ТОМСКИЙ
ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ

На правах рукописи

Емельянова Екатерина Юрьевна

АГРЕГИРОВАНИЕ ПРЕДПОЧТЕНИЙ НА ОСНОВЕ ТОЧНОГО РЕШЕНИЯ ЗАДАЧИ О РАНЖИРОВАНИИ КЕМЕНИ

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

2.3.1 – Системный анализ, управление и обработка информации

Томск – 2022

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования "Национальный исследовательский Томский политехнический университет"

Научный руководитель:	**Муравьев Сергей Васильевич**
	доктор технических наук, профессор, Национальный исследовательский Томский политехнический университет
Официальные оппоненты:	**Захарова Алёна Александровна**
	доктор технических наук, главный научный сотрудник, Институт проблем управления им. В.А. Трапезникова РАН
	Ходашинский Илья Александрович
	доктор технических наук, профессор, Томский государственный университет систем управления и радиоэлектроники

Защита состоится 27.06.2022 г. в 15 часов 00 мин на заседании диссертационного совета ДС.ТПУ.15 при федеральном государственном автономном образовательном учреждении высшего образования "Национальный исследовательский Томский политехнический университет" по адресу: Россия, 634034, г. Томск, ул. Советская, 84/3, ауд. 214.

С диссертацией можно ознакомиться в научно-технической библиотеке ФГАОУ ВО "Национальный исследовательский Томский политехнический университет" и на сайте dis.tpu.ru.

Автореферат разослан "___" ________ 2022 г.

Ученый секретарь
диссертационного совета ДС.ТПУ.15
к.т.н. А.Я. Пак

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. *Агрегирование предпочтений* – это классическая задача определения ранжирования консенсуса для заданного исходного профиля предпочтений, состоящего из m ранжирований, возможно, включающих 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 нетранзитивности 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 полиномиальным 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 $N >> m$, 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 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.
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 агрегированием 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 123 страницы основного текста, включая 33 рисунка и 29 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

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

В первой главе "Алгоритм точного решения задачи о ранжировании Кемени" 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.

Пусть дано некоторое множество $A = \{a_1, a_2, ..., a_n\}$, $n \geq 2$, состоящее из элементов (альтернатив, кандидатов, объектов и т.п.) a_i, $i = 1, \ldots, n$, подлежащих сравнению по степени проявления *некоторого признака* или *свойства* k из списка $K = \{1, 2, ..., m\}$. По каждому k-му признаку 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. $a_i \sim_k a_j$. Тогда k-ое *ранжирование* (или *слабый порядок*) – это бинарное отношение предпочтения λ_k на множестве A в виде цепочки

$$\lambda_k = (a_1 \succ_k a_2 \succ_k \ldots \sim_k a_s \sim_k a_t \succ_k \ldots \sim_k a_n),\ k = 1, \ldots, m. \tag{1}$$

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

Профилем предпочтения называется множество $\Lambda(m, n) = \{\lambda_1, \lambda_2, ..., \lambda_m\}$, состоящее из m ранжирований n альтернатив, определенных на множестве A.

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

Каждое ранжирование в профиле предпочтения $\Lambda(m, n)$ можно рассматривать как *турнир*, представляющий собой процедуру *парного сравнения альтернатив*. Если альтернативы сравниваются m раз, то турнир называется m-взвешенным. Тогда профиль $\Lambda(m, n)$ можно представить m-взвешенной турнирной $(n \times n)$ матрицей $S = [s_{ij}]$, где:

$$s_{ij} = \sum\nolimits_{k=1}^{m} \mathbf{I}_k(a_i, a_j),\ i \neq j = 1, ..., n; \tag{2}$$

$$\mathbf{I}_k(a_i, a_j) = \begin{cases} 1 & \text{если } a_i \succ_k a_j \\ 0{,}5 & \text{если } a_i \sim_k a_j \\ 0 & \text{если } a_i \prec_k a_j \end{cases}. \tag{3}$$

Для турнирной матрицы S справедливо следующее свойство:

$$s_{ij} + s_{ji} = m \text{ для } i \neq j;\ i, j = 1, \ldots, n, \tag{4}$$

где значение s_{ij} равно доле значения m, которую альтернатива a_i выигрывает (приобретает) в результате сравнения с альтернативой a_j.

Цель агрегирования предпочтений состоит в определении *ранжирования консенсуса* β, которое является единственным отношением предпочтения, дающим интегральную характеризацию профиля $\Lambda(m, n)$. Способ нахождения ранжирования консенсуса определяется конкретным аксиоматически обоснованным *правилом агрегирования предпочтений* (или функцией социального выбора). В диссертационной работе рассмотрены четыре основных правила агрегирования: правило простого большинства, правило Борда, правило Кондорсе и правило Кемени.

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

$w(a_i) = |\{k \mid r_i^k = 1, k = 1,...,m\}|$. Альтернатива a_i называется *победителем* $SM(a_i)$ *по правилу простого большинства*, который определяется выражением:

$$SM(a_i) = \{a_i \mid \max[w(a_i)], i = 1, ..., n\}. \tag{5}$$

Это правило часто незаслуженно объявляет победителем кандидата, которому большинство избирателей предпочли другого кандидата, если общее число кандидатов $n \geq 2$, что связано с недостаточным использованием информации о взаимном доминировании альтернатив.

Правило Борда. Пусть сумма элементов i-й строки турнирной матрицы S профиля $\Lambda(m, n)$ называется *выигрышем* $z_i = \sum_{j=1}^{n} s_{ij}$ альтернативы a_i для $i = 1, ..., n$. Выигрыш z_i показывает общий доход от всех выигрышей альтернативы a_i во всех $m(n-1)$ сравнениях. Альтернатива a_i называется *победителем* $Bo(a_i)$ *по правилу Борда*, который определяется выражением:

$$Bo(a_i) = \{a_i \mid \max(z_i), i = 1, ..., n\}. \tag{6}$$

Это правило не всегда приводит к наилучшему отношению консенсуса, однако победитель по правилу Борда с высокой вероятностью может оказаться победителем по правилу Кондорсе, когда он существует.

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

$$Co(a_i) = \{a_i \mid s_{ij} > s_{ji}, i \neq j, i, j = 1,...,n\}. \tag{7}$$

Правило Кондорсе с полным основанием могло бы претендовать на роль наилучшего правила агрегирования, т.к. оно опирается на полную информацию о парных сравнениях альтернатив, содержащуюся в турнирной матрице S. Однако, это правило страдает существенным недостатком: *победитель Кондорсе может не существовать* из-за *парадокса Кондорсе*, когда построение ранжирования консенсуса β в соответствии с выражением (7) приводит к одновременному выполнению трех отношений строгого предпочтения $(a_i \succ a_j) \in \beta$, $(a_j \succ a_k) \in \beta$, $(a_k \succ a_i) \in \beta$. Вероятность возникновения парадокса Кондорсе превышает 50 %.

Правило Кемени. Введем понятие расстояния между ранжированиями в терминах *числа парных несоответствий* (инверсий альтернатив). *Расстоянием Кемени* называется расстояние между двумя ранжированиями λ_k и λ_l, определяемое выражением

$$d(\lambda_k, \lambda_l) = \sum_{i<j} |\operatorname{sgn}(a_i^k, a_j^k) - \operatorname{sgn}(a_i^l, a_j^l)|, \tag{8}$$

где $\operatorname{sgn}(a_i^k, a_j^k) = \begin{cases} 1, & a_i \succ_k a_j \\ 0, & a_i \sim_k a_j \\ -1, & a_i \prec_k a_j \end{cases}$ и $\operatorname{sgn}(a_i^l, a_j^l) = \begin{cases} 1, & a_i \succ_l a_j \\ 0, & a_i \sim_l a_j \\ -1, & a_i \prec_l a_j \end{cases}$.

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

$$D(\lambda, \Lambda) = \sum_{k=1}^{m} d(\lambda, \lambda_k) = \sum_{i<j} \sum_{k=1}^{m} |\operatorname{sgn}(a_i^k, a_j^k) - \operatorname{sgn}(a_i, a_j)|. \tag{9}$$

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

$$p_{ij} = \sum\nolimits_{k=1}^{m}[1-\mathrm{sgn}(a_i^k, a_j^k)],\; i, j = 1, ..., n\,. \tag{10}$$

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 $\mathbb{N}_n = \{1, 2, \ldots, n\}$.

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.

$$\beta = \arg\min_{\rho\in\Pi_n} D(\rho, \Lambda) = \arg\min \sum\nolimits_{i<j} p_{ij}\,. \tag{11}$$

Это правило представляет собой оптимизационную задачу, которую будем называть *задачей о ранжировании Кемени* (ЗРК). В терминах матрицы профиля правило Кемени заключается в нахождении такой перестановки строк и столбцов матрицы P, что сумма элементов её верхней треугольной подматрицы минимальна.

Правило Кемени сохраняет положительные свойства правила Кондорсе и не приводит к парадоксу, однако, при его применении в агрегировании предпочтений существуют две проблемы: ЗРК (1) является *NP*-трудной и (2) может приводить к множественным ранжированиям консенсуса.

Принимая во внимание эти трудности, для получения надежного отношения консенсуса при применении правила Кемени в диссертационной работе предложено придерживаться следующей *стратегии*: (1) всегда находить полный список всех N отношений консенсуса, которые являются точными оптимальными решениями, достижимыми при размерности задачи $n < 20$, и (2) находить для этого списка единственное подходящее итоговое отношение консенсуса.

Все возможные ранжирования Кемени можно найти только теми методами, которые генерируют частичные решения при построении дерева поиска оптимального решения. В данной диссертационной работе для этой цели используется разработанный на основе *метода ветвей и границ* в научной группе под руководством проф. Муравьева С.В. *рекурсивный алгоритм* RECURSALL.

Входом алгоритма служит матрица P профиля предпочтения. Основное тело алгоритма содержит всего два этапа: (1) *инициализацию*, при которой вычисляется *наименьшее возможное расстояние Кемени* $D_{\mathrm{least}} = \sum\nolimits_{i<j} \min(p_{ij}, p_{ji})$ для матрицы P и все необходимые переменные приобретают исходные значения, и (2) *вызов рекурсивной процедуры*, принимающей два параметра K и D, где K – номер уровня дерева поиска, а D – составляющая нижней границы, определяемая фиксированным порядком элементов текущего *частичного* решения.

Рекурсивная процедура содержит основной цикл по номеру частичного решения на K-ом уровне. На каждом этапе цикла формируется частичное решение, которое всякий раз задает новое положение строк и столбцов матрицы P, и минимально возможное значение (нижнюю границу) D_{low} суммы элементов верхней треугольной подматрицы как функции расстояния. В качестве верхней границы D_u используется наименьшее значение D_{low} для полученных к данному моменту *полных* решений, определяющих порядок всех альтернатив.

Если $D_{low} \le D_u$ и текущее решение неполное (т.е. $K < n - 1$), то вызывается рекурсивная процедура с новыми параметрами, обеспечивающими *ветвление* дерева поиска. Если $D_{low} > D_u$, то все текущие частичные решения не могут быть оптимальными и соответствующая часть дерева поиска не проверяется, т.е. *отсекается*. Если при $D_{low} \le D_u$ текущее решение является полным, то оно запоминается как пара β_N и $D_u = D_{low}$. После нахождения первого оптимального решения при $N = 1$ каждое новое найденное оптимальное решение запоминается при соответствующем инкрементированном N. Таким образом, последнее значение N указывает на общее число оптимальных решений. Поиск других неполных и полных решений продолжается, пока все неперспективные решения не будут отсечены. Результатом работы алгоритма являются N оптимальных решений ЗРК; соответствующее этим решениям значение верхней границы функции расстояния D_u и характеризующее матрицу профиля P наименьшее возможное расстояние D_{least}.

Будем называть множество ранжирований Кемени $\mathrm{B}(N, n) = \{\beta_1, \beta_2, \ldots, \beta_N\} \subset \Pi$ *выходным профилем* ЗРК. Тогда профиль предпочтения $\Lambda(m, n) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ можно называть *входным профилем* ЗРК, см. рисунок 1.

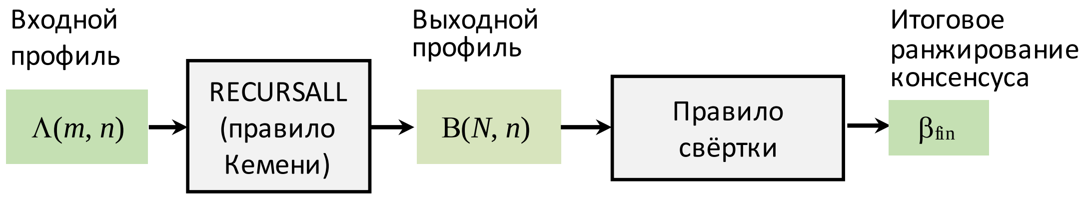

Рисунок 1 – Входной и выходной профили алгоритма RECURSALL

Во второй главе "Множественность решений задачи о ранжировании Кемени" 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 $\mathrm{B}(N, n)$.

Учитывая тот факт, что множество Π_n всех отношений строгого предпочтения (перестановок элементов множества A) обладает конечной мощностью n, ранжирование Кемени всегда существует, однако оно не всегда является единственным, т.к. значение оптимального расстояния $D(\beta, \Lambda)$ может соответствовать не единственной перестановке строк и столбцов матрицы профиля P, т.е.

$$D(\beta_1, \Lambda) = \sum\nolimits_{i<j} p_{ij}^1 = D(\beta_2, \Lambda) = \sum\nolimits_{i<j} p_{ij}^2 = \ldots = D(\beta_N, \Lambda) = \sum\nolimits_{i<j} p_{ij}^N, \qquad (12)$$

Оптимальные решения могут ранжировать альтернативы по-разному, что приводит к гораздо большей неоднозначности, чем та, которая изначально присуща исходному профилю $\Lambda(m, n)$.

Выделяем два основных фактора, приводящие к неединственности оптимальных решений ЗРК.

1. *Несоответствие друг другу мощностей множества возможных решений и множества исходных ранжирований.* Каждое оптимальное (точное) решение 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 нескомпенсированная несогласованность исходных ранжирований между собой.* 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.

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.

$$p_{ik} \leq p_{ki}, \text{ если } p_{ij} \leq p_{ji} \text{ и } p_{jk} \leq p_{kj},\ i \neq j \neq k = 1, \ldots, n. \tag{13}$$

Матрица P всякий раз будет оказываться *нетранзитивной*, если условия (13) нарушаются. Таким образом, транзитивность любого входного профиля, после определения ранжирования консенсуса β по правилу Кемени (11), может быть установлена с использованием следующего критерия:

$$\text{Профиль } \Lambda(m,n) \begin{cases} \text{транзитивен}, & \text{если } D_{\text{least}} = D(\beta,\Lambda) \\ \text{нетранзитивен}, & \text{если } D_{\text{least}} < D(\beta,\Lambda) \end{cases}, \tag{14}$$

где наименьшее расстояние D_{least} достижимо, если матрица P транзитивна.

Транзитивность входного профиля и ее связь с множественностью решений ЗРК были экспериментально исследованы в диссертационной работе. Входные матрицы профиля формировались из ранжирований, полученных путем объединения псевдослучайных строгих порядков ρ и толерантностей τ, сгенерированных по отдельности на основе библиотечной функции Microsoft randomize() и распределенных по *равномерному закону*. Функция randomize() принимает произвольное целое число, так называемую *стартовую точку* S_p, которая служит номером генерируемого профиля. В ходе экспериментов генерировались по 100 индивидуальных профилей (задач) со стартовыми точками $S_p = 1, \ldots, 100$ для каждого набора параметров $m = 3, 4, 5, 6, 7, 10, 15, 49, 50$; $n = 7, 10, 15, 18, 20$ и трех фиксированных значений плотности толерантностей $t \in \{t_0, t_1, t_2\}$. Сгенерированные профили подавались на вход алгоритма RECURSALL. Таким образом, были решены более 13500 индивидуальных задач.

Ранее проведенные экспериментальные исследования множественности 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 (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.

Рисунок 2 иллюстрирует изменение числа оптимальных решений N (вертикальная ось) для групп из 100 индивидуальных задач (горизонтальная ось) при m = 4, 5 и 49 для различных значений t (сплошная линия – t_0, штрих-пунктирная t_1; пунктирная t_2) и n = 10, 15, 20. Заметим, что значения S_p отсортированы в порядке убывания значений N.

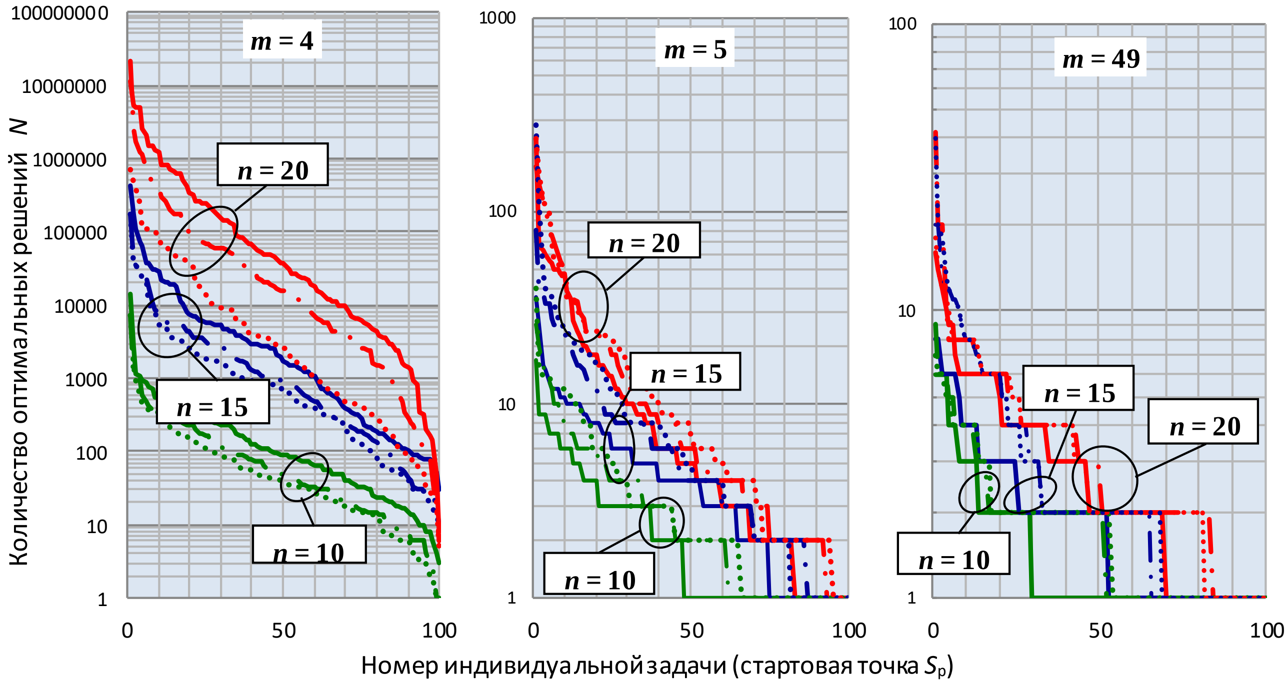

Рисунок 2 – Множественность оптимальных решений N для m = 4; 5 и 49 при t_0, t_1, t_2

Пример результатов исследований транзитивности/нетранзитивности в соответствии с критерием (14) входного профиля при различных m и n показан на рисунке 3 при m = 4, 5 и 49 и n = 7, 10, 15, 18, 20. Кривые, соответствующие расстоянию D_{least}, показаны ромбами, а $D(\beta, \Lambda)$ – квадратами.

Получены следующие результаты экспериментальных исследований транзитивности/нетранзитивности входного профиля: (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) минимальная степень нетранзитивности достигается при малых чётных значениях m при отсутствии толерантностей; степень нетранзитивности уменьшается по мере уменьшения числа оптимальных решений ЗРК.

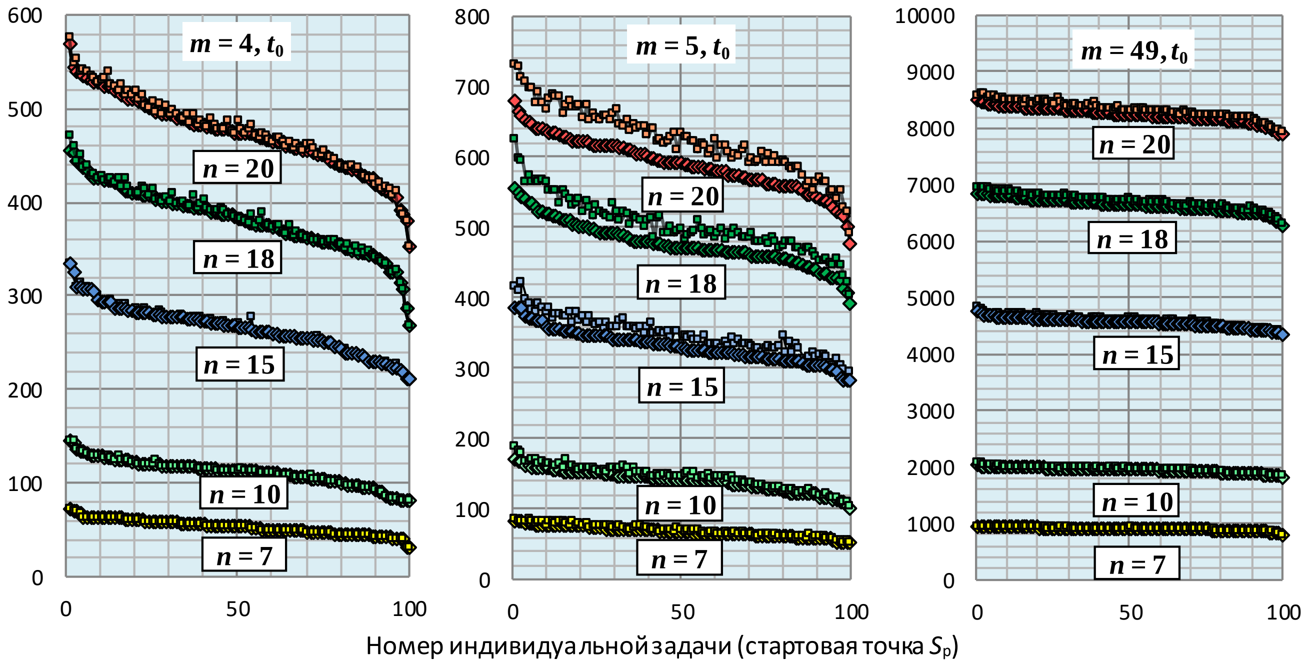

Рисунок 3 – Расстояния D_{least} и $D(\beta, \Lambda)$ для $m = 4$; 5; 49 при t_0, t_1, t_2

Числовые данные, полученные в ходе экспериментов, позволили 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 может привести к интересным результатам.

Пусть $P(m, n)$ – вероятность победителя Кондорсе; $Q(m, n) = 1 - P(m, n)$ – вероятность парадокса Кондорсе; $P_{\text{tr}}(m, n)$ – вероятность появления транзитивного профиля; $Q_{\text{tr}}(m, n) = 1 - P_{\text{tr}}(m, n)$ – вероятность появления нетранзитивного профиля. Вероятность $P_{\text{tr}}(m, n)$ для группы из 100 индивидуальных задач может быть оценена на основе полученных экспериментальных данных следующим образом: $P_{\text{tr}}(m, n)$ = число транзитивных профилей / число задач в группе (= 100).

Графики полученных оценок вероятностей $P_{\text{tr}}(m, n)$ для нечетных и четных m приведены на рисунке 4. Из графиков следует, что чем больше m, тем меньше вероятность транзитивности $P_{\text{tr}}(m, n)$. При этом:

- при любых нечетных m вероятность $P_{\text{tr}}(m, n)$ близка к нулю уже при числе n альтернатив, близким к 10 (см. рисунок 4,а);
- вероятность транзитивности значительно выше при четных значениях m, чем при нечетных, и плавно уменьшается с ростом n (см. рисунок 4,б); при четном значении $m = 4$ вероятность транзитивности довольно высока даже для $n = 15$; уже при $m = 10$ вероятность падает и приближается к вероятности для нечетных m, однако при росте четных m остается заметной разница между графиками вероятности для нечетных и четных m.

Выявлено также, что 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.

Полученные экспериментальные оценки вероятностей определенно согласуются с известными аналитическими расчетами вероятности парадокса Кондорсе и позволяют сформулировать следующие утверждения, см. рисунок 5:

- при *нечётных m* вероятность появления нетранзитивного профиля всегда больше или равна вероятности парадокса Кондорсе, т.е. $Q_{tr}(m, n) \geq Q(m, n)$; это 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;
- при *чётных m* вероятность появления нетранзитивного профиля всегда меньше вероятности парадокса Кондорсе, т.е. $Q_{tr}(m, n) < Q(m, n)$.

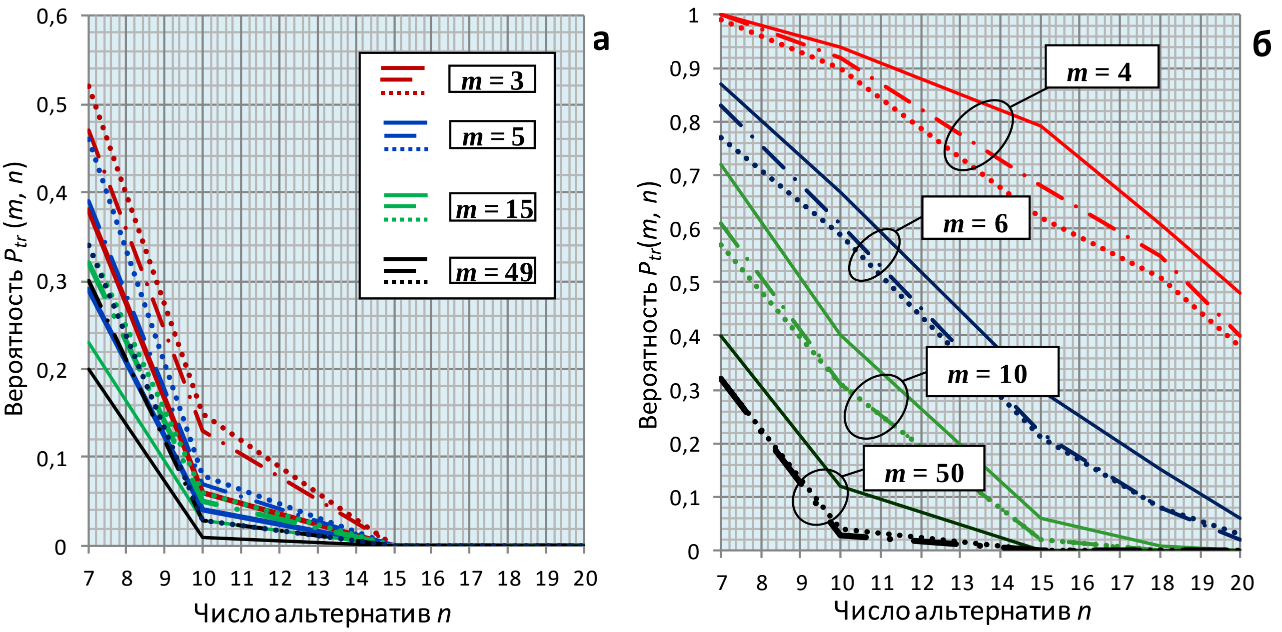

Рисунок 4 – Оценка вероятности $P_{tr}(m, n)$ для нечётных (а) и чётных m (б)

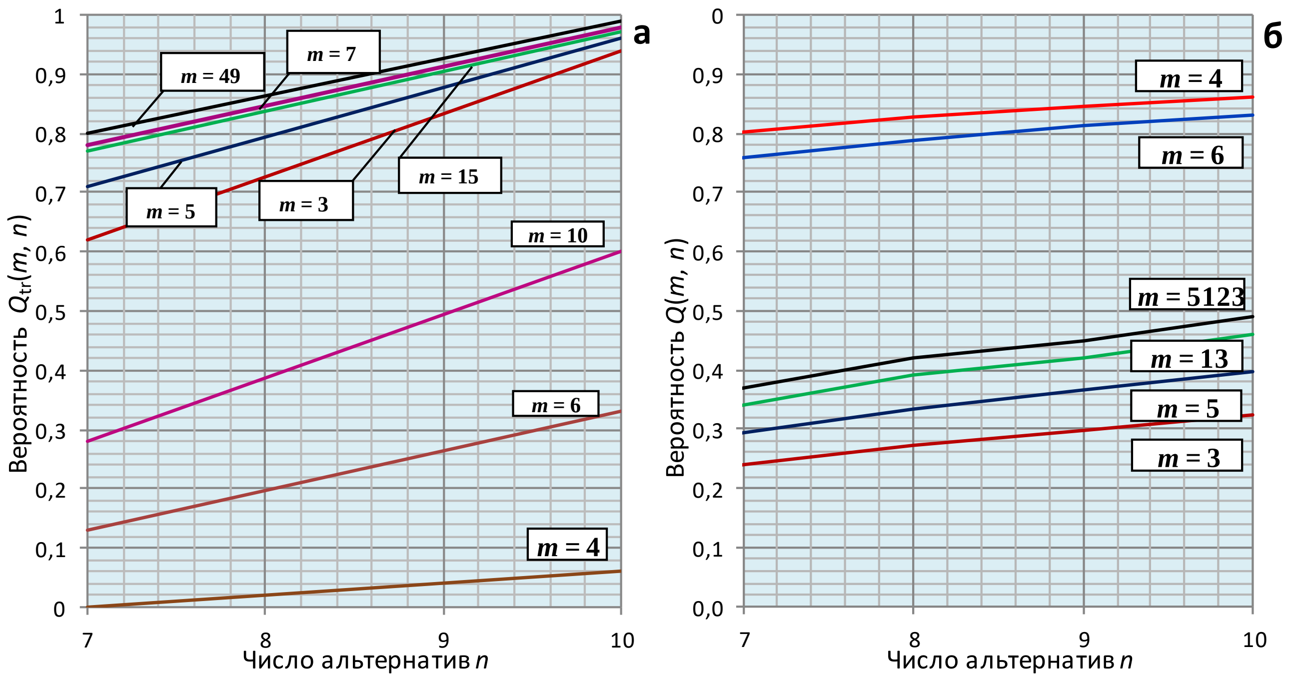

Рисунок 5 – Оценка вероятности нетранзитивного профиля $Q_{tr}(m, n)$ (а) и аналитический расчет вероятности парадокса Кондорсе $Q(m, n)$ (б)

В третьей главе "Преобразование выходного профиля в единственное оптимальное решение" исследуются свойства взвешенной турнирной матрицы, построенной для выходного профиля ЗРК; в терминах этой матрицы введены понятия выигрыша, проигрыша и ранга альтернативы, показана их взаимосвязь. Сформулированы правила преобразования (свертки) выходного профиля в единственное оптимальное решение. Справедливость предложенных правил свёртки демонстрируется на основе счета Борда и понятия отношения "между", заданного на ранжированиях.

Пусть $B(N, n) = \{\beta_1, \beta_2, \ldots, \beta_N\}$, $B \subset \Pi_n$, является выходным профилем ЗРК с входным профилем $\Lambda(m, n) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$, заданным на некотором наборе альтернатив $A = \{a_1, a_2, \ldots, a_n\}$. Каждое ранжирование консенсуса β_k, $k = 1, \ldots, N$, является антисимметричным и антирефлексивным отношением, т.е. отношением строгого порядка, не содержащим никаких толерантностей. Представим $B(N, n)$ *N-взвешенной турнирной матрицей* $S = [s_{ij}]$, которая определяется выражением (2), но для случая N строгих ранжирований, где

$$s_{ij} = \sum_{k=1}^{N} \mathbf{I}_k(a_i, a_j),\ i \neq j = 1, \ldots, n; \qquad (15)$$

$$\mathbf{I}_k(a_i, a_j) = \begin{cases} 1, \text{ если } (a_i, a_j) \in \beta_k \\ 0, \text{ если } (a_i, a_j) \notin \beta_k \end{cases}. \qquad (16)$$

В диссертационной работе доказаны два предложения.

Предложение 1. Для любого выходного профиля $B(N, n)$, представленного соответствующей турнирной матрицей $[s_{ij}]$, справедливо следующее выражение:

$$y_j + z_i = N(n - 1),\ i = 1, \ldots, n, \qquad (17)$$

где *проигрыш* $y_j = \sum_{i=1}^{n} s_{ij}$ альтернативы a_j – это сумма элементов j-го столбца матрицы S; *выигрыш* $z_i = \sum_{j=1}^{n} s_{ij}$ альтернативы a_j – это сумма элементов i-й строки матрицы S.

Предложение 2. Для любого выходного профиля $B(N, n)$, представленного соответствующей турнирной матрицей $[s_{ij}]$, справедливо следующее выражение:

$$r_i = nN - z_i,\ i = 1, \ldots, n, \qquad (18)$$

где общий ранг $r_i = \sum_{k=1}^{N} r_i^k$ альтернативы a_i – это сумма ее рангов r_i^k во всех ранжированиях консенсуса $\beta_k \in B(N, n)$, $k = 1, \ldots, N$.

Показано, что выходной профиль $B(N, n)$ имеет следующие свойства:

(i) каждое ранжирование $\beta_k \in B(N, n)$ является оптимальным решением для входного профиля $\Lambda(m, n)$ по правилу Кемени (11); это означает, что все ранжирования находятся на одинаковом расстоянии $D(\beta_k, \Lambda)$ от профиля $\Lambda(m, n)$, 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;

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

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

На основе аксиоматического подхода были сформулированы *требования к правилу свёртки* всех ранжирований консенсуса в итоговое ранжирование β_{fin}: (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.

Правило свёртки 1. Для заданных выигрышей z_i, вычисленных по взвешенной турнирной матрице S выходного профиля $B(N, n)$, для всех $i < j$, $i, j = 1, \ldots, n$, справедливы следующие соотношения: $z_i > z_j \Rightarrow a_i \succ a_j$ и $z_i = z_j \Rightarrow a_i \sim a_j$, где оба отношения $\succ$ и $\sim$ принадлежат единственному итоговому отношению консенсуса β_{fin}. ♦

Таким образом, итоговое ранжирование консенсуса может включать толерантности, 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). Ясно, что Правило свертки 1 удовлетворяет условию (P4), поскольку оценка его вычислительной сложности равна $O(n^2)$, что следует из простого рассмотрения уравнения (17). Выигрыши z_i, реализующие баллы Борда $\text{Bo}(a_i)$, гарантируют, что Правило свертки 1 преобразует выходной профиль $B(N, n)$ в единственное ранжирование консенсуса β_{fin}, удовлетворяющее свойствам (P1) – (P4).

Учитывая тот факт, что правило Кемени основано на понятии расстояния Кемени, можно утверждать, что итоговое ранжирование консенсуса β_{fin} находится между ранжированиями $\beta_1, \beta_2, \ldots, \beta_N$ выходного профиля $B(N, n)$, если и только если

$$\bigcap_{k=1}^{N} \beta_k \subseteq \beta_{\text{fin}} \subseteq \bigcup_{k=1}^{N} \beta_k . \tag{19}$$

Из выражения (19) 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.

$$D(\beta_{\text{fin}}, \Lambda) \le D(\beta_k, \Lambda), \; k = 1, \ldots, N, \tag{20}$$

откуда следует справедливость условия (P3).

Правило свёртки 2. Для заданных общих рангов r_i, вычисленных по взвешенной турнирной матрице S выходного профиля $B(N, n)$, для всех $i < j$, $i, j = 1, \ldots, n$, справедливы следующие соотношения: $r_i > r_j \Rightarrow a_i \succ a_j$ и $r_i = r_j \Rightarrow a_i \sim a_j$, где оба отношения $\succ$ и $\sim$ принадлежат единственному итоговому отношению консенсуса β_{fin}. ♦

Из Предложения 2 следует, что баллы Борда могут быть рассчитаны как $\text{Bo}(a_i) = nN - r_i$. Таким образом, Правила свёртки 1 и 2 дополняют друг друга и каждое из них гарантированно приводит к одному и тому же итоговому ранжированию консенсуса β_{fin}. Выбор одного из правил можно осуществлять, исходя из практических соображений, например, вычислительных или алгоритмических.

В четвертой главе "Практические аспекты агрегирования предпочтений на основе правила Кемени" 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 инражирования, которые являются частным случаем отношения слабого порядка (или ранжирования) на множестве n дискретных значений, принадлежащих к этим интервалам.

Под *комплексированием интервальных данных* понимается процедура определения такого *результирующего интервала* $[x^* \pm \varepsilon^*]$, который согласован с максимальным количеством исходных интервалов $\{I_k\}$, $k = 1, \ldots, m$, (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^*.

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 $I_k = [x_k - \varepsilon_k, x_k + \varepsilon_k]$; $x_k, \varepsilon_k \in \mathbb{R}$. Для представления исходных интервалов ранжированиями введем понятие *диапазона актуальных значений* (ДАЗ), представляющего собой множество $A = \{a_1, a_2, \ldots, a_n\}$ строго упорядоченных дискретных значений $a_1 < a_2 < \ldots < a_n$. Дискретное множество A может быть получено, исходя из заданных непрерывных интервалов $\{I_k\}_{k=1}^{m}$, в три следующих этапа, см. рисунок 6.

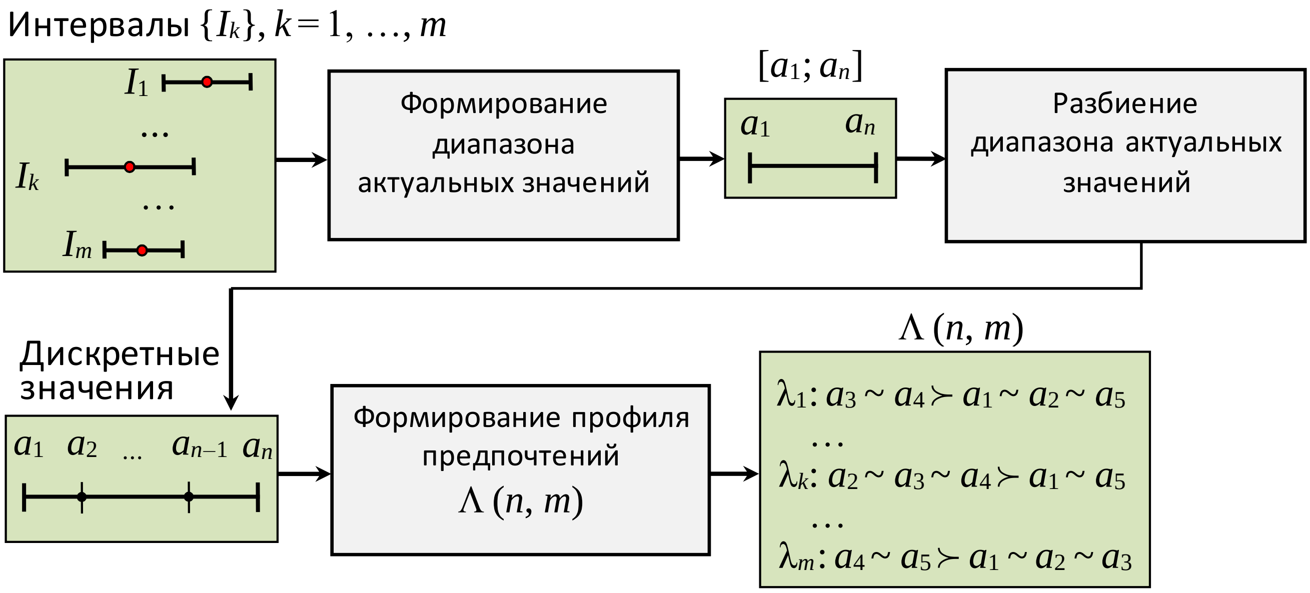

Рисунок 6 – Три этапа перехода от непрерывного пространства значений к дискретному при формировании ДАЗ

Этап 1. *Формирование диапазона актуальных значений*. Наименьшая нижняя граница для всех интервалов выбирается в качестве нижней границы ДАЗ $a_1 = \min \{x_k - \varepsilon_k | \; k = 1, \ldots, m\}$; в качестве верхней границы a_n берется наибольшая верхняя граница для этих интервалов $a_n = \max \{x_k - \varepsilon_k | \; k = 1, \ldots, m\}$.

Этап 2. *Разбиение диапазона актуальных значений*. Для получения элементов $a_2, a_3, \ldots, a_{n-1}$ разбиваем полученный интервал $[a_1, a_n]$ на $n - 1$ равных подынтервалов длиной (нормой) $h = (a_n - a_1)/(n - 1)$, где i-й элемент ДАЗ определяется формулой $a_i = a_{i-1} + h$, $i = 2, \ldots, n$.

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

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, $k = 1, \ldots, m$, имеем разбиение множества A на два класса эквивалентности A_k и $\bar{A}_k$, т.е. $A = A_k \cup \bar{A}_k$, $A_k \cap \bar{A}_k = \varnothing$.

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

$$a_i \in A_k \wedge a_j \notin A_k \Rightarrow a_i \succ a_j; \quad (21)$$

$$a_i, a_j \in A_k \vee a_i, a_j \notin A_k \Rightarrow a_i \sim a_j; \quad (22)$$

$$a_i \notin A_k \wedge a_j \in A_k \Rightarrow a_i \prec a_j; \quad (23)$$

$$a_i, a_j \in A_k \text{ соседние элементы} \Rightarrow j \equiv i + 1. \quad (24)$$

Условие (24) 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*. Заметим, что структура любого k-го инранжирования определяется выражением $\lambda_k = A_k \succ \bar{A}_k$, т.е. каждое инранжирование содержит *единственный* символ строгого порядка $\succ$ и $n - 2$ символов толерантности $\sim$.

Следующие три этапа завершают процедуру IF&PA.

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

Этап 5. *Проведение свертки профиля* $\mathrm{B}(N, n)$ *в* β_{fin}. Преобразуем профиль $\mathrm{B}(N, n)$ в единственное итоговое ранжирование консенсуса β_{fin} по Правилу свёртки 1 или 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^*.

Метод обеспечивает получение более робастного, точного и достоверного результата комплексирования x^* по сравнению с традиционными методами обработки интервальных данных.

Введенное новое понятие инражирования появляется в результате ограничений, накладываемых на обычные ранжирования интервальным характером исходных данных. В рамках данной диссертационной работы были исследованы комбинаторные свойства инранжирований и связанных с ними подпространств слабых порядков.

Пусть 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.

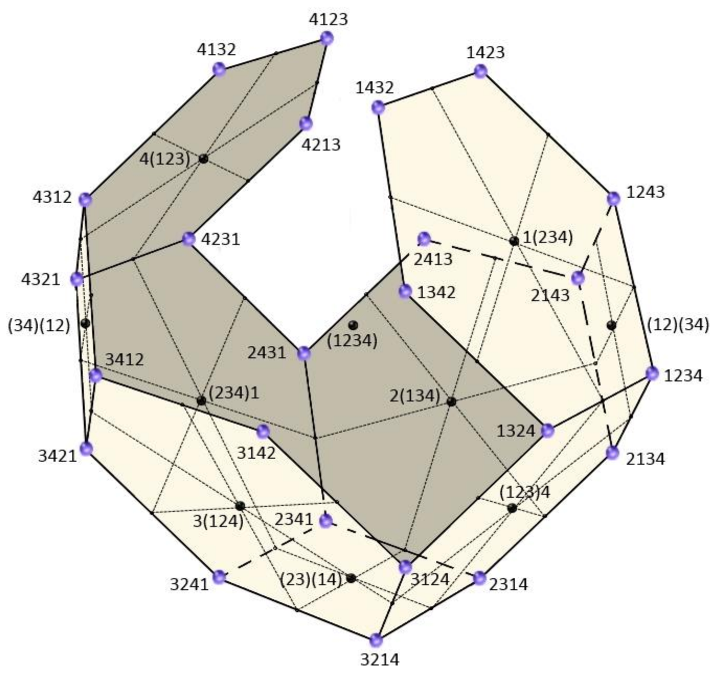

Рисунок 7 – Пространство инражирований для n = 4

Подпространство Ω_2 образовано слабыми порядками, которые удовлетворяют условиям (21)-(24), см. рис. 7. Нарушение условия (24) приводит к существованию подпространства Ω_f *запрещенных слабых порядков*, которые, однако, удовлетворяют условиям (21)-(23). Ясно, что

$$\Omega_f = \bar{\Omega}_2 \cap \Omega_1. \qquad (25)$$

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:

$$\Omega_2 \subseteq \Omega_1 \subseteq \Omega_0. \qquad (26)$$

Для этих пространств, очевидно, справедливы также следующие тождества:

$$\Omega_1 = \Omega_2 \cup \Omega_f,\ \Omega_1 \cap \bar{\Omega}_1 = \varnothing,\ \Omega_2 \cap \bar{\Omega}_2 = \varnothing,\ |\Omega_1| = |\Omega_2| + |\Omega_f|. \qquad (27)$$

Известно, что 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:

$$|\Omega_0| = \sum_{q=0}^{n} q!S_{n,q}. \qquad (28)$$

Показано, что мощность пространства Ω_2 инранжирований равна треугольному числу T_n для заданного числа n дискретных элементов, которое определяется по формуле:

$$\left|\Omega_2\right| = T_n = \frac{n(n+1)}{2} = \binom{n+1}{2} = \binom{n+1}{n-1} = (S_{n+1,n}) = \sum_{i=1}^{n}\left|A_k\right|_i . \qquad (29)$$

Мощность F_n пространства Ω_f запрещенных слабых порядков определяется как:

$$|\Omega_f| = F_n = 2^n - 1 - n(n+1)/2 . \qquad (30)$$

Очевидно, что мощность подпространства Ω_1 определяется по формуле:

$$\left|\Omega_1\right| = T_n + F_n = 2S_{n,2} + 1 = S_{n+1,2} = 2^n - 1 . \qquad (31)$$

Получена формула для перечисления элементов класса A_k при заданном n:

$$A_k \in \bigcup_{j=0}^{n-1} \bigcup_{i=1}^{n-j} \{a_{i+j}\} . \qquad (32)$$

Инранжирования, как и любые слабые порядки, формируются из множеств строгих порядков. Пусть $\Pi = \{\pi_1, ..., \pi_{|A_k|!}\}$ – это множество перестановок элементов класса A_k и $\Pi' = \{\pi'_1, ..., \pi'_{|\bar{A}_k|!}\}$ – множество перестановок элементов $\bar{A}_k$.

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

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

$$N(A_k) = |A_k|! \cdot |\bar{A}_k|! \cdot (n - |A_k| + 1) . \qquad (33)$$

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

- 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.
- 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.

- 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.
- 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.
- 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.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, рекомендованных ВАК

1. 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. (Q1, WoS, Scopus).
2. 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 (Scopus).
3. 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 (Scopus).
4. 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. (WoS, Scopus).
5. 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 (Q1, WoS, Scopus).

Статьи в других изданиях

6. 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.

7. 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.
8. 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.
9. 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.
10. 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.
11. 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.