

Эффективные вычислительные схемы расчета манипулируемости процедур агрегирования*

А. А. Иванов

Институт проблем управления РАН, г. Москва, Россия
Национальный исследовательский университет «Высшая школа экономики», г. Москва, Россия

Аннотация. Рассматриваются алгоритмы оценки степени манипулируемости процедур агрегирования. Рассматривается 27 известных процедур агрегирования, а также обобщённые скоринговые правила, когда приписывается различный вес второй или второй и третьей альтернатив. Рассматриваются случаи Impartial Culture и Impartial Anonymus Culture, определяющие вероятности различных профилей. Рассматриваются 6 индексов манипулируемости, включая классический индекс Нитцана-Келли. Оценивается алгоритмическая сложность отдельных процедур агрегирования и всех вычислений. Показывается, что за счёт оптимизации алгоритмов возможно вычислить индексы манипулируемости для случаев 3, 4 и 5 альтернатив при количестве участников голосования от 3 до 100. Рассматривается архитектура разработанного программного обеспечения.

Ключевые слова: процедуры агрегирования, манипулирование, индексы манипулируемости, скоринговые правила.

DOI 10.14357/20718632200204

Введение

Манипулирование – это ситуация, когда при заданной процедуре агрегирования один или группа участников голосования могут голосовать неискренне, чтобы получить более выгодный для себя результат.

В работах [1, 2] было показано, что не существует процедуры агрегирования, которая была бы одновременно и недиктаторской, и неманипулируемой. Поэтому одна из главных задач исследования манипулируемости – найти наименее манипулируемую процедуру.

В целом ряде работ рассматривалась проблема манипулируемости для различных процедур агрегирования. В одной из первых работ по сравнению различных процедур агрегирования [3] были рассмотрены четыре процедуры: правило Борда, процедура Хара, правило относительного большинства и правило Кумбса. Дальнейшие работы в этой области обычно включали большее количество процедур агрегирования. Самое большое их количество – 26 процедур - можно найти в [4], где рассматривалось индивидуальное манипулирование в основном для небольшого количества участни-

* Статья подготовлена в результате проведения исследования/работы в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ) и с использованием средств субсидии в рамках государственной поддержки ведущих университетов Российской Федерации "5-100".

ков. В [5] были рассмотрены случаи большего количества участников, но для индивидуального манипулирования и трёх альтернатив. В [6] были рассмотрены и случаи для 4 и 5 альтернатив в индивидуальном манипулировании.

Частные случаи, когда выбираются процедуры какого-то типа для какого-то количества участников или альтернатив, могут описываться аналитически. Например, в [7] можно увидеть аналитическое исследование для скоринговых процедур агрегирования для трех альтернатив. Аналитический вывод формул можно увидеть и в [8], где рассматриваются также и разные модели манипулирования. Проблема в том, что для всех процедур агрегирования (не только скоринговых, но и мажоритарных и т.д.) для случаев с разными количествами альтернатив (не только 3, но и 4 и т.д.), а также не только для индивидуального, но и для коалиционного манипулирования вывести аналитические формулы представляется невозможным, именно поэтому для анализа большего количества процедур и всех возможных случаев мы применяем компьютерное моделирование.

Степень манипулируемости процедуры агрегирования оценивается как доля ситуаций, когда манипулирование возможно, к общему количеству ситуаций (профилей). Так определяется индекс Нитцана-Келли, который был введен в [9, 10].

Вычисление степени манипулируемости заданной процедуры агрегирования – непростая алгоритмическая задача, так как невозможно сгенерировать все возможные профили из-за их огромного количества. Уже при небольшом количестве участников голосования, например, при 20 участниках, комбинаторная сложность рассмотрения всех возможных попыток манипулирования может превышать вычислительные мощности современных ЭВМ. Поэтому используется генерация определенного количества случайных профилей, чтобы для них посчитать индексы манипулируемости.

Статья построена следующим образом. Сначала мы рассмотрим основные обозначения, список процедур агрегирования, алфавитное устранение несравнимости и механизмы построения расширенных предпочтений, а также

индекс манипулируемости и разницу между индивидуальным и коалиционным манипулированием. Затем мы обсудим индексы манипулируемости и алгоритм вычисления степени манипулируемости процедур агрегирования. Наконец, будет рассмотрена архитектура программного обеспечения, оценка времени работы, а также приведены некоторые результаты.

1. Основные обозначения

Мы используем обозначения из [5]. Количество участников голосования обозначаем через n , количество альтернатив (кандидатов) – m . Каждый кандидат (альтернатива) обозначается строчной буквой латинского алфавита. У каждого участника голосования есть предпочтения на множестве альтернатив, выраженное линейным порядком, т.е. у каждого участника одно из $m!$ возможных предпочтений.

Профиль – это совокупность участников голосования и их предпочтений на множестве кандидатов, то есть множество из n линейных порядков на множестве из m альтернатив. Процедура агрегирования – это отображение множества профилей на множество всех возможных результатов.

Для представления результата процедуры агрегирования мы используем две концепции. Обе дают ответ на вопрос, что будет, если несколько альтернатив получат одинаковое количество голосов. Первая – множественный выбор. Предполагается, что если у нескольких альтернатив одинаковое количество голосов, то все они войдут в результат процедуры агрегирования. В этом подходе всего возможно $2^m - 1$ исходов, т.е. результат – это подмножество множества альтернатив за вычетом пустого множества. А как определить, какой выбор будет для участника лучше? Например, если до манипулирования результат процедуры агрегирования был $\{b\}$, а после он станет равным $\{a, c\}$, это будет вариант лучше или хуже? В литературе используются методы построения расширенных предпочтений, т.е. предпочтений участника не на множестве альтернатив, а на множестве всех возможных подмножеств множества альтернатив (всех возможных множественных выборов).

Для 3 альтернатив известно 4 способа построения расширенных предпочтений: Лексимин, Лексимакс, Рискофил, Рискофоб [5]. Например, в Лексимаксе при выборе между двумя результатами (подмножествами множества альтернатив) участник предпочтёт тот, в котором самая хорошая альтернатива будет как можно лучше дня него. То есть в Лексимаксе $\{a, c\}$ будет предпочтительнее, чем $\{b\}$. Если лучшие альтернативы в выборе одинаковые, то сравниваются вторые наилучшие и т.д. Поэтому для Лексимаксе $\{a, b, c\}$ будет лучше чем $\{a, c\}$.

Для 4 альтернатив известно 10 способов построения расширенных предпочтений, для 5 альтернатив – 12 методов. В работе мы рассматриваем подсчёт результатов для каждого метода построения расширенных предпочтений. Итак, в концепции множественного выбора процедура агрегирования – это отображение множества профилей на множество исходов, где множество исходов – это множество подмножеств множества альтернатив за вычетом пустого множества (его мощность $2^m - 1$).

Вторая концепция – это алфавитное устранение несравнимости. Если несколько альтернатив набрали одинаковый результат, то побеждает та, которая меньше лексикографически. Например, если альтернативы a и b набрали по 5 голосов в процедуре относительного большинства, а другие альтернативы – строго меньше, то исходом будет $\{a\}$. В случае алфавитного устранения несравнимости процедура агрегирования – это отображение множества профилей на множество исходов, где множество исходов – это множество альтернатив (его мощность m).

Манипулирование происходит, если существует участник (или группа участников), который записывают в бюллетень неискренние предпочтения, после чего выбор оказывается для него (них) лучше, чем если бы он вписал в бюллетень свои искренние предпочтения.

Задача исследования – найти, какие процедуры агрегирования являются наименее манипулируемыми для различных ситуаций.

2. Процедуры агрегирования

В литературе известно немало процедур агрегирования, и разные работы исследуют различные процедуры по отдельности или в группе из

нескольких. Нашей задачей было найти наименее манипулируемые процедуры агрегирования из наибольшего возможного количества известных процедур, поэтому мы взяли за основу список процедур агрегирования из [4, 5], так как именно этих работах приводится один из наиболее полных списков известных процедур агрегирования.

Таким образом, первая группа процедур – это процедуры агрегирования в том виде, в котором они изложены в [4, 5]. Список процедур агрегирования, а также примерные алгоритмические оценки времени расчета результатов процедур агрегирования для одного профиля указаны в Табл. 1.

При реализации каждой процедуры использовались их описания из [4]. При реализации использовались оптимизации, связанные с исключением расчёта одних и тех же вспомогательных данных несколько раз. Например, если речь идёт о мажоритарных правилах, для которых нужно высчитывать мажоритарную матрицу, то она высчитывалась для заданного профиля один раз, затем использовалась для всех мажоритарных процедур. За счёт этого достигается небольшая экономия во времени расчётов.

Вторая группа процедур агрегирования – это обобщённые скоринговые процедуры агрегирования. Это более широкое рассмотрение группы скоринговых правил, т.е. правил, где каждая альтернатива получает определенное количество баллов за нахождение на определенном месте в бюллетене. Примерами скоринговых правил являются правило Борда, правило относительного большинства и т.д. Например, правило относительного большинства – это получение 1 балла за нахождение на первом месте и получение 0 баллов за нахождение на любых других местах в бюллетене. Таким образом, мы можем рассмотреть не только несколько скоринговых процедур (правила Борда, относительного большинства), а рассмотреть намного большее количество различных процедур с разными весами альтернатив. Такой подход также уже встречался в литературе, например, в [11].

Мы можем обобщить такие правила. Для случая 3 альтернатив:

1. Первая альтернатива в бюллетене/предпочтениях получает 1 балл.
2. Вторая альтернатива получает α баллов.

Табл. 1. Процедуры агрегирования и асимптотическое время их расчёта

Процедура агрегирования	Асимптотическое время расчета
Правило относительного большинства	$O(n+m)$
Одобряющее голосование с $q=2$	$O(n+m)$
Правило Борда	$O(n^*m)$
Процедура Блэка	$O(n^*m^*m)$
Обратное правило Борда	$O(n^*m^*m)$
Пороговое правило	$O(m^*(n+m))$
Процедура Хара	$O(n^*m^*m)$
Обратное правило относительного большинства	$O(n+m)$
Процедура Нансона	$O(n^*m^*m)$
Процедура Кумбса	$O(n^*m^*m)$
Минимальное доминирующее множество	$O(m^*m^*(n+2^m))$
Минимальное недоминируемое множество	$O(m^*m^*(n+2^m))$
Непокрытое множество I	$O(n^*m^*m)$
Непокрытое множество II	$O(n^*m^*m)$
Правило Ричельсона	$O(n^*m^*m)$
Минимальное слабоустойчивое множество	$O(m^*m^*(n+2^m))$
Правило Фишберна	$O(n^*m^*m)$
Правило Коупланда I	$O(n^*m^*m)$
Правило Коупланда II	$O(n^*m^*m)$
Правило Коупланда III	$O(n^*m^*m)$
Процедура Симпсона	$O(n^*m^*m)$
Процедура минмакс	$O(n^*m^*m)$
Сильное q -Паретовское правило простого большинства	$O(m^*(n+m))$
Сильное q -Паретовское правило относительного большинства	$O(m^*(n+m))$
1-устойчивое множество	$O(m^*m^*(n+m+2^m))$
2-устойчивое множество	$O(m^*m^*(n+m+2^m))$
3-устойчивое множество	$O(m^*m^*(n+m+2^m))$

3. Третья альтернатива получает 0 баллов.

При этом $0 \leq \alpha \leq 1$

Для случая 4 альтернатив:

1. Первая альтернатива в бюллетене/предпочтениях получает 1 балл.
2. Вторая альтернатива получает α_1 баллов.
3. Третья альтернатива получает α_2 баллов.
4. Четвертая альтернатива получает 0 баллов.

При этом $0 \leq \alpha_2 \leq \alpha_1 \leq 1$

В рамках модели мы перебирали веса с шагом в 0.01. Для случая 3 альтернатив это перебор 101 разного значения веса второй альтернативы (α), так как он может принимать значения от 0 до 1 включительно. Для случая 4 альтернатив это перебор $\frac{101*(101+1)}{2}$ различных комбинаций весов второй и третьей альтернатив (α_1 и α_2).

Помимо рассмотрения различных весов мы также рассмотрели 4 вариации процедур агрегирования, которые применяются для каждой комбинации весов.

1. Простое скоринговое правило.

Суммируются набранные баллы для каждой альтернативы. Результат процедуры агрегирования – альтернативы, которые набрали наибольшее количество баллов. Частные случаи для случая трёх альтернатив: если вес второй альтернативы равен 0.5, это правило Борда. Если вес второй альтернативы равен 0, то это правило относительного большинства. Если вес второй альтернативы равен 1, то это одобряющее голосование с $q=1$.

2. Скоринговое правило с выбыванием, вариант 1.

Суммируются баллы для каждой альтернативы, также считается среднее количество баллов. Затем из процедуры выбывают альтернативы, у которых количество баллов строго меньше среднего значения. Процедура повторяется, пока не определён победитель.

3. Скоринговое правило с выбыванием, вариант 2.

Суммируются баллы для каждой альтернативы, также считается среднее количество баллов. Затем из процедуры выбывают альтернативы, у которых количество баллов меньше или равен среднему значению. Процедура повторяется, пока не определён победитель.

4. Скоринговое правило с выбыванием, вариант 3.

Суммируются баллы для каждой альтернативы. Затем из процедуры выбывает альтернатива, у которой наименьшее количество баллов. Процедура повторяется, пока не определён победитель.

Учитывая, что всего мы рассматривали разные веса второй альтернативы для случая трёх альтернатив и второй и третьей альтернатив для случая четырёх альтернатив, а также четыре разные вариации процедур, то для случая трёх альтернатив речь идёт о $101 \cdot 4 = 404$ различных процедурах, для случая четырёх альтернатив о $\frac{101 \cdot (101+1)}{2} \cdot 4 = 20604$ различных процедурах агрегирования.

3. Impartial Culture (IC) и Impartial Anonymous Culture (IAC)

При оценке степени манипулируемости процедур агрегирования возникает вопрос: каковы вероятности появления тех или иных профилей голосования, т.е. наборов множеств участников голосования, у каждого из которых есть предпочтения на множестве альтернатив?

Две наиболее популярные в литературе концепции – это Impartial Culture (IC) и Impartial Anonymous Culture (IAC).

В рамках IC предполагается, что предпочтения участников независимы и равновероятны. С точки зрения моделирования это эквивалентно генерации одного из $m!$ различных линейных порядков в качестве предпочтений для каждого из n участников. Всего получается $m!^n$ различных возможных профилей голосования.

В рамках модели IAC предполагается, что профили голосования равновероятны с учетом анонимности, т.е. все возможные профили делятся на классы эквивалентности, где два профиля попадают в один класс, если они могут быть получены переименованием избирателей. И в IAC предполагается, что каждый класс эк-

вивалентности равновероятен. Общее количество различных профилей в IAC равно C_{m+n-1}^n .

При оценке степени манипулируемости процедур агрегирования мы рассматриваем для каждого правила оба случая: и IC, и IAC. Из результатов видно, что степени манипулируемости для правил в IC и IAC различаются, например, индексы манипулируемости могут отличаться в два раза, хотя есть и много сходств по тому, какие процедуры оказываются наименее манипулируемыми для разных ситуаций.

Общее количество профилей для случаев IC и IAC очень велико ($m!^n$ и C_{m+n-1}^n , соответственно), и за разумное время сгенерировать и определить манипулируемость всех возможных профилей невозможно.

Поэтому мы генерируем 1 миллион профилей для каждой возможной ситуации, после чего рассчитываем коэффициенты манипулируемости для данных профилей. В [12] показано, что при оценке индекса манипулируемости при случайной генерации 1 миллиона профилей погрешность при вычислении индекса Нитцана-Келли не превышает 0.001. Такая погрешность даёт возможность оценить, какие процедуры агрегирования являются наименее манипулируемыми для различных ситуаций.

4. Индивидуальное и коалиционное манипулирование

Какое количество участников голосования может пытаться манипулировать? Если только один участник предьявляет неискренние предпочтения, то это будет индивидуальное манипулирование. Если группа участников объединяется, то это коалиционное манипулирование.

Важно подчеркнуть, что профиль является манипулируемым, если существует хотя бы один способ для какого-то участника (в случае индивидуального манипулирования) или у группы участников (случай коалиционного манипулирования), при котором предьявление неискренних предпочтений повлечет за собой более выгодный результат голосования. Отсюда следует, что если мы проверяем профиль на манипулируемость, то мы должны рассмотреть все возможные попытки манипулирования. И если мы находим хотя бы одну успешную воз-

можную попытку, мы можем сразу пометить профиль как манипулируемый (для заданной процедуры агрегирования и рассматриваемого случая). Но если рассмотренная попытка не была успешной, мы должны продолжать рассматривать все возможные попытки.

В случае с индивидуальным манипулированием надо рассмотреть всех участников голосования и все возможные способы предъявления неискренних предпочтений для каждого участника голосования. Для случая m альтернатив получится $m!-1$ возможных попыток манипулирования для каждого участника. Если всего участников n , то всего получится $n*(m!-1)$ попыток для всех участников, которые мы должны перебрать.

Здесь возможно провести оптимизацию перебора, которая будет особенно эффективна для случаев небольшого количества альтернатив и большого количества участников. Например, для случая $n=100$ и $m=3$, мы будем перебирать $100*(3!-1)=500$ попыток манипулирования. Но если вспомнить предпосылку, что для нас все участники голосования эквивалентны, то очевидно, что нет смысла перебирать попытки манипулирования для двух участников голосования, если у них абсолютно одинаковые предпочтения (одинаковые линейные порядки на множестве альтернатив). Например, если для какого-то участника $a>b>c$ и мы сгенерировали все возможные попытки манипулирования, то мы можем не перебирать все возможные попытки манипулирования для других участников, у которых такие же предпочтения. Это означает, что максимальное количество участников, для которых мы будем перебирать все возможные попытки манипулирования ограничено числом $m!$, то есть сложность перебора будет равна $(m!-1)*\min(n, m!)$. И для случая $n=100$ и $m=3$ это будет равно не 500 попыткам (как в полном переборе), а только 30.

В случае коалиционного манипулирования ключевым аспектом являются принципы, по которым участники могут объединяться в коалиции. Здесь должны быть точные определения, участники с какими типами предпочтений могут объединяться, есть ли ограничения на размер возможных коалиций: могут ли члены коалиции предъявлять разные неискренние предпочтения или только одинаковые и т.д.

Если никаких принципов нет, то есть любой участник может объединяться с любыми в коалиции, то это будет самый тяжелый с вычислительной точки зрения случай, так как перебор по количеству коалиций может достигать 2^n-1 , т.е. все возможные подмножества участников, исключая пустую коалицию. Однако в реальной ситуации объединение участников, вероятно, должно происходить вокруг чего-то общего: общих искренних предпочтений и/или общих желаний «продвинуть» какую-то альтернативу с помощью неискренних предпочтений.

Мы будем рассматривать два принципа формирования коалиций. Первый принцип – это когда в коалиции могут объединяться те и только те участники, у которых одинаковые искренние предпочтения. Кроме того, при голосовании все они могут предъявить только одинаковые неискренние предпочтения. Дополнительно могут вводиться ограничения на максимальный размер коалиции. Второй принцип – это объединение в коалиции вокруг продвижения какого-то единого кандидата. Здесь будут объединяться те участники, которые хотят продвинуть единую альтернативу, при этом их предпочтения могут быть разными.

В первом случае оптимизация вычислений может быть достигнута за счёт того, что нам не нужно рассматривать все подмножества участников, достаточно рассматривать отдельно группы участников с одинаковыми предпочтениями, а уже внутри этой группы смотреть возможные коалиции. Например, если участников n , а участников с предпочтениями $a>b>c$ в профиле ровно n_1 , то мы можем рассматривать не 2^n-1 коалиций, а $2^{n_1}-1$. Затем, если у n_2 участников предпочтения $a>c>b$, то для них будет $2^{n_2}-1$ возможных коалиций.

Однако и здесь будут получаться весьма большие размеры вычислений, поэтому можно использовать следующее наблюдение. Для нас неважно, какие именно участники будут объединяться в коалиции среди группы с одинаковыми предпочтениями, для нас важно их количество. Например, если у пяти участников предпочтения $a>b>c$, то случаи, когда первый и второй объединятся будут эквивалентны случаю, когда третий и четвертый объединятся. Тогда размер перебора внутри группы участни-

ков с одинаковыми предпочтениями сводится не к перебору по подмножествам, а к перебору размера коалиции. Если n_1 участников имеют предпочтения $a > b > c$, то мы будем перебирать n_1 возможных коалиций, в каждую из которых будет входить от 1 до n_1 участников. Тогда получится, что общий размер перебора сводится к $n_1 + n_2 + \dots + n_6 = n$. Для случая 4-х и большего количества альтернатив также получится перебор размера n , так как будет перебираться каждая из $m!$ групп участников с одинаковыми предпочтениями от 1 до мощности этого множества, а суммарная мощность всех множеств будет равна количеству участников голосования. Таким образом от экспоненциальной сложности перебора коалиций мы пришли к линейной сложности.

Второй способ формирования коалиций – это предположение, что участники коалиций объединяются, чтобы вместе продвинуть некоего кандидата. Это более гибкая предпосылка по сравнению с первым методом формирования коалиций: здесь объединяться могут участники с разными предпочтениями, а также записывать в бюллетени они могут разные искренние или неискренние предпочтения. Из-за высокой комбинаторной сложности для данного типа формирования коалиций мы рассматриваем только случай трёх альтернатив.

В данном случае вновь мы могли бы иметь экспоненциальную сложность перебора, однако она превращается в квадратичную для случая трёх альтернатив. По определению этого типа коалиционного манипулирования, из $m! = 3! = 6$ групп участников только две могут объединяться, чтобы продвинуть единую альтернативу. Тогда будет два цикла перебора: от 0 до n_1 (сколько участников с предпочтениями первого типа входят в коалицию) и от 0 до n_2 (сколько участников с предпочтениями второго типа входят в коалицию). В итоге – квадратичная сложность вместо экспоненциальной.

5. Коэффициенты манипулируемости

В исследованиях по манипулируемости используется целый ряд различных коэффициентов. Перечислим те из них, которые рассматриваются ниже.

Один из самых популярных индексов манипулируемости – это индекс Нитцана-Келли

(NK), который был впервые введён в [9 и 10]. Индекс равен доле манипулируемых профилей в общем количестве профилей. Формула выглядит следующим образом:

$$NK = \frac{d_0}{d_{\text{общее}}},$$

где d_0 – количество манипулируемых профилей, т.е. профилей, где существует хотя бы один успешный способ манипулирования, а $d_{\text{общее}}$ – общее количество профилей.

В случае с генерацией 1 миллиона профилей в знаменателе будет находиться 1 миллион, а в числителе – количество профилей из сгенерированного 1 миллиона, для которых существует хотя бы одна успешная попытка манипулирования.

Помимо индекса NK в исследованиях встречались и другие индексы. Например, в [4] был исследован индекс свободы манипулирования. Рассмотрим все возможные попытки манипулирования, которые может предпринять участник голосования. Их количество равно $m! - 1$. При каждой попытке манипулирования может получиться одна из трёх ситуаций:

1. После искажения предпочтений выбор стал лучше для манипулирующего участника.
2. После искажения предпочтений выбор остался таким же для манипулирующего участника.
3. После искажения предпочтений выбор стал хуже для манипулирующего участника.

Тогда были предложены три коэффициента свободы манипулирования: I_1^+ , I_1^0 и I_1^- , где каждый будет равен доле случаев, которые приведут к улучшению, неизменности или ухудшению выбора. Иначе говоря, если задан профиль, то для каждой попытки манипулирования смотрим, каким будет выбор. Если лучше, чем начальный, то I_1^+ увеличивается на 1. Если эквивалентный начальному, то I_1^0 увеличивается на 1. Если хуже начального, то I_1^- увеличивается на 1. Индексы нормируются на количество попыток манипулирования, т.е. каждый делится на $m! - 1$. В итоге получится, что $I_1^+ + I_1^0 + I_1^- = 1$.

Помимо свободы манипулирования в [4] был введен индекс эффективности манипулирования. Рассмотрим все попытки предъявить неискренние предпочтения ($m! - 1$), и из них возьмём те, которые приводят к улучшению результата процедуры агрегирования для мани-

пулирующего участника. В каждой такой попытке мы можем рассчитать, на какое количество мест в его расширенных предпочтениях улучшится итоговый выбор. Возьмём среднее значение по каждой успешной возможной попытке манипулирования. Усредним эти значения по всем участникам в профиле, затем усредним по всем профилям. В итоге получится индекс I_2 , являющийся индексом эффективности манипулирования.

Есть вторая вариация индекса эффективности манипулирования – индекс I_3 . В этом случае мы будем не усреднять все возможные выигрыши, измеряемые в количестве мест продвижения в предпочтениях участника, а брать максимальное. Иначе говоря, схема выглядит следующим образом:

1. берём профиль;
2. для каждого участника рассматриваем все попытки манипулирования;
3. берем максимальное улучшение в терминах количества мест, на которое может улучшиться выбор в расширенных предпочтениях участника при манипулировании;
4. усредняем по всем участникам внутри профиля;
5. усредняем по всем профилям.

Итак, в данной работе мы рассчитываем индексы NK , I_1^+ , I_1^0 , I_1^- , I_2 и I_3 .

6. Алгоритмическая схема оценки степени манипулируемости

Алгоритм расчета индексов манипулируемости состоит из нескольких этапов. На первом этапе задаются параметры модели.

1. IAC или IC.
2. Количество альтернатив, от 3 до 5.
3. Количество участников голосования (от 3 до 100).
4. Индивидуальное или коалиционное манипулирование (если коалиционное, то какого из двух описанных выше типов).

Эти параметры подаются на вход программе на следующий этап. Для оценки процедуры агрегирования на предмет манипулируемости требуется рассмотреть случаи от 3 до 100 участников голосования, для случаев IAC и IC.

Дальнейшая стадия – это генерация профилей. Для каждого профиля рассматриваются следующие ситуации и наборы параметров.

1. Алфавитное устранение несравнимости и каждый возможный тип построения расширенных предпочтений (3 способа для 3 альтернатив, 10 способов для 4 альтернатив и 12 способов для 5 альтернатив) для случая множественного выбора.

2. Каждая процедура агрегирования из списка рассматриваемых процедур агрегирования.

Для каждого возможного набора параметров проверяются все возможные попытки манипулирования.

Если речь идёт об индивидуальном манипулировании, то рассматриваются все возможные способы предъявления неискренних предпочтений одним участником. Это перебор по n участникам (но как было указано выше, не более чем по $m!$ участникам), у каждого из которых рассматриваются $m!-1$ попыток предъявить неискренние предпочтения.

Для случая коалиционного манипулирования рассматриваются два различных варианта в зависимости от типа коалиционного манипулирования. В первом типе делаем перебор величины n , как описывалось в пункте про коалиционное манипулирование.

Для случая коалиций, которые продвигают одного кандидата, мы ограничиваемся расчетами по 3 альтернативам, так как здесь сложность перебора вырастает существенно: нам нужно рассмотреть все возможные комбинации участников с разными предпочтениями. Однако здесь мы используем прежнюю оптимизацию, основанную на том, что для нас участники равноценны. Так, если n_1 участников имеют предпочтения $a > b > c$, а n_2 участников имеют предпочтения $a > c > b$, и требуется рассмотреть все возможные коалиции из k участников, которые состояли бы только из участников с такими предпочтениями, то нам не нужно делать полный перебор размера $C_{n_1+n_2}^k$, нам достаточно провести перебор размера $k + 1$, так как мы будем перебирать, сколько (от 0 до k включительно) из участников будут из первой группы (остальные из второй). Тот факт, что мы избегаем количества сочетаний и сложной комбинаторной сложности в рассмотрении коалиций, позволяет

сделать оценку степени коалиционной манипулируемости за разумное время.

Рассматривая все возможные попытки манипулирования, мы делаем расчет коэффициентов манипулируемости. При этом в силу особенности коэффициентов свободы и эффективности манипулирования надо для каждого профиля рассмотреть все возможные попытки манипулирования (а не ограничиться нахождением хотя бы одной успешной), чтобы рассчитать точные коэффициенты. После того как рассмотрены все попытки по всем профилям, рассчитываются значения индексов манипулируемости для сгенерированных профилей и получаем итоговые значения.

Очевидно, что вычислительная сложность схемы манипулирования огромна, так как сначала генерируется около тысячи комбинаций параметров (разные количества участников, альтернатив, рассматриваются различные методы построения расширенных предпочтений), затем генерируется миллион профилей, после чего следует генерация всех возможных попыток манипулирования.

Примерная сложность описанной схемы составляет:

$$\begin{aligned} & \sum_{m=3,4,5} \sum_{n=3}^{100} \sum_{k=1}^n 4 * 27 * 1,000,000 * C_n^k * k * m \\ & \quad * (m! - 1) * n \\ = & \sum_{m=3,4,5} \sum_{n=3}^{100} \sum_{k=1}^n 1.08 * 10^8 * C_n^k * k * n * m \\ & \quad * (m! - 1) \end{aligned}$$

операций, что для случая с 5 альтернативами ($m=5$) и 100 участниками голосования ($n=100$) делает невозможным полные расчеты на современных суперкомпьютерах даже за несколько сотен лет.

Именно поэтому важны описанные выше методики оптимизации, особенно для случая коалиционной манипулируемости. Благодаря эквивалентности участников с одинаковыми предпочтениями мы не рассматриваем все возможные коалиции (чисто сочетаний), а рассматриваем по одному случаю из множества эквивалентных ситуаций.

После применения этих и нескольких других алгоритмических оптимизаций мы добились следующего времени работы алгоритма (в операциях).

Итоговая сложность вычислений

$$\begin{aligned} & = \sum_{m=3,4,5} \sum_{n=3}^{100} \sum_{k=1}^n 4 * 1,000,000 * opt * n * m! \\ & \quad * \sum_{i=1}^{27} O_i(C(\vec{P})) \\ = & \sum_{m=3,4,5} \sum_{n=3}^{100} \sum_{k=1}^n 4 * 10^6 * opt * n * m! \\ & \quad * \sum_{i=1}^{27} O_i(C(\vec{P})) \end{aligned}$$

Такая сложность позволяет получить оценку манипулируемости 27 правил коллективного выбора для случая 5 альтернатив и 100 участников голосования всего лишь за несколько недель.

7. Архитектура программного обеспечения

В рамках проекта по оценке степени манипулируемости процедур агрегирования было разработано программное обеспечение, которое позволяет рассчитывать различные индексы манипулируемости для различных входных параметров.

Ключевые параметры системы:

1. Гибкость к добавлению новых процедур агрегирования. Были получены результаты по наименее манипулируемым процедурам, но если потребуется оценить манипулируемость какой-то новой процедуры, текущее ПО позволит это быстро сделать

2. Гибкость в настройке параметров. Помимо различных процедур агрегирования есть и другие параметры, которые могут меняться. Например, могут добавляться новые виды построения расширенных предпочтений. Изменение параметров запуска программы становится очень простым и позволяет в короткие сроки проводить новые исследования.

3. Возможность задавать параметры в конфигурационных файлах и/или в интерфейсе программы, чтобы непрограммист мог осуществлять запуск расчетов при требуемых ему параметрах.

4. Возможность запуска расчетов на нескольких ЭВМ. Запуски программы проводились на нескольких компьютерах (в среднем на пяти) одновременно для ускорения расчетов. В программе должны быть предусмотрены воз-

возможности параллельных запусков и сбора расчетов из нескольких пакетов вычислений в единый файл.

5. Устойчивость программного обеспечения к сбою операционной системы, отключению электричества и т.д. Для большого количества участников голосования (например, для $n=100$, $m=5$) задача вычисления индексов манипулируемости может занимать несколько недель, даже на нескольких компьютерах. В случае внезапного сбоя на компьютере должна быть возможность перезапустить расчет с минимальными потерями.

В качестве языка программирования был выбран C# - высокоуровневый язык программирования с возможностью создавать удобные интерфейсы и разрабатывать сложные методы для расчетов индексов манипулируемости.

8. Время работы алгоритмов

В Табл. 2 представлено примерное время работы программного обеспечения для расчета индексов манипулируемости для различных значений параметров.

Если мы возьмём один из самых времязатраных случаев, то есть случай $n = 100$, $m = 4$ и разобьём вычисления на несколько параллельно работающих ЭВМ, то процесс займёт всего лишь несколько дней, когда для случая полного перебора потребовалось бы несоизмеримо большее время (для случая $n = 100$, $m = 4$ полный перебор потребовал бы примерно

Табл. 2. Время работы алгоритмов для различных случаев

Количество альтернатив (m)	Количество участников (n)	Время, требуемое для расчетов индексов манипулируемости на одной ЭВМ (ч)
3	10	2
3	20	5
3	50	19
3	80	42
3	100	63
4	10	13
4	20	33
4	50	138
4	80	311
4	100	467

$6.3 \cdot 10^{43}$ операций для оценки степени манипулируемости 27 процедур агрегирования). В итоге все вычисления были проделаны суммарно за несколько месяцев работы нескольких ЭВМ в ИПУ РАН.

9. Проверка результатов

Для проверки корректности работы алгоритмов мы сравнили полученные результаты с известными в литературе. Чаще всего, это какие-то частные случаи, которые были подсчитаны в предыдущих работах на тему манипулируемости.

Для индивидуального манипулирования проверка результатов была произведена с работами [4, 5, 6]. Для коалиционного манипулирования проверка результатов проводилась по частным случаям, так как индивидуальное манипулирование – это случай, когда коалиция состоит ровно из 1 участника.

Для случаев расчета обобщенных скоринговых правил проверялось, что частные случаи совпадают с другими схемами расчета индивидуального манипулирования. Так, если в случае обобщенного скорингового правила для трех альтернатив вес второй альтернативы равен 0.5, то это эквивалентно правилу Борда. Вес второй альтернативы равный 0 эквивалентен правилу относительного большинства. Вес второй альтернативы равный 1 эквивалентен одобряющему голосованию с $q=2$.

Проведенные проверки показали, что полученные результаты с точностью до погрешности сошлись с другими известными в литературе и приведенными в разделе «Литература».

10. Примеры результатов

Рассчитанные индексы степени манипулируемости процедур агрегирования были опубликованы в нескольких работах. Для каждого набора входных параметров (количество альтернатив, метод построения расширенных предпочтений, тип манипулирования) мы можем построить график различных индексов манипулируемости, рассмотренных выше. Так как всего комбинаций параметров очень много, то все возможные графики привести в одной статье невозможно. Поэтому мы ограничимся

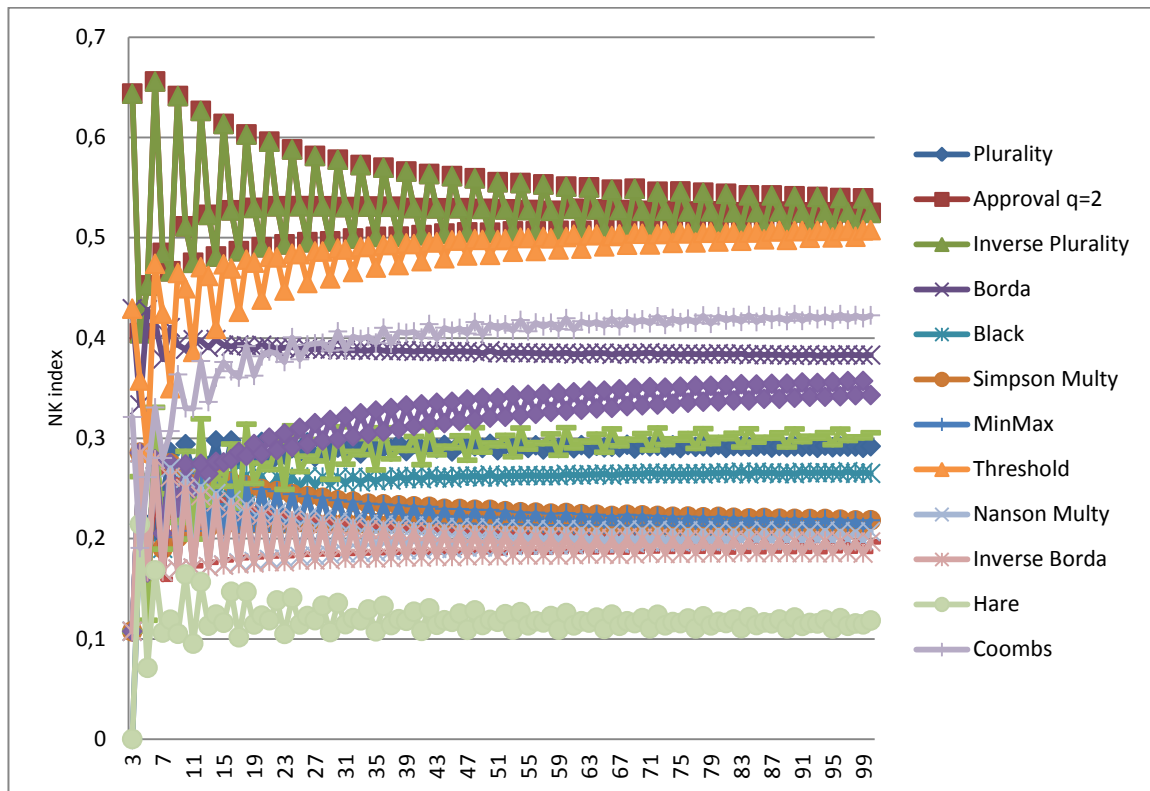


Рис. 1. Индекс NK для случая: Лексимакс, 3 альтернативы, IAC

примером графика для случая индивидуального манипулирования для 3 альтернатив, метод построения расширенных предпочтений Лексимакс, IAC. Приводятся значения индекса Нитцана-Келли (Рис.1).

На графике мы видим, что процедура Хара является наименее манипулируемой для этого случая. Индекс NK для неё находится в пределах 0.1-0.15, что означает, что от 10 до 15% профилей являются манипулируемыми. Одобряющее голосование, когда выбираются ровно две кандидатуры от каждого участника, наоборот, является одним из самых манипулируемых правил.

Из результатов для разных случаев мы можем увидеть, что не существует единой процедуры агрегирования, которая была бы наименее манипулируемой для всех случаев, т.е. комбинаций параметров. Для различных случаев разные процедуры являются наименее манипулируемыми. Однако отметим, что чаще других наименее манипулируемыми оказываются процедура Хара (Рис. 1) и правило Нансона.

Заключение

Мы рассмотрели подробные схемы расчета нескольких коэффициентов манипулируемости для различных ситуаций. При экспоненциальном общем количестве всех возможных профилей голосования полученная модель позволяет подсчитывать коэффициенты манипулируемости за разумное время с определенной точностью.

Было рассмотрено большое количество различных ситуаций: случаи Impartial Culture и Impartial Anonymous Culture, различные способы построения расширенных предпочтений, различные количества альтернатив (3, 4 и 5 альтернатив), случаи от 3 до 100 участников голосования. Для каждой комбинации параметров были произведены расчеты, показывающие, какая процедура агрегирования будет наименее манипулируемой для данного случая. В качестве основного индекса манипулируемости был использован индекс Нитцана-Келли, также рассмотрен расчет индексов свободы и эффективности манипулирования. Разработан-

ные алгоритмы и программное обеспечение позволяют получать результаты для большего количества различных комбинаций параметров, чем встречалось в отдельных предыдущих исследованиях. Чаще других наименее манипулируемыми были процедуры Хара и правило Нансона, однако то какая процедура будет являться наименее манипулируемой, зависит от параметров (количество участников голосования, количество альтернатив, метод построения расширенных предпочтений и т.д.).

Описанная схема является гибкой к добавлению новых процедур агрегирования. Несмотря на проведенные расчеты для 27 известных процедур агрегирования и обобщённых скоринговых правил, даже если потребуется оценить степень манипулируемости какой-то новой процедуры агрегирования, это можно будет сделать с помощью описанной схемы. Схема реализована в виде программного обеспечения на языке C#.

Автор благодарит Ф. Т. Алескерова за постановку задачи и консультации по разработке и осуществлению плана исследований, а также Д. С. Карабекяна и В. И. Якубу за консультации во время проведения исследования.

Литература

1. Gibbard A. Manipulation of voting schemes // *Econometrica*. 1973. Vol. 41. P. 587–601.
2. Satterthwaite M. Strategy-proofness and Arrow's conditions: existence and correspondence theorems for

- voting procedures and social welfare functions // *Journal of Economic Theory*. 1975. Vol. 10. P. 187–217.
3. Chamberlin J. R. An investigation into the relative manipulability of four voting systems // *Behavioral Science*. 1985. Vol.30(4). P.195–203.
4. Aleskerov F., Kurbanov E. Degree of manipulability of social choice procedures // Alkan A. et al. (eds.) *Current Trends in Economics. Studies in Economic Theory*. Berlin Heidelberg, N.Y.: Springer, 1999. Vol.8. P.13-27.
5. Aleskerov F., Karabekyan D., Sanver R., Yakuba V. On manipulability of positional voting rules // *SERIEs: Journal of the Spanish Economic Association*. 2011. Vol. 2 (4). P. 431-446.
6. Aleskerov F., Karabekyan D., Sanver R., Yakuba V. On the manipulability of voting rules: Case of 4 and 5 Alternatives // *Mathematical Social Sciences*. 2012. Vol. 64 (1). P. 67–73.
7. Pritchard G., Wilson M. Exact results on manipulability of positional voting rules// *Social Choice and Welfare*. 2007. Vol. 29. P. 487–513.
8. Favardin P., Lepelley D. Some further results on the manipulability of social choice rules // *Social Choice and Welfare*. 2006. Vol. 26. P. 485–509.
9. Nitzan S. The vulnerability of point-voting schemes to preference variation and strategic manipulation // *Public Choice*. 1985. Vol. 47. P. 349–370.
10. Kelly J. Almost all social choice rules are highly manipulable, but few aren't // *Social Choice and Welfare*. 1993. Vol. 10. P. 161–175.
11. Xia, L., Conitzer, V. Generalized scoring rules and the frequency of coalitional manipulability // *Proceedings of the 9th ACM conference on Electronic commerce*. 2008, P. 109-118.
12. Карабекян Д.С. «Манипулирование в задаче коллективного принятия решений», диссертация на соискание ученой степени кандидата экономических наук, Москва, 2012.

Иванов Александр Александрович. аспирант ИПУ РАН, стажёр-исследователь МЦАВР НИУ ВШЭ. Количество печатных работ: 12. Область научных интересов: манипулируемость процедур агрегирования, свойства правил коллективного выбора, программирование. E-mail: ivanovalex@gmail.com

On Efficient Schemes of Estimating the Degree of Manipulability of Aggregation Procedures

A. A. Ivanov

Institute of Control Sciences RAS, Moscow, Russia
National Research University Higher School of Economics, Moscow, Russia

Abstract. We study the algorithms for evaluation of manipulability of aggregation procedures. We consider 27 known aggregation procedures and generalized scoring rules, when the weights of second or second and third alternatives are different. Both Impartial Culture and Impartial Anonymous Culture are studied. We calculate 6 manipulability indices including well-known Nitzan-Kelly index. We provide asymptotic estimations of the computational complexity. It is shown that optimization allows

to calculate manipulability indices for aggregation procedures for the number of agents between 3 and 100 and for 3, 4 and 5 alternatives. The algorithms and their structures are discussed.

Keywords: Manipulation, manipulability indices, aggregation procedures, scoring social choice rules.

DOI 10.14357/20718632200204

References

1. Gibbard A. Manipulation of voting schemes // *Econometrica*. 1973. Vol. 41. P. 587–601.
2. Satterthwaite M. Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions // *Journal of Economic Theory*. 1975. Vol. 10. P. 187–217.
3. Chamberlin J. R. An investigation into the relative manipulability of four voting systems // *Behavioral Science*. 1985. Vol.30(4). P.195–203.
4. Aleskerov F., Kurbanov E. Degree of manipulability of social choice procedures // Alkan A. et al. (eds.) *Current Trends in Economics. Studies in Economic Theory*. Berlin Heidelberg, N.Y.: Springer, 1999. Vol.8. P.13-27.
5. Aleskerov F., Karabekyan D., Sanver R., Yakuba V. On manipulability of positional voting rules // *SERIEs: Journal of the Spanish Economic Association*. 2011. Vol. 2 (4). P. 431-446.
6. Aleskerov F., Karabekyan D., Sanver R., Yakuba V. On the manipulability of voting rules: Case of 4 and 5 Alternatives // *Mathematical Social Sciences*. 2012. Vol. 64 (1). P. 67–73.
7. Pritchard G., Wilson M. Exact results on manipulability of positional voting rules// *Social Choice and Welfare*. 2007. Vol. 29. P. 487–513.
8. Favardin P., Lepelley D. Some further results on the manipulability of social choice rules // *Social Choice and Welfare*. 2006. Vol. 26. P. 485–509.
9. Nitzan S. The vulnerability of point-voting schemes to preference variation and strategic manipulation // *Public Choice*. 1985. Vol. 47. P. 349–370.
10. Kelly J. Almost all social choice rules are highly manipulable, but few aren't // *Social Choice and Welfare*. 1993. Vol. 10. P. 161–175.
11. Xia, L., Conitzer, V. Generalized scoring rules and the frequency of coalitional manipulability // *Proceedings of the 9th ACM conference on Electronic commerce*. 2008, P. 109-118.
12. Karabekyan, D.S., «Manipulation in collective decision making», Candidate of Science Thesis, Moscow, 2012/

Ivanov A. A. PhD candidate ICS RAS, research assistant NRU HSE, Moscow, Russia. E-mail: ivanovalexalex@gmail.com