достаточно большом числе исходных ранжирований.

Значения n взяты из диапазона от 7 до 20 с небольшим примерно одинаковым шагом. Максимальное значение $n = 20$ обусловлено NP -полнотой ЗРК, см. п. 1.4.2.

План экспериментов представлен в таблице 2.2. Строки, выделенные серым цветом, соответствуют чётным значениям m .

Как и ранее, исходные матрицы профиля рассчитывались по ранжированиям, полученным путем объединения псевдослучайных строгих порядков p и толерантностей t , сгенерированных по отдельности на основе библиотечной функции Microsoft Visual C++ `randomize()` и распределенных по *равномерному закону*.

Тем самым, в экспериментальных исследованиях будем придерживаться так называемого условия *беспристрастной культуры* (*impartial culture*), подразумевающего равномерное распределение мнений избирателей, которое в теории голосования рассматривается как наилучший (следовательно, наиболее надежный) случай при моделировании профилей предпочтений [61,93,94].

Таблица 2.2 – План экспериментов: генерировались по 100 индивидуальных задач для каждого набора параметров m, n, t

m	t_0	t_1	t_2
3	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$
4	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$
5	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$
6	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$
7	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$
10	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$
15	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$
49	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$
50	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$	$n = 7, 10, 15, 18, 20$

Таким образом, в ходе экспериментов были сгенерированы 13500 индивидуальных профилей (задач), каждый из которых служил входными данными для алгоритма RECURSALL. Затем фиксировались значения расстояний D_{least} и $D(\beta, \Lambda)$, рассчитанные для каждой индивидуальной задачи S_p . Пример необработанных результатов решения 100 индивидуальных задач для набора параметров $m = 49, n = 18$ и t_1 приведен в таблице 2.3.

Введем понятие *степени нетранзитивности*, которое вытекает из критерия (2.5), и определяется как абсолютная разность

$$\Delta = (D(\beta, \Lambda) - D_{\text{least}}). \quad (2.7)$$

Степень нетранзитивности можно определить также в относительной форме

$$\delta = (D(\beta, \Lambda) - D_{\text{least}}) / D_{\text{least}}, \quad (2.8)$$

которую, при необходимости, можно выражать в процентах. Соответствующие столбцы для Δ и δ показаны в таблице 2.3.

Ясно, что минимально возможная степень нетранзитивности Δ входного профиля $\Lambda(m, n)$ равна 0, в этом случае профиль полностью транзитивен, и признак транзитивности $\alpha = 1$.

Таблица 2.3 – Пример необработанных результатов решения 100 индивидуальных задач для $m = 49$, $n = 18$ и t_1

S_p	N	D_{least}	$D(\beta, \Lambda)$	α	Λ	$\delta, \%$	S_p	N	D_{least}	$D(\beta, \Lambda)$	α	Λ	$\delta, \%$
1	2	6711	6789	0	78	1,2	44	2	6580	6638	0	58	0,9
2	3	6496	6552	0	56	0,9	45	1	6861	6943	0	82	1,2
3	32	6722	6780	0	58	0,9	46	42	6429	6471	0	42	0,7
4	1	6661	6747	0	86	1,3	47	2	6522	6586	0	64	1,0
5	3	6725	6781	0	56	0,8	48	2	6758	6832	0	74	1,1
6	3	6817	6903	0	86	1,3	49	6	6847	6927	0	80	1,2
7	1	6769	6853	0	84	1,2	50	5	6585	6649	0	64	1,0
8	1	6830	6922	0	92	1,3	51	3	6673	6719	0	46	0,7
9	3	6523	6571	0	48	0,7	52	1	6788	6904	0	116	1,7
10	3	6773	6847	0	74	1,1	53	2	6830	6904	0	74	1,1
11	2	6715	6815	0	100	1,5	54	1	6749	6805	0	56	0,8
12	1	6528	6584	0	56	0,9	55	2	6636	6684	0	48	0,7
13	4	6523	6563	0	40	0,6	56	2	6484	6526	0	42	0,6
14	1	6675	6725	0	50	0,7	57	6	6638	6700	0	62	0,9
15	6	6819	6905	0	86	1,3	58	9	6779	6845	0	66	1,0
16	4	6677	6729	0	52	0,8	59	1	6523	6555	0	32	0,5
17	1	6695	6771	0	76	1,1	60	4	6696	6778	0	82	1,2
18	1	6523	6551	0	28	0,4	61	1	6859	6935	0	76	1,1
19	2	6739	6823	0	84	1,2	62	2	6551	6609	0	58	0,9
20	2	6558	6608	0	50	0,8	63	6	6631	6697	0	66	1,0
21	2	6717	6777	0	60	0,9	64	4	6574	6620	0	46	0,7
22	1	6641	6733	0	92	1,4	65	6	6633	6731	0	98	1,5
23	6	6647	6725	0	78	1,2	66	1	6625	6703	0	78	1,2
24	4	6651	6723	0	72	1,1	67	1	6345	6375	0	30	0,5
25	4	6562	6660	0	98	1,5	68	16	6512	6534	0	22	0,3
26	4	6665	6719	0	54	0,8	69	2	6597	6679	0	82	1,2
27	15	6543	6629	0	86	1,3	70	10	6742	6814	0	72	1,1
28	4	6652	6722	0	70	1,1	71	2	6591	6667	0	76	1,2
29	2	6754	6828	0	74	1,1	72	4	6836	6932	0	96	1,4
30	16	6644	6712	0	68	1,0	73	2	6638	6728	0	90	1,4
31	2	6764	6812	0	48	0,7	74	1	6601	6635	0	34	0,5
32	1	6701	6759	0	58	0,9	75	4	6535	6601	0	66	1,0
33	6	6731	6791	0	60	0,9	76	1	6585	6635	0	50	0,8
34	3	6706	6774	0	68	1,0	77	10	6804	6868	0	64	0,9
35	1	6746	6816	0	70	1,0	78	2	6660	6720	0	60	0,9
36	2	6597	6651	0	54	0,8	79	6	6266	6308	0	42	0,7
37	8	6747	6795	0	48	0,7	80	6	6638	6698	0	60	0,9
38	2	6663	6773	0	110	1,7	81	3	6723	6791	0	68	1,0
39	2	6595	6667	0	72	1,1	82	3	6797	6893	0	96	1,4
40	2	6676	6722	0	46	0,7	83	12	6575	6617	0	42	0,6
41	6	6850	6968	0	118	1,7	84	1	6644	6704	0	60	0,9
42	1	6596	6634	0	38	0,6	85	2	6566	6626	0	60	0,9
43	3	6640	6698	0	58	0,9	86	1	6617	6665	0	48	0,7

Продолжение таблицы 2.3

S_p	N	D_{least}	$D(\beta, \Lambda)$	α	Δ	$\delta, \%$	S_p	N	D_{least}	$D(\beta, \Lambda)$	α	Δ	$\delta, \%$
87	8	6658	6752	0	94	1,41	94	3	6680	6762	0	82	1,2
88	8	6728	6834	0	106	1,58	95	1	6817	6925	0	108	1,6
89	4	6467	6505	0	38	0,59	96	2	6536	6590	0	54	0,8
90	2	6670	6756	0	86	1,3	97	1	6448	6474	0	26	0,4
91	10	6831	6917	0	86	1,3	98	2	6728	6804	0	76	1,1
92	4	6695	6765	0	70	1,0	99	1	6598	6656	0	58	0,9
93	8	6369	6433	0	64	1,0	100	1	6718	6792	0	74	1,1

2.3.3. Транзитивность входного профиля при различных t и n

В ходе экспериментальных исследований для каждого прогона из 100 индивидуальных задач строился график, на котором совместно изображены две зависимости $D_{\text{least}} = f(S_p)$ и $D(\beta, \Lambda) = f(S_p)$ для каждого значения n (см. рисунки 2.6 - 2.8). Значения S_p сортировались в порядке убывания значений D_{least} .

На рисунках 2.6 - 2.8 приведены кривые, показывающие изменение значений расстояний D_{least} и $D(\beta, \Lambda)$ (вертикальная ось) для каждой группы из 100 индивидуальных задач (горизонтальная ось) для $t = 3; 4; 5; 6; 7; 10; 49$ и 50 при t_0, t_1 и t_2 . Кривые, соответствующие расстоянию D_{least} , показаны ромбами, а $D(\beta, \Lambda)$ – квадратами. Цветом выделены различные значения n : красный цвет для $n = 20$; зеленый цвет для $n = 18$; синий цвет для $n = 15$; светло-зелёный цвет для $n = 10$; жёлтый цвет для $n = 7$.

На рисунках 2.9 - 2.12 приведены кривые, показывающие изменение относительной степени нетранзитивности $\delta = f(S_p)$ входного профиля (вертикальная ось) для каждой группы из 100 индивидуальных задач (горизонтальная ось) для $t = 3; 4; 49$ и 50 при различных значениях t (сплошная линия – t_0 , штрихпунктирная – t_1 ; пунктирная – t_2 , см. таблица 2.1). Цветом выделены различные значения n : красный цвет для $n = 20$; синий цвет для $n = 15$; зелёный цвет для $n = 10$. Значения S_p сортировались в порядке убывания значений δ .

На основе рисунков 2.6 - 2.12 был проведен анализ полученных результатов экспериментальных исследований нетранзитивности.

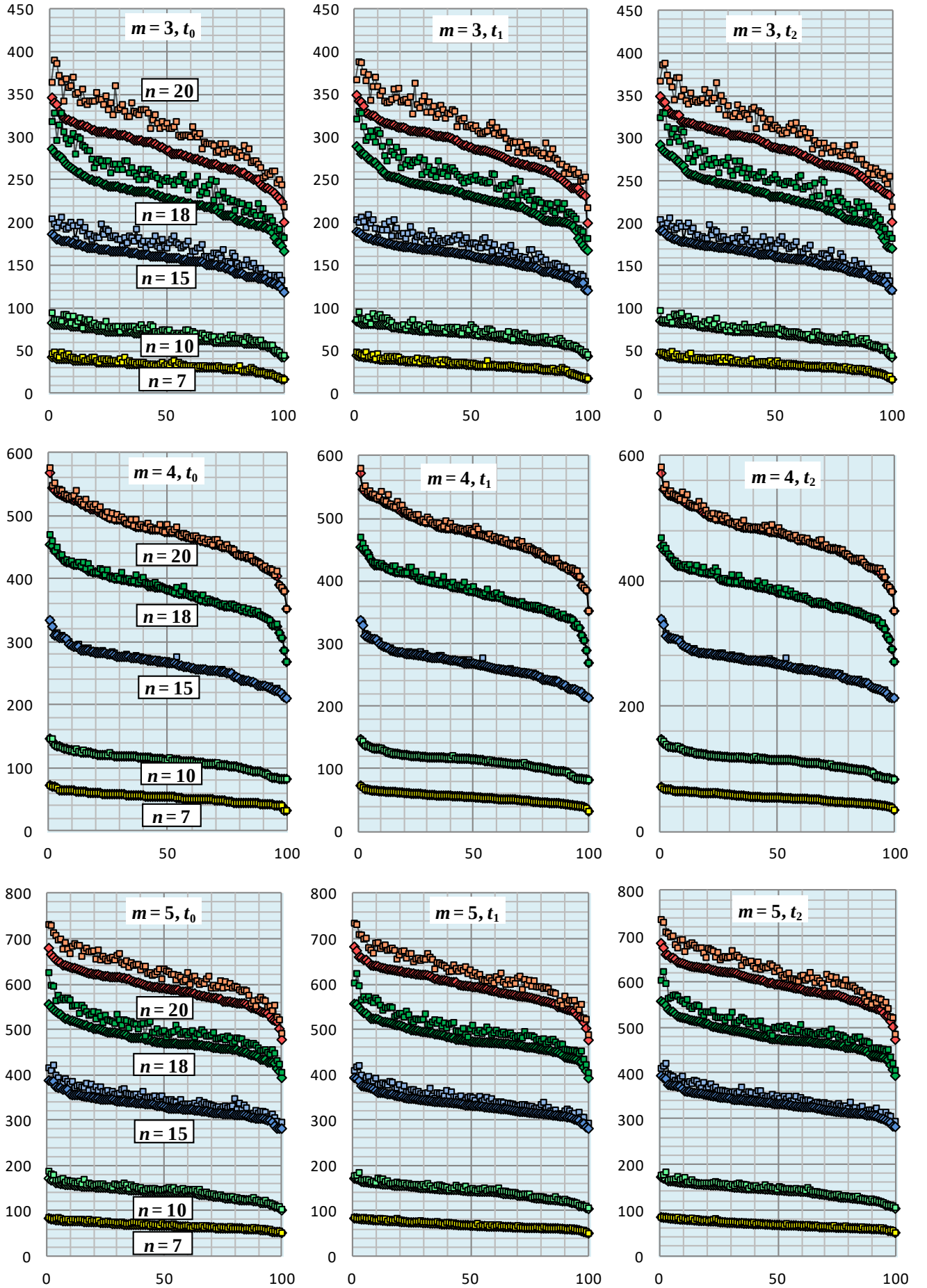


Рисунок 2.6 – Расстояния D_{least} и $D(\beta, \Lambda)$ для $m = 3; 4; 5$

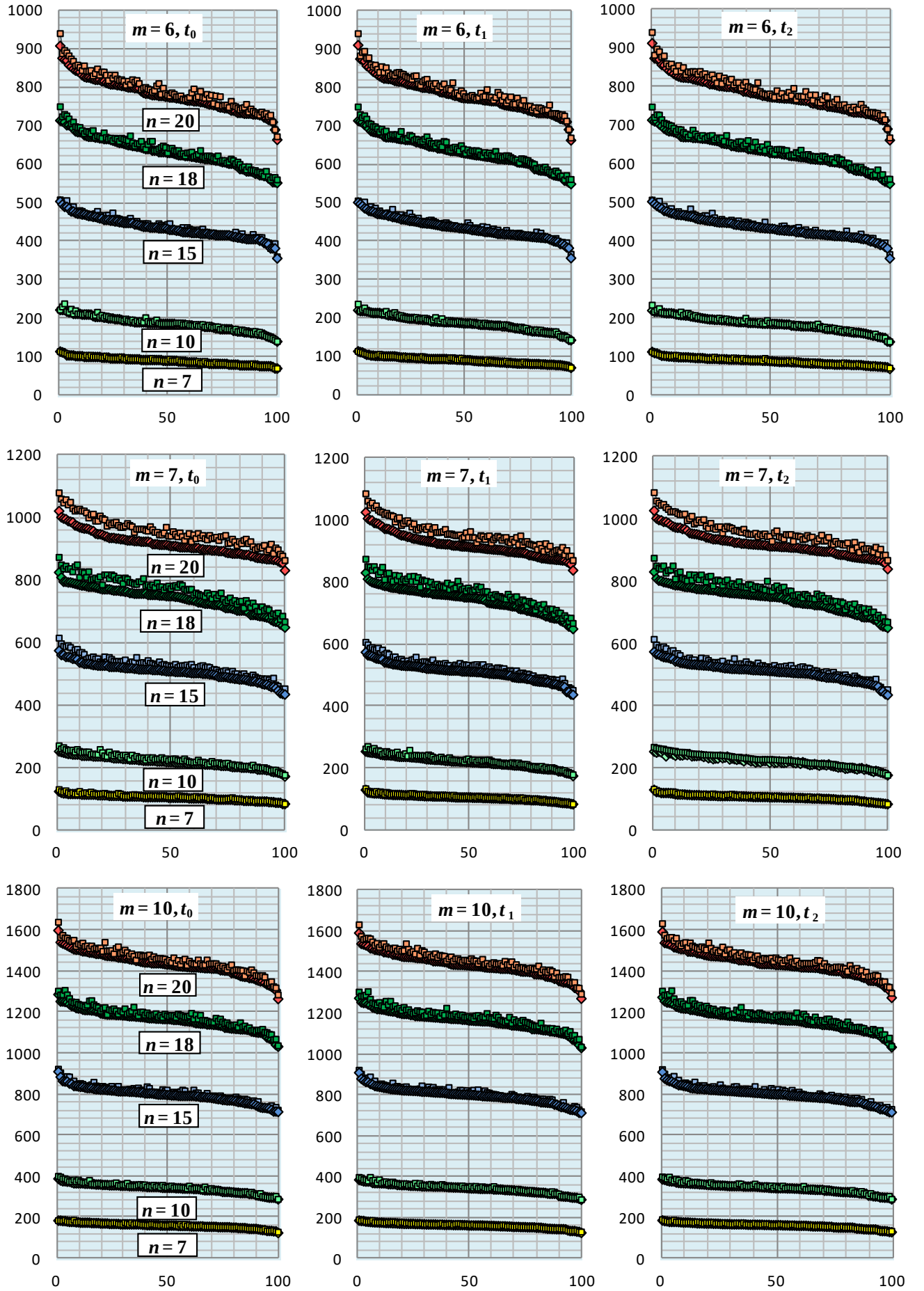


Рисунок 2.7 – Расстояния D_{least} и $D(\beta, \Lambda)$ для $m = 6; 7; 10$

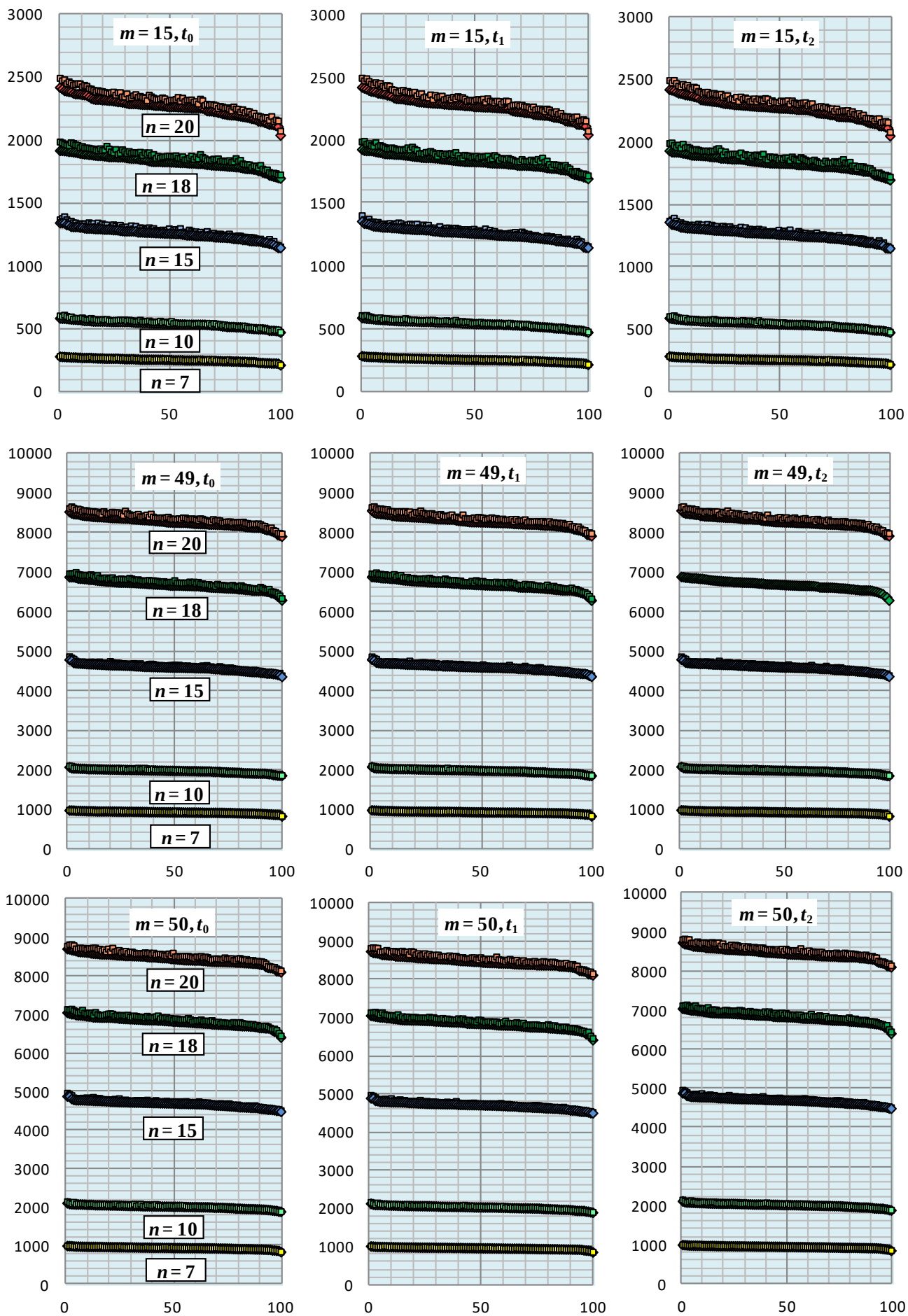


Рисунок 2.8 – Расстояния D_{least} и $D(\beta, \Lambda)$ для $m = 15; 49; 50$

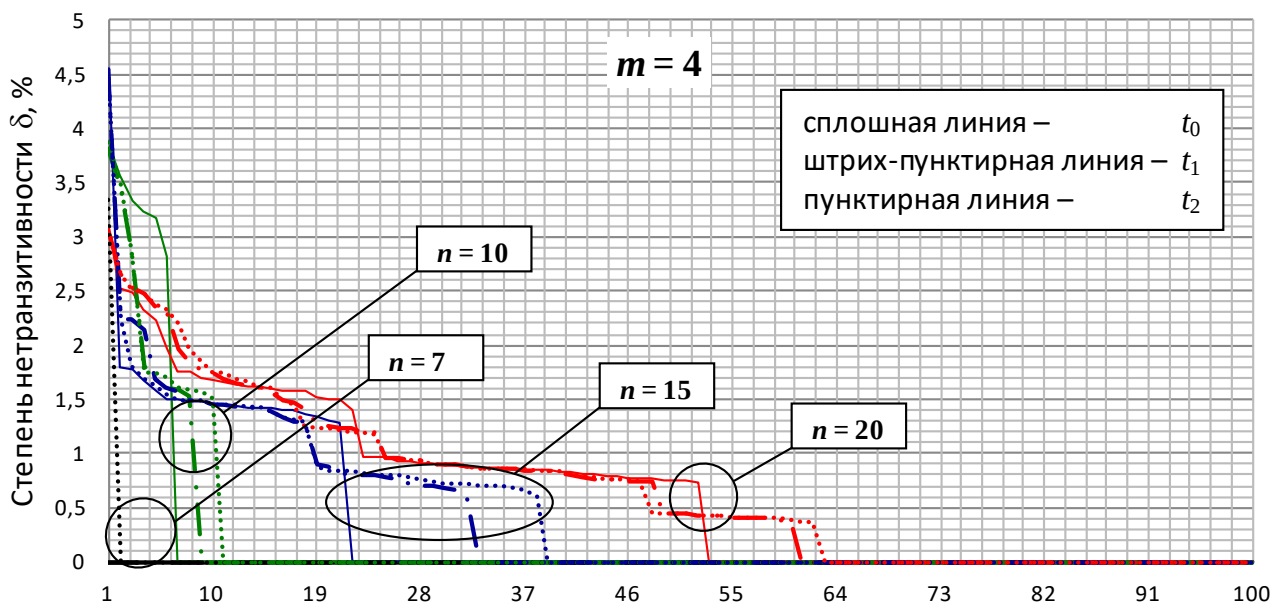


Рисунок 2.9 – Степень нетранзитивности профиля $\Lambda(m, n)$ для $m = 4$ при t_0, t_1, t_2

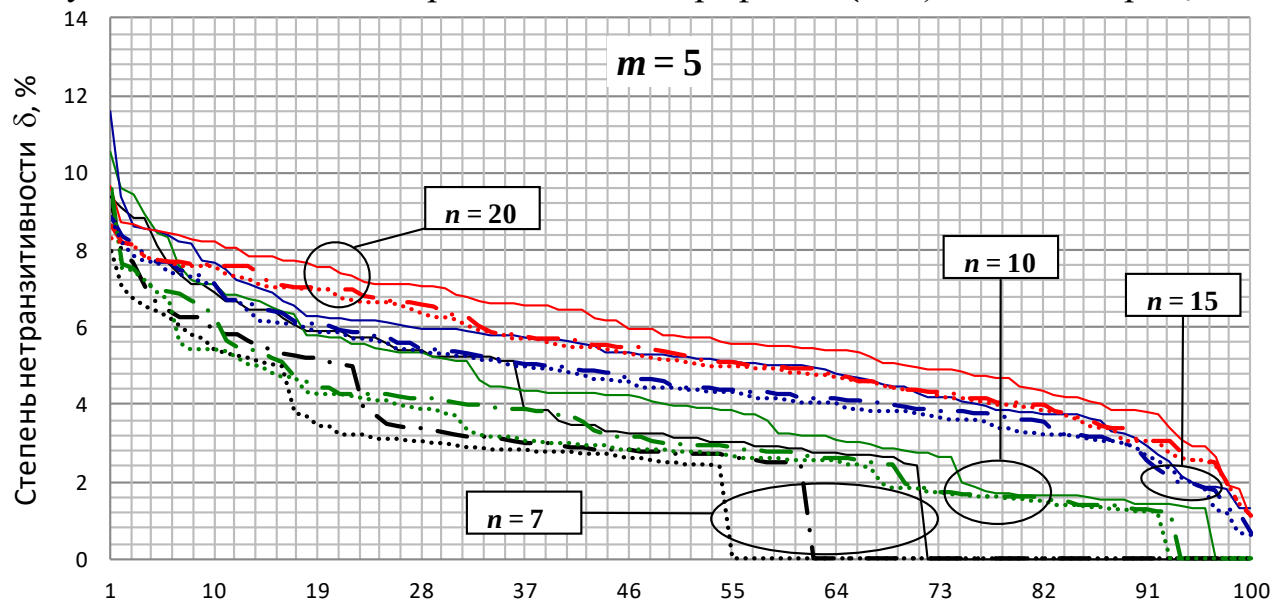


Рисунок 2.10 – Степень нетранзитивности профиля $\Lambda(m, n)$ для $m = 5$ при t_0, t_1, t_2

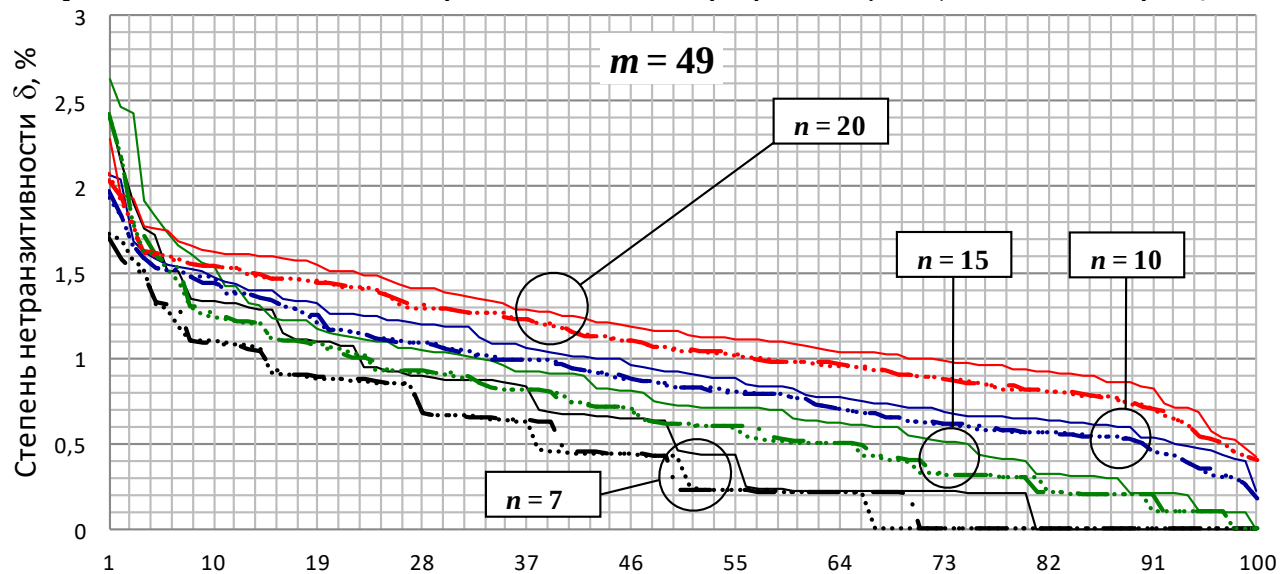


Рисунок 2.11 – Степень нетранзитивности профиля $\Lambda(m, n)$ для $m = 49$ при t_0, t_1, t_2

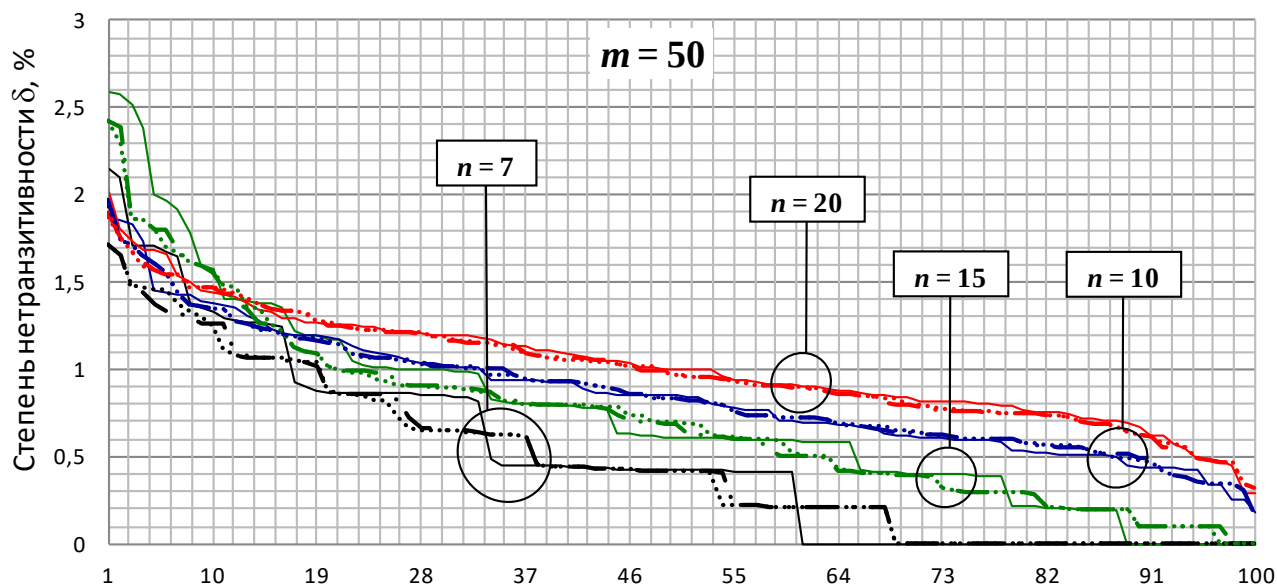


Рисунок 2.12 – Степень нетранзитивности профиля $\Lambda(m, n)$ для $m = 50$ при t_0, t_1, t_2

Из рисунков 2.6. - 2.12 следуют следующие утверждения.

1. Абсолютные значения расстояний $D(\beta, \Lambda)$ и D_{least} увеличиваются с увеличением m и n .

2. Для заданных n и m , наиболее часто оказывались транзитивными профили *только* при *четных* значениях m . *Нечетные* значения m практически всегда приводили к нетранзитивным входным профилям. При этом во всех случаях число транзитивных профилей уменьшалось с возрастанием числа n альтернатив.

3. Изменение плотности толерантностей t крайне незначительно влияет на расстояния $D(\beta, \Lambda)$ и D_{least} . При этом увеличение t при чётных m способствует незначительному уменьшению разности между $D(\beta, \Lambda)$ и D_{least} .

4. Минимальная степень нетранзитивности достигается быстрее при малых чётных значениях m при минимальной плотности толерантностей t_0 . Степень нетранзитивности уменьшается по мере уменьшения числа оптимальных решений ЗРК.

2.3.4. Оценка вероятности транзитивного входного профиля

Числовые данные, полученные в ходе экспериментов, позволяют построить оценки вероятностей появления транзитивного профиля при различных наборах параметров m, n, t . Сравнение этих оценок с известными из литературы в области теории социального выбора расчетами вероятности парадокса Кондорсе, см.

п. 1.3.3., может привести к интересным результатам.

Известно, что парадокс Кондорсе, возникающий из-за нарушения свойства транзитивности входного профиля, может возникать довольно часто, см. [51,54, 61,71,94]. Существование альтернативы (победителя) Кондорсе является событием, противоположным появлению парадокса Кондорсе. В литературе приняты следующие обозначения:

- $P(m, n)$ – вероятность альтернативы Кондорсе и
- $Q(m, n) = 1 - P(m, n)$ – вероятность парадокса Кондорсе.

На рисунках 2.13 и 2.14 показаны графики изменения вероятностей альтернативы Кондорсе (см. рисунок 2.13) и парадокса Кондорсе (см. рисунок 2.14) в зависимости от m и n при t_0 , построенные по данным работы [61], в которой значения вероятностей рассчитаны на основе рекурсивных аналитических выражений, выведенных из комбинаторных соображений.

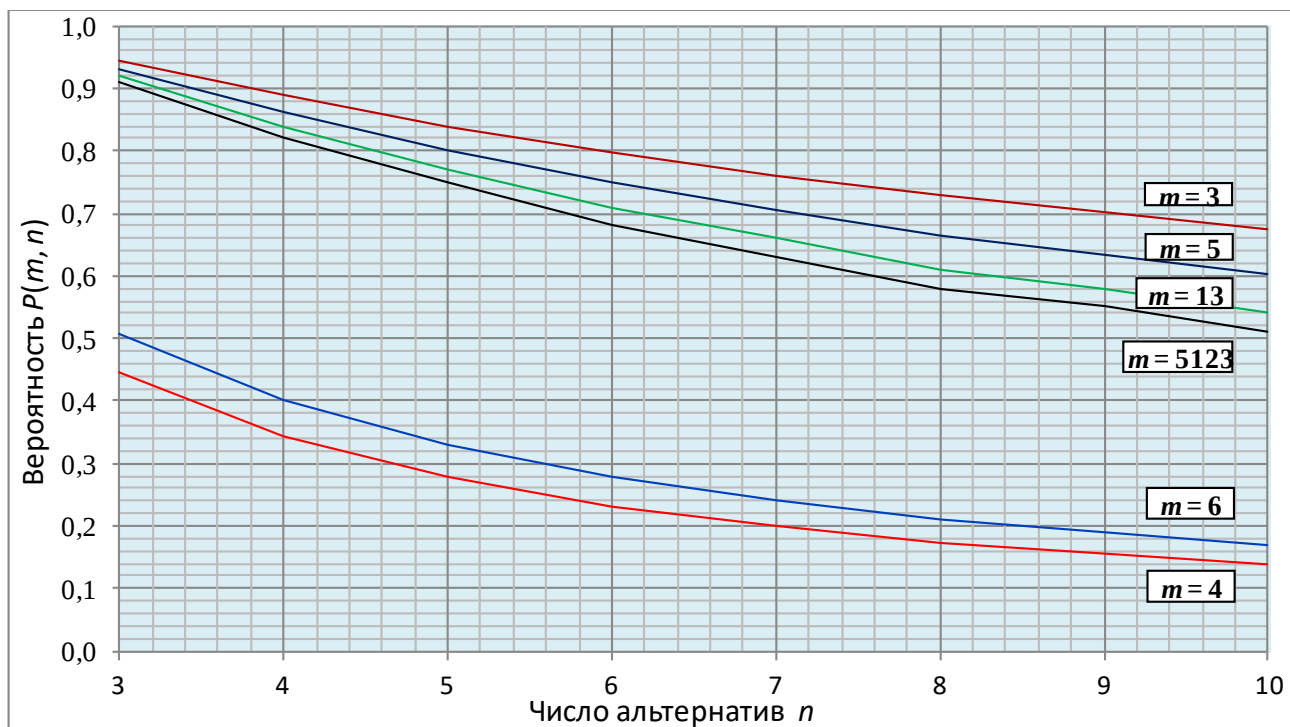


Рисунок 2.13 – Вероятность победителя Кондорсе $P(m, n)$, построенная по данным [61]

Из рисунка 2.13 видно, что вероятность альтернативы Кондорсе значительно выше при нечётных значениях m по сравнению с чётными m и уменьшается с ростом m и n . Ясно, что для вероятности парадокса Кондорсе справедливы обратные утверждения.

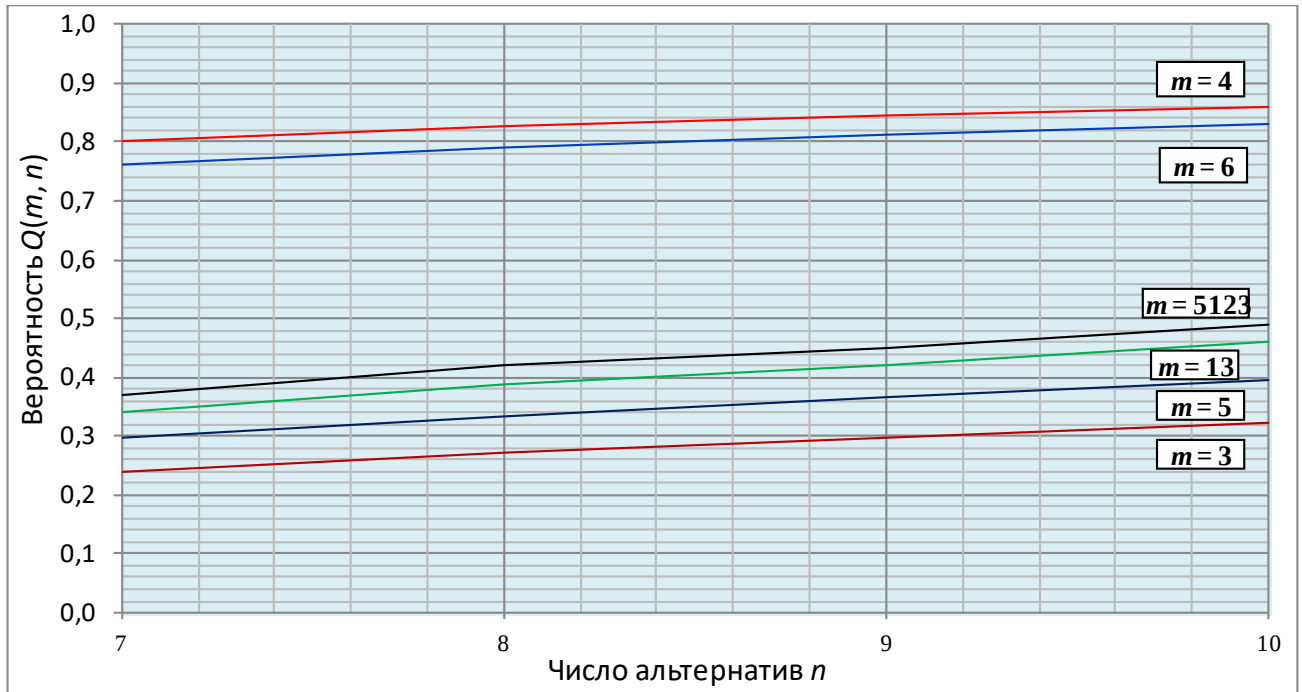


Рисунок 2.14 – Вероятность парадокса Кондорсе $Q(m, n)$, построенная по данным [61]

Результаты экспериментальных исследований множественности, полученные в работе [71], позволили получить оценки вероятности единственного решения $P_1(m, n) = P(N = 1)$ и, следовательно, вероятности противоположного события – появления множественности – $Q_1(m, n) = 1 - P_1(m, n) = P(N > 1)$.

Приведенные в [71] графики $P_1(m, n)$ находятся в практически полном согласии с данными графика вероятности $P(m, n)$ альтернативы Кондорсе, см. рисунок 2.13.

Введем обозначения для оценок вероятностей, связанных с нетранзитивностью, построенных по полученным в ходе экспериментальных исследований транзитивности данным, см., например, таблицу 2.4:

- $P_{tr}(m, n)$ – вероятность появления транзитивного профиля и
- $Q_{tr}(m, n) = 1 - P_{tr}(m, n)$ – вероятность появления нетранзитивного профиля.

Вероятность $P_{tr}(m, n)$ для прогона из 100 индивидуальных задач может быть оценена на основе полученных экспериментальных данных по следующей формуле:

$$P_{tr}(m, n) = \text{число транзитивных профилей} / \text{число задач в прогоне}. \quad (2.9)$$

Числа транзитивных профилей для каждого набора параметров m , n и t сведены в таблицу 2.4.

Используя данные таблицы 2.4, можно оценить вероятность транзитивности $P_{tr}(m, n)$ для любого набора значений m , n и t по формуле (2.10), где число задач в прогоне равно 100. Например, при $m = 4$, $n = 20$ и $t = t_0$ имеем $P_{tr}(m, n) = 48 / 100 = 0,48$.

Таблица 2.4 – Количество задач из 100, обладающих транзитивным профилем для каждого набора параметров m , n , t

m	t_0					t_1					t_2				
	n					n					n				
	7	10	15	18	20	7	10	15	18	20	7	10	15	18	20
3	38	6	0	0	0	47	13	0	0	0	52	15	0	0	0
4	100	94	79	61	48	100	92	68	55	40	99	90	62	51	38
5	29	4	0	0	0	39	7	0	0	0	46	8	0	0	0
6	87	67	30	15	6	83	61	22	8	2	77	59	21	8	3
7	22	2	0	0	0	36	3	0	0	0	42	6	1	0	0
10	72	40	6	1	0	61	31	2	0	0	57	31	2	0	0
15	23	3	0	0	0	32	5	0	0	0	34	6	0	0	0
49	20	1	0	0	0	30	3	0	0	0	34	3	0	0	0
50	40	12	0	0	0	32	3	0	0	0	32	4	0	0	0

На основе вычисленных указанным образом оценок вероятности $P_{tr}(m, n)$ были построены соответствующие графики для нечётных значений m (см. рисунок 2.15) и для четных значений m (см. рисунок 2.16).

В статье [54] Б. Джоунсом (B. Jones) на основе комбинаторных соображений аналитически показано, что при нечетных m вероятность появления нетранзитивного профиля всегда больше или равна вероятности парадокса Кондорсе, т.е.

$$Q_{tr}(m, n) \geq Q(m, n) \text{ при нечётных } m. \quad (2.10)$$

Справедливость (2.10) объясняется тем, что при расчете вероятности $Q_{tr}(m, n)$ следует учитывать наличие в ранжированиях не только нетранзитивных циклов, а также и противоположных порядков (инверсий) альтернатив (см. п. 2.1).

Полученные нами экспериментальные оценки вероятностей определенно согласуются с утверждением (2.10), что следует из сравнения графиков вероятностей парадокса Кондорсе (рисунок 2.14) и нетранзитивности входного профиля (рисунок 2.17) при нечетных m .

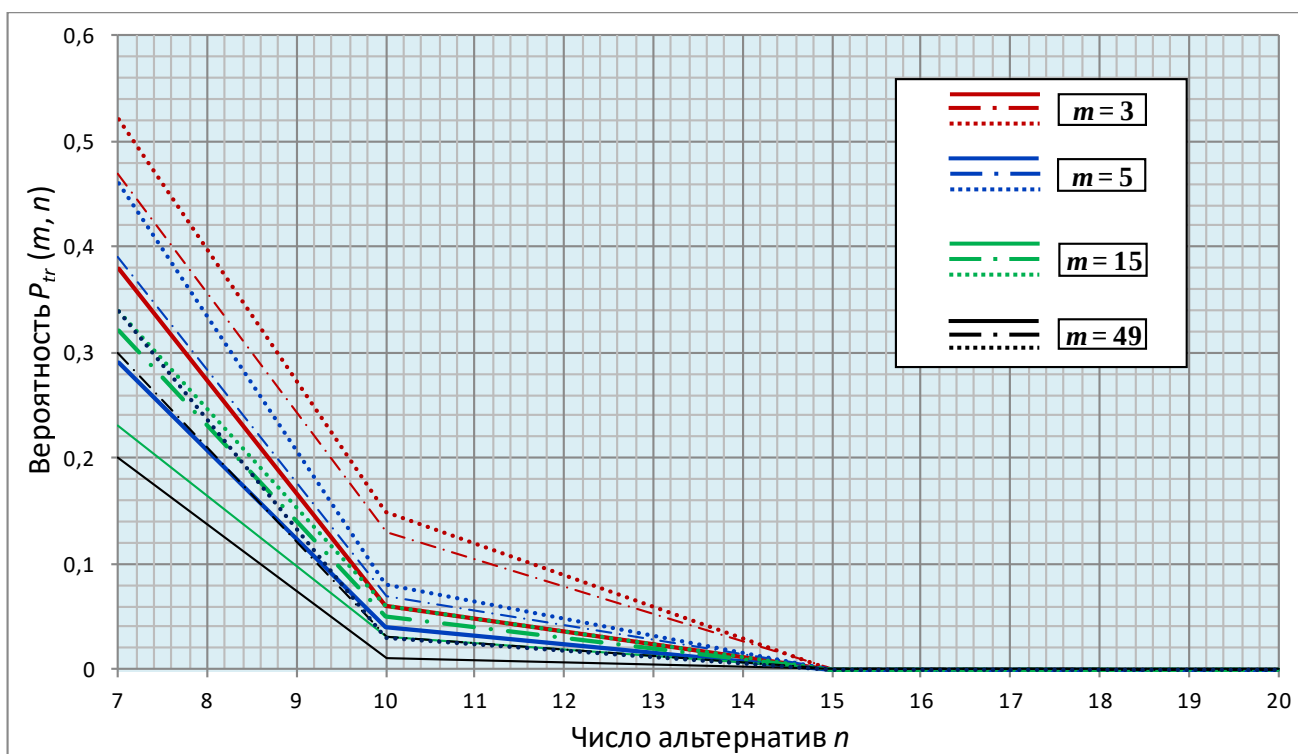


Рисунок 2.15 – Оценка вероятности $P_{tr}(m, n)$ для нечётных m

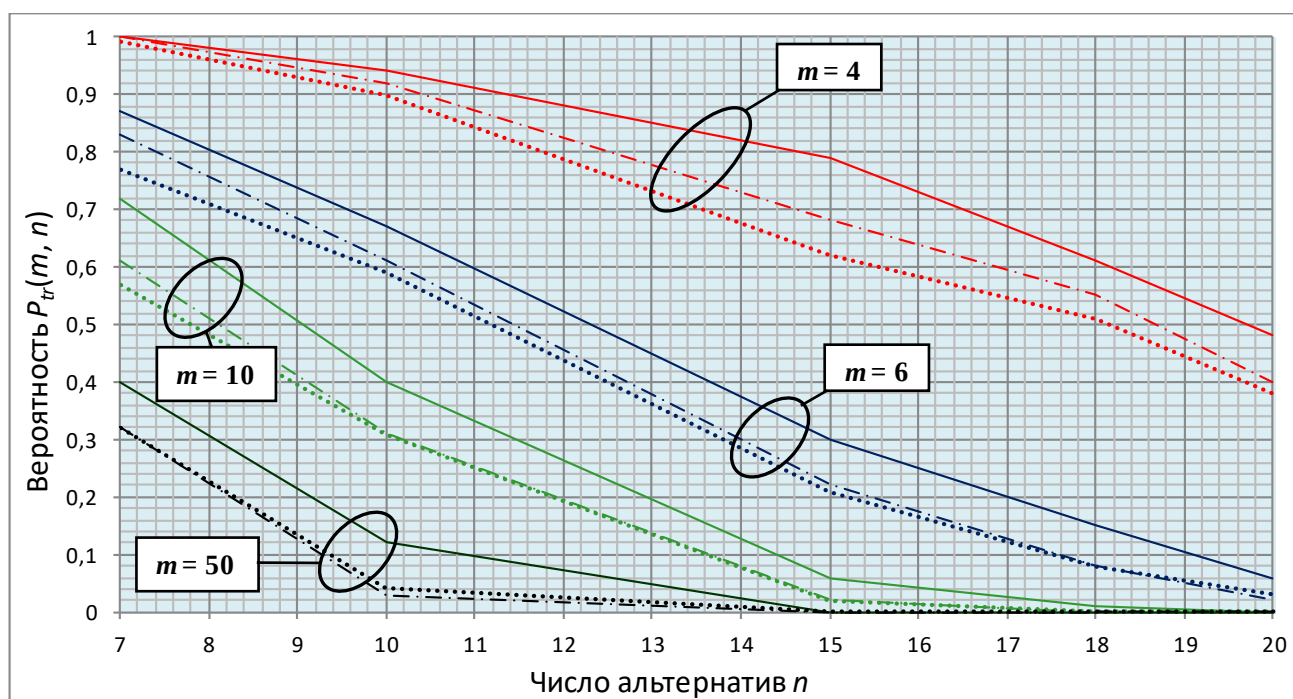


Рисунок 2.16 – Оценка вероятности $P_{tr}(m, n)$ для чётных m

По поводу четных m из рассмотрения тех же графиков (рисунки 2.14 и 2.17) можно сделать вывод о том, что в этом случае вероятность появления нетранзитивного профиля всегда меньше вероятности парадокса Кондорсе, т.е.

$$Q_{tr}(m, n) < Q(m, n) \text{ при чётных } m. \quad (2.11)$$

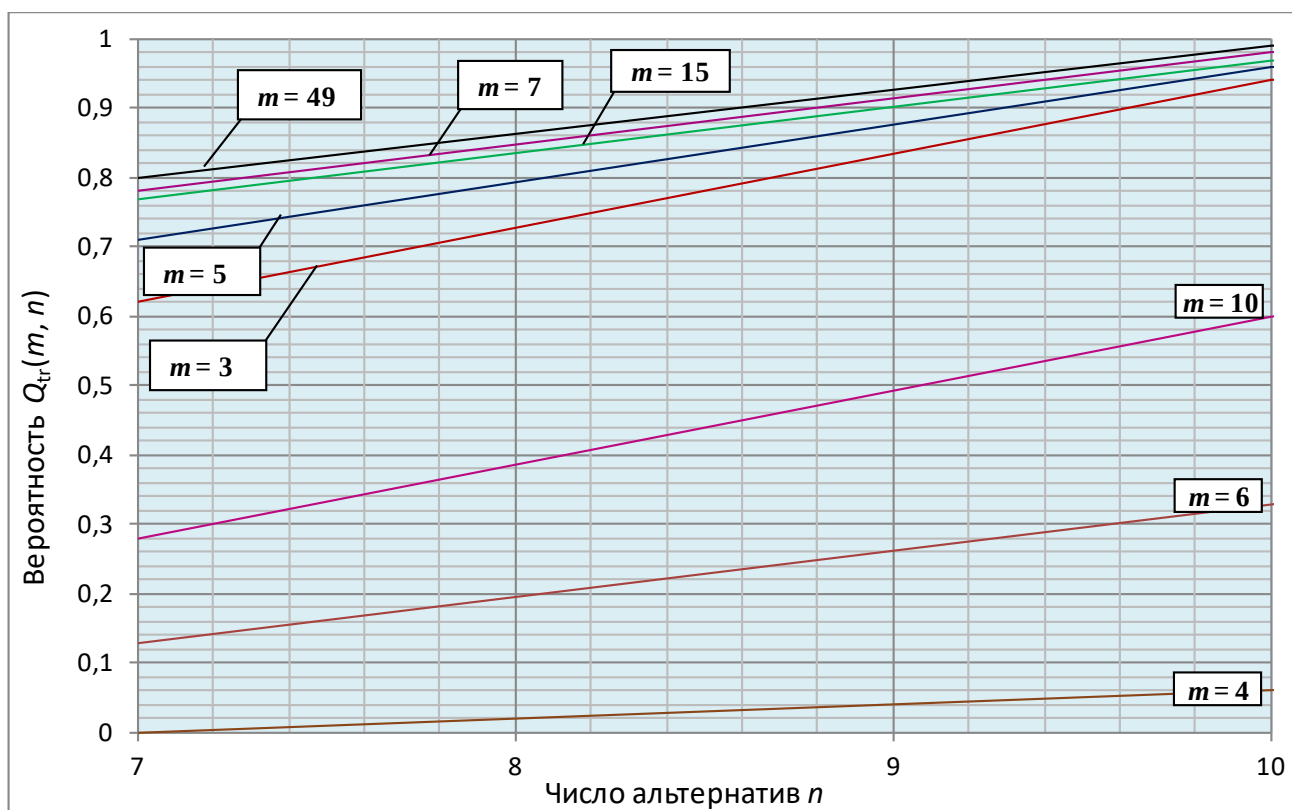


Рисунок 2.17 – Оценка вероятности нетранзитивного профиля $Q_{tr}(m, n)$

На основе полученных результатов можно сформулировать следующие утверждения.

1. Чем больше m , тем меньше вероятность транзитивности. При этом
 - при любых нечетных m вероятность $P_{tr}(m, n)$ транзитивности входного профиля близка к нулю уже при числе n альтернатив, близким к 10 (см. рисунок 2.15);
 - вероятность транзитивности значительно выше при чётных значениях m , чем при нечётных, и плавно уменьшается с ростом n (см. рисунок 2.16); при чётном значении $m = 4$ вероятность транзитивности довольно высока даже для $n = 15$; уже при $m = 10$ вероятность падает и приближается к вероятности для нечётных m , однако при росте четных m остается заметной разница между графиками вероятности для нечётных и чётных m .

2. Наличие толерантности незначительно увеличивает вероятность транзитивности при малых значениях n ($n \leq 10$) для нечётных m и также незначительно уменьшает вероятность транзитивности для чётных m .

3. Полученные оценки вероятности появления нетранзитивности входного

профиля находятся в полном согласии с известными аналитическими оценками вероятности альтернативы Кондорсе и оценками вероятности появления множественности решений ЗРК.

4. Полученные оценки соответствуют утверждениям, что при нечётных m вероятность появления нетранзитивного профиля всегда больше или равна вероятности парадокса Кондорсе, а при чётных m вероятность появления нетранзитивного профиля всегда меньше вероятности парадокса Кондорсе.

2.4 Транзитивность выходного профиля

Вопрос исследования транзитивности/нетранзитивности выходного профиля $B(N, n)$ является нетривиальным в силу сложности задачи поиска всех циклов на графе, представляющем множество всех строгих ранжирований, являющихся оптимальными решениями ЗРК.

Ясно, что необходимым условием возможности появления нетранзитивных циклов на множестве ранжирований Кемени (выходном профиле ЗРК) является существование более, чем двух оптимальных решений, т.е. $N \geq 3$.

Рассмотрим пример, иллюстрирующий наличие циклов (парадоксов Кондорсе) в выходном профиле достаточно большого размера. Число $m = 49$ выберем нечетным, поскольку четные числа m приводят к гораздо большему числу N оптимальных решений ЗРК [71]. Число $n = 18$ возьмем близким к верхней границе в соответствии с принятой стратегией, см. п. 1.4.2.

Пример 2.1. Пусть дан входной профиль $\Lambda(m, n) = \Lambda(49, 18)$, см. таблицу 2.5. Алгоритм RECURSALL построил выходной профиль $B(N, n) = B(32, 18)$. Расстояния оказались равны следующим значениям: $D_{\text{least}} = 6722$, $D(\beta, \Lambda) = 6780$, т.е. $D_{\text{least}} \neq D(\beta, \Lambda)$, откуда матрица профиля P – нетранзитивна.

Выходной профиль $B(32, 18)$ представлен в таблице 2.6. Он, очевидно, содержит только строгие порядки. Визуальный анализ таблицы 2.6 позволяет обнаружить восемь (нетранзитивных) циклов в ранжированиях с β_9 по β_{32} . Циклы состоят из альтернатив 2, 11 и 17, расположенных на первых трех позициях всех

ранжирований. Каждая из трех перестановок цикла повторяется восемь раз и показана в таблице разными цветами: перестановка $\{2, 11, 17\}$ в ранжированиях с β_9 по β_{16} – зеленый цвет; перестановка $\{11, 17, 2\}$ в ранжированиях с β_{17} по β_{24} – желтый, перестановка $\{17, 2, 11\}$ в ранжированиях с β_{25} по β_{32} – синий.

Таблица 2.5 – Входной профиль $\Lambda(m, n) = \Lambda(49, 18)$ для примера 2.1

λ_1 : 1 4 5 (16 17) 10 7 8 9 14 2 (12 18) 3 15 6 11 13	λ_{26} : 12 10 8 3 2 (6 13) 7 18 14 4 1 17 16 11 9 5 15
λ_2 : 5 11 17 6 1 15 16 (9 14) 13 2 10 3 18 8 7 12 4	λ_{27} : 7 5 (11 17) (2 6) 3 15 14 12 10 (9 4) 8 13 18 16 1
λ_3 : 17 (11 13) 16 12 10 2 8 7 14 6 18 1 3 4 5 15 9	λ_{28} : 11 7 4 6 (2 17) 9 18 1 8 3 15 10 14 13 5 16 12
λ_4 : 7 8 (1 18) 6 12 10 5 4 16 14 2 3 9 15 13 17 11	λ_{29} : 3 13 12 1 2 8 6 (7 16) 17 18 9 11 5 10 15 14 4
λ_5 : 4 11 5 7 3 14 6 13 8 9 18 17 16 12 2 15 1 10	λ_{30} : 6 5 7 2 8 9 4 12 1 15 16 14 11 10 17 13 3 18
λ_6 : 6 12 5 16 13 2 18 9 15 7 8 1 (4 11 14) 10 3 17	λ_{31} : 11 17 9 3 1 14 13 2 15 7 (4 8) 16 5 10 12 18 6
λ_7 : 18 12 1 14 4 3 7 (5 9 13) 15 2 16 11 8 6 10 17	λ_{32} : 18 13 11 4 (15 7) 14 3 10 5 1 17 8 (6 12 16) 2 9
λ_8 : 10 8 7 2 12 5 11 13 3 18 14 (4 16) 9 17 1 15 6	λ_{33} : (14 17) 5 12 (1 18) 6 9 15 (4 7) 8 2 10 16 13 11 3
λ_9 : 14 17 9 13 10 11 12 4 18 2 (7 15) 1 16 5 8 6 3	λ_{34} : 17 4 2 3 11 18 13 9 10 (7 14) 6 5 12 1 8 15 16
λ_{10} : 15 3 13 2 14 18 17 6 1 8 11 16 10 9 5 12 7 4	λ_{35} : 11 17 15 10 18 7 5 (2 12) 1 (3 8) (4 9) 14 16 13 6
λ_{11} : 4 8 2 7 (3 12) 16 17 11 15 18 9 5 10 13 6 1 14	λ_{36} : 16 4 3 5 14 6 15 12 (7 8) 13 18 2 11 10 1 9 17
λ_{12} : 9 16 3 (1 5 11) 15 2 17 8 18 4 12 13 (7 14) 10 6	λ_{37} : 8 1 9 17 18 7 16 11 12 5 2 3 15 6 10 4 13 14
λ_{13} : 17 5 (12 15) 1 3 14 2 9 4 18 6 10 7 13 11 16 8	λ_{38} : 17 2 15 (10 14) 8 11 4 (7 16) (6 12) 18 5 3 1 (9 13)
λ_{14} : 1 13 17 5 (4 9 10 16) 11 18 7 3 8 14 12 2 6 15	λ_{39} : 14 8 13 7 10 2 (5 4) (11 16) 1 18 12 3 15 9 6 17
λ_{15} : 4 (3 13) 9 12 8 17 7 (6 10) 1 11 2 5 15 16 14 18	λ_{40} : 11 8 7 17 (1 10) 18 15 2 13 9 5 12 16 14 4 6 3
λ_{16} : 15 9 2 6 3 12 17 (4 14) 5 (11 18) 8 16 13 7 1 10	λ_{41} : 9 12 6 18 11 1 4 (10 13) 17 2 16 3 5 15 7 8 14
λ_{17} : 1 2 11 5 (6 7 8 10) 17 4 14 3 18 13 16 15 9 12	λ_{42} : (3 9) 8 14 11 15 6 10 17 4 1 (5 16) 7 13 2 18 12
λ_{18} : 16 15 8 5 12 (14 17) 7 10 4 11 6 18 2 13 3 9 1	λ_{43} : 12 8 10 5 11 3 15 (18 13) (6 14) 17 16 4 2 7 9 1
λ_{19} : 8 17 6 14 15 13 3 7 5 (9 12) 2 4 18 16 10 11 1	λ_{44} : 18 4 8 13 (7 17) 3 9 1 (12 14) 15 5 10 16 (2 11) 6
λ_{20} : 5 10 11 1 12 14 4 16 15 13 6 2 7 9 3 17 18 8	λ_{45} : 2 10 17 3 15 (9 16) 12 5 4 11 13 1 6 7 18 14 8
λ_{21} : 2 4 1 8 10 3 14 11 18 7 5 6 15 (17 13) 9 12 16	λ_{46} : 4 1 9 15 16 11 13 2 (12 3) 7 18 17 8 10 6 14 5
λ_{22} : 12 7 8 9 10 16 14 17 3 5 11 (6 18) 13 15 4 2 1	λ_{47} : 2 17 15 (1 3 16) 6 11 12 18 13 10 (9 14) 5 8 7 4
λ_{23} : 11 17 6 4 2 7 18 14 16 9 8 15 1 3 13 12 10 5	λ_{48} : 11 4 1 6 2 3 10 9 18 15 16 5 (7 13) 17 14 8 12
λ_{24} : 2 5 9 16 4 7 10 13 14 11 6 8 17 18 1 15 3 12	λ_{49} : 10 2 14 1 15 13 (8 12) (3 11) 6 16 4 9 18 5 7 17
λ_{25} : 9 15 (5 8) (12 18) 2 14 16 13 17 10 (6 11) (3 4) 1 7	

Наличие упомянутых выше циклов не приводит к парадоксу Кондорсе для всего рассматриваемого профиля $V(N, n)$, т.к. в ранжированиях с β_1 по β_8 на первых четырех позициях расположены элементы 2, 11, 8, 17, которые обеспечивают превалирование альтернатив перестановки $\{2, 11, 17\}$ над остальными двумя перестановками из цикла. Если бы ранжирования $\beta_1, \dots, \beta_8$ отсутствовали в профиле $V(N, n)$, то порядок первых трех альтернатив во всех оставшихся 24 ранжированиях демонстрировал бы парадокс Кондорсе, который мог бы быть разрешен путем введения толерантностей между альтернативами, принадлежащими циклу, т.е. $\{2,$

11, 17} $\rightarrow$ {2 ~ 11 ~ 17}.

Таблица 2.6 – Выходной профиль $V(N, n) = V(32, 18)$ для примера 2.1

β	Номера альтернатив																	
β_1	2	11	8	17	4	7	1	3	5	9	15	12	10	14	6	18	13	16
β_2	2	11	8	17	4	7	1	3	9	5	15	12	10	14	6	18	13	16
β_3	2	11	8	17	4	7	3	5	9	12	18	1	15	10	14	6	13	16
β_4	2	11	8	17	4	7	3	5	9	12	18	1	15	10	14	13	6	16
β_5	2	11	8	17	4	7	3	5	9	12	18	1	15	10	14	13	16	6
β_6	2	11	8	17	4	7	3	9	5	12	18	1	15	10	14	6	13	16
β_7	2	11	8	17	4	7	3	9	5	12	18	1	15	10	14	13	6	16
β_8	2	11	8	17	4	7	3	9	5	12	18	1	15	10	14	13	16	6
β_9	2	11	17	4	7	8	1	3	5	9	15	12	10	14	6	18	13	16
β_{10}	2	11	17	4	7	8	1	3	9	5	15	12	10	14	6	18	13	16
β_{11}	2	11	17	4	7	8	3	5	9	12	18	1	15	10	14	6	13	16
β_{12}	2	11	17	4	7	8	3	5	9	12	18	1	15	10	14	13	6	16
β_{13}	2	11	17	4	7	8	3	5	9	12	18	1	15	10	14	13	16	6
β_{14}	2	11	17	4	7	8	3	9	5	12	18	1	15	10	14	6	13	16
β_{15}	2	11	17	4	7	8	3	9	5	12	18	1	15	10	14	13	6	16
β_{16}	2	11	17	4	7	8	3	9	5	12	18	1	15	10	14	13	16	6
β_{17}	11	17	2	4	7	8	1	3	5	9	15	12	10	14	6	18	13	16
β_{18}	11	17	2	4	7	8	1	3	9	5	15	2	10	14	6	18	13	16
β_{19}	11	17	2	4	7	8	3	5	9	12	18	1	15	10	14	6	13	16
β_{20}	11	17	2	4	7	8	3	5	9	12	18	1	15	10	14	13	6	16
β_{21}	11	17	2	4	7	8	3	5	9	12	18	1	15	10	14	13	16	6
β_{22}	11	17	2	4	7	8	3	9	5	12	18	1	15	10	14	6	13	16
β_{23}	11	17	2	4	7	8	3	9	5	12	18	1	15	10	14	13	6	16
β_{24}	11	17	2	4	7	8	3	9	5	12	18	1	15	10	14	13	16	6
β_{25}	17	2	11	4	7	8	1	3	5	9	15	12	10	14	6	18	13	16
β_{26}	17	2	11	4	7	8	1	3	9	5	15	12	10	14	6	18	13	16
β_{27}	17	2	11	4	7	8	3	5	9	12	18	1	15	10	14	6	13	16
β_{28}	17	2	11	4	7	8	3	5	9	12	18	1	15	10	14	13	6	16
β_{29}	17	2	11	4	7	8	3	5	9	12	18	1	15	10	14	13	16	6
β_{30}	17	2	11	4	7	8	3	9	5	12	18	1	15	10	14	6	13	16
β_{31}	17	2	11	4	7	8	3	9	5	12	18	1	15	10	14	13	6	16
β_{32}	17	2	11	4	7	8	3	9	5	12	18	1	15	10	14	13	16	6

Из дальнейшего анализа таблицы 2.6 следует, что выходной профиль $V(N, n)$ содержит много инверсий (противоположных порядков) пар альтернатив, которые трудно выявить визуально даже при не слишком большой размерности профиля.

Очевидно, что существует настоятельная необходимость в обоснованной

процедуре свёртки, т.е. преобразования всех N ранжирований выходного профиля $B(N, n)$ (оптимальных решений ЗРК) в единственное итоговое оптимальное решение β_{fin} . Решению этой проблемы посвящена третья глава диссертационной работы.

Для завершения рассмотрения примера 2.1 применим правило свёртки (см. п. 3.2.). Полученное итоговое ранжирование β_{fin} имеет следующий вид: $a_2 \succ a_{11} \succ a_{17} \succ a_4 \succ a_7 \sim a_8 \succ a_3 \succ a_5 \sim a_9 \succ a_{12} \succ a_1 \succ a_{18} \succ a_{15} \succ a_{10} \succ a_{14} \succ a_6 \sim a_{13} \succ a_{16}$.

Короткая форма записи β_{fin} выглядит следующим образом:

2 11 17 4 (7 8) 3 (5 9) 12 1 18 15 10 14 (6 13) 16.

Выводы к главе 2

1. Ранее проведенные экспериментальные исследования показали, что (1) число оптимальных решений N увеличивается с увеличением количества альтернатив n ; (2) множественность проявляется в значительно большей степени при чётных значениях m , чем при нечетных; (3) множественность резко увеличивается при уменьшении чётных значений m , но плавно увеличивается при уменьшении нечётных значений m ; (4) наличие толерантностей во входном профиле $\Lambda(m, n)$ влияет на множественность по-разному: увеличение плотности толерантностей t при нечётных m приводит к плавному увеличению N , в то время как для чётных m наоборот, N плавно уменьшается.

2. Предметом проведенных в ходе диссертационной работы экспериментальных исследований являлся фактор множественности, связанный с проявлением транзитивности/нетранзитивности входного профиля. Были получены следующие результаты: (1) абсолютные значения расстояний $D(\beta, \Lambda)$ и D_{least} увеличиваются с увеличением m и n ; (2) для заданных n и m , наиболее часто оказывались транзитивными профили *только* при *четных* значениях m ; *нечетные* значения m практически всегда приводили к нетранзитивным входным профилям; при этом во всех случаях число транзитивных профилей уменьшалось с возрастанием чис-

ла n альтернатив; (3) минимальная степень нетранзитивности достигается при малых чётных значениях m при отсутствии толерантностей; степень нетранзитивности уменьшается по мере уменьшения числа оптимальных решений ЗРК.

3. Числовые данные, полученные в ходе экспериментов, позволили построить оценки вероятностей появления транзитивного профиля при различных наборах параметров m , n , t , откуда следует, что чем больше m , тем меньше вероятность транзитивности. При этом (1) при любых нечетных m вероятность транзитивности входного профиля близка к нулю уже при числе альтернатив $n \approx 10$; (2) вероятность транзитивности значительно выше при *четных* значениях m , чем при нечетных, и плавно уменьшается с ростом n .

4. Полученные оценки вероятности появления нетранзитивности входного профиля находятся в полном согласии с известными аналитическими оценками вероятности альтернативы Кондорсе и оценками вероятности появления множественности решений ЗРК. Полученные оценки соответствуют утверждениям, что при нечетных m вероятность появления нетранзитивного профиля всегда больше или равна вероятности парадокса Кондорсе, а при четных m вероятность появления нетранзитивного профиля всегда меньше вероятности парадокса Кондорсе.

ГЛАВА 3

ПРЕОБРАЗОВАНИЕ ВЫХОДНОГО ПРОФИЛЯ В ЕДИНСТВЕННОЕ ОПТИМАЛЬНОЕ РЕШЕНИЕ

Выходной профиль представляет собой набор всех ранжирований Кемени, который может быть представлен взвешенной турнирной матрицей. В этой главе исследуются свойства взвешенной турнирной матрицы, построенной для выходного профиля; а также соответствующие понятия выигрыша, проигрыша и ранга. Формулируются правила свёртки, т.е. преобразования выходного профиля в единственное оптимальное решение. Обсуждаются справедливость и обоснованность предлагаемых правил свёртки на основе правила Борда и понятия отношения "между", заданного на ранжированиях.

3.1 Взвешенный турнир выходного профиля

3.1.1 Матрица взвешенного турнира

Пусть $V(N, n) = \{\beta_1, \beta_2, \dots, \beta_N\}$, $V \subset \Pi_n$, является выходным профилем ЗРК с входным профилем $\Lambda(m, n) = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$, заданным на некотором наборе альтернатив $A = \{a_1, a_2, \dots, a_n\}$. Каждое ранжирование консенсуса β_k , $k = 1, \dots, N$, является полным и асимметричным бинарным отношением (т.е. антисимметричным и антирефлексивным, см. рисунок 1.1), т.е. отношением строгого предпочтения, представленным соответствующей перестановкой первых n натуральных чисел $\mathbb{N}_n = \{1, 2, \dots, n\}$. Ясно, что ранжирование β_k не может содержать никаких толерантностей [42].

Как и при рассмотрении турнирной матрицы входного профиля в п. 1.2.3., каждое строгое ранжирование β_k в выходном профиле можно рассматривать как круговой турнир.

Если игроки встречаются N раз, турнир называется N -взвешенным. Выходной профиль $V(N, n)$ может быть представлен N -взвешенной турнирной $(n \times n)$ матрицей $S = [s_{ij}]$, которая определяется выражением (1.13), но для случая N стро-

гих ранжирований, т.е.

$$s_{ij} = \sum_{k=1}^N \mathbf{I}_k(a_i, a_j), i \neq j = 1, \dots, n, \quad (3.1)$$

является элементом матрицы,

$$\mathbf{I}_k(a_i, a_j) = \begin{cases} 1, & \text{если } (a_i, a_j) \in \beta_k \\ 0, & \text{если } (a_i, a_j) \notin \beta_k \end{cases} \quad (3.2)$$

является индикаторной функцией. Для индикаторной функции (3.2) справедливы следующие утверждения для всех $k = 1, \dots, m$:

$$(a_i, a_j) \in \beta_k \Leftrightarrow a_i \succ_k a_j \text{ и } (a_i, a_j) \notin \beta_k \Leftrightarrow a_i \prec_k a_j. \quad (3.3)$$

Очевидно, что для турнирной матрицы S выполняется следующее свойство:

$$s_{ij} + s_{ji} = N \text{ для } i \neq j; i, j = 1, \dots, n, \quad (3.4)$$

где значение s_{ij} равно доле N , которую альтернатива a_i приобретает, когда превалирует (доминирует или побеждает) в сравнении с альтернативой a_j . Таким образом, матрица S содержит информацию о том, насколько сильна альтернатива a по сравнению с любой другой альтернативой после проведения N турниров β_k , $k = 1, \dots, N$ [39].

3.1.2 Параметры взвешенного турнира

Определение 3.1. Сумма элементов i -й строки матрицы S называется *выигрышем* z_i альтернативы a_i для $i = 1, \dots, n$ и определяется по формуле:

$$z_i = \sum_{j=1}^n s_{ij}. \quad (3.5)$$

Выигрыш z_i показывает общий доход от всех выигрышей альтернативы a_i во всех $N(n-1)$ сравнениях.

Определение 3.2. Сумма элементов j -го столбца матрицы S называется *проигрышем* y_j альтернативы a_j для $j = 1, \dots, n$ и определяется как

$$y_j = \sum_{i=1}^n s_{ij}. \quad (3.6)$$

Проигрыш y_j показывает общий убыток от всех проигрышей альтернативы a_j во всех $N(n-1)$ сравнениях.

Выигрыш и проигрыш альтернативы связаны друг с другом, как показано в следующем предложении.

Предложение 3.1. Для любого выходного профиля $V(N, n)$, представленного соответствующей турнирной матрицей $[s_{ij}]$, справедливо следующее выражение:

$$y_j + z_i = N(n - 1), i = 1, \dots, n. \quad (3.7)$$

Доказательство. Используя условие (3.2) и утверждение (3.3) для некоторой фиксированной альтернативы $a_q \in A$, из выражения (3.1) можно записать:

$$\underbrace{(s_{1q} + s_{q1}) + (s_{2q} + s_{q2}) + \dots + (s_{iq} + s_{qi}) + \dots + (s_{nq} + s_{qn})}_{(n-1) \text{ раз}} = \underbrace{N + N + \dots + N}_{(n-1) \text{ раз}} = N(n - 1). \quad (3.8)$$

В выражении (3.8), при суммировании компонентов в скобках, случай, где $i = j = q$ должен быть исключен ввиду антирефлексивности β_k , поэтому число n этих компонентов следует уменьшить на 1. С учетом этого, выражение (3.8) доказывает справедливость выражения (3.7). ♦

Определение 3.3. Пусть ранг r_i^k является позицией альтернативы a_i в ранжировании консенсуса $\beta_k \in V(N, n)$, $k = 1, \dots, N$. Будем называть *общий ранг* r_i альтернативы a_i сумму ее рангов во всех ранжированиях консенсуса, т.е.

$$r_i = \sum_{k=1}^N r_i^k. \quad (3.9)$$

Ясно, что любой выходной профиль $V(N, n)$ может быть однозначно представлен таблицей рангов по форме, представленной в таблице 3.1.

Выигрыш и проигрыш альтернативы связаны с общим рангом альтернативы как показано в следующем предложении 3.2.

Предложение 3.2. Для любого выходного профиля $V(N, n)$, представленного соответствующей турнирной матрицей $[s_{ij}]$, справедливо следующее выражение:

$$r_i = nN - z_i, i = 1, \dots, n. \quad (3.10)$$

Доказательство. С учетом свойства (3.4), достаточно показать справедливость следующего выражения:

$$r_i = N + y_i, i = 1, \dots, n \quad (3.11)$$

Из определения ранга r_i^k как позиции альтернативы a_i в ранжировании β_k следует, что

$$r_i^k = |\{a_j \in A \mid a_j \succ_k a_i\}| + 1, a_i \in A. \quad (3.12)$$

Тогда, с учетом выражений (3.2) и (3.3), для любой фиксированной альтернативы $a_q \in A$ справедливо следующее равенство (см. таблицу 3.2):

$$r_q^k - 1 = \sum_{i \neq q}^n \mathbf{I}_k(a_i, a_q). \quad (3.13)$$

Таблица 3.1 – Таблица рангов выходного профиля $V(N, n)$

Ранг r_i^k	Альтернативы					
	a_1	a_2	...	a_i	...	a_n
r_i^1	r_1^1	r_2^1	...	r_i^1	...	r_n^1
r_i^2	r_1^2	r_2^2	...	r_i^2	...	r_n^2
...	...	...	...	...	...	...
r_i^N	r_1^N	r_2^N	...	r_i^N	...	r_n^N
Общий ранг $r_i = \sum_{k=1}^N r_i^k$	$r_1 = \sum_{k=1}^N r_1^k$	$r_2 = \sum_{k=1}^N r_2^k$	...	$r_i = \sum_{k=1}^N r_i^k$	...	$r_n = \sum_{k=1}^N r_n^k$

Таблица 3.2 – К обоснованию выражения (3.11) для фиксированной альтернативы a_q

$r_i^k - 1$	$i \neq q$					
	$\mathbf{I}_k(a_1, a_q)$	$\mathbf{I}_k(a_2, a_q)$	...	$\mathbf{I}_k(a_i, a_q)$	...	$\mathbf{I}_k(a_n, a_q)$
$r_i^1 - 1$	$\mathbf{I}_1(a_1, a_q)$	$\mathbf{I}_1(a_2, a_q)$	...	$\mathbf{I}_1(a_i, a_q)$	...	$\mathbf{I}_1(a_n, a_q)$
$r_i^2 - 1$	$\mathbf{I}_2(a_1, a_q)$	$\mathbf{I}_2(a_2, a_q)$	...	$\mathbf{I}_2(a_i, a_q)$	...	$\mathbf{I}_2(a_n, a_q)$
...	...	...	...	...	...	...
$r_i^N - 1$	$\mathbf{I}_N(a_1, a_q)$	$\mathbf{I}_N(a_2, a_q)$	...	$\mathbf{I}_N(a_i, a_q)$	...	$\mathbf{I}_N(a_n, a_q)$
s_{iq}	$s_{1q} = \sum_{k=1}^N \mathbf{I}_k(a_1, a_q)$	$s_{2q} = \sum_{k=1}^N \mathbf{I}_k(a_2, a_q)$	...	$s_{iq} = \sum_{k=1}^N \mathbf{I}_k(a_i, a_q)$	...	$s_{nq} = \sum_{k=1}^N \mathbf{I}_k(a_n, a_q)$

Просуммируем левую и правую части выражения (3.13) по всем ранжированиям β_k . Получаем выражение $\sum_{k=1}^N (r_q^k - 1) = \sum_{k=1}^N \sum_{i \neq q}^n \mathbf{I}_k(a_i, a_q)$, откуда с учетом (3.1), (3.2), (3.6) и (3.9) следует, что $r_q - N = \sum_{i \neq q}^n s_{iq} = y_q$, что доказывает справед-

ливость (3.11). ♦

Проиллюстрируем справедливость предложений 3.1 и 3.2 на примере.

Пример 3.1. Пусть задан входной профиль $\Lambda(m, n) = \Lambda(3, 5)$, представленный таблице 3.3. в короткой форме (1.2). В этой же таблице показан соответствующий выходной профиль $B(N, n) = B(12, 5)$, найденный алгоритмом RECURSALL, реализующим правило Кемени. Оказалось, что расстояние $D(\beta, \Lambda) = 24$ для всех $k = 1, \dots, N$.

Таблица 3.3 – Пример входного профиля $\Lambda(m, n) = \Lambda(3, 5)$ и соответствующий ему выходной профиль $B(N, n) = B(12, 5)$

Входной профиль $\Lambda(3, 5)$		Выходной профиль $B(12, 5)$			
λ_k	Ранжирование	β_k	Ранжирование	β_k	Ранжирование
λ_1	(23)(145)	β_1	2 3 1 4 5	β_7	3 2 1 4 5
λ_2	(12)(345)	β_2	2 3 1 5 4	β_8	3 2 1 5 4
λ_3	(345)(12)	β_3	2 3 4 1 5	β_9	3 2 4 1 5
		β_4	2 3 4 5 1	β_{10}	3 2 4 5 1
		β_5	2 3 5 1 4	β_{11}	3 2 5 1 4
		β_6	2 3 5 4 1	β_{12}	3 2 5 4 1

В таблице 3.4 показана турнирная матрица S для выходного профиля $B(12, 5)$ и соответствующие значения выигрышей z_i и проигрышей y_j .

Таблица 3.4 – Турнирная матрица для $B(N, n) = B(12, 5)$ для примера 3.1

	a_1	a_2	a_3	a_4	a_5	z_i
a_1	0	0	0	6	6	12
a_2	12	0	6	12	12	42
a_3	12	6	0	12	12	42
a_4	6	0	0	0	6	12
a_5	6	0	0	6	0	12
y_j	36	6	6	36	36	

Таблица рангов выходного профиля $B(N, n)$ показана в таблице 3.5. Параметры турнирной матрицы показаны в таблице 3.6, откуда следует справедливость предложения 3.1. Справедливость Предложения 3.2 проиллюстрирована в таблице 3.7 для фиксированной альтернативы $a_q = a_1$, где $r_i - 12 = s_{21} + s_{31} + s_{41} + s_{51} = 12 + 12 + 6 + 6 = 36 = y_i$.

Таблица 3.5 – Таблица рангов выходного профиля $B(12, 5)$ для примера 3.1

r_i^k	a_1	a_2	a_3	a_4	a_5
r_i^1	3	1	2	4	5
r_i^2	3	1	2	5	4
r_i^3	4	1	2	3	5
r_i^4	5	1	2	3	4
r_i^5	4	1	2	5	3
r_i^6	5	1	2	4	3
r_i^7	3	2	1	4	5
r_i^8	3	2	1	5	4
r_i^9	4	2	1	3	5
r_i^{10}	5	2	1	3	4
r_i^{11}	4	2	1	5	3
r_i^{12}	5	2	1	4	3
$r_i = \sum_{k=1}^N r_i^k$	48	18	18	48	48

Таблица 3.6 – Параметры выходного профиля $B(N, n) = B(12, 5)$ для примера 3.1

Альтернативы	a_1	a_2	a_3	a_4	a_5	Сумма
Выигрыш z_i	12	42	42	12	12	$\sum_{i=1}^n z_i = 120$
Проигрыш y_i	36	6	6	36	36	$\sum_{i=1}^n y_i = 120$
Общий ранг r_i	48	18	18	48	48	$\sum_{i=1}^n r_i = 180$
$nN = z_i + r_i$	60	60	60	60	60	$\sum_{i=1}^n r_i - nN = 180 - 60 = 120$
$y_i + z_i = N(n - 1)$	48	48	48	48	48	$n^2 N - \sum_{i=1}^n r_i = 300 - 180 = 120$

Из таблицы 3.7 можно также видеть, что каждый элемент второго столбца равен элементу третьего столбца, который, в свою очередь, равен сумме элементов столбцов с четвертого по седьмой. Очевидна справедливость Предложения 3.2 для всех остальных альтернатив $a_2, \dots, a_5$. Дальнейшее рассмотрение примера представлено в п. 3.2.4 и в п. 4.2.2. ♦

Таблица 3.7 – Демонстрация справедливости Предложения 3.2 для фиксированной альтернативы $a_q = a_1$ для примера 3.1

k	$r_1^k - 1$	$\sum_{i=2}^5 \mathbf{I}_k(a_i, a_1)$	$\mathbf{I}_k(a_2, a_1)$	$\mathbf{I}_k(a_3, a_1)$	$\mathbf{I}_k(a_4, a_1)$	$\mathbf{I}_k(a_5, a_1)$
1	$r_1^1 - 1$	2	1	1	0	0
2	$r_1^2 - 1$	2	1	1	0	0
3	$r_1^3 - 1$	3	1	1	1	0
4	$r_1^4 - 1$	4	1	1	1	1
5	$r_1^5 - 1$	3	1	1	0	1
6	$r_1^6 - 1$	4	1	1	1	1
7	$r_1^7 - 1$	2	1	1	0	0
8	$r_1^8 - 1$	2	1	1	0	0
9	$r_1^9 - 1$	3	1	1	1	0
10	$r_1^{10} - 1$	4	1	1	1	1
11	$r_1^{11} - 1$	3	1	1	0	1
$N = 12$	$r_1^{12} - 1$	4	1	1	1	1
Сумма	$\sum_{k=1}^{12} (r_1^k - 1)$	$s_{i1} = \sum_{k=1}^{12} \mathbf{I}_k(a_i, a_1)$	$s_{21} = 12$	$s_{31} = 12$	$s_{41} = 6$	$s_{51} = 6$
	$p_1 - 12$	$y_1 = \sum_{i=2}^5 s_{i1}$	36			

Предложения 3.1 и 3.2 устанавливают взаимосвязь между характеристиками отдельных альтернатив, информация о которых содержится в турнирной матрице и таблице рангов. Приведенные ниже очевидные следствия из этих предложений указывают на существование аналогичной взаимосвязи между суммарными характеристиками всех альтернатив.

Следствие 1. Для любого выходного профиля $B(N, n)$, представленного соответствующей турнирной матрицей $[s_{ij}]$, справедливы следующие выражения:

$$\sum_{i=1}^n z_i = n^2 N - \sum_{i=1}^n r_i = \sum_{i=1}^n y_i = \sum_{i=1}^n r_i - nN. \quad (3.14)$$

Следствие 2. Сумма всех элементов турнирной матрицы $[s_{ij}]$ определяется следующим образом:

$$\sum_{i=1}^n \sum_{j=1}^n s_{ij} = nN(n-1)/2. \quad (3.15)$$

Следствие 3. Сумма всех общих рангов любого выходного профиля $B(N, n)$

выглядит следующим образом:

$$\sum_{i=1}^n r_i = nN(n+1)/2. \quad (3.16)$$

3.2 Построение правила свёртки

3.2.1 Характеризация выходного профиля ЗРК

Входной профиль предпочтений $\Lambda(m, n)$, очевидно, носит случайный характер. В отличие от него, выходной профиль $V(N, n)$, который выполняет роль входных данных для правила свёртки (см. рисунок 1.6), характеризуется рядом особых свойств:

(i) каждое ранжирование $\beta_k \in V(N, n)$ является оптимальным решением для входного профиля $\Lambda(m, n)$ по правилу Кемени (1.31); это значит, что все ранжирования находятся на одинаковом расстоянии $D(\beta_k, \Lambda)$ от профиля $\Lambda(m, n)$, см. выражение (1.25), однако они могут находиться на разном расстоянии друг от друга;

(ii) каждое ранжирование β_k является строгим порядком, следовательно, оно не включает толерантности; однако, как отношение слабого порядка, оно обладает свойством рефлексивности, т.е. его матрица L (см. п. 1.2.1) имеет все «единицы» на главной диагонали;

(iii) профиль $V(N, n)$ не содержит *клонов*, т.е. не имеет двух одинаковых ранжирований;

(iv) все ранжирования β_k в профиле $V(N, n)$ лексикографически упорядочены, поскольку они были сформированы в процессе построения узлов дерева поиска методом ветвей и границ.

Из положений теории голосования, см. например, [57,101] известно, что каждое ранжирование β_k , как результат применения правила Кемени к профилю $\Lambda(m, n)$, обладает следующим набором свойств (в литературе по теории голосования их называют также условиями, аксиомами, принципами, критериями или

требованиями¹):

- полнота выбора (non imposition, citizen sovereignty),
- критерий Кондорсе (Condorcet criterium),
- критерий большинства (majority criterium),
- отсутствие диктатура (non-dictatorship),
- универсальность (universality, unrestricted domain),
- эффективность по Парето или единогласие (Pareto efficiency, unanimity),
- монотонность (строгая) (monotonicity),
- критерий Смита (Smith criterium),
- независимость от проигравших по Смиту (independence of Smith-dominated alternatives, independence of Smith losers),
- пополнение, согласованность (reinforcement, consistency),
- обратная симметрия (reversal symmetry).

В то же время, каждое ранжирование β_k , полученное по правилу Кемени, *не удовлетворяет* следующим известным свойствам:

- независимость от посторонних альтернатив (independence of irrelevant alternatives),
- участие (participation),
- последующее отсутствие ущерба (later-no-harm),
- независимость от клонов (independence of clones),
- неуязвимость к выталкиванию (invulnerability to push-over),
- Принцип Шварца (Schwartz criterium).

Ясно, что правило Кемени, как и все другие известные правила агрегирования предпочтений, подтверждает справедливость хорошо известной теоремы о невозможности Эрроу [30,50], которая утверждает, что *не существует правила агрегирования предпочтений, которое одновременно удовлетворяло бы трём естественным условиям*: (1) эффективность по Парето (или единогласие), (2) отсутствие диктатуры и (3) независимость от посторонних альтернатив.

¹ Для краткости здесь не приводятся известные формулировки этих критериев, их можно найти, например, в [65,99].

Заметим, что в работе [97] Г. Янг (H.P. Young) показал, что, хотя решение по правилу Кемени не удовлетворяет условию независимости от посторонних альтернатив, вместо этого оно удовлетворяет более слабое свойство локальной независимости от посторонних альтернатив.

Также следует обратить внимание на то, что нет никаких гарантий того, что выполнение вышеуказанных критериев для каждого оптимального решения β_k распространяется и на результат их свёртки, т.е. итоговое ранжирование консенсуса β_{fin} . Очевидно, что свойства β_{fin} будут определяться особенностями правила, которое мы выбираем для свёртки. В связи с этим рассмотрим требования к правилу свёртки.

3.2.2 Желательные свойства правила свёртки

Опираясь на идеи аксиоматического подхода, развитого Эрроу [65], будем полагать, что правило свёртки, которое преобразует выходной профиль $V(N, n)$ в итоговое ранжирование консенсуса β_{fin} , должно иметь следующие естественные свойства:

- (P1) ранжирование β_{fin} может включать толерантности;
- (P2) правило свёртки должно быть совместимым с правилами Кондорсе и Кемени;
- (P3) итоговое ранжирование консенсуса β_{fin} должно находиться *между* оптимальными ранжированиями β_k и
- (P4) вычислительная сложность процедуры свёртки должна быть полиномиальной.

Вероятно, лучшим кандидатом на роль правила свёртки было бы правило Кондорсе. Однако от этого варианта приходится отказаться, поскольку вероятность парадокса Кондорсе сохраняется не только для входного профиля $\Lambda(m, n)$, но и для выходного профиля $V(N, n)$, как показано в п. 2.2.4.

Рассмотрим возможность использования правила Борда в качестве правила свёртки.

Г. Янг в работе [98] показал, что правило Борда – это единственное правило агрегирования предпочтений, которое удовлетворяет следующим свойствам:

- анонимность (anonymity)
- нейтральность (neutrality)
- пополнение (reinforcement)
- достоверность (faithfulness)
- отмена (cancellation).

Свойство отмены означает, что если все сравнения по правилу большинства приводят к толерантности, то в качестве победителей выбираются все соответствующие альтернативы [13,65]. Это означает, что свойство отмены согласуется с нашим условием (P1).

Заметим, что для правила Борда, также как и для правила Кемени, не выполняется условие независимости от клонов (independence of clones). Однако из свойства (iii) выходного профиля $B(N, n)$ следует, что этот факт не оказывает влияния на получаемое решение.

Известно, см., например, [57,65], что балл Борда (Borda score), см. п. 1.3.2, $Bo(a_i)$ можно вычислить, используя результаты парных сравнений альтернатив, содержащихся в матрице взвешенного турнира S , т.е.

$$Bo(a_i) = z_i, \quad (3.17)$$

значит балл Борда и выигрыш z_i совпадают, см. формулу (3.7).

Очевидно, что победитель Борда может не совпадать с победителем Кондорсе. Но весьма благоприятным фактом для правила Борда в контексте его использования для свёртки множества оптимальных решений является то, что победитель Кондорсе никогда не является проигравшим по Борда. Кроме того, проигравший по Кондорсе, т.е. альтернатива, уступившая в парном сравнении всем другим альтернативам, никогда не будет победителем по Борда [49]. Таким образом, можно утверждать, что правило Борда и правило Кондорсе согласованы. В работе [90] доказано, что аналогичное соответствие существует между правилами Борда и Кемени: счет Борда всегда ранжирует победителя по Кемени выше про-

игравшего по Кемени, и, наоборот, правило Кемени всегда ранжирует победителя по Борда выше проигравшей альтернативы по Борда. Эта согласованность является уникальным свойством счета Борда среди всех других позиционных правил. На основании сказанного, можно сделать вывод, что условие (P2) выполняется, если для свёртки используется правило Борда.

Выполнение условия (P3) будет рассмотрено в п. 3.2.5. Ясно, что правило Борда удовлетворяет условию (P4), поскольку оценка его вычислительной сложности равна $O(n^2)$, что следует из простого рассмотрения уравнения (3.7).

Учитывая свойства (P1) - (P4), сформулируем два правила, каждое из которых позволяет обоснованно свести все N оптимальных решений β_k к некоторому подходящему единственному итоговому ранжированию β_{fin} . Для этой цели будем использовать характеристики выходного профиля, описанные в п. 3.1.2.

3.2.3 Свёртка на основе выигрышей z_i

Заметим, что альтернативы множества A могут быть упорядочены в соответствии с их "силой", которая выражается для каждой альтернативы выигрышем z_i , показывающим общий доход от всех выигрышей альтернативы a_i во всех $N(n-1)$ сравнениях. Чем больше выигрыш, тем сильнее альтернатива. Две альтернативы толерантны друг к другу, если они имеют одинаковый выигрыш.

Тогда правило свёртки множества оптимальных перестановок $B(N, n)$ можно сформулировать следующим образом.

Правило свёртки 1. Для заданных выигрышей z_i , вычисленных по взвешенной турнирной матрице S выходного профиля $B(N, n)$, для всех $i < j$, $i, j = 1, \dots, n$, справедливы следующие соотношения:

$$z_i > z_j \Rightarrow a_i \succ a_j \text{ и } z_i = z_j \Rightarrow a_i \sim a_j, \quad (3.18)$$

где оба отношения $\succ$ и $\sim$ принадлежат единственному итоговому отношению консенсуса β_{fin} . ♦

Таким образом, итоговое ранжирование консенсуса может включать толе-

рантности, также как произвольное исходное ранжирование $\lambda_k \in \Lambda$, т.е., вообще говоря, $\beta_{\text{fin}} = \rho \cup \tau$ и $\beta_{\text{fin}} \notin \Pi_n$ (это означает что единственное итоговое ранжирование консенсуса может не быть строгим порядком). Выигрыши z_i , реализующие баллы Борда $\text{Bo}(a_i)$, гарантируют, что правило (3.18) преобразует выходной профиль $V(N, n)$ в единственное ранжирование консенсуса β_{fin} , удовлетворяющее свойствам (P1) – (P4).

3.2.4 Свёртка с использованием общих рангов r_i

Из предложения 3.2 следует, что для каждой альтернативы a_i ее общий ранг r_i является *дополнением* до значения nN на величину выигрыша z_i . Следовательно, можно построить правило свёртки множества оптимальных перестановок $V(N, n)$ следующим образом.

Правило свёртки 2. Для заданных общих рангов r_i , вычисленных по взвешенной турнирной матрице S выходного профиля $V(N, n)$, для всех $i < j$, $i, j = 1, \dots, n$, справедливы следующие соотношения:

$$r_i > r_j \Rightarrow a_i \succ a_j \text{ и } r_i = r_j \Rightarrow a_i \sim a_j, \quad (3.19)$$

где оба отношения $\succ$ и $\sim$ принадлежат единственному итоговому отношению консенсуса β_{fin} . ♦

Из уравнения (3.10) следует, что баллы Борда могут быть рассчитаны как $\text{Bo}(a_i) = nN - r_i$. Понятно, что Правила свёртки 1 и 2 гарантированно приводят к одному и тому же итоговому ранжированию консенсуса β_{fin} . При реализации процедуры свёртки можно выбрать одно из правил, исходя из любых практических соображений, например, вычислительных или алгоритмических.

Пример 3.1 (продолжение). Из таблицы 3.6 для выходного профиля $V(12, 5)$ имеем выигрыши $z_1 = 12$, $z_2 = 42$, $z_3 = 42$, $z_4 = 12$, $z_5 = 12$, и общие ранги $r_1 = 48$, $r_2 = 18$, $r_3 = 18$, $r_4 = 48$, $r_5 = 48$. В соответствии с правилом свёртки 1, уравнением (3.18) или его дополнением (правило свёртки 2), уравнением (3.19) итоговым решением ЗРК для входного профиля $\Lambda(3, 5)$ и его выходного профиля $V(12,$

5) для таблицы 3.3 является $\beta_{\text{fin}} = (23)(145) = a_2 \sim a_3 \succ a_1 \sim a_4 \sim a_5$. ♦

3.2.5 Расстояние Кемени и итоговое ранжирование консенсуса

Поскольку правило Кемени основано на понятии расстояния Кемени (см. 1.3.4), которое, в свою очередь, происходит от расстояний Кендалла и Хэмминга [34,55], обсудим выполнение условия (Р3) и исследуем результат β_{fin} введенных выше правил свёртки в терминах понятия отношения "между".

Рассмотрим ранжирования λ , λ_1 и λ_2 , которые являются строгими порядками, т.е. не содержат симметричной части. Их можно рассматривать как точки пространства слабого порядка, характеризуемого расстоянием Кемени (см. выражение (1.25)). Следуя классическому понятию отношения "между", заданного на ранжированиях [16,47,60,67,68], с учетом справедливости метрического свойства $d(\lambda_1, \lambda_2) = d(\lambda_1, \lambda) + d(\lambda, \lambda_2)$ можно утверждать, что ранжирование λ находится между ранжированиями λ_1 и λ_2 , т.е. $\lambda \in [\lambda_1, \lambda_2]$, если и только если

$$\lambda_1 \cap \lambda_2 \subseteq \lambda \subseteq \lambda_1 \cup \lambda_2. \quad (3.20)$$

Пример 3.2. Пусть даны два ранжирования $\lambda_1 = a_1 \succ a_2 \succ a_3$ и $\lambda_2 = a_1 \succ a_3 \succ a_2$. Рассмотрим ранжирование $\lambda = a_1 \succ a_2 \sim a_3$, которое лежит между λ_1 и λ_2 . Соответствующие точки в пространстве слабого порядка при $n = 3$ показаны на рисунке 1.3. В таблице 3.8 приведены матрицы $L(\lambda_1)$, $L(\lambda_2)$ и $L(\lambda)$ (матрица L определена выражением (1.4) в п. 1.2.1) рассматриваемых отношений λ_1 , λ_2 и λ . Из таблицы 3.8 видно, что $\lambda = \lambda_1 \cup \lambda_2$ и $\lambda_1 \cap \lambda_2 \subseteq \lambda$, т.е. выполняется условие (3.20). Очевидно также, что $d(\lambda_1, \lambda_2) = 2$ и $d(\lambda, \lambda_1) = d(\lambda, \lambda_2) = 1$. ♦

Таблица 3.8 – Матрицы ранжирований λ_1 , λ_2 и λ из примера 3.2, и матрицы их объединения и пересечения

$L(\lambda_1)$	$L(\lambda_2)$	$L(\lambda)$	$L(\lambda_1 \cap \lambda_2)$	$L(\lambda_1 \cup \lambda_2)$
$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

В работе [16] показано, что определение отношения "между" естественным образом обобщается на случай произвольного числа p ранжирований, т.е.

$$\lambda \in [\lambda_1, \dots, \lambda_p] \bigcap_{k=1}^p \lambda_k \subseteq \lambda \subseteq \bigcup_{k=1}^p \lambda_k, \quad (3.21)$$

откуда следует, что итоговое ранжирование консенсуса β_{fin} находится между ранжированиями $\beta_1, \beta_2, \dots, \beta_N$ выходного профиля $V(N, n)$, если и только если

$$\bigcap_{k=1}^N \beta_k \subseteq \beta_{\text{fin}} \subseteq \bigcup_{k=1}^N \beta_k. \quad (3.22)$$

Из определения правила Кемени (1.31) следует, что выражение, подобное выражению (3.22), также справедливо для входного профиля $\Lambda(m, n)$, т.е.:

$$\bigcap_{k=1}^m \lambda_k \subseteq \beta_{\text{fin}} \subseteq \bigcup_{k=1}^m \lambda_k. \quad (3.23)$$

Продemonстрируем справедливость уравнений (3.22) и (3.23), используя данные из примера 3.1.

Пример 3.3. В таблице 3.9 показаны матрицы всех ранжирований входного профиля $\Lambda(m, n) = \Lambda(3, 5)$, а также матрицы их объединения и пересечения и матрица для β_{fin} . Видно, что условие (3.23) выполнено. В таблице 3.10 представлены матрицы всех ранжирований выходного профиля $V(N, n) = V(12, 5)$, а также матрицы их объединения и пересечения. Заметим, что, как и в примере 3.2, где $\lambda = \lambda_1 \cup \lambda_2$, в этом примере мы имеем подобный результат для всех оптимальных ранжирований β_k , т.е. $\beta_{\text{fin}} = \bigcup_{k=1}^N \beta_k$, и условие (3.22) выполняется. ♦

Из выражения (3.22) следует, что расстояние Кемени от итогового отношения консенсуса β_{fin} до входного профиля Λ не превышает оптимального (минимального) расстояния от любого из строгих ранжирований β_k , т. е.

$$D(\beta_{\text{fin}}, \Lambda) \leq D(\beta_k, \Lambda), \quad k = 1, \dots, N. \quad (3.24)$$

С учетом выражения (1.26), расстояние от β_{fin} до входного профиля Λ определяется формулой:

$$D(\beta_{\text{fin}}, \Lambda) = \sum_{k=1}^m \sum_{i < j} |\text{sgn}(a_i^{\text{fin}}, a_j^{\text{fin}}) - \text{sgn}(a_i^k, a_j^k)|. \quad (3.25)$$

Проверим справедливость выражения (3.24) с использованием данных из примера 3.1.

Таблица 3.9 – Матрицы для λ_1 , λ_2 и λ_3 из Примера 3.1 и матрицы их объединения и пересечения, а также β_{fin}

$L(\lambda_1)$	$L(\lambda_2)$	$L(\lambda_3)$	$L\left(\bigcap_{k=1}^3 \lambda_k\right)$	$L\left(\bigcup_{k=1}^3 \lambda_k\right)$	$L(\beta_{\text{fin}})$
$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}$

Таблица 3.10 – Матрицы отношения консенсуса $\beta_1, \dots, \beta_N$ для Примера 3.1, а также матрицы их пересечения и объединения

$L(\beta_1)$	$L(\beta_2)$	$L(\beta_3)$	$L(\beta_4)$	$L(\beta_5)$	$L(\beta_6)$	$L(\beta_7)$
$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$
$L(\beta_8)$	$L(\beta_9)$	$L(\beta_{10})$	$L(\beta_{11})$	$L(\beta_{12})$	$L\left(\bigcap_{k=1}^{12} \beta_k\right)$	$L\left(\bigcup_{k=1}^{12} \beta_k\right)$
$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}$

Пример 3.4. Значения функций $\text{sgn}(a_i^k, a_j^k)$ для входного профиля $\Lambda(3, 5)$ сводятся к Таблице 3.11. Расчеты с использованием (3.25) дают: $d(\beta_{\text{fin}}, \lambda_1) = 0$; $d(\beta_{\text{fin}}, \lambda_2) = 8$; $d(\beta_{\text{fin}}, \lambda_3) = 10$; $D(\beta_{\text{fin}}, \Lambda) = d(\beta_{\text{fin}}, \lambda_1) + d(\beta_{\text{fin}}, \lambda_2) + d(\beta_{\text{fin}}, \lambda_3) = 0 + 8 + 10 = 18$. Поскольку $D(\beta_k, \Lambda) = 24$, видно, что условие (3.25) выполнено. ♦

Продemonстрируем пример выполнения свёртки при решении задачи о ранжировании Кемени (ЗРК) для входного профиля достаточно большой размерности. Число $m = 49$ выбрано нечетным, так как известно, что четные числа m приводят к слишком большим количествам N оптимальных решений ЗРК [71]. Число $n = 18$ принято таким, чтобы быть достаточно близким к верхней границе (равной 20) числа альтернатив, при которой NP -трудная ЗРК может быть решена за

приемлемое время (см. п. 1.4.2).

Таблица 3.11 – К расчету $D(\beta_{\text{fin}}, \Lambda)$ для входного профиля $\Lambda(m, n)$ из Примера 3.1.

$\text{sgn}(a_i^k, a_j^k)$			$\text{sgn}(a_i^{\text{fin}}, a_j^{\text{fin}})$
$\text{sgn}(a_i^1, a_j^1)$	$\text{sgn}(a_i^2, a_j^2)$	$\text{sgn}(a_i^3, a_j^3)$	
$\begin{bmatrix} 0 & -1 & -1 & 0 & 0 \\ & 0 & 0 & 1 & 1 \\ & & 0 & 1 & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ & 0 & 1 & 1 & 1 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & -1 & -1 & -1 \\ & 0 & -1 & -1 & -1 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & -1 & 0 & 0 \\ & 0 & 0 & 1 & 1 \\ & & 0 & 1 & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}$

Пример 3.5. Входной профиль $\Lambda(m, n) = \Lambda(49, 18)$; полученный алгоритмом RECURSALL выходной профиль $B(N, n) = B(32, 18)$. Окончательное решение β_{fin} , полученное по правилам свёртки 1 и 2, имеет вид

2 11 17 4 (7 8) 3 (5 9) 12 1 18 15 10 14 (6 13) 16.

Расстояния $D(\beta_k, \Lambda) = 6780$, $D(\beta_{\text{fin}}, \Lambda) = 6764$, т.е. условие (3.24) выполняется. Параметры выходного профиля представлены в таблицах 2.5, 2.6 и 3.12.

Таблица 3.12 – Параметры выходного профиля $B(N, n) = B(32, 18)$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}
z_i	232	520	344	440	296	48	408	408	296	136	512	240	48	104	176	8	496	184
y_i	312	24	200	104	248	496	136	136	248	408	32	304	496	440	368	536	48	360
r_i	344	56	232	136	280	528	168	168	280	440	64	336	528	472	400	568	80	392

Видно, что все рассмотренные выше свойства взвешенного турнира выполняются и в этом случае. В частности, мы имеем:

$$\sum_{i=1}^n z_i = 4896;$$

$$\sum_{i=1}^n y_i = 4896;$$

$$\sum_{i=1}^n r_i = 5472;$$

$$\sum_{i=1}^n r_i - nN = 5472 - 18 \cdot 32 = 4896;$$

$$n^2 N - \sum_{i=1}^n r_i = 10368 - 5472 = 4896;$$

$$nN = z_i + r_i = 576;$$

$$y_i + z_i = N(n - 1) = 544. \spadesuit$$

Выводы к главе 3

1. Показано, что множество оптимальных перестановок, т.е. выходной профиль $B(N, n) = \{\beta_1, \beta_2, \dots, \beta_N\}$, состоящий из решений ЗРК при входном профиле $\Lambda(m, n)$ m ранжирований n альтернатив, включая толерантности, может быть эффективно преобразован в точное единственное итоговое ранжирование β_{fin} , которое также может содержать толерантности, с помощью предложенных в диссертационной работе правил свёртки.

2. Предложенное Правило свёртки 1 состоит в следующем: в итоговом ранжировании консенсуса β_{fin} альтернативы располагаются в порядке убывания *выигрышей* (строковых сумм) z_i , вычисленных для N -взвешенной турнирной матрицы S выходного профиля $B(N, n)$; две альтернативы толерантны друг другу, если они имеют одинаковые выигрыши z_i в β_{fin} .

3. Предложенное Правило свёртки 2, являющееся дополнением к Правилу 1, состоит в следующем: в итоговом ранжировании консенсуса β_{fin} альтернативы располагаются в порядке возрастания их *общих рангов* r_i , вычисленных для выходного профиля $B(N, n)$; две альтернативы толерантны друг другу, если они имеют одинаковые общие ранги в β_{fin} .

4. Справедливость построенных правил свёртки показана с помощью правила Борда. Правила свёртки также демонстрируют полную совместимость с известным понятием отношения "между" на множестве ранжирований.

ГЛАВА 4

ПРАКТИЧЕСКИЕ АСПЕКТЫ АГРЕГИРОВАНИЯ ПРЕДПОЧТЕНИЙ НА ОСНОВЕ ПРАВИЛА КЕМЕНИ

В данной главе рассмотрены две основные парадигмы применения правила Кемени: измерение в многомерной порядковой шкале и комплексирование интервалов агрегированием предпочтений. Основное внимание уделено рассмотрению процедуры комплексирования интервальных данных агрегированием предпочтений (IF&PA), которая оперирует так называемыми "инранжированиями" – слабыми порядками, наведенными входными интервалами. Процедура IF&PA преобразует заданные m исходных интервалов на вещественной числовой оси в инранжирования, которые являются частным случаем отношения слабого порядка (или ранжирования) на множестве n дискретных значений, принадлежащих к этим интервалам. Новое понятие инранжирования появляется в результате ограничений, накладываемых на обычные ранжирования интервальным характером исходных данных. Исследуются свойства пространства инранжирований с точки зрения комбинаторной теории. Показано, что пространство инранжирований является подмножеством множества всех слабых порядков с единственным символом строгого порядка. Мощность пространства инранжирований определяется треугольным числом для заданного числа n дискретных элементов. Также рассматриваются мощности других смежных пространств.

4.1 Две основные парадигмы применения правила Кемени

Будем придерживаться двух основных парадигм применения правила Кемени, которые могут названы следующим образом: (1) *измерение в многомерной порядковой шкале* и (2) *комплексирование интервалов агрегированием предпочтений*. Рассмотрим особенности каждой из этих парадигм.

4.1.1 Измерение в многомерной порядковой шкале

Агрегирование предпочтений можно рассматривать как измерение в многомерной порядковой шкале [79,80,95], где n альтернатив ранжируются по m разно-

родным свойствам в соответствии с проявлениями этих свойств объектами, а результатом такого рода измерения является некоторое ранжирование консенсуса (n альтернатив), которое в наибольшей степени согласуется со всеми m исходными ранжированиями. Упрощенный пример практической задачи такого рода приведен в главе 1 (см. пример 1.5).

Агрегирование предпочтений является популярным инструментом при многокритериальном (или групповом) принятии решений [68]. В последние годы агрегирование предпочтений и обработка ранжированных данных привлекают значительный интерес в рамках гораздо более широкого рассмотрения, связанного с такой быстро развивающейся областью искусственного интеллекта, как машинное обучение [32]. Соответствующие методы используются для решения таких проблем, как краудсенсинг/лэйбелинг, анализ настроений, метапоисковые системы и другие [36,52,62,84].

Проблема, сформулированная в рамках парадигмы 1, может формулироваться в различных контекстах, таких например, как агрегирование предпочтений экспертной группы в практике оценивания качества продукции [12,63], в инженерном проектировании [48], интерпретация данных энергетического обследования магистральных электрических сетей [78], при цифровом цветометрическом анализе состава веществ [7] и т.д.

4.1.2 Комплексирование интервалов агрегированием предпочтений

Научный коллектив под руководством проф. Муравьева С.В. с 2011 года развивает метод, названный *комплексирование интервальных данных агрегированием предпочтений* (*Interval Fusion with Preference Aggregation, IF&PA*). Метод, подробно описанный в [43,73,74,79,81], работает на множестве m исходных интервалов $\{I_k\}$. Центрами интервалов являются измеренные значения, а их нижние и верхние границы определяются заданной погрешностью используемых средств измерений. Эти интервалы $\{I_k\}$ преобразуются в m ранжирований, сформированных из n дискретных значений, принадлежащих объединению всех исходных ин-

тервалов. Наилучшее дискретное значение в ранжировании Кемени (ранжировании консенсуса), найденном для профиля из m ранжирований, принимается как результат x^* (опорное значение) процедуры IF&PA.

Метод имеет большие перспективы для использования при обработке гетероскедастичных, т.е. характеризующихся неравной дисперсией, измерительных данных в таких важных практических ситуациях как: межлабораторные и/или ключевые сличения [70], согласование значений фундаментальных физических констант [69], точное оценивание значения измеряемой величины в условиях гетероскедастичности [79] и др.

Благодаря переходу от шкалы отношения к порядковой шкале, основным преимуществом метода IF&PA является его робастность, т.е. независимость его выходного опорного значения x^* от вида закона распределения вероятностей входных интервальных данных [43]. Кроме того, в статье [43] представлены результаты вычислительных экспериментов таких показателей качества, как достоверность и точность результатов процедуры IF&PA. Оказалось, что по этим показателям процедура IF&PA, как правило, превосходит традиционные методы обработки данных, полученных несколькими разными средствами измерения.

Дальнейшее содержание данной главы посвящено рассмотрению особенностей метода IF&PA.

4.2 Метод комплексирования интервальных данных IF&PA

4.2.1 Этапы метода IF&PA

Под *комплексированием интервальных данных* [18,26,43,73] понимается процедура определения такого *результатирующего интервала* $[x^* \pm \varepsilon^*]$, который согласован с максимальным количеством исходных интервалов $\{I_k\}$, $k = 1, \dots, m$, (не обязательно согласованных между собой) и с максимальной степенью правдоподобия содержит значение x^* , которое может служить представителем всех этих интервалов. *Результатом комплексирования* x^* является средняя точка результирующего интервала с соответствующей неопределенностью ε^* .

Исходные интервалы

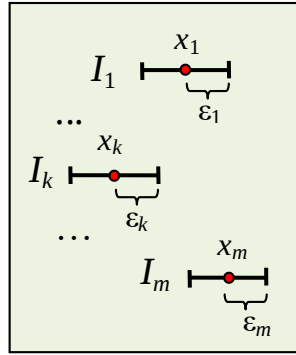
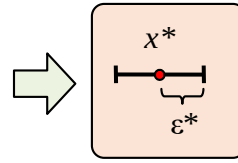
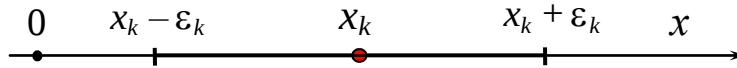
Результат
комплексирования

Рисунок 4.1 – К определению комплексирования интервальных данных

Рассмотрим набор из m замкнутых интервалов $\{I_k\}_{k=1}^m$ на вещественной числовой оси, где каждый интервал характеризуется *средней точкой* x_k , *нижней границей* $x_k - \varepsilon_k$ и *верхней границей* $x_k + \varepsilon_k$, так что $I_k = [x_k - \varepsilon_k, x_k + \varepsilon_k]$; $x_k, \varepsilon_k \in \mathbb{R}$. (см. рисунок 4.2).

Рисунок 4.2 – Интервал I_k на оси действительных чисел

Для представления исходных интервалов ранжированиями введем понятие *диапазона актуальных значений* (ДАЗ), представляющего собой множество $A = \{a_1, a_2, \dots, a_n\}$ строго упорядоченных дискретных значений $a_1 < a_2 < \dots < a_n$.

Дискретное множество A может быть получено, исходя из заданных непрерывных интервалов $\{I_k\}_{k=1}^m$, в три этапа, как показано на рисунке 4.3.

Этап 1. *Формирование диапазона актуальных значений*. Наименьшая нижняя граница для всех интервалов выбирается в качестве нижней границы ДАЗ a_1 , т.е.

$$a_1 = \min \{x_k - \varepsilon_k \mid k = 1, \dots, m\}; \quad (4.1)$$

в качестве верхней границы a_n берется наибольшая верхняя граница для этих интервалов, т.е.

$$a_n = \max \{x_k + \varepsilon_k \mid k = 1, \dots, m\}. \quad (4.2)$$

Интервалы $\{I_k\}$, $k = 1, \dots, m$

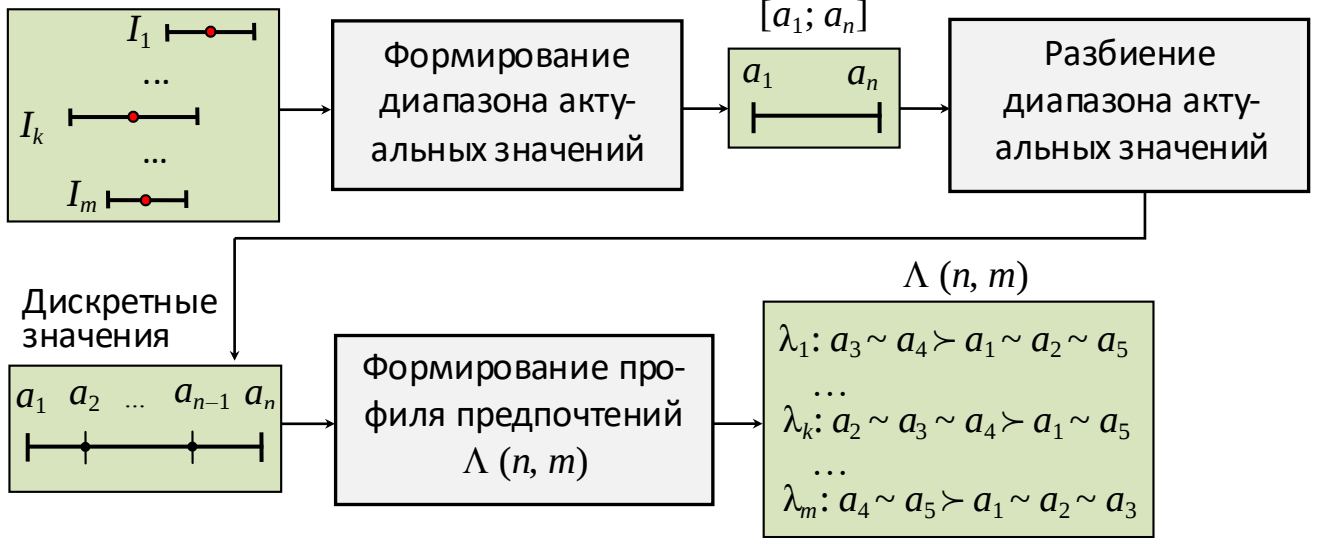


Рисунок 4.3 – Три этапа перехода от непрерывного пространства значений к дискретному при формировании ДАЗ

Этап 2. *Разбиение диапазона актуальных значений.* Для получения элементов $a_2, a_3, \dots, a_{n-1}$ разбиваем полученный интервал $[a_1, a_n]$ на $n - 1$ равных подынтервалов длиной

$$h = (a_n - a_1)/(n - 1). \quad (4.3)$$

Длину h будем называть *нормой* разбиения. Ясно, что после разбиения норма определяется формулой

$$h = |a_i - a_{i-1}|, i = 2, \dots, n, \quad (4.4)$$

а i -й элемент ДАЗ определяется формулой

$$a_i = a_{i-1} + h, i = 2, \dots, n. \quad (4.5)$$

Этап 3. *Формирование профиля предпочтений $\Lambda(m, n)$.* Множество $A = \{a_1 < a_2 < \dots < a_n\}$ строго упорядоченных дискретных значений a_i , $i = 1, \dots, n$, теперь можно использовать для формирования ранжирований, представляющих интервалы $\{I_k\}_{k=1}^m$.

Условия формирования профиля предпочтений будут рассмотрены в следующем параграфе 4.2.2.

Следующие три этапа завершают процедуру IF&PA.

Этап 4. *Вычисление выходного профиля* $V(N, n)$. Находим все возможные ранжирования консенсуса $\{\beta_1, \beta_2, \dots, \beta_N\}$ для профиля $\Lambda(m, n)$ по правилу Кемени рекурсивным алгоритмом RECURSALL.

Этап 5. *Проведение свертки профиля* $V(N, n)$ в β_{fin} . Преобразуем профиль $V(N, n)$ в единственное итоговое ранжирование консенсуса β_{fin} по правилу свёртки (см. п. 3.2).

Этап 6. *Определение результата комплексирования* x^* . Определяем результат комплексирования x^* как наиболее предпочтительную альтернативу в итоговом ранжировании консенсуса β_{fin} и его неопределенность ε^* .

Наиболее предпочтительная альтернатива a_i , которая занимает крайнюю левую позицию в полученном итоговом ранжировании β_{fin} принимается за результат комплексирования x^* . Это означает, что ранг r_i^{fin} этой альтернативы в β_{fin} равен 1, т.е.

$$x^* = a_i \text{ при } a_i \in \beta_{\text{fin}}, r_i^{\text{fin}} = 1. \quad (4.6)$$

Нередки случаи, когда β_{fin} содержит несколько наилучших альтернатив. В этом случае, имеется множество $A_b = \{a_i \mid a_i \in \beta_{\text{fin}}, r_i^{\text{fin}} = 1\}$ мощности $v = |A_b| > 1$, т.е.

$$\underbrace{\{a_i \sim a_j \sim \dots \sim a_k\}}_v \subseteq \beta_{\text{fin}}, \quad (4.7)$$

где $r_i^{\text{fin}} = r_j^{\text{fin}} = \dots = r_k^{\text{fin}} = 1$. Тогда единственным естественным представителем всех этих альтернатив (которые уже отсортированы в порядке возрастания) является *выборочная медиана*, см., например [53], откуда:

$$x^* = \begin{cases} a_{(v+1)/2}, & \text{если } v \text{ четно;} \\ (a_{v/2} + a_{1+v/2})/2, & \text{если } v \text{ нечетно.} \end{cases} \quad (4.8)$$

Неопределенность ε^* результата комплексирования x^* зависит от нормы h , см. Этап 2 и формулу (4.3). Ее можно оценить по очевидной формуле (см., например [58,96]):

$$\varepsilon^* = \pm 0.5h. \quad (4.9)$$

4.2.2 Инранжирования

Рассмотрим теперь условия формирования профиля предпочтений $\Lambda(m, n)$, состоящего из ранжирований, представляющих интервалы $\{I_k\}_{k=1}^m$, которые обладают интересными особенностями, позволяющими ввести для работы с ними специальное понятие "инранжирования" [10,41].

Обозначим через A_k множество всех элементов из A , принадлежащих интервалу I_k , т.е. $A_k = \{a_i \mid a_i \in I_k \wedge a_i \in A\}$. Ясно, что его дополнение $\bar{A}_k$ будет включать все остальные элементы из A , т.е. $\bar{A}_k = \{a_i \mid a_i \notin I_k \wedge a_i \in A\}$. Тогда для любого интервала I_k , $k = 1, \dots, m$, мы имеем разбиение множества A на два класса A_k и $\bar{A}_k$, т.е.

$$A = A_k \cup \bar{A}_k, A_k \cap \bar{A}_k = \emptyset. \quad (4.10)$$

Определение 4.1. Ранжирование, наведенное интервалом I_k , будет называть *инранжированием* λ_k , если оно удовлетворяет следующим четырем условиям для $i, j = 1, \dots, n$:

$$a_i \in A_k \wedge a_j \notin A_k \Rightarrow a_i \succ a_j; \quad (4.11)$$

$$a_i, a_j \in A_k \vee a_i, a_j \notin A_k \Rightarrow a_i \sim a_j; \quad (4.12)$$

$$a_i \notin A_k \wedge a_j \in A_k \Rightarrow a_i \prec a_j; \quad (4.13)$$

$$a_i, a_j \in A_k \text{ соседние элементы} \Rightarrow j \equiv i + 1. \quad (4.14)$$

Условие (4.14) является необходимым, т.к. последовательность элементов $\{a_i\}$ множества A является *строго монотонной*, т. е. $a_i < a_{i+1}$ для всех $i \in \mathbb{N}_n$. Класс $A_k \subseteq A$ может включать только последовательные наборы элементов из A без пропусков, т.е. список индексов этих элементов должен соответствовать *отрезку натурального ряда*. Это означает, что разность индексов любых двух соседних элементов a_i и a_j из A_k не может быть больше 1, т.е. $j \equiv i + 1$.

Заметим, что k -е инранжирование состоит из двух *классов эквивалентности*, образованных элементами введенных выше множеств A_k и $\bar{A}_k$. При этом

элементы класса A_k всегда строго предпочитаются элементам класса $\bar{A}_k$, т.е. структура любого k -го инранжирования определяется выражением

$$\lambda_k = A_k \succ \bar{A}_k. \quad (4.15)$$

Это означает, что каждое инранжирование содержит *единственный* символ строгого порядка $\succ$ и $n - 2$ символов толерантности $\sim$.

Пример 4.1. Для $n = 5$ одним из возможных инранжирований может быть следующее:

$$\lambda_k = \{a_2 \sim a_3 \sim a_4 \succ a_1 \sim a_5\}$$

или в короткой форме

$$\lambda_k = (234) (15),$$

где $A_k = \{a_2 \sim a_3 \sim a_4\}$, $\bar{A}_k = \{a_1 \sim a_5\}$ и $A_k \succ \bar{A}_k = \{a_2 \sim a_3 \sim a_4\} \succ \{a_1 \sim a_5\}$. Видно, что список индексов элементов из класса A_k имеет вид 2, 3, 4 и является строго монотонным без пропусков. ♦

Проиллюстрируем работу процедуры IF&PA на специальном практическом примере из области межлабораторных сличений, см. пример 4.2.

Пример 4.2. Пусть в ходе *межлабораторных сличений* три участвующие лаборатории оценивают одно и то же номинальное значение $x_{\text{ном}} = 4.0$ некоторой измеряемой величины x [19]. Целью сличений является определение опорного значения измеряемой величины x^* , которое характеризует согласованные результаты измерений, предоставленные соответствующими лабораториями в виде интервалов $I_1 = [3,98; 4,24]$, $I_2 = [3,80; 4,02]$ и $I_3 = [4,13; 4,60]$, см. рисунок 4.4.

В соответствии с методом IF&PA исходные интервалы преобразуются в инранжирования, образованные из дискретных значений, принадлежащих интервалам, следующим образом:

- формируем *диапазон актуальных значений* $A = \{a_1, a_2, \dots, a_5\}$ как объединение исходных интервалов $\{I_k\}$, $k = 1, 2, 3$, такое, что минимальная нижняя граница $a_1 = 3,8$ и максимальная верхняя граница $a_5 = 4,6$ среди всех I_k ста-

новятся нижней и верхней границами ДАЗ соответственно;

- разбиваем ДАЗ на $n - 1 = 4$ равных подынтервала длины $h = (a_5 - a_1) / (5 - 1) = (4,6 - 3,8) / 4 = 0,2$ для получения дискретных значений $a_2 = 4,0$; $a_3 = 4,2$ и $a_4 = 4,4$;
- представляем интервалы *инранжированиями*, т.е. интервалы I_1 , I_2 и I_3 преобразуются во входной профиль $\Lambda(3, 5)$, который состоит, соответственно, из трех инранжирований, удовлетворяющих условиям (4.10) - (4.13): $\lambda_1: a_2 \sim a_3 \succ a_1 \sim a_4 \sim a_5$, $\lambda_2: a_1 \sim a_2 \succ a_3 \sim a_4 \sim a_5$ и $\lambda_3: a_3 \sim a_4 \sim a_5 \succ a_1 \sim a_2$.

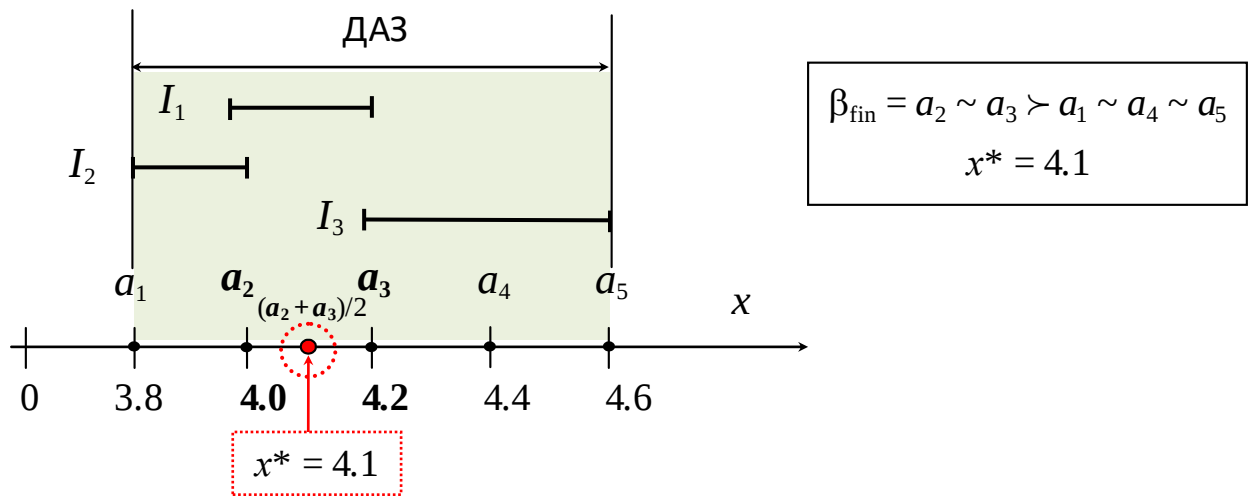


Рисунок 4.4 – Исходные интервальные данные и результат комплексирования x^* для примера 4.4

Полученный входной профиль $\Lambda(3, 5)$ уже был нами использован в примере 3.1, см. таблицу 3.3, для демонстрации работы алгоритма свертки, когда после обработки входного профиля алгоритмом RECURSALL был получен выходной профиль $B(12, 5)$, свертка 12 оптимальных решений которого дала итоговое ранжирование консенсуса $\beta_{\text{fin}} = a_2 \sim a_3 \succ a_1 \sim a_4 \sim a_5$.

Теперь мы можем определить опорное значение x^* как наиболее предпочтительную альтернативу в отношении консенсуса β_{fin} . Поскольку у нас есть две такие альтернативы, а именно $a_2 = 4.0$ и $a_3 = 4.2$ (см. рисунок 4.4), выборочная медиана в соответствии с формулой (4.9) позволяет получить результат комплексирования $x^* = 4,1$ с неопределенностью $\varepsilon^* = \pm 0,1$. ♦

Заметим, что в недавних работах [25,69,79] был предложен ряд значительных усовершенствований метода IF&PA, обеспечивающих многократное повышение его точности.

4.3 Свойства инранжирований и связанных с ними пространств

Исследуем комбинаторные свойства введенных в предыдущем параграфе инранжирований и связанных с ними подпространств слабых порядков. Исследование свойств этих пространств может помочь улучшить качество алгоритмов определения ранжирования консенсуса для заданного профиля $\Lambda(m, n)$.

4.3.1 Пространства инранжирований

В методе IF&PA конкретный входной профиль предпочтений $\Lambda(m, n)$ всегда состоит из инранжирований, взятых из пространства всех возможных инранжирований. В данном разделе мы постараемся ответить на вопрос, сколько различных инранжирований существует для заданного n .

Введем обозначения:

- Ω_0 – пространство слабых порядков (ранжирований);
- Ω_1 – подпространство слабых порядков с единственным символом $\succ$;
- Ω_2 – подпространство инранжирований;
- Ω_f – подпространство запрещенных слабых порядков.

Подпространство Ω_2 образовано слабыми порядками, которые удовлетворяют условиям (4.11)-(4.14) по определению 4.1. Подпространство Ω_1 образуют такие ранжирования, которые содержат *один символ* строгого порядка $\succ$, см. п. 4.2.2 [41]. Нарушение условия (4.14) приводит к существованию подпространства Ω_f *запрещенных слабых порядков*, которые, однако, удовлетворяют условиям (4.11)-(4.13). Ясно, что

$$\Omega_f = \bar{\Omega}_2 \cap \Omega_1. \quad (4.16)$$

Пространство Ω_0 играет роль универсального множества для подпространств Ω_1 и Ω_2 . Пространство Ω_1 состоит из подпространства инранжирований

Ω_2 и подпространства запрещенных слабых порядков Ω_f .

Ясно, что для рассматриваемых пространств выполняются следующие отношения вложенности (см. рисунок 4.5):

$$\Omega_2 \subseteq \Omega_1 \subseteq \Omega_0. \quad (4.17)$$

Для этих пространств, очевидно, справедливы также следующие тождества:

$$\begin{aligned} \Omega_1 &= \Omega_2 \cup \Omega_f, \quad \Omega_1 \cap \bar{\Omega}_1 = \emptyset, \quad \Omega_2 \cap \bar{\Omega}_2 = \emptyset, \\ |\Omega_1| &= |\Omega_2| + |\Omega_f|. \end{aligned} \quad (4.18)$$

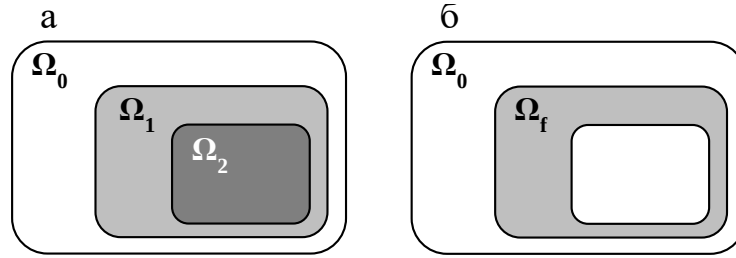


Рисунок 4.5 – Отношение вложенности на пространствах:

а) Ω_0 , Ω_1 и Ω_2 , и б) Ω_0 и Ω_f

Мощности рассматриваемых пространств $|\Omega_0|$, $|\Omega_1|$ и $|\Omega_2|$ связаны с числами Стирлинга 2-го рода $S_{n,q}$, которые определяют количество неупорядоченных разбиений n -элементного множества на q непустых подмножеств [9,11]. В частности, количество элементов пространства всех слабых порядков Ω_0 определяется [2] по формуле:

$$|\Omega_0| = \sum_{q=0}^n q! S_{n,q}. \quad (4.19)$$

Состав пространств Ω_1 , Ω_2 и Ω_f для $n = 1, \dots, 5$ при соответствующих мощностях пространства Ω_0 и класса A_k представлен в таблице 4.1.

Из таблицы 4.1 видно, что значения мощности подпространства инранжирований Ω_2 в зависимости от n представляют собой последовательность целых чисел

$$1, 3, 6, 10, 15, 21, 28, 36, 45, 55, \dots \quad (4.20)$$

Последовательность (4.20) имеет номер A000217 в онлайн-энциклопедии целочисленных последовательностей (OEIS – The On-Line Encyclopedia of Integer Sequences) [91]. Элементы последовательности (4.20) называются *треугольными числами* T_n .

Таблица 4.1 – Состав пространств Ω_1 , Ω_2 и Ω_f для $n = 1, \dots, 5$

$ \Omega_0 $	$ A_k $	Ω_1	$ \Omega_1 $	Ω_2	$ \Omega_2 $	Ω_f	$ \Omega_f $	$ \Psi_k $	$n - A_k + 1$
1	2	3	4	5	6	7	8	9	10
$n = 1$									
1	1	1	1	1	1	$\emptyset$	0	1	1
$n = 2$									
3	1	12, 21	3	12, 21	3	$\emptyset$	0	1	2
	2	(12)		(12)		$\emptyset$		2	1
$n = 3$									
13	1	1(23), 2(13), 3(12)	7	1(23), 2(13), 3(12)	6	$\emptyset$	1	2	3
	2	(12)3, (23)1, (13)2		(12)3, (23)1		(13)2		2	2
	3	(123)		(123)		$\emptyset$		6	1
$n = 4$									
75	1	1(234), 2(134), 3(124), 4(123)	15	1(234), 2(134), 3(124), 4(123)	10	$\emptyset$	5	6	4
	2	(12)(34), (23)(14), (13)(24), (24)(13), (14)(23), (34)(12)		(12)(34), (23)(14), (34)(12)		(14)(23), (13)(24), (24)(13)		4	3
	3	(123)4, (134)2, (124)3, (234)1		(123)4, (234)1		(124)3, (134)2		6	2
	4	(1234)		(1234)		$\emptyset$		24	1
$n = 5$									
541	1	1(2345), 2(1345), 3(1245), 4(1235), 5(1234)	30	1(2345), 2(1345), 3(1245), 4(1235), 5(1234)	15	$\emptyset$	16	24	5
	2	(12)(345), (13)(245), (14)(235), (15)(234), (23)(145), (24)(135), (25)(134), (34)(125), (35)(124), (45)(123)		(12)(345), (23)(145), (34)(125), (45)(123)		(13)(245), (14)(235), (15)(234), (24)(135), (25)(134), (35)(124)		12	4
	3	(123)(45), (124)(35), (125)(34), (134)(25), (135)(24), (145)(23), (234)(15), (235)(14), (245)(13), (345)(12)		(123)(45), (234)(15), (345)(12)		(124)(35), (125)(34), (134)(25), (135)(24), (145)(23), (235)(14), (245)(13)		12	3
	4	(1234)5, (1235)4, (1245)3, (1345)2, (2345)1		(1234)5, (2345)1		(1235)4, (1245)3, (1345)2		24	2
	5	(12345)		(12345)		$\emptyset$		120	1

Треугольное число и, следовательно, мощность Ω_2 определяются по известной формуле:

$$|\Omega_2| = T_n = \frac{n(n+1)}{2} = \binom{n+1}{2} = \binom{n+1}{n-1} = (S_{n+1,n}) = \sum_{i=1}^n |A_k|_i. \quad (4.21)$$

Все допустимые интервалы и соответствующие треугольные числа для различных $n = 1, \dots, 5$ показаны на рисунке 4.6.

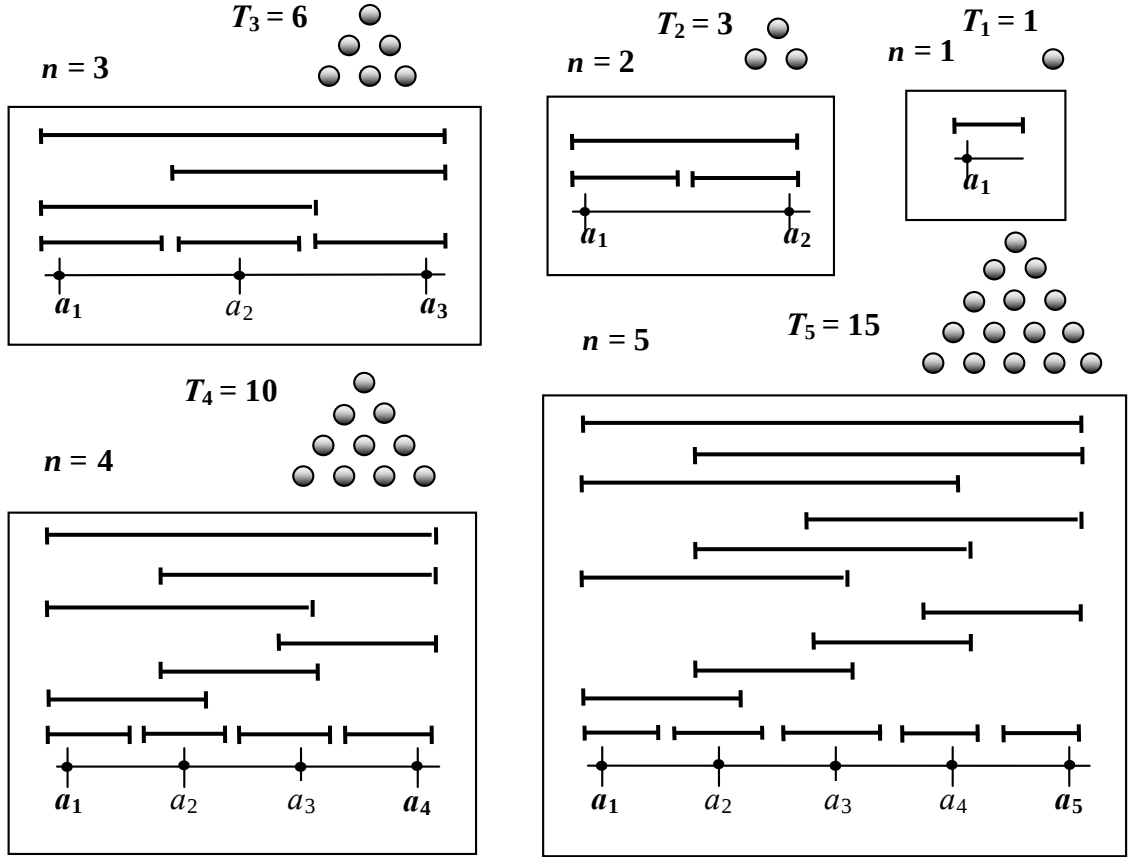


Рисунок 4.6 – Допустимые интервалы и соответствующие треугольные числа для $n = 1, \dots, 5$

В свою очередь, мощности F_n пространства Ω_f запрещенных слабых порядков в зависимости от значения n соответствует последовательность

$$0, 0, 1, 5, 16, 42, 99, 219, 466, 968, 1981, 4017, 8100, 16278, \dots, \quad (4.22)$$

которая имеет номер A002662 в OEIS [2]. Мощность F_n определяется по формуле:

$$|\Omega_f| = F_n = 2^n - 1 - n(n+1)/2. \quad (4.23)$$

Очевидно, что мощность подпространства Ω_1 определяется по формуле:

$$|\Omega_1| = T_n + F_n = 2S_{n,2} + 1 = S_{n+1,2} = 2^n - 1. \quad (4.24)$$

Из уравнений (4.21), (4.23) и (4.24) следует, что числа T_n могут быть обнаружены среди чисел Стирлинга второго рода и треугольнике Паскаля. Связь мощно-

сти подпространства Ω_1 с числами Стирлинга 2-го рода $S_{n,q}$ показана в таблице 4.3, а связь мощности подпространства Ω_2 и треугольника Паскаля показана в таблице 4.4 [2,29].

Крайние правые столбцы в таблицах 4.3 и 4.4 содержат суммы элементов соответствующих строк. Последовательность чисел

$$1, 1, 2, 5, 15, 52, 203, 877, 4140, 21147, \dots, \quad (4.25)$$

имеющая номер A000110 в OEIS, состоит из чисел Белла B_n , определяющих числа всех неупорядоченных разбиений n -элементного множества. Другими словами, число Белла определяет число отношений эквивалентности на множестве.

Таблица 4.2 – Числа Стирлинга 2-го рода $S_{n,q}$ и их связь с мощностью подпространства Ω_1

n	q											B_n
	0	1	2	3	4	5	6	7	8	9	10	
0	1	0	0	0	0	0	0	0	0	0	0	1
1	0	1	0					0	0	0	0	1
2	0	1	1					0	0	0	0	2
3	0	1	3					0	0	0	0	5
4	0	1	7	6	1	0	0	0	0	0	0	15
5	0	1	15	25	10	1	0	0	0	0	0	52
6	0	1	31	90	65	15	1				0	203
7	0	1	63	301	350	140	21	1	0	0	0	877
8	0	1	127	966	1701	1050	266	28	1	0	0	4140
9	0	1	255	3025	7770	6951	2646	462	36	1	0	21147
10	0	1	511	9330	34105	42525	22827	5880	750	45	1	115975

Значение $|P(A)|$ в крайнем правом столбце таблицы 4.3 связано с биномиальными коэффициентами (сочетаниями) $\binom{n}{m}$, характеризует мощность множества во всех подмножествах n -элементного множества и определяется по формуле:

$$|P(A)| = 2^{|A|} = 2^n = \sum_{i=0}^n \binom{n}{i}. \quad (4.26)$$

Ясно, что $S_{n+1,2} = |P(A)| - 1$.

Таблица 4.3 – Треугольник Паскаля и его связь с мощностью подпространства Ω_2

n	q											$ P(A) = 2^n$
	0	1	2	3	4	5	6	7	8	9	10	
0	1	0	0	0	0	0	0	0	0	0	0	1
1	1	1	0	0	0	0	0	0	0	0	0	2
2	1	2	1	0	0	0	0	0	0	0	0	4
3	1	3	3	1	0	0	0	0	0	0	0	8
4	1	4	6	4	1	0	0	0	0	0	0	16
5	1	5	10	10	5	1	0	0	0	0	0	32
6	1	6	15	20	15	6	1	0	0	0	0	64
7	1	7	21	35	35	21	7	1	0	0	0	128
8	1	8	28	56	70	56	28	8	1	0	0	256
9	1	9	36	84	126	126	84	36	9	1	0	512
10	1	10	45	120	210	252	210	120	45	10	1	1024

Числа Стирлинга 2-го рода $S_{n+1,2}$, определяющие мощность пространства Ω_1 , образуют последовательность чисел (см. таблицу 4.2, столбец для $q = 2$):

$$0, 1, 3, 7, 15, 31, 63, 127, \dots, \quad (4.27)$$

имеющую номер A000225 в OEIS. Эта последовательность известна под названием "биномиальный коэффициент Гаусса".

4.3.2 Мощность классов A_k

При анализе состава пространств инранжирований полезно пользоваться формулой для перечисления элементов класса A_k при фиксированном n :

$$A_k \in \bigcup_{j=0}^{n-1} \bigcup_{i=1}^{n-j} \{a_{i+j}\}. \quad (4.28)$$

Справедливость выражения (4.28) при $n = 5$ обосновывается с учетом рассмотрения содержания таблицы 4.4. Очевидно, что выражение (4.28) справедливо для любого n , см. также таблицу 4.1.

Таблица 4.4 – К обоснованию выражения (4.28) при $n = 5$

$j = A_k $	Элементы класса A_k
1	$\{a_i\}, i = 1, \dots, n,$
2	$\{a_i\} \cup \{a_{i+1}\}, i = 1, \dots, (n-1),$
3	$\{a_i\} \cup \{a_{i+1}\} \cup \{a_{i+2}\}, i = 1, \dots, (n-2),$
4	$\{a_i\} \cup \{a_{i+1}\} \cup \{a_{i+2}\} \cup \{a_{i+3}\}, i = 1, \dots, (n-3),$
5	$\{a_i\} \cup \{a_{i+1}\} \cup \{a_{i+2}\} \cup \{a_{i+3}\} \cup \{a_{i+4}\}, i = 1, \dots, (n-4)$

Из выражения (4.28) для заданного n вытекает следующее общее выражение, описывающее все возможные инранжирования:

$$\lambda_k \in \bigcup_{j=0}^{n-1} \left\{ \left(\bigcup_{i=1}^{n-j} \{a_{i+j}\} \right) \succ \left(A_k \setminus \bigcup_{i=1}^{n-j} \{a_{i+j}\} \right) \right\}. \quad (4.29)$$

На рисунке 4.7 показаны пространства инранжирований, принадлежащие *пермutoэдрам* (перестановочным многогранникам) [27,34] порядка $n = 1, \dots, 4$. Можно видеть, что число измерений пространства увеличивается с увеличением n , а именно: пермutoэдр порядка 1 является точкой (0D); порядок 2 дает отрезок прямой линии (1D); порядок 3 приводит к правильному шестиугольнику (2D); порядок 4 приводит к усеченному октаэдру (3D) и т.д.

На рисунке 4.7 запрещенные ранжирования не показаны, но строгие порядки (хотя они и не принадлежат подпространству инранжирований Ω_2 , а входят лишь в пространство слабых порядков Ω_0) показаны, т.к. они участвуют в формировании инранжирований. В принципе, все слабые порядки происходят от некоторых комбинаций строгих порядков. В этом пространстве два слабых порядка соединяются отрезком прямой линии (ребром), если они различаются инверсией двух элементов.

4.3.3 Порождающие инранжирования множества

Инранжирования, как и любые слабые порядки, формируются из множеств строгих порядков. Продемонстрируем это на примере.

Пример 4.3. Пусть дано два отношения строгого порядка $\rho_1 = 213$ и $\rho_2 = 231$. Тогда лежащим между ними (*порожденным* ими) отношением слабого порядка является $\lambda_k = 2(13)$, или в расширенной записи $\lambda_k = \{a_2 \succ a_1 \sim a_3\}$. ♦

Введем обозначения:

- $\Pi = \{\pi_1, \dots, \pi_{|A_k|}\}$ – множество перестановок элементов класса A_k и
- $\Pi' = \{\pi'_1, \dots, \pi'_{|\bar{A}_k|}\}$ – множество перестановок элементов класса $\bar{A}_k$.

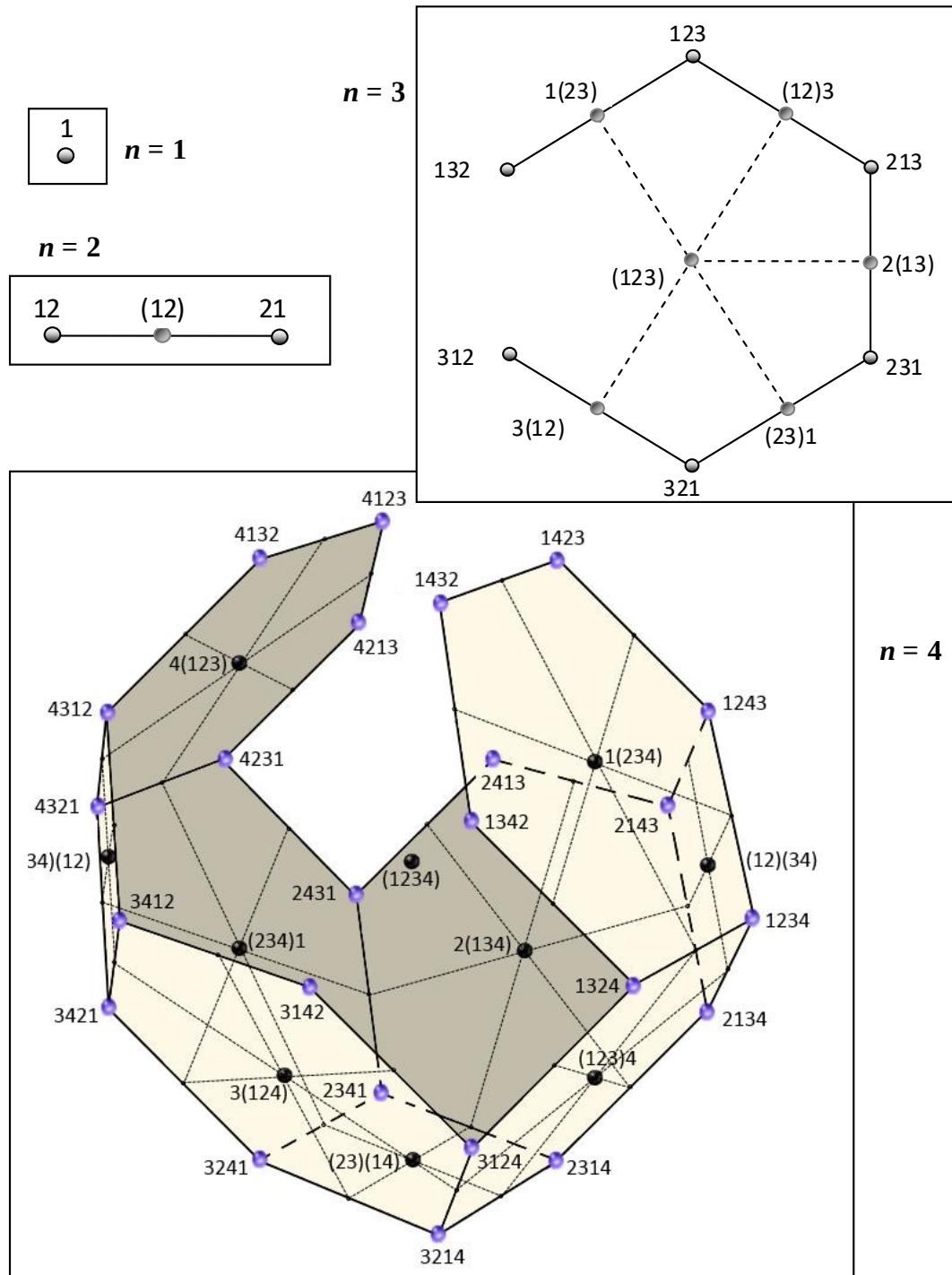


Рисунок 4.7 – Пространства инранжирований, принадлежащих перестановочным многогранникам порядка $n = 1, \dots, 4$

Введем определение.

Определение 4.2. Множество строгих порядков $\Psi_k = \{\rho_g^k\}$, $g = 1, \dots, |A_k|! \cdot |\bar{A}_k|!$, будем называть *порождающим множеством* для инранжирования $\lambda_k = A_k \succ \bar{A}_k$, если $\Psi_k = \Pi \times \Pi' = \{(\pi_u, \pi'_v) \mid \pi_u \in \Pi, \pi'_v \in \Pi'\}$, где под элементом (π_u, π'_v) декартова произведения понимается результат операции конкатенации ρ

$\mid = \pi_u \div \pi'_v$ двух перестановок π_u и π'_v .

Пример 4.4. Пусть имеется инранжирование $\lambda_k = (234)(15)$. Тогда его порождающее множество Ψ_k состоит из $|\Psi_k| = 3! \cdot 2! = 12$ следующих строгих порядков ρ_g , $g = 1, \dots, 12$: 23415, 23451, 24315, 24351, 32415, 32451, 34215, 34251, 42315, 42351, 43215, 43251. ♦

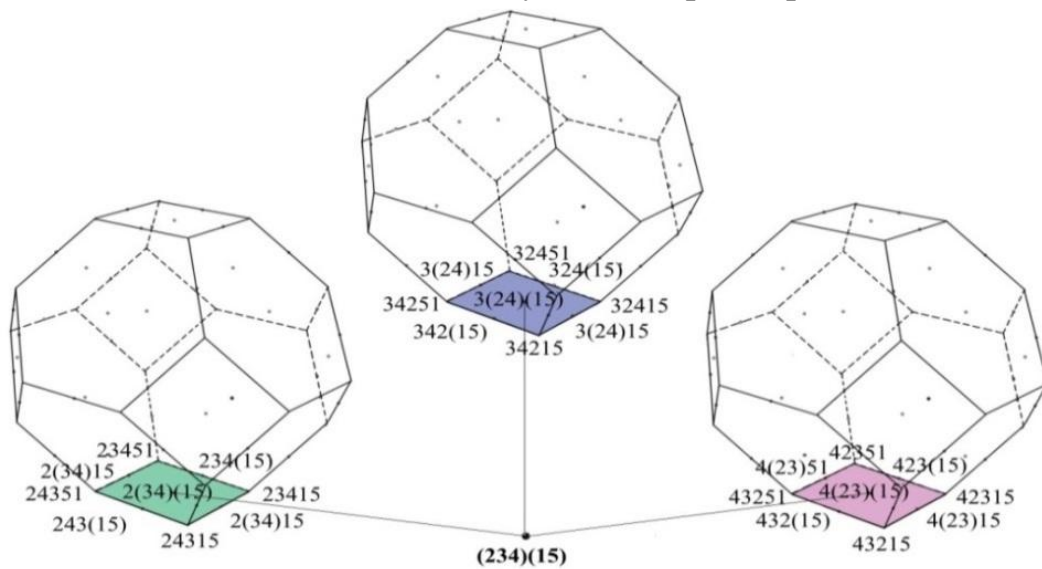
Получим выражение для числа порождающих множеств в зависимости от числа n .

Для заданной мощности $|A_k|$ при любом n число порождающих множеств определяется выражением $(n - |A_k| + 1)$. Следовательно, общее число порождающих строгих порядков для всех возможных инранжирований может быть рассчитано как

$$N(A_k) = |A_k|! \cdot |\bar{A}_k|! \cdot (n - |A_k| + 1). \quad (4.30)$$

Можно утверждать, что для заданного n объединение порождающих множеств для $|A_k| = 1$ совпадает с единственным порождающим множеством для $|A_k| = n!$. Все возможные мощности $\Psi_k = |A_k|! \cdot |\bar{A}_k|!$ и $(n - |A_k| + 1)$ для различных $|A_k|$ и $n = 1, \dots, 5$ показаны соответственно в девятой и десятой колонках таблицы 4.1.

Введенные понятия могут быть наглядно проиллюстрированы на перестановочном многограннике. На рисунке 4.8 показано пространство всех строгих порядков для $n = 5$ и пространство Ω_2 всех соответствующих инранжирований. Для наглядности в качестве перестановочных многогранников использованы усеченные октаэдры с центральными узлами, соответствующими инранжированиям 1(2345), 2(1345), 3(1245), 4(1235) и 5(1234). На рисунке показаны только связи между инранжированиями и центрами граней многогранника, содержащего слабые порядки. Обозначения инранжирований показаны жирным шрифтом (например, (23)(145), (34)(125), 4(1235) и т.д.). Каждый из усеченных октаэдров представляет собой порождающее множество при $|A_k| = 1$. Подмножества порождающих множеств, принадлежащих многограннику для инранжирований при $|A_k| = 2, 3, 4$, показаны разными цветами [9].



На рисунке 4.9 показано порождающее множество для инранжирования $\lambda_k = (234)(15)$, рассмотренного в Примере 4.4. Можно заметить, что центры ребер и граней многогранника соответствуют слабым порядкам. Как и на рисунке 4.8, показаны только связи между центрами граней многогранника и инранжированием $(234)(15)$. Число многогранников определяется мощностью $|A_k| = 3$.

Выводы к главе 4

1. Рассмотрены две основные парадигмы применения правила Кемени: измерение в многомерной порядковой шкале и комплексирование интервалов агрегированием предпочтений. Основное внимание уделено рассмотрению процедуры комплексирования интервальных данных агрегированием предпочтений (IF&PA), которая преобразует заданные m исходных интервалов на вещественной числовой оси в "инранжирования", которые являются частным случаем отношения слабого порядка (или ранжирования) на множестве n дискретных значений, принадлежащих к этим интервалам.

2. Введено новое понятие инранжирования как слабого порядка, наведенного входными интервалами, которое появляется в результате ограничений, накладываемых на обычные ранжирования интервальным характером исходных данных.

3. Исследованы свойства пространства инранжирований с точки зрения комбинаторной теории. Показано, что пространство инранжирований является подмножеством пространства всех слабых порядков с единственным символом строгого порядка. Получено выражение для расчета общего числа порождающих строгих порядков для всех возможных инранжирований заданной мощности для любого n .

4. Показано, что мощность пространства инранжирований определяется треугольным числом для заданного числа n дискретных элементов. Также рассматриваются мощности других смежных пространств.

Заключение

1. Предложена стратегия получения итогового ранжирования консенсуса при применении правила Кемени, заключающаяся в нахождении всех ранжирований консенсуса, являющихся точными оптимальными решениями, достижимыми при размерности задачи $n < 20$, и нахождении для них единственного подходящего итогового ранжирования консенсуса. Стратегия обеспечивает возможность применения правила Кемени для широкого круга практических задач агрегирования предпочтений с максимально возможной точностью.
2. Экспериментальные исследования фактора множественности, связанного с проявлением транзитивности/нетранзитивности входного профиля предпочтения, показали, что вероятность транзитивности значительно выше при четных значениях m , чем при нечетных, и плавно уменьшается с ростом n , тогда как вероятность множественности резко увеличивается при уменьшении чётных значений m , но плавно увеличивается при уменьшении нечётных значений m . Полученные оценки согласуются с известными аналитическими расчетами.
3. Предложено правило свёртки всех N оптимальных решений задачи о ранжировании Кемени в точное единственное итоговое ранжирование: в итоговом ранжировании консенсуса альтернативы располагаются в порядке убывания строковых сумм, вычисленных для N -взвешенной турнирной матрицы выходного профиля; две альтернативы толерантны друг другу, если они имеют одинаковые строковые суммы. Предложенная процедура свёртки имеет естественное обоснование на основе правила Борда.
4. В рамках применения правила Кемени в методе комплексирования интервалов агрегированием предпочтений IF&PA введено новое понятие инражирования; показано, что мощность пространства инражирований определяется треугольным числом для заданного числа n дискретных элементов.
5. Результаты диссертационной работы были использованы при выполнении НИР по гранту РНФ 18-19-00203, а также в научно-исследовательской лаборатории "Безопасность и электромагнитная совместимость радиоэлектронных

средств" Томского государственного университета систем управления и радиоэлектроники; в лаборатории службы релейной защиты и испытаний электрооборудования ООО "Горсети" г. Томска; в отделении автоматизации и робототехники Инженерной школы информационных технологий и робототехники ТПУ.

Список сокращений и обозначений

β – ранжирование консенсуса

n – число альтернатив

$A = \{a_1, a_2, \dots, a_n\}$ – множество альтернатив, множество дискретных значений ДАЗ

m – число избирателей (исходных ранжирований (разнородных свойств, интервалов))

λ_k – k -ое исходное ранжирование/инранжирование

ЗРК – задача нахождения ранжирования Кемени

$\Lambda(m, n) = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$ – входной профиль ЗРК

ρ – отношение строгого порядка

Π_n – пространство строгих порядков

τ – отношение толерантности

L – матрица отношения $[l_{ij}]$

R – матрица отношения $[r_{ij}]$

r_i – общий ранг альтернативы a_i

S – $S = [s_{ij}]$ – $(n \times n)$ взвешенная турнирная матрица

$I_k(a_i, a_j)$ – элемент матрицы S , индикаторная функция

$w(a_i)$ – число случаев, когда альтернатива a_i заняла первое место в k -ом ранжировании профиля $\Lambda(m, n)$

$SM(a_i)$ – победитель по правилу простого большинства

z_i – выигрыш альтернативы a_i , сумма элементов i -й строки матрицы S

$Bo(a_i)$ – победитель по правилу Борда балл Борда (счёт Борда)

$Co(a_i)$ – победителем Кондорсе

$d(\lambda_k, \lambda_l)$ – расстояние Кемени между двумя ранжированиями λ_k и λ_l

$D(\beta, \Lambda)$ – расстояние Кемени от произвольного отношения консенсуса β до $\Lambda(m, n)$

Ω_0 – множество (пространство) всех слабых порядков (ранжирований)

$|\Omega_0|$ – Мощность пространства слабых порядков

$S_{n,q}$ – числами Стирлинга 2-го рода

q – непустых подмножеств

P – матрица профиля предпочтения $[p_{ij}]$

$\mathbb{N}_n$ – множество n натуральных чисел (множество перестановок)

$S = \{s_1, s_2, \dots, s_K\}$ – частичное решение (лидер)

$T = \{t_1, t_2, \dots, t_{K-n}\}$ – дополнение S до полного порядка

K – номер уровня дерева поиска

D_u – верхняя граница

D_{low} – нижняя граница

D_{least} – наименьшее возможное расстояние Кемени до $\Lambda(m, n)$

β_N – N -ое ранжирование консенсуса

$B(N, n) = \{\beta_1, \beta_2, \dots, \beta_m\}$ – выходной профиль ЗРК

β_k – ранжирование консенсуса (ранжирование Кемени)

N – количество оптимальных решений ЗРК (оптимальных перестановок)

β_{fin} – единственное итоговое оптимальное решение, итоговое ранжирование консенсуса

RECURSALL – алгоритм, реализующий рекурсивный метод ветвей и границ

$t(t_0, t_1, t_2)$ – плотность толерантностей

n_τ – количество толерантностей в $\Lambda(m, n)$

$D(\beta_{fin}, \Lambda)$ – расстояние Кемени от итогового отношения консенсуса β_{fin} до $\Lambda(m, n)$

S_p – стартовая точка, номер (идентификатор) генерируемого профиля (соответствующего индивидуальной задаче о ранжировании Кемени)

α – признак транзитивности входного профиля

Δ – степень нетранзитивности входного профиля

δ – степень нетранзитивности входного профиля в относительной форме

$P(m, n)$ – вероятность альтернативы Кондорсе

$Q(m, n)$ – вероятность парадокса Кондорсе

$P_1(m, n)$ – вероятность единственного решения ЗРК

$P_{tr}(m, n)$ – вероятность появления транзитивного профиля

$Q_{tr}(m, n)$ – вероятность появления нетранзитивного профиля

y_i – проигрыш альтернативы a_j , сумма элементов j -го столбца матрицы S

IF&PA (Interval Fusion with Preference Aggregation) – комплексирование интервальных данных агрегированием предпочтений

$\{I_k\}$ – набор исходных интервалов

x^* – результат комплексирования, опорное значение процедуры IF&PA

ε^* – неопределенность результата комплексирования

x_k – средняя точка интервала

$\mathbb{R}$ – множество вещественных (действительных) чисел

x_{nom} – номинальное значение

ДАЗ – диапазон актуальных значений, объединение исходных интервалов $\{I_k\}$

a_1 – нижняя граница ДАЗ

a_n – верхняя граница ДАЗ

h – норма разбиения ДАЗ

r_i^{fin} – ранг альтернативы a_i в β^{fin}

A_b – множество наилучших альтернатив

ν – мощность множества A_b

A_k – множество всех элементов A , принадлежащих интервалу I_k

$\bar{A}_k$ – дополнение множества A_k

Ω_1 – подпространство слабых порядков с единственным символом $\succ$

Ω_2 – подпространство инранжирований

Ω_f – подпространство запрещенных слабых порядков

$S_{n,q}$ – числа Стирлинга 2-го рода

q – число непустых подмножеств неупорядоченных разбиений

OEIS (On-Line Encyclopedia of Integer Sequences) – Онлайн-энциклопедия целочисленных последовательностей

T_n – треугольное число

F_n – мощность запрещенных ранжирований

B_n – число Белла

$|P(A)|$ – мощность множества всех подмножеств

$\Pi = \{\pi_1, \dots, \pi_{|A_k|}\}$ – – множество перестановок элементов класса A_k

$\Pi' = \{\pi'_1, \dots, \pi'_{|\bar{A}_k|}\}$ – – множество перестановок элементов класса $\bar{A}_k$

Ψ_k – порождающее множество для инранжирования λ_k

$N(A_k)$ – число порождающих строгих

Список используемой литературы

1. Авен, П.О. Функциональное шкалирование / П.О. Авен, И.Б. Мучник, А.А. Ослон. – М.: Наука. – 1988. – 182 с.
2. Айгнер, М. Комбинаторная теория / М. Айгнер. – М.: Мир. – 1982. – 558 с.
3. Вентцель, Е.С. Исследование операций: задачи, принципы, методология / Е.С. Вентцель. – М.: Наука. – 1980. – 208 с.
4. Вирт, Н. Алгоритмы и структуры данных / Н. Вирт. – М.: Мир, 1989. – 360 с.
5. Гэри, М. Вычислительные машины и труднорешаемые задачи / М. Гэри, Д. Джонсон. – М.: Мир. – 1982. – 416 с.
6. Дидэ, Э. Методы анализа данных: Подход, основанный на методе динамических сгущений / Э. Дидэ и др. – М.: Финансы и статистика. – 1985. – 367 с.
7. Емельянова, Е.Ю. Агрегирование предпочтений при выборе цветовой модели для цифрового цветометрического анализа состава веществ / С.В. Муравьев, А.С. Спиридонова, Е.Ю. Емельянова // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2020. – Том 331. – № 8. – С. 167-179.
8. Емельянова, Е.Ю. Влияние нетранзитивности на множественность решений при агрегировании предпочтений с использованием правила Кемени / Е.Ю. Емельянова // Сборник трудов XIV Всероссийской научной конференции молодых ученых "Наука. Технологии. Инновации" (НТИ-2020), 30 ноября - 04 декабря 2020 г. – Часть 2: Информационные технологии математического моделирования и обработки данных. – Новосибирск: НГТУ, 2020. – С. 30-36.
9. Емельянова, Е.Ю. Графическое представление пространств ранжирований и инранжирований с учетом свойств слабых порядков комплексированных данных / Е.Ю. Емельянова // Молодёжь и современные информационные технологии: сборник трудов XV Международной научно-практической конференции студентов, аспирантов и молодых учёных, г. Томск, 4-7 декабря 2017 г. – Томск: Изд-во ТПУ, 2017. – С. 110-111.

10. Емельянова, Е.Ю. Инранжирования и геометрия пространств слабых порядков / Е.Ю. Емельянова // Молодежь и современные информационные технологии. Сборник трудов XVI Международной научно-практической конференции студентов, аспирантов и молодых ученых (3-7 декабря 2018 г., г. Томск). Томск: Изд-во ТПУ, 2019. – С. 50-51.
11. Емельянова, Е.Ю. Комбинаторные свойства слабых порядков, наведенных интервалами / Е.Ю. Емельянова // Интеллектуальный анализ сигналов, данных и знаний: методы и средства: Сборник статей II всероссийской научно-практической конференции с международным участием им. В.В. Губарева. Новосибирск: НГТУ, 2018. – С. 143-149.
12. Емельянова, Е.Ю. Модель бизнес-процессов предприятия / Е.Ю. Емельянова // Современные техника и технологии, сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых, Томск, 24-28 марта 2008 г. Томск: ТПУ. – 2008. – Т. 3 . С. 214-216.
13. Емельянова, Е.Ю. Функция выбора при агрегировании инранжирований / Е.Ю. Емельянова // Молодежь и современные информационные технологии. Сборник трудов XVII Международной научно-практической конференции студентов, аспирантов и молодых учёных, 17-20 февраля 2020 г., г. Томск. – Томск: Изд-во ТПУ. – 2020 . – С. 210-212.
14. Кемени Дж. Кибернетическое моделирование / Дж. Кемени, Дж. Снелл. – М.: Сов. Радио, 1972. – 192 с.
15. Компьютер и задачи выбора. – М.: Наука. – 1989. – 208 с.
16. Кузьмин, В.Б. Построение групповых решений в пространствах четких и нечетких двоичных отношений / В.Б. Кузьмин. – М.: Наука, 1982. – 168 с.
17. Литвак, Б.Г. Экспертная информация: Методы получения и анализа / Б.Г. Литвак. – М.: Радио и связь, 1982. – 184 с.
18. Маринушкина, И.А. Обработка данных сличений электрических величин методом агрегирования предпочтений: дис. ... канд. техн. наук: 05.11.01 / Ирина Александровна Маринушкина. – Томск, 2015. – 147 с.
19. Муравьев, С.В. Обработка данных межлабораторных сличений мето-

дом агрегирования предпочтений / С.В. Муравьев, И.А. Маринушкина // Измерительная техника. – 2015. – № 12. – С. 3-7.

20. Муравьев, С.В. Элементы формальной теории измерительных информационных систем и ее применение к программированию измерительных процедур : дис. ... доктора техн. наук: 05.11.01, 05.13.11 / Сергей Васильевич Муравьев. – Томск, 1998. – 302 с.

21. Рейнгольд, Э. Комбинаторные алгоритмы. Теория и практика / Э. Рейнгольд, Ю. Нивергельт, Н. Део. – М.: Мир, 1980. – 476 с.

22. Робертс, Ф.С. Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам / Ф.С. Робертс. – М.: Наука, 1986. – 452 с.

23. Фридман, А.А. Дискретные задачи и метод ветвей и границ / А.А. Фридман, А.А. Вотяков А.А. // Экономика и математические методы. – 1974. – № 3. – С. 611-621.

24. Харари, Ф. Теория графов / Ф. Харари. – М.: Мир. – 1973. – 300 с.

25. Хо, М.Д. Повышение точности анализа гетероскедастичных измерительных данных : дис. ... канд. техн. наук: 05.11.01 / Дай Минь Хо. – Томск, 2021. – 123 с.

26. Худоногова, Л.И. Комплексование интервальных измерительных данных методом агрегирования предпочтений: дис. ... канд. техн. наук: 05.11.13 / Худоногова Людмила Игоревна. – Томск, 2017. – 142 с.

27. Циглер, Г. М. Теория многогранников / Г.М. Циглер. – М.: МЦНМО. – 2014. – 568 с.

28. Ali, A. Experiments with Kemeny ranking: What works when? / A. Ali, M. Meilă // Mathematical Social Sciences. – 2012. – Vol. 64. – No. 1. – P. 28-40.

29. Anderson, J.A. Discrete Mathematics with Combinatorics / J.A. Anderson // Upper Saddle River: Prentice Hall. – 2003. – 928 p.

30. Arrow, K.J. Social Choice and Individual Values / K.J. Arrow // New York: Wiley. –1963.

31. Barthelemy J.P. Median linear orders: Heuristics and a branch and bound

algorithm / J.P. Barthelemy, A. Guenoche, O. Hudry // European Journal of Operational Research. – 1989. – Vol. 42. – No. 3. – P. 313-325.

32. Blum, A. Foundations of Data Science / A. Blum, J. Hopcroft, R. Kannan // New York: Cambridge University Press. – 2020.

33. Blumenthal, S. Estimating the Binomial Parameter n / S. Blumenthal, R.C. Dahiya // Journal of the American Statistical Association. – 1981. – Vol. 76. – No. 376. – P. 903-909.

34. Can, B. A re-characterization of the Kemeny distance / B. Can, T. Storcken // Journal of Mathematical Economics. – 2018. – Vol. 79. – P. 112-116.

35. Charon, I. An updated survey on the linear ordering problem for weighted or un-weighted tournaments / I. Charon, O. Hudry // Annals of Operations Research. – 2010. – Vol. 175. – P. 107-158.

36. Chen, X. Pairwise ranking aggregation in a crowdsourced setting / X. Chen, P.N. Bennett, K. Collins-Thompson, E. Horvitz // Proceedings of the 6th ACM International Conference on Web Search and Data Mining (February 04 –08, 2013 Rome, Italy). – 2013. – P. 193-202.

37. Cygan, M. Parameterized Algorithms / M. Cygan, F.V. Fomin, Ł. Kowalik, D. Lokshtanov et al. – New York: Springer, 2015.

38. Davenport, A. A computational study of the Kemeny rule for preference aggregation / A. Davenport, J. Kalagnanam // Association for the Advancement of Artificial Intelligence. – 2004. – P. 697-702.

39. De Donder, P. Choosing from a weighted tournament / P. De Donder, M. Le Breton, M. Truchon // Mathematical Social Sciences. – 2000. – Vol. 40. – P. 85-109.

40. Dwork, C. Rank aggregation methods for the web / C. Dwork, R. Kumar, M. Naor, D. Sivakumar. – 2001. – P. 613-622.

41. Emelyanova, E.Y. Combinatorial characterization of inrankings as weak orders induced by intervals / S.V. Muravyov, E.Y. Emelyanova // Journal of Physics: Conference Series. – 2019. – Vol. 1379. – P. 012052 (Proceedings of Joint IMEKO TC1-TC7-TC13-TC18 Symposium, 2-5 July 2019, Saint-Petersburg, Russia).

42. Emelyanova, E.Y. How to transform all multiple solutions of the Kemeny

Ranking Problem into a single solution / S.V. Muravyov, P.F. Baranov, E.Y. Emelyanova // *Journal of Physics: Conference Series*. – 2019. – Vol. 1379. – P. 012053 (Proceedings of Joint IMEKO TC1-TC7-TC13-TC18 Symposium, 2-5 July 2019, Saint-Petersburg, Russia).

43. Emelyanova, E.Y. Interval data fusion with preference aggregation // S.V. Muravyov, L.I. Khudonogova, E.Y. Emelyanova // *Measurement*. – 2018. – Vol. 116. – P. 621-630.

44. Emelyanova, E.Y. Kemeny rule for preference aggregation: reducing all exact solutions to a single one / S.V. Muravyov, E.Y. Emelyanova // *Measurement*. – 2021. – Vol. 182. – P.109403.

45. Finkelstein, L. A review of the fundamental concepts of measurement / L. Finkelstein, M.S. Leaning // *Measurement*. – 1984. – Vol. 2. – No. 2 – P. 25–34.

46. Fischer, F. Weighted Tournament Solutions / F. Fischer, O. Hudry, R. Niedermeier. In: F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. Procaccia (Eds.) // *Handbook of Computational Social Choice*. – New York: Cambridge University Press, 2016. – P. 85-102.

47. Fishburn, P.C. Betweenness, orders and interval graphs / P.C. Fishburn // *Journal of Pure and Applied Algebra*. – 1971. – Vol. 1. – No. 2 – P. 159-178.

48. Franceschini, F. Aggregating multiple ordinal rankings in engineering design: the best model according to the Kendall's coefficient of concordance / F. Franceschini, D. Maisano // *Research in Engineering Design*. – Vol. 32. – 2021. – P. 91-103.

49. Gaertner, W. A Primer in Social Choice Theory / W. Gaertner. – Oxford: Oxford University Press, 2009. – 232 p.

50. Geanakoplos, J. Three brief proofs of Arrow's Impossibility Theorem / J. Geanakoplos // *Economic Theory*. – 2005. – Vol. 26. – P. 211-215.

51. Gehrlein, W.V. Condorcet's paradox / W.V. Gehrlein. – Berlin: Springer, 2006. – 298 p.

52. Grandi, U. A Borda count for collective sentiment analysis / U. Grandi, A. Loreggia, F. Rossi, V. Saraswat // *Annals of Mathematics and Artificial Intelligence*. –

2016. – Vol. 77. – No. 3-4. – P. 281-302.

53. Hogg, R.V. Introduction to Mathematical Statistics / R.V. Hogg, A.T. Craig. – New York: Macmillan. – 1995. – 438 p.

54. Jones, B. Condorcet Winners and the Paradox of Voting: Probability Calculations for Weak Preference Orders / B.Jones, B.Radcliff, C.Taber, R.Timpone // The American Political Science Review. – 1995. – No. 89. – P. 137-144.

55. Kemeny, J.G. Mathematical Models in the Social Sciences / J.G. Kemeny, J.L. Snell. – New York: Blaisdell, 1962.

56. Kemeny, J.G. Mathematics without numbers / J.G. Kemeny // Cambridge: Daedalus. – 1959. – Vol. 88 – No. 4. – P. 577-591.

57. Kondratev, A.Yu. Positional voting methods satisfying the criteria of weak mutual majority and Condorcet loser / A.Yu. Kondratev // Automation and Remote Control. – 2018. – Vol. 79. – No. 8 – P. 1489-1514.

58. Kwan, A. Vernier scales and other early devices for precise measurement / A. Kwan // American Journal of Physics. – 2011. – Vol. 79. – P. 368-373.

59. List, C. Social Choice Theory, The Stanford Encyclopedia of Philosophy, E.N. Zalta (ed.), <https://plato.stanford.edu/archives/win2013/entries/social-choice/>.

60. Luce, R.D. Psychophysical scaling / R.D. Luce, R.R. Bush, E. Galanter (Eds.) / Handbook of mathematical psychology // New York: Wiley. – 1963. – Vol. 1.

61. Maassen, H. Generating random weak orders and the probability of a Condorcet winner / H. Maassen, T. Bezembinder // Social Choice and Welfare. – 2002. – Vol. 19. – No. 3. – P. 517-532.

62. Mandal, M. Multiobjective PSO-based rank aggregation: Application in gene ranking from microarray data / M. Mandal, A. Mukhopadhyay // Information Sciences. – 2017. – P. 385-386.

63. Marmor, Y.N. Accuracy of multi-experts' prioritization under Mallows' model of errors creation / Y.N. Marmor, T. Gadrich, E. Bashkansky // Quality Engineering. – 2020. – DOI: 10.1080/08982112.2020.1830419.

64. Maul, A. The quality of measurement results in terms of the structural features of the measurement process / A. Maul, L. Mari, D.T. Irribarra, M. Wilson // Meas-

urement. – 2018. – Vol. 116. – P. 611-620.

65. Merlin, V. The axiomatic characterization of majority voting and scoring rules / V. Merlin // *Mathematical Social Sciences*. – 2003. – Vol. 41. – No. 161. – P. 87-109.

66. Michell, J. Representational measurement theory: Is its number up? / J. Michell // *Theory Psychol.* – 2020. – doi:10.1177/0959354320930817.

67. Mirkin, B. Distance and consensus for preference relations corresponding to ordered partitions / B. Mirkin, T.I. Fenner // *Journal of Classification*. – Vol. 36. – 2019. – P. 350-367.

68. Mirkin, B. *Group Choice* / B. Mirkin. – Washington: Halsted Press. – 1979. – 252 p.

69. Muravyov, S.V. Adjustment of fundamental physical constant values using the interval fusion with preference aggregation / S.V. Muravyov, L.I. Khudonogova, M.D. Ho // *Measurement*. – 2020. – Vol. 163. – P. 108037.

70. Muravyov, S.V. Analysis of heteroscedastic measurement data by the self-refining method of interval fusion with preference aggregation –IF&PA / S.V. Muravyov, L.I. Khudonogova, M.D. Ho // *Measurement*. – 2021. – Vol. 183. – 109851. doi: 10.1016/j.measurement.2021.109851

71. Muravyov, S.V. Dealing with chaotic results of Kemeny ranking determination / S.V. Muravyov // *Measurement*. – 2014. – Vol. 51. – P. 328-334.

72. Muravyov, S.V. Discrete-mathematical approach to formal description of measurement procedure / S.V. Muravyov, V. Savolainen // *Measurement*. – 1996. – Vol. 18. – No. 2. – P. 71-80.

73. Muravyov, S.V. Largest consistent subsets in interlaboratory comparisons: preference aggregation approach / S.V. Muravyov, I.A. Marinushkina // *Proc. 14th Joint Intern. IMEKO TC1, TC7, TC13*. – 2011. – P. 69-73.

74. Muravyov, S.V. Numerical experimental investigation of comparison data evaluation method using preference aggregation / S.V. Muravyov, I.A. Marinushkina, D.D. Garif // *Acta IMEKO*. – 2017. – Vol. 6. – No. 1. – P. 20-26.

75. Muravyov, S.V. Some structural properties of a formal model of measure-

ment procedure / S. V. Muravyov, V. Savolainen // *Measurement*. – 1996. – Vol. 18. – No. 2. – P. 81-87.

76. Muravyov, S.V. Precise measurand value estimating by interval fusion with preference aggregation: heteroscedasticity case / S.V. Muravyov, L.I. Khudonogova, M.D. Ho // *Global trends in Testing, Diagnostics & Inspection for 2030 Proceedings 17th IMEKO TC 10 and EUROLAB Virtual Conference, Dubrovnik, October 20-22, 2020: 2nd Conference jointly organized by IMEKO and EUROLAB aisbl: . — Budapest : IMEKO, 2020. – P. 208-213.*

77. Muravyov, S.V. Ordinal measurement, preference aggregation and interlaboratory comparisons / S.V. Muravyov // *Measurement*. – 2013. – Vol. 46. – P. 2927–2935.

78. Muravyov, S.V. Preference aggregation in interpretation of energy survey data / S.V. Muravyov, M.A. Borisova // *Tomsk: TPU, Geo Assets Engineering*. – 2018. – Vol. 329. – No. 12. – P. 155-163.

79. Muravyov, S.V. Processing data from interlaboratory comparisons by the method of preference aggregation / S.V. Muravyov, I.A. Marinushkina // *Measurement Techniques*. – 2016. – Vol. 58. – No. 12. – P. 1285–1291.

80. Muravyov, S.V. Rankings as ordinal scale measurement results / S.V. Muravyov // *Metrology and Measurement Systems*. – 2007. – Vol. XIV. – No. 1 – P. 9-23.

81. Muravyov, S.V. Representation of interval data by weak orders yields robustness of the data fusion outcomes / S.V. Muravyov, L.I. Khudonogova, I.A. Marinushkina // *Journal of Physics: Conference Series*. – 2016. – Vol. 772. – No. 1. – P. 012064.

82. Muravyov, S.V. Representation theory treatment of measurement semantics for ratio, ordinal and nominal scales / S.V. Muravyov, V. Savolainen // *Measurement*. – 1997. – Vol. 22. – No. 1-2. – P. 37-46.

83. Muravyov, S.V. Scales and covers in general measurement problem / S.V. Muravyov, V. Savolainen // *IEEE Transactions on Instrumentation and Measurement*. – 1997. – Vol. 46. – No. 4. – P. 961-965.

84. Mustapää, T. Metrological challenges in collaborative sensing: applicability of digital calibration certificates / T. Mustapää, P. Nikander, D. Hutzschenreuter, R. Viitala // *Sensors*. – 2020. – Vol. 20. – No. 17. – P. 4730.
85. Nielsen, J. A mathematical model of the finding of usability problems / J. Nielsen, T.K. Landauer // *Proceedings of the ACM/IFIP INTERCHI'93 Conference*, IOS Press, The Netherlands: Amsterdam. – 1993. – P. 206-213.
86. Nurmi, H. Preferences, choices, tournaments: alternative foundations for the evaluation of voting schemes / H. Nurmi // *Kluwer Academic Publishers*. – 1991. – Vol. 25. – P.393-405.
87. Patty, J.W. Measuring Fairness, Inequality, and Big Data: Social Choice Since Arrow / J.W. Patty, E.M. Penn // *Annual Review of Political Science*. – 2019. – Vol. 22. P. 435–460.
88. Roberts, F.S. *Measurement Theory* / F.S. Roberts // Reading, Addison-Wesley – New York: Cambridge University Press (digitally reprinted). – 2009.
89. Rossi, G.B. A formal theory of the measurement system / G.B. Rossi, F. Crenna // *Measurement*. – 2018. – Vol.116. – P. 644–651.
90. Saari, D.G. A geometric examination of Kemeny's rule / D.G. Saari, V.R. Merlin // *Social Choice and Welfare*. – 2000. – Vol.17. – No. 3. – P. 403-438.
91. Sloane, N.J.A. The On-Line Encyclopedia of Integer Sequences. The OEIS Foundation, <https://oeis.org>.
92. Smith, A.F.M. An algorithm for determining Slater's i and all nearest adjoining orders / A.F.M. Smith, C.D. Payne // *British Journal of Mathematical and Statistical Psychology*. – 1974. – Vol. 27. – P. 49-52.
93. Tsetlin, I. The impartial culture maximizes the probability of majority cycles / I. Tsetlin, M. Regenwetter, B. Grofman // *Social Choice and Welfare*. – 2003. – Vol. 21. – P. 387-398.
94. Van Deemen, A.M.A. The probability of the paradox of voting for weak preference orderings / A.M.A. Van Deemen // *Social Choice and Welfare*. – 1999. – No. 16. – P. 171-182.
95. Vanacore, A. Some metrological aspects of preferences expressed by prior-

itization of alternatives / A. Vanacore, Y.N. Marmor, E. Bashkansky // *Measurement*. – 2019. – Vol. 135. – P. 520-526.

96. Wilrich, P.-Th. Rounding of measurement values or derived values / P.-Th. Wilrich // *Measurement*. – 2005. – Vol. 37. – P. 21-30.

97. Young, H.P. A Consistent Extension of Condorcet's Election Principle / H.P. Young, A. Levenglick // *SIAM Journal on Applied Mathematics*. – 1978. – Vol. 35. – P. 285–300. – No. 9. – P. 51-64.

98. Young, H.P. An axiomatization of Borda's rule / H.P. Young // *Journal of Economic Theory*. – 1974. – Vol. 9. – P. 43-52.

99. Young, H.P. Condorcet's theory of voting / H.P. Young // *American Political Science Review*. – 1988. – Vol. 82. – No. 2. – P. 1231–1244.

100. Young, H.P. Optimal voting rules / H.P. Young // *Journal of Economic Perspectives*. – 1995. – No. 9. – P. 51-64.

101. Zwicker, W. Introduction to the theory of voting / W. Zwicker, In: F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. Procaccia (Eds.) // *Handbook of Computational Social Choice*, New York: Cambridge University Press. – 2016. – P. 23-56.

ПРИЛОЖЕНИЕ А

Акты внедрения результатов диссертационной работы



"УТВЕРЖДАЮ"

Проректор ТУСУРа по научной работе
и инновациям, к.т.н.

А.Г. Лоцилов

"15" 03 2022 г.

А К Т**об использовании результатов****диссертации Емельяновой Екатерины Юрьевны на тему:**

**"Агрегирование предпочтений на основе точного решения задачи о ранжировании Кемени",
представленной на соискание ученой степени кандидата технических наук
по специальности 2.3.1 – Системный анализ, управление и обработка информации,
в научно-исследовательской лаборатории "Безопасность и электромагнитная
совместимость радиоэлектронных средств" (НИЛ «БЭМС РЭС») ТУСУРа**

Комиссия в составе: Куксенко С.П., заведующий НИЛ «БЭМС РЭС», д.т.н., Комнатнов М.Е.,
с.н.с. НИЛ «БЭМС РЭС», к.т.н., Демаков А.В., м.н.с. НИЛ «БЭМС РЭС» – составила настоящий
акт о том, что результаты диссертационной работы Емельяновой Е.Ю.:

- подход к преобразованию множества оптимальных решений задачи о ранжировании Кемени в единственное итоговое решение и
- метод комплексирования интервальных данных агрегированием предпочтений (IF&PA)

используются в НИЛ «БЭМС РЭС» Томского государственного университета систем управления и радиоэлектроники при обработке неравноточных результатов измерений радиотехнических величин.

Метод IF&PA позволяет гарантированно повысить точность и робастность определяемых с его помощью оценок измеряемых величин и может использоваться для обработки интервальных данных произвольной природы.

Зав. НИЛ «БЭМС РЭС»

Куксенко С.П.

Ст. научный сотрудник НИЛ «БЭМС РЭС»

Комнатнов М.Е.

Мл. научный сотрудник НИЛ «БЭМС РЭС»

Демаков А.В.

ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ
«Г О Р С Е Т И»

Россия, 634012 г.Томск, ул. Шевченко,62а,
тел. (3822) 999-883, факс (3822) 999-677, ИНН 7017081040, КПП 701701001,
Р/с 40702810664000045810 Отделение №8616 Сбербанка России г. Томска, БИК 046902606

"УТВЕРЖДАЮ"

Технический директор ООО "Горсети",

Р.Х. Валитов

"__" ____ 2022 г.

А К Т

о внедрении результатов

кандидатской диссертации Емельяновой Е.Ю.

в лаборатории службы релейной защиты и испытаний электрооборудования (СРЗиИЭ)
ООО "Горсети"

Комиссия в составе: Карпов Л.С., начальник лаборатории СРЗиИЭ, Тянь В.Ю., инженер лаборатории СРЗиИЭ, Семенова О.Л., инженер-метролог лаборатории СРЗиИЭ – составила настоящий акт о том, что следующие результаты диссертационной работы Емельяновой Е.Ю. используются при испытаниях электрооборудования для обработки неравноточных результатов измерений электрических величин:

- метод комплексирования интервальных измерительных данных агрегированием предпочтений;
- алгоритмическое и программное обеспечение выбора единственного отношения консенсуса для выходного профиля задачи о ранжировании Кемени.

Метод комплексирования интервальных измерительных данных агрегированием предпочтений позволяет повысить точность и достоверность оценок измеряемых величин по сравнению с традиционными методами.

Начальник лаборатории СРЗиИЭ

Инженер лаборатории СРЗиИЭ

Инженер-метролог лаборатории СРЗиИЭ

Карпов Л.С.

Тянь В.Ю.

Семенова О.Л.





Ministry of Science and Higher Education of the Russian Federation
Federal State Autonomous Educational Institution of Higher Education
«National Research Tomsk Polytechnic University» (TPU)
30, Lenin ave., Tomsk, 634050, Russia
Tel. +7-3822-606333, +7-3822-701779,
Fax +7-3822-606444, e-mail: tpu@tpu.ru, tpu.ru
OKPO (National Classification of Enterprises and Organizations):
02069303,
Company Number: 027000890168,
VAT/KPP (Code of Reason for Registration)
7018007264/701701001, BIC 016902004

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное
учреждение высшего образования
«Национальный исследовательский
Томский политехнический университет» (ТПУ)
Ленина, пр., д. 30, г. Томск, 634050, Россия
тел.: +7-3822-606333, +7-3822-701779,
факс +7-3822-606444, e-mail: tpu@tpu.ru, tpu.ru
ОКПО 02069303, ОГРН 1027000890168,
ИНН/КПП 7018007264/701701001, БИК 016902004

"УТВЕРЖДАЮ"

И.о. директора ИШИТР, к.т.н.,

Демин А.Ю.

" " 2022 г.

Акт

о внедрении в учебный процесс результатов
диссертации Емельяновой Е.Ю. на тему: "Агрегирование предпочтений на основе точного
решения задачи о ранжировании Кемени", представленной на соискание ученой степени
кандидата технических наук

Комиссия в составе: председателя – заведующего кафедрой-руководителя отделения
автоматизации и робототехники на правах кафедры, к.т.н., доцента Филипас А.А., и членов
– к.т.н., доцента Спиридоновой А.С. и к.т.н., доцента Кузьминской Е.В. – составила
настоящий акт в том, что результаты диссертационной работы Емельяновой Е.Ю.:

- алгоритм и программное обеспечение решения задачи о ранжировании Кемени
- метод комплексирования интервалов агрегированием предпочтений IF&PA для
повышения точности оценок интервальных данных различного происхождения;
- способ выбора единственного отношения консенсуса для множества оптимальных
решений задачи о ранжировании Кемени –

используются при проведении практических и лабораторных занятий по дисциплинам
"Квалиметрия и управление качеством", "Статистические методы в измерениях".
Результаты обеспечивают возможность применения инновационных технологий
преподавания и повышают качество учебного процесса.

Заведующий кафедрой-руководитель
отделения автоматизации и робототехники
на правах кафедры ИШИТР

Доцент ОАР

Доцент ОАР

Филипас А.А.

Спиридонова А.С.

Кузьминская Е.В.