ИЗВЕСТИЯ РАН. ТЕОРИЯ И СИСТЕМЫ УПРАВЛЕНИЯ, 2022, № 6, с. 123–132

СИСТЕМНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

УДК 519.81

СОГЛАСОВАНИЕ ИНДИВИДУАЛЬНЫХ РАНЖИРОВОК МЕТОДОМ ВЕТВЕЙ И ГРАНИЦ

© 2022 г. В. Н. Нефедов[a],*, В. А. Осипова[a],**

[a] МАИ (национальный исследовательский ун-т), Москва, Россия

*e-mail: nefedovvn54@yandex.ru

**e-mail: victoria.a.osipova@gmail.com

Поступила в редакцию 02.03.2022 г.
После доработки 03.04.2022 г.
Принята к публикации 30.05.2022 г.

В задаче коллективного выбора предлагается алгоритм построения оптимальной агрегированной ранжировки, наиболее близкой ко всем индивидуальным строгим ранжировкам в смысле медианы Кемени. Применение метода ветвей и границ позволяет решать задачи большой размерности, а также может использоваться в случае, когда индивидуальные предпочтения являются произвольными бинарными отношениями.

DOI: 10.31857/S0002338822060142

Введение. Рассматривается реляционная модель коллективного выбора, состоящая в отыскании результирующей ранжировки, в максимальной степени учитывающей индивидуальные предпочтения, которые могут быть описаны индивидуальными ранжировками.

Различные подходы к агрегированию индивидуальных предпочтений в рамках реляционной модели группового выбора описаны в многочисленных работах, начиная с основанных на классических идеях Ш. Борда и Н. Кондорсе, затем на концепции рационального коллективного выбора, реализованной в аксиоматическом подходе К. Эрроу, а также методах, использующих понятие близости бинарных отношений [1–5].

В работе для нахождения оптимальной агрегированной ранжировки предлагается применить метод ветвей и границ для построения медианы Кемени, являющейся наиболее близкой ко всем индивидуальным ранжировкам и тем самым отражающей консенсус между индивидуальными предпочтениями. Для нахождения медианы Кемени известны эвристические методы и точные методы дискретной оптимизации, разработанные, в частности, в [4, 6–8]. При этом точные переборные методы хорошо применимы в ограниченной постановке, а эвристические методы (например, алгоритм Кука–Сейфорда, медиана Литвака в пространстве векторов предпочтений) используют различные эвристики и способы определения ранга вариантов. Окончательное ранжирование может дать разные результаты в зависимости от выбора эвристического метода и, в частности, отличаться от результата точного метода.

Предлагаемый алгоритм значительно уменьшает перебор в этой экспоненциальной задаче, позволяет решать задачи большой размерности, что иллюстрируется приведенными результатами. Некоторые другие идеи метода, на котором основан этот алгоритм, можно найти в [9].

Для работы алгоритма не требуется специальных ограничений на исходные бинарные отношения, задающие индивидуальные предпочтения. Это могут быть не только линейные порядки или квазипорядки, но и произвольные бинарные отношения, заданные как совокупности упорядоченных пар. В данной работе авторы не рассматривают вопросы, связанные с интерпретацией и возможном практическом приложении подобных обобщений.

1. Рациональный коллективный выбор в реляционной модели агрегирования индивидуальных предпочтений. Проблема рационального коллективного выбора в рамках реляционной модели агрегирования индивидуальных предпочтений формулируется следующим образом, например в [4]. Группа из m лиц рассматривает несколько возможных вариантов (альтернатив) $A = \{a_1, \ldots, a_n\}$

решения некоторой проблемы. Индивидуальные предпочтения членов группы выражаются ранжированиями, строгими или нестрогими, и задают профиль $\langle \rho_1, \ldots, \rho_m \rangle$ бинарных отношений на A.

Обозначим через $LO\,[n]$ совокупность всех линейных порядков на A, т.е. бинарных отношений на A, являющихся одновременно рефлексивными, антисимметричными, транзитивными и линейными. Под строгим ранжированием понимается любое отношение из $LO\,[n]$, для нестрогого ранжирования необязательно условие антисимметричности. Расстояние между бинарными отношениями $\rho, \rho' \subseteq A \times A$ с матрицами $R(\rho) = \|r_{ij}\|_{n\times n}$, $R(\rho') = \|r'_{ij}\|_{n\times n}$, где r_{ij}, $r'_{ij} \in \{0,1\}$, будем задавать метрикой Хемминга:

$$d(\rho, \rho') = \sum_{i=1}^{n} \sum_{j=1}^{n} |r_{ij} - r'_{ij}|.$$

В качестве реляционного правила агрегирования предлагается выбрать медиану Кемени, т.е результирующую строгую ранжировку, “наиболее близкую” в смысле метрики Хемминга к индивидуальным ранжировкам. Таким образом, рассматриваемая задача сводится к нахождению линейной ранжировки $\hat{\rho} \in LO\,[n]$, минимизирующей функцию

$$D(\rho) = \sum_{t=1}^{m} d(\rho, \rho_t),$$

т.е.

$$\hat{\rho} \in \operatorname{Arg} \min_{\rho \in LO[n]} D(\rho). \tag{1.1}$$

В общем случае такая ранжировка может оказаться неединственной и потребуется нахождение всего множества

$$\operatorname{Arg} \min_{\rho \in LO[n]} D(\rho). \tag{1.2}$$

2. Применение метода ветвей и границ для получения результирующей ранжировки. Опишем метод ветвей и границ и предложенный алгоритм для нахождения результирующей ранжировки (1.1). Целевой функцией является $D(\rho)$, и мы ищем либо любое $\hat{\rho}$, удовлетворяющее (1.1), либо все множество (1.2). При этом, как показано в [9], в случае $\left|\operatorname{Arg} \min_{\rho \in LO[n]} D(\rho)\right| > 1$ при наличии дополнительной информации можно выделить единственную ранжировку либо множественность решений указывает на эквивалентность некоторых групп вариантов.

Для дальнейшего понадобятся некоторые обозначения и утверждения. Пусть $R^t(\rho) = \|r^t_{ij}\|_{n\times n}$ – матрица бинарного отношения ρ_t, $t = \overline{1,m}$,

$$P = \sum_{t=1}^{m} R^t = \|p_{ij}\|_{n\times n}, \quad P^{(1)} = \left\|p^{(1)}_{ij}\right\|_{n\times n} = \|m - p_{ij}\|_{n\times n}, \quad P^* = \left\|p^*_{ij}\right\|_{n\times n} = \left\|\min(p_{ij}, p^{(1)}_{ij})\right\|_{n\times n},$$

т.е. $p^{(1)}_{ij} = m - p_{ij}$, $p^*_{ij} = \min(p_{ij}, p^{(1)}_{ij})$, $i, j = \overline{1,n}$. Тогда, очевидно, что

$$\sum_{t=1}^{m} |0 - r^t_{ij}| = \sum_{t=1}^{m} r^t_{ij} = p_{ij}, \quad \sum_{t=1}^{m} |1 - r^t_{ij}| = \sum_{t=1}^{m} (1 - r^t_{ij}) = m - \sum_{t=1}^{m} r^t_{ij} = m - p_{ij} = p^{(1)}_{ij}. \tag{2.1}$$

Рассмотрим бинарное отношение $\tilde{\rho}$ с матрицей $R(\tilde{\rho}) = \|\tilde{r}_{ij}\|_{n\times n}$, построенное из $\rho_1, \rho_2, \ldots, \rho_m$ по следующему правилу: $\tilde{r}_{ij} = 1 \Leftrightarrow p_{ij} \geq m/2$, $i, j = \overline{1,n}$. Тогда в силу (2.1)

$$\sum_{t=1}^{m} |\tilde{r}_{ij} - r^t_{ij}| = \min_{r \in \{0,\,1\}} \sum_{t=1}^{m} |r - r^t_{ij}| = p^*_{ij}, \quad i, j = \overline{1,n}. \tag{2.2}$$

Для доказательства (2.2) достаточно рассмотреть случаи:

а) $\tilde{r}_{ij} = 1$. Тогда $p_{ij} \geq m/2 \Rightarrow 2p_{ij} \geq m \Rightarrow p_{ij} \geq m - p_{ij} = p_{ij}^{(1)}$, откуда, используя (2.1), получаем

$$\sum_{t=1}^{m} |\tilde{r}_{ij} - r_{ij}^t| = p_{ij}^{(1)} = \min(p_{ij}, p_{ij}^{(1)}) = p_{ij}^*.$$

б) $\tilde{r}_{ij} = 0$. Тогда $p_{ij} \leq m/2 \Rightarrow 2p_{ij} \leq m \Rightarrow p_{ij} \leq m - p_{ij} = p_{ij}^{(1)}$, откуда, используя (2.1), получаем

$$\sum_{t=1}^{m} |\tilde{r}_{ij} - r_{ij}^t| = p_{ij} = \min(p_{ij}, p_{ij}^{(1)}) = p_{ij}^*.$$

Определим на множестве произвольных бинарных отношений ρ на множестве A с $R(\rho) = \|r_{ij}\|_{n\times n}$ (т.е. при любом выборе $r_{ij} \in \{0,1\}$, $i, j = \overline{1,n}$) минимально возможное значение величины

$$D_i(\rho) = \sum_{t=1}^{m} \sum_{j=1}^{n} |r_{ij} - r_{ij}^t| = \sum_{j=1}^{n} \sum_{t=1}^{m} |r_{ij} - r_{ij}^t|.$$

В силу (2.2), это значение достигается при $r_{ij} = \tilde{r}_{ij}$, $j = \overline{1,n}$. Обозначив минимально возможное значение величины $D_i(\rho)$ через α_i, получаем, что

$$\alpha_i = \sum_{t=1}^{m} \sum_{j=1}^{n} |\tilde{r}_{ij} - r_{ij}^t| = D_i(\tilde{\rho}), \quad i = \overline{1,n}, \quad D(\tilde{\rho}) = \sum_{i=1}^{n} D_i(\tilde{\rho}) = \sum_{i=1}^{n} \alpha_i.$$

При этом $D(\tilde{\rho})$ – минимально возможное значение величины $D(\rho)$ для всех бинарных отношений ρ на множестве A. Используя (2.2), имеем

$$\alpha_i = \sum_{t=1}^{m} \sum_{j=1}^{n} |\tilde{r}_{ij} - r_{ij}^t| = \sum_{j=1}^{n} \sum_{t=1}^{m} |\tilde{r}_{ij} - r_{ij}^t| = \sum_{j=1}^{n} p_{ij}^*, \quad i = \overline{1,n}.$$

Допустимыми решениями (последовательностями) являются последовательности альтернатив $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_n} \rangle$, где $\{a_{i_1}, a_{i_2}, \ldots, a_{i_n}\} = A$, т.е. перестановки множества A, каждая из которых соответствует некоторой ранжировке, т.е. линейному порядку вида $a_{i_1} \succ a_{i_2} \succ \ldots \succ a_{i_{n-1}} \succ a_{i_n}$ (или, что то же самое, вида $a_{i_n} \prec a_{i_{n-1}} \prec \ldots \prec a_{i_2} \prec a_{i_1}$) на A, где a_{i_1} (соответственно a_{i_n}) – наибольший (наименьший) элемент в этом линейном порядке. Множество возможных решений описывается n деревьями, каждое из которых имеет начальную вершину (1-го уровня), соответствующую некоторой альтернативе a_{i_1}, $i_1 = 1, 2, \ldots, n$. Каждая вершина a_{i_1} 1-го уровня соединяется ребрами со всеми вершинами $a_{i_2} \in A \backslash \{a_{i_1}\}$ 2-го уровня. Каждая вершина $a_{i_2} \in A \backslash \{a_{i_1}\}$ 2-го уровня соединяется ребрами со всеми вершинами из множества $A \backslash \{a_{i_1}, a_{i_2}\}$ 3-го уровня и т.д., пока не дойдем до вершин последнего n-го уровня. Объединение этих деревьев дает лес $Л$ допустимых решений. Единственная цепь, соединяющая в $Л$ любую вершину a_{i_n} n-го уровня с одной из вершин 1-го уровня, дает нам одно из допустимых решений $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_n} \rangle$, где a_{i_1} – вершина 1-го уровня, а a_{i_n} – вершина n-го уровня. Кроме того, каждой промежуточной вершине a_{i_k} k-го уровня, где $k = \overline{2, n-1}$, аналогичным образом ставим в соответствие последовательность $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_k} \rangle$, однозначно определяемую единственной цепью, соединяющей a_{i_k} с некоторой (единственной) вершиной a_{i_1} 1-го уровня. Такие последовательности также будем называть допустимыми. Тем самым каждая вершина a_{i_k} k-го уровня, где $k = \overline{1, n-1}$, определяет множество допустимых решений, а именно последовательностей вида

$$\left\langle \underbrace{a_{i_1}, a_{i_2}, \ldots, a_{i_k}}_{\text{1-я часть}}, \underbrace{a_{i_{k+1}}, \ldots, a_{i_n}}_{\text{2-я часть}} \right\rangle,$$

продолжающих (т.е. с одинаковой первой частью последовательности длины k) допустимую последовательность $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_k} \rangle$. Воспользуемся величинами α_i для получения нижних оценок для $D(\rho)$ на множествах допустимых решений, соответствующих каждой промежуточной вершине в $Л$. Очевидно, что величина

$$D(\tilde{\rho}) = \sum_{i=1}^{m} \alpha_i$$

является нижней оценкой для множества всех допустимых решений. Приведем теперь нижние оценки величины $D(\rho)$ для вершин 1-го уровня. Нижние оценки будут уточняться по мере увеличения номера уровня вершин. Каждая вершина a_i 1-го уровня соответствует выбору альтернативы a_i в качестве наибольшего элемента. Это полностью определяет величину

$$D_i(\rho) = \sum_{t=1}^{m} \sum_{j=1}^{n} |r_{ij} - r_{ij}^t|$$

для всех линейных порядков ρ на A с наибольшим элементом a_i, которую обозначим через β_i, поскольку в силу такого выбора для $R(\rho) = \|r_{ij}\|_{n \times n}$ выполняется $r_{ii} = 1$, $r_{ij} = 0$ при $i \neq j$. Следовательно, используя (2.1), получаем

$$\beta_i = D_i(\rho) = \sum_{t=1}^{m} \sum_{j=1}^{n} |r_{ij} - r_{ij}^t| = \sum_{j=1}^{n} \sum_{t=1}^{m} |r_{ij} - r_{ij}^t| =$$
$$= \sum_{t=1}^{m} |1 - r_{ii}^t| + \sum_{j \neq i} \sum_{t=1}^{m} |0 - r_{ij}^t| = p_{ii}^{(1)} + \sum_{j \neq i} p_{ij}, \quad i = 1, \ldots, n.$$

Но тогда можно уточнить нижние оценки величины $D(\rho)$ для допустимых решений, соответствующих выбору a_i в качестве вершины 1-го уровня, т.е. решений, соответствующих дереву с начальной вершиной a_i. Это будет величина

$$\beta_i + \sum_{j=1, j \neq i}^{n} \alpha_j, \quad i = \overline{1, n}$$

(т.е. в качестве нижней оценки для

$$D(\rho) = D_i(\rho) + \sum_{j=1, j \neq i}^{n} D_j(\rho)$$

взяли сумму точного значения величины $\beta_i = D_i(\rho)$ для всех решений, соответствующих этому дереву, и сумму нижних оценок α_j для остальных $D_j(\rho)$).

Заметим далее, что выбор некоторой альтернативы a_i в качестве вершины последнего n-го уровня соответствует выбору альтернативы a_i в качестве наименьшего элемента, а это полностью определяет величину

$$D_i(\rho) = \sum_{t=1}^{m} \sum_{j=1}^{n} |r_{ij} - r_{ij}^t|$$

для всех линейных порядков ρ на A с наименьшим элементом a_i, которую обозначим через γ_i, поскольку в силу такого выбора для $R(\rho) = \|r_{ij}\|_{n \times n}$ выполняется $r_{ij} = 1$, $j = \overline{1, n}$. Тогда, используя (2.1), получаем

$$\gamma_i = D_i(\rho) = \sum_{t=1}^{m} \sum_{j=1}^{n} |r_{ij} - r_{ij}^t| = \sum_{j=1}^{n} \sum_{t=1}^{m} |r_{ij} - r_{ij}^t| = \sum_{j=1}^{n} \sum_{t=1}^{m} |1 - r_{ij}^t| = \sum_{j=1}^{n} p_{ij}^{(1)}, \quad i = \overline{1, n}.$$

Обозначим $\eta_i = \gamma_i - \alpha_i$, $i = \overline{1,n}$. Учитывая то, что в любом из допустимых решений какая-нибудь из альтернатив a_j является наименьшим элементом, а следовательно, даст вклад в $D(\rho)$, равный $D_j(\rho) = \gamma_j = \alpha_j + \eta_j$, получаем еще одно уточнение для нижней оценки величины $D(\rho)$ для допустимых решений, соответствующих выбору a_i в качестве вершины 1-го уровня, т.е. решений, соответствующих дереву с начальной вершиной a_i. Это будет величина, которую обозначаем через ν_i и определяем по формуле

$$\nu_i = \beta_i + \sum_{j=1, j\neq i}^{n} \alpha_j + \min_{j\in\{1,\ldots,n\}\setminus\{i\}} \eta_j. \tag{2.3}$$

Аналогично находятся нижние оценки и для каждой промежуточной вершины a_{i_k} k-го уровня, где $k = \overline{2, n-1}$, которая, как отмечалось выше, задает некоторое множество допустимых решений. Ранее этой вершине была поставлена в соответствие допустимая последовательность $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_k}\rangle$, однозначно определяемая единственной цепью, соединяющей a_{i_k} с некоторой (единственной) вершиной a_{i_1} 1-го уровня. Поставим теперь этой последовательности в соответствие длину $\beta_{i_1 i_2\ldots i_k}$, равную ее вкладу в $D(\rho)$ для всех допустимых решений ρ, продолжающих ее, т.е.

$$\beta_{i_1 i_2\ldots i_k} = D_{i_1}(\rho) + \ldots + D_{i_k}(\rho), \tag{2.4}$$

которую удобно определять по мере увеличения уровня k, согласно рекуррентному соотношению

$$\beta_{i_1 i_2\ldots i_k} = \beta_{i_1 i_2\ldots i_{k-1}} + D_{i_k}(\rho) = \beta_{i_1 i_2\ldots i_{k-1}} + \sum_{t=1}^{m}\sum_{j=1}^{n} |r_{i_k,j} - r_{i_k,j}^t|,$$

где начальная величина β_{i_1} была определена ранее. Для организации вычислительного процесса последовательного определения этих величин остается уточнить порядок вычисления величины

$$D_{i_k}(\rho) = \sum_{t=1}^{m}\sum_{j=1}^{n} |r_{i_k,j} - r_{i_k,j}^t|.$$

Любое допустимое решение ρ, продолжающее допустимую последовательность $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_k}\rangle$, имеет вид $\langle a_{i_1}, \ldots, a_{i_k}, a_{i_{k+1}}, \ldots, a_{i_n}\rangle$, где $a_{i_n} \prec \ldots \prec a_{i_{k+1}} \prec a_{i_k} \prec \ldots \prec a_{i_1}$. Следовательно, для $R(\rho) = \|r_{ij}\|_{n\times n}$ справедливы равенства

$$r_{i_k,j} = 0, \quad j \in \{1,\ldots,n\}\setminus\{i_1,\ldots,i_k\}; \quad r_{i_k,j} = 1, \quad j \in \{i_1,\ldots,i_k\}, \tag{2.5}$$

дающие возможность однозначно определить значение $D_{i_k}(\rho)$ для указанных допустимых решений ρ. Такая однозначность позволяет обозначить величину $D_{i_k}(\rho)$ для любого допустимого решения ρ, продолжающего допустимую последовательность $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_k}\rangle$, иначе (без указания бинарного отношения ρ, используемого в ней) $D_{i_k}(\rho) = D_{i_1 i_2\ldots i_k}$. Тогда равенство (2.4) можно переписать в виде (исключаем вхождение неопределенного параметра ρ)

$$\beta_{i_1 i_2\ldots i_k} = D_{i_1} + D_{i_1 i_2} + \ldots + D_{i_1 i_2\ldots i_k}.$$

По аналогии с формулой (2.3), приведенной ранее для величины ν_i, соответствующей вершине a_i 1-го уровня, приведем теперь формулу для величины $\nu_{i_1 i_2\ldots i_k}$, соответствующей промежуточной вершине a_{i_k}, принадлежащей k-му уровню, где $k = \overline{2, n-1}$. Эта формула является нижней оценкой для $D(\rho)$ на множестве всех допустимых решений ρ, продолжающих допустимую последовательность $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_k} \rangle$, однозначно определяемую вершиной a_{i_k}:

$$\nu_{i_1 i_2 \ldots i_k} = \beta_{i_1 i_2 \ldots i_k} + \sum_{j \in \{1,\ldots,n\} \setminus \{i_1,\ldots,i_k\}} \alpha_j + \min_{j \in \{1,\ldots,n\} \setminus \{i_1,\ldots,i_k\}} \eta_j.$$

Приведем, кроме того, некоторые формулы для вычисления величины $D_{i_1 i_2 \ldots i_k}$, основанные на равенствах (2.5). Например, из (2.5) следует, что число $D_{i_1 i_2 \ldots i_k}$ равно сумме расстояний от i_k-х строк матриц $R^t = R(\rho_t) = \|r_{ij}^t\|_{n \times n}$, $t = \overline{1,m}$, до вектор-строки $(r_1, \ldots, r_n)$, где

$$r_j = \begin{cases} 0, & j \in \{1,\ldots,n\} \setminus \{i_1,\ldots,i_k\}, \\ 1, & j \in \{i_1,\ldots,i_k\}. \end{cases}$$

Кроме того, для большей простоты вычислений можно также воспользоваться матрицами P, $P^{(1)}$, а также равенствами (2.1):

$$D_{i_1 i_2 \ldots i_k} = \sum_{j \in \{i_1,\ldots,i_k\}} p_{i_k,j}^{(1)} + \sum_{j \in \{1,\ldots,n\} \setminus \{i_1,\ldots,i_k\}} p_{i_k,j}, \quad (2.6)$$

поскольку в силу (2.5)

$$D_{i_1 i_2 \ldots i_k} = \sum_{t=1}^{m} \sum_{j=1}^{n} |r_{i_k,j} - r_{i_k,j}^t| = \sum_{j \in \{i_1,\ldots,i_k\}} \sum_{t=1}^{m} |1 - r_{i_k,j}^t| + \sum_{j \in \{1,\ldots,n\} \setminus \{i_1,\ldots,i_k\}} \sum_{t=1}^{m} |0 - r_{i_k,j}^t|.$$

Отметим при этом, что

$$D_{i_1 i_2 \ldots i_n} = \gamma_{i_n}, \quad \beta_{i_1 i_2 \ldots i_n} = D(\rho), \quad (2.7)$$

где ρ – линейный порядок, соответствующий допустимой последовательности $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_n} \rangle$.

Замечание 1. Из формулы (2.6) следует, что при $k = \overline{3,n}$ в случае $\{a_{i'_1}, a_{i'_2}, \ldots, a_{i'_{k-1}}\} = \{a_{i_1}, a_{i_2}, \ldots, a_{i_{k-1}}\}$ (т.е. при $\{i_1, \ldots, i_{k-1}\} = \{i'_1, \ldots, i'_{k-1}\}$) выполняется $D_{i_1 i_2 \ldots i_k} = D_{i'_1 i'_2 \ldots i'_{k-1} i_k}$, что дает экономию в общем объеме вычислений. Например, $D_{2,1,3} = D_{1,2,3}$, $D_{2,1,4} = D_{1,2,4}$, $D_{3,1,4} = D_{1,3,4}$, $D_{3,2,4} = D_{2,3,4}$. Учитывая этот факт, можно сократить количество вычисляемых величин $D_{i_1 i_2 \ldots i_k}$ в $(k-1)!$ раз.

Описанный лес допустимых решений $Л$ дает схему ветвления для поставленной задачи. Заметим, что при достаточно большом n количество вершин в $Л$ может оказаться чрезмерно большим (например, в $Л$ имеется $n!$ вершин одного только n-го уровня). В связи с этим будем при обходе вершин из $Л$ использовать так называемое правило отсечения, позволяющее отсекать в $Л$ вершины вместе с цепями, соединяющими эти вершины с вершинами n-го уровня. Чтобы сформулировать это правило, понадобится величина $\tilde{D}$, которую будем называть рекордом. Текущее значение этой величины определяется при первом достижении любой вершины n-го уровня и уточняется при каждом прохождении любой другой вершины n-го уровня. В процессе обхода вершин из $Л$ будем для каждой текущей вершины a_{i_k}, принадлежащей k-му уровню, $1 \le k \le n$, где $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_k} \rangle$ – допустимая последовательность, находить величины $\beta_{i_1 i_2 \ldots i_k}, \nu_{i_1 i_2 \ldots i_k}$. Тогда при $k = n$ величина $\beta_{i_1 i_2 \ldots i_k} = \beta_{i_1 i_2 \ldots i_n}$ дает точное значение величины $D(\rho)$ для линейного порядка ρ, определяемого последовательностью $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_n} \rangle$. При каждом прохождении вершины n-го уровня уточним значение $\tilde{D}$ так, чтобы оно равнялось минимальному значению среди всех вычисленных на данном этапе величин вида $\beta_{i_1 i_2 \ldots i_n}$. В этом случае для любой вершины a_{i_k}, принадлежащей k-му уровню, где $1 \le k \le n-2$, в случае выполнения неравенства

$$\nu_{i_1 i_2 \ldots i_k} > \tilde{D} \quad (2.8)$$

вершина a_{i_k} может быть исключена из $Л$ вместе с цепями, соединяющими эту вершину с вершинами n-го уровня. При выполнении этого условия все линейные порядки ρ, которые соответствуют допустимым решениям, продолжающим последовательность $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_k} \rangle$, имеют значение $D(\rho)$, превышающее $\tilde{D}$, а следовательно, не могут принадлежать $\operatorname{Arg} \min_{\rho \in LO[n]} D(\rho)$. В случае, если ищем хотя бы одно решение из $\operatorname{Arg}\min\limits_{\rho\in LO[n]} D(\rho)$, условие отсечения (2.8) можно заменить на более слабое $\nu_{i_1 i_2 \ldots i_k} \geq \tilde{D}$. При выполнении этого условия все линейные порядки, которые соответствуют допустимым последовательностям, продолжающим $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_k}\rangle$, не дадут улучшения рекорда $\tilde{D}$.

Замечание 2. В силу (2.7) для любой вершины $(n-1)$-го уровня, соответствующей допустимой последовательности $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_{n-1}}\rangle$, выполняется равенство $\nu_{i_1 i_2 \ldots i_{n-1}} = \beta_{i_1 i_2 \ldots i_{n-1}} + \alpha_{i_n} + \eta_{i_n} = \beta_{i_1 i_2 \ldots i_{n-1}} + \gamma_{i_n} = \beta_{i_1 i_2 \ldots i_{n-1}} + D_{i_1 i_2 \ldots i_n} = \beta_{i_1 i_2 \ldots i_n}$, где $\{i_n\} = \{1, 2, \ldots, n\}\backslash\{i_1, \ldots, i_{n-1}\}$. Другими словами, $\nu_{i_1 i_2 \ldots i_{n-1}} = \beta_{i_1 i_2 \ldots i_n} = D(\rho)$, где ρ – линейный порядок, соответствующий допустимой последовательности $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_n}\rangle$. Отсюда следует, что уточнение рекорда можно производить уже при достижении любой вершины $(n-1)$-го уровня. Допустимая последовательность $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_{n-1}}\rangle$, соответствующая этой вершине, однозначно определяет допустимую последовательность $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_n}\rangle$, и при этом $\nu_{i_1 i_2 \ldots i_{n-1}} = \beta_{i_1 i_2 \ldots i_n}$. Таким образом, переход от вершин $(n-1)$-го уровня к вершинам n-го уровня является излишним.

Опишем теперь порядок прохождения вершин в *Л*. В [10] описаны три возможных способа продолжения ветвления. Применим к решению нашей задачи следующий способ. При выборе вершины для очередного ветвления (т.е. для прохождения продолжающих эту вершину цепей леса *Л*) берется не исключенная вершина a_{i_k} максимального уровня k, где $1 \leq k \leq n-2$, из множества всех достигнутых к этому моменту вершин леса *Л* и имеющая минимальное значение $\nu_{i_1 i_2 \ldots i_k}$ среди всех вершин этого уровня. Выбор вершины a_{i_k} k-го уровня однозначно определяет множество вершин $A\backslash\{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\}$ $(k+1)$-го уровня. Если таких вершин несколько, то выбираем среди них вершину с минимальным значением $\beta_{i_1 i_2 \ldots i_k}$, а если и таковых несколько, то – с наименьшим номером i_k.

3. Ускорение метода. Построим по матрице $P = R^1 + \ldots + R^m = \|p_{ij}\|_{n\times n}$ матрицу $\tilde{R} = \|\tilde{r}_{ij}\|_{n\times n}$, где $\tilde{r}_{ii} = 1$, $i = \overline{1,n}$, а в случае $i \neq j$ выполняется: $\tilde{r}_{ij} = 1$, если $p_{ij} > p_{ji}$, $\tilde{r}_{ij} = 0$, если $p_{ij} < p_{ji}$, и $\tilde{r}_{ij} = 1/2$, если $p_{ij} = p_{ji}$. Рассмотрим для каждой альтернативы $a_i \in A$ величину

$$\kappa_i = \kappa_i(a_i) = \sum_{j=1}^{n} \tilde{r}_{ij}, \quad i = \overline{1,n}.$$

Нетрудно показать (см., например, [1]), что если числа κ_i, $i = \overline{1,n}$, можно линейно упорядочить таким образом, что $\kappa_{i_1} = 1, \kappa_{i_2} = 2, \ldots, \kappa_{i_n} = n$, то $\operatorname{Arg}\min\limits_{\rho\subseteq A\times A} D(\rho)$ (здесь минимум берется на множестве всех бинарных отношений на A) состоит из единственного бинарного отношения, являющегося линейным порядком, и в этом случае $\operatorname{Arg}\min\limits_{\rho\subseteq A\times A} D(\rho) = \operatorname{Arg}\min\limits_{\rho\in LO[n]} D(\rho)$. Понятно, что такой случай является очень редким, но всегда по числам $\kappa_1, \ldots, \kappa_n$ может быть построен квазипорядок $\breve{\rho}$ на A: $\langle a_i, a_j\rangle \in \breve{\rho} \Leftrightarrow \kappa_i \leq \kappa_j$. Указанный квазипорядок определяет множество всех линейных порядков $\lambda \in LO\,[n]$, согласованных с $\breve{\rho}$, т.е. удовлетворяющих условию $\langle a_i, a_j\rangle \in \lambda \Rightarrow \langle a_i, a_j\rangle \in \breve{\rho}$. Выделив все согласованные $\lambda \in LO\,[n]$, можно перед началом работы метода ветвей и границ предварительно посчитать значение $D(\lambda)$ для выделенных таким образом $\lambda \in LO\,[n]$ и получить некоторое предварительное используемое в алгоритме значение рекорда $\tilde{D}$. Тогда применение этого значения в самом начале работы алгоритма (в случае его близости к $D^{LO}_{\min} = \min\{D(\rho) \mid \rho \in LO\,[n]\}$) может увеличить число отсечений и тем самым сократить время работы алгоритма. В тех случаях, когда найдутся $\lambda \in \operatorname{Arg}\min\limits_{\rho\in LO[n]} D(\rho)$, согласованные с $\breve{\rho}$, в результате совершения описанных предварительных вычислений будет уже в самом начале работы алгоритма выполняться $\tilde{D} = D^{LO}_{\min}$, что может привести к значительному уменьшению времени работы алгоритма.

Таблица 1. Данные для первых семи лидирующих альтернатив

i	12	13	14	15	1	2	3
a_i	a_{12}	a_{13}	a_{14}	a_{15}	a_1	a_2	a_3
$D^i_{\min}$	1124	1134	1140	1142	1156	1170	1184
$D^i_{\min} - D_{\min}$	50	60	66	68	82	96	110
RT_i	1	0.83	0.76	0.735	0.61	0.52	0.45

4. Случай, когда каждой индивидуальной ранжировке приписывается некоторый вес. Описанный метод построения агрегированной ранжировки легко переносится на случай, когда каждой индивидуальной ранжировке приписывается некоторый вес, отражающий ее относительную важность. Тогда вместо целевой функции $D(\rho)$ используется функция более общего вида:

$$\tilde{D}(\rho) = \sum_{t=1}^{m} c_t d(\rho, \rho_t), \quad \text{где} \quad c_t > 0, \quad t = 1,\ldots,m, \quad \sum_{t=1}^{m} c_t = m,$$

т.е. учитывается относительная важность бинарных отношений $\rho_1, \rho_2, \ldots, \rho_m$. При этом вместо матриц P, $P^{(1)}$, P^* соответственно используем матрицы

$$\tilde{P} = \sum_{t=1}^{m} c_t R^t = \|\tilde{p}_{ij}\|_{n\times n},$$

$\tilde{P}^{(1)} = \|\tilde{p}^{(1)}_{ij}\|_{n\times n} = \|m - \tilde{p}_{ij}\|_{n\times n}$, $\tilde{P}^* = \|\tilde{p}^*_{ij}\|_{n\times n} = \|\min\{\tilde{p}_{ij}, \tilde{p}^{(1)}_{ij}\}\|_{n\times n}$. После такой замены описанный выше метод ветвей и границ переносится на этот случай практически без изменений. Бинарное отношение $\tilde{\rho}$, с матрицей $R(\tilde{\rho}) = \|\tilde{r}_{ij}\|_{n\times n}$, $\tilde{r}_{ij} \in \{0,1\}$, строится, исходя из $\rho_1, \rho_2, \ldots, \rho_m$, по правилу: $\tilde{r}_{ij} = 1 \Leftrightarrow \tilde{p}_{ij} \geq m/2$, $i, j = \overline{1,n}$. Аналогично (2.2) выполняются равенства

$$\forall i, j \in \{1,2,\ldots,n\} \quad \sum_{t=1}^{m} c_t |\tilde{r}_{ij} - r^t_{ij}| = \min\left\{\sum_{t=1}^{m} c_t |r - r^t_{ij}| \,|\, r \in \{0,1\}\right\} = \tilde{p}^*_{ij},$$

а все последующие рассуждения основываются только на этих равенствах.

5. Пример. Приведем примеры работы описанного алгоритма.

Группа из $m = 9$ лиц строго ранжирует варианты из $A = \{a_1,\ldots,a_n\}$, где $n = 20$. Для простоты заменим обозначение варианта a_i числом i, т.е. $A = \{1,2\ldots,20\}$. Пусть индивидуальные ранжировки имеют следующий вид:

$$\rho_1 : 1 \prec 2 \prec 3 \prec 4 \prec 5 \prec 6 \prec 7 \prec 8 \prec 9 \prec 10 \prec 11 \prec 12 \prec 13 \prec 14 \prec 15 \prec 16 \prec 17 \prec 18 \prec 19 \prec 20,$$

$$\rho_2 : 16 \prec 15 \prec 20 \prec 19 \prec 18 \prec 17 \prec 14 \prec 13 \prec 12 \prec 11 \prec 10 \prec 9 \prec 8 \prec 7 \prec 6 \prec 5 \prec 4 \prec 3 \prec 2 \prec 1,$$

$$\rho_3 : 19 \prec 14 \prec 20 \prec 13 \prec 12 \prec 11 \prec 16 \prec 17 \prec 10 \prec 9 \prec 8 \prec 7 \prec 6 \prec 5 \prec 4 \prec 3 \prec 2 \prec 1 \prec 15 \prec 18,$$

$$\rho_4 : 18 \prec 13 \prec 12 \prec 11 \prec 16 \prec 10 \prec 9 \prec 8 \prec 7 \prec 6 \prec 20 \prec 17 \prec 5 \prec 4 \prec 3 \prec 2 \prec 1 \prec 15 \prec 14 \prec 19,$$

$$\rho_5 : 12 \prec 20 \prec 17 \prec 19 \prec 11 \prec 10 \prec 9 \prec 8 \prec 7 \prec 6 \prec 5 \prec 4 \prec 3 \prec 2 \prec 1 \prec 15 \prec 16 \prec 14 \prec 13 \prec 18,$$

$$\rho_6 : 20 \prec 11 \prec 16 \prec 10 \prec 18 \prec 17 \prec 9 \prec 8 \prec 7 \prec 6 \prec 5 \prec 4 \prec 3 \prec 2 \prec 1 \prec 15 \prec 14 \prec 13 \prec 12 \prec 19,$$

$$\rho_7 : 19 \prec 18 \prec 17 \prec 10 \prec 9 \prec 8 \prec 7 \prec 6 \prec 5 \prec 4 \prec 3 \prec 2 \prec 1 \prec 15 \prec 14 \prec 16 \prec 13 \prec 12 \prec 11 \prec 20,$$

$$\rho_8 : 18 \prec 19 \prec 9 \prec 20 \prec 17 \prec 8 \prec 7 \prec 6 \prec 5 \prec 4 \prec 3 \prec 2 \prec 1 \prec 16 \prec 15 \prec 14 \prec 13 \prec 12 \prec 11 \prec 10,$$

$$\rho_9 : 8 \prec 7 \prec 6 \prec 19 \prec 5 \prec 4 \prec 3 \prec 2 \prec 20 \prec 1 \prec 15 \prec 14 \prec 16 \prec 13 \prec 12 \prec 18 \prec 17 \prec 11 \prec 10 \prec 9.$$

В результате произведенных вычислений было выделено множество $\operatorname{Arg}\min\limits_{\rho\in LO[n]} D(\rho)$, состоящее из единственной ранжировки

$$\lambda : 19 \prec 20 \prec 16 \prec 18 \prec 17 \prec 11 \prec 10 \prec 9 \prec 8 \prec 7 \prec 6 \prec 5 \prec 4 \prec 3 \prec 2 \prec 1 \prec 15 \prec 14 \prec 13 \prec 12.$$

Таблица 2. Зависимость времени расчетов от количества вершин n

n	20	21	22	23	24	25	26
$t(n)$	48.45 с	131.28 с	≅4 мин	≅6.9 мин	≅13.7 мин	≅65.2 мин	≅5 ч 17 мин
$K(n)$	88 208	200 881	303 046	461 831	777 970	3 298 727	14 404 423
$K(n)/n!$	$\cong 3.63\times10^{-14}$	$\cong 3.93\times10^{-15}$	$\cong 2.7\times10^{-16}$	$\cong 1.79\times10^{-17}$	$\cong 1.25\times10^{-18}$	$\cong 2.13\times10^{-19}$	$\cong 3.57\times10^{-20}$

В случае простого перебора пришлось бы перебрать $20! = 1.28047...\times10^{17}$ вариантов, т.е. крайне большое число. В результате применения описанного метода вычисления на компьютере Asus N73S (4-ядерный процессор Intel Core i7-2630QM, 2 ГГц) заняли 49 с. Программа написана на языке Python 3.8.3.

Были также подсчитаны значения $D_{\min} = D(\tilde{\rho}) = \min\{D(\rho) \mid \rho \subseteq A \times A\} = 1074$, $D_{\min}^{LO} = \min\{D(\rho) | \rho \in LO[n]\} = 1124 = D_{\min} + 50$, т.е. $D_{\min}^{LO}$ превышает $D_{\min}$ всего на 4.88%. Кроме того, были просчитаны минимальные значения $D(\rho)$ на множестве $LO[n]$ при фиксированном наибольшем элементе a_i, $i = \overline{1,20}$. Полученные в результате значения, которые обозначены через $D_{\min}^{i}$, можно использовать для формирования весов альтернатив, определяемых, например, по формуле $RT_i = (D_{\min}^{i} - D_{\min})/(D_{\min}^{LO} - D_{\min}), i = \overline{1,20}$. В табл. 1 приведены данные для первых лидирующих (относительно линейного порядка λ) семи альтернатив.

Вычисления производились и для случаев с $n > 20$. Последовательно добавлялись новые альтернативы (включались в каждый из девяти линейных порядков случайным образом). Приведем численные результаты для примера с $n = 25$, $m = 9$. Запишем элементы исходных линейных порядков по возрастанию их предпочтительности:

1, 25, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 21, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24;
24, 22, 16, 25, 23, 15, 20, 19, 21, 18, 17, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1;
19, 14, 20, 24, 13, 12, 22, 23, 11, 25, 16, 17, 10, 9, 8, 7, 6, 21, 5, 4, 3, 2, 1, 15, 18;
18, 24, 13, 25, 21, 12, 23, 11, 16, 10, 9, 8, 7, 6, 20, 17, 22, 5, 4, 3, 2, 1, 15, 14, 19;
12, 20, 23, 17, 24, 22, 19, 11, 10, 9, 8, 7, 6, 5, 4, 3, 25, 2, 1, 15, 16, 21, 14, 13, 18;
20, 11, 22, 16, 23, 21, 10, 24, 18, 17, 9, 8, 7, 6, 5, 4, 3, 2, 1, 15, 14, 25, 13, 12, 19;
22, 19, 24, 23, 18, 17, 25, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 15, 21, 14, 16, 13, 12, 11, 20;
21, 18, 25, 19, 9, 20, 17, 8, 7, 6, 5, 4, 3, 2, 1, 22, 16, 15, 14, 23, 13, 12, 11, 24, 10;
8, 7, 6, 19, 5, 4, 3, 2, 20, 1, 15, 14, 24, 16, 13, 12, 22, 18, 17, 11, 10, 23, 9, 21, 25.

Количество пройденных вершин составляет 3 298 727.

В результате произведенных вычислений было выделено множество $\mathrm{Arg}\min_{\rho \in LO[n]} D(\rho)$, состоящее из трех ранжировок:

19, 20, 24, 22, 23, 25, 16, **18, 17, 11, 10, 9, 8, 7, 6, 21, 5, 4, 3, 2, 1, 15, 14, 13, 12;**
19, 20, 24, 22, 25, 16, 23, **18, 17, 11, 10, 9, 8, 7, 6, 21, 5, 4, 3, 2, 1, 15, 14, 13, 12;**
20, 24, 22, 25, 19, 16, 23, **18, 17, 11, 10, 9, 8, 7, 6, 21, 5, 4, 3, 2, 1, 15, 14, 13, 12**.

Во всех трех найденных линейных порядках прослеживается одинаковый строгий порядок относительно 18 лидирующих альтернатив.

Результаты с временем расчетов $t(n)$, где $n = 20,\ldots,26$, представлены в табл. 2. В процессе вычислений определялось также $K(n)$ – количество пройденных вершин, приведенное в третьей строке. В последней строке приводится отношение $K(n)/n!$, где $n!$ – количество вычислений целевой функции $D(\rho)$, которое бы потребовалось в случае использования полного перебора.

В связи с этим отметим, что для $n = 26$ при простом переборе пришлось бы рассматривать $26! = 403\,291\,461\,126\,605\,635\,584\,000\,000 = 4.03 \times 10^{26}$ вариантов, что невозможно для любого компьютера за любой мыслимый промежуток времени.

Работу приведенного алгоритма можно условно разбить на два этапа: вычисление матрицы P и дальнейшее применение метода ветвей и границ, базирующееся только на элементах этой матрицы. При этом на первом этапе применяется простая процедура, имеющая сложность порядка

$O(mn^3)$ (т.е. вклад числа экспертов имеет здесь линейный тип зависимости). Таким образом зависимость алгоритма от количества экспертов m несущественна. Увеличение m даже в несколько раз дает изменения общего времени счета на доли секунды.

Заключение. Применение метода ветвей и границ позволяет найти точное решение поставленной задачи, а в случае его не единственности – все множество решений. Предложенный алгоритм может использоваться и при предположении, что индивидуальные ранжировки учитываются с некоторыми коэффициентами относительной важности. Применение алгоритма при наложении на множество ранжировок некоторых дополнительных условий (например, при рассмотрении задачи с закрепленными концами) позволяет устанавливать относительные веса альтернатив или производить попарное сравнение любых двух альтернатив.

СПИСОК ЛИТЕРАТУРЫ

1. *Миркин Б.Г.* Проблема группового выбора. М.: Наука, 1974. 256 с.
2. *Мулен Э.* Кооперативное принятие решений: Аксиомы и модели. М.: Мир, 1991. 464 с.
3. *Young H.P.* Condorcet's Theory of Voting // American Political Science Review. 1988. № 82. P. 1231–1244.
4. *Петровский А.Б.* Теория принятия решений. М.: ИЦ "Академия", 2009. 391 с.
5. *Нефедов В.Н., Осипова В.А., Смерчинская С.О., Яшина Н.П.* Непротиворечивое агрегирование отношений строгого порядка // Изв. вузов. Математика. 2018. № 5. С. 71–85.
6. *Литвак Б.Г.* Экспертная информация: методы получения и анализа. М.: Радио и связь, 1982.
7. *Корнеенко В.П.* Методы многокритериального оценивания объектов с многоуровневой структурой показателей эффективности. М.: МАКС Пресс, 2018. 296 с.
8. *Cook W.D.* Distance–based and Adhoc Consensus Models in Ordinal Preference Ranking // Europ. J. of Operational Research. 2006. № 172. P. 369–385.
9. *Нефедов В.Н.* Некоторые свойства линейно упорядоченной медианы для нечетного числа линейных асимметричных отношений. М.: МАИ,2021.50 с. – Деп. в ВИНИТИ 08.11.2021, № 62 – В2021.
10. *Нефедов В.Н.* Задачи дискретной оптимизации. М.: Изд-во МАИ, 1993. 60 с.