



Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ «Робототехника и комплексная автоматизация»

КАФЕДРА «Системы автоматизированного проектирования»

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
К ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ
НА ТЕМУ:

*Сравнительный анализ методов решения
задачи линейного упорядочения*

Студент РК6-41М
(Группа)

(Подпись, дата)

А.С. Антонов
(И.О.Фамилия)

Руководитель ВКР

(Подпись, дата)

А. Н. Божко
(И.О.Фамилия)

Нормоконтролер

(Подпись, дата)

С.В. Грошев
(И.О.Фамилия)

РЕФЕРАТ

Расчетно-пояснительная записка содержит 66 с., 22 рис., 0 табл., 21 источник, 1 приложение.

ЛИНЕЙНОЕ УПОРЯДОЧЕНИЕ, LOP, NP-ТРУДНЫЕ ЗАДАЧИ, КОМБИНАТОРНАЯ ОПТИМИЗАЦИЯ, МЕТАЭВРИСТИКИ, СРАВНИТЕЛЬНЫЙ АНАЛИЗ, ПРОГРАММНЫЙ КОМПЛЕКС.

Объект исследования – алгоритмы решения задачи линейного упорядочения (Linear Ordering Problem, LOP): точные, эвристические и метаэвристические методы.

Цель выпускной квалификационной работы – проведение сравнительного анализа эффективности различных классов алгоритмов решения задачи линейного упорядочения и разработка программного инструмента для их тестирования.

Методы проведения исследования – теоретический анализ NP-трудных задач комбинаторной оптимизации; разработка модульного программного комплекса, реализующего алгоритмы решения LOP; планирование и проведение вычислительных экспериментов на стандартных бенчмарках; статистический анализ результатов; оценка алгоритмов по критериям качества решения, времени выполнения и устойчивости.

Основные результаты выпускной квалификационной работы:

- проведен аналитический обзор современных методов решения LOP (точные, эвристические, метаэвристические) и оценена их теоретическая сложность;
- разработан программный комплекс с модульной архитектурой, реализующий ключевые алгоритмы и обеспечивающий генерацию тестовых данных, визуализацию и сравнение результатов;
- проведены вычислительные эксперименты, установившие границы эффективной применимости методов;
- разработаны практические рекомендации по выбору алгоритма решения LOP в зависимости от размерности матрицы и требований к точности/скорости.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
1. Постановка задачи линейного упорядочения	6
1.1. Пример решений для матричной перестановки	6
1.2. Альтернативные постановки	7
1.3. Модификации задачи	7
1.4. Связанные задачи	8
2. Анализ проблемы линейного упорядочения	9
3. Окрестности вставок и локальная оптимальность	13
3.1. Окрестность вставок	13
3.2. Локальная оптимальность	15
4. Ограниченная окрестность вставок	17
5. Методы решения задачи линейного упорядочения	20
5.1. Метод границ и ветвей с LP-релаксацией	20
5.2. Комбинированный алгоритм Митчелла и Борчерса	22
5.3. Алгоритм Беккера	23
5.4. Жадный алгоритм последовательной вставки	23
5.5. Парное жадное упорядочивание	25
5.6. Деструктивные эвристики	27
5.7. Комбинированные конструктивно-деструктивные методы	28
5.8. Иерархическое упорядочивание	30
5.9. Алгоритм локального поиска	31
5.10. Алгоритм поиска с запретами	33
5.11. Алгоритм поиска с рассеиванием	34
5.12. Итеративный локальный поиск	35
5.13. Локальный поиск на ограниченной окрестности вставки	37
5.14. Меметический алгоритм	37
5.15. Алгоритм великого потопа	38
6. Применения	40
7. Практическая часть	42
7.1. Бенчмарки	43
7.1.1. LOLIB	43
7.1.2. Случайно сгенерированные матрицы	44
7.1.3. XLOLIB	44

7.2. Архитектура программного комплекса.....	45
7.2.1. Архитектура классов и иерархия алгоритмов.....	45
7.2.2. Принцип работы	46
7.3. Вычислительный эксперимент.....	47
7.4. Анализ результатов	49
ЗАКЛЮЧЕНИЕ	58
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	60
ПРИЛОЖЕНИЕ А.....	62

ВВЕДЕНИЕ

Задача линейного упорядочения (Linear Ordering Problem, LOP) занимает центральное место в комбинаторной оптимизации благодаря своей универсальности и широкому спектру приложений. В условиях роста объема данных [1] и потребности в эффективном управлении сложными системами LOP становится инструментом для решения задач ранжирования, планирования и анализа зависимостей.

Гэри и Джонсон [2] продемонстрировали, что LOP является NP-сложной задачей. Однако, благодаря её многочисленным применениям в различных областях, таких как археология [3], экономика [4], теория графов [5], машинный перевод [6] мы можем найти большое количество работ, в которых бы LOP решался с помощью точных, эвристических и метаэвристических методов.

В экономике LOP используется для ранжирования инвестиционных проектов на основе их потенциальной доходности и рисков. В машинном обучении задача помогает строить консенсусные рейтинги в рекомендательных системах [7]. В спортивной аналитике с её помощью определяют рейтинги команд, учитывая исторические результаты матчей [8]. В биоинформатике LOP применяется для упорядочения геномных последовательностей, что критично для понимания эволюционных процессов [9]. В археологии LOP нашёл применение в задаче стратификации для определения наиболее вероятного хронологического порядка образцов, найденных в разных местах. Матрица, описывающая эту проблему, известна как матрица Харриса.

Рост интереса к Big Data и искусственному интеллекту усиливает потребность в алгоритмах, способных работать с высокоразмерными данными, что делает исследование LOP особенно актуальным.

Цель работы состоит в изучении теоретических основ задачи линейного упорядочения, исследовании областей применения и проведении сравнительного анализа методов её решения.

В теоретической части раскрывается постановка задачи и её связь с другими проблемами оптимизации. Далее анализируются методы решения, а в практическом разделе рассматриваются области применения данной задачи.

1. Постановка задачи линейного упорядочения

Для заданной матрицы B размера $n \times n$ задача линейного упорядочения (LOP) формулируется как задача нахождения перестановки строк и столбцов π , которая бы максимизировала функцию $f(\pi)$ в формуле (1):

$$f(\pi) = \sum_{i=1}^n \sum_{j=i+1}^n B_{\pi(i)\pi(j)}, \quad (1)$$

где $\pi(i)$ обозначает индекс строки (и столбца), занимающей позицию i в решении π . Иными словами, цель заключается в нахождении такой перестановки строк и столбцов матрицы B , которая бы максимизировала сумму элементов, находящихся выше главной диагонали. В данной постановке задача линейного упорядочивания известна как *triangulation problem of input-output matrices* [10]. Тем не менее существуют и альтернативные варианты представления LOP. Так, Марти и Рейнелт [11] интерпретируют LOP в терминах теории графов как поиск ациклического турнира $T = (V, A)$ в полном ориентированном графе K_n с весами дуг w_{ij} , максимизирующего целевую функцию:

$$\sum_{(i,j) \in A} w_{ij}, \text{ где } A \text{ — дуги, индуцированные } \pi.$$

Значения $B_{i,j}$ отображают силу предпочтения i -го элемента j -му.

1.1. Пример решений для матричной перестановки

Рассмотрим пример для $n = 5$, который будет использоваться в дальнейшем. На рисунке 1 представлены три различные решения: e , σ и σ^* . Исходная матрица представлена перестановкой $e = (1,2,3,4,5)$ (рис. 1a) и значением её фитнес функции $f(e)$, равным 138. Решение $\sigma = (2,3,1,4,5)$ (рис. 1b) иллюстрирует другой вариант решения, имеющий более оптимальное значение фитнеса $f(\pi) = 158$. Наилучшее возможное решение представлено перестановкой $\sigma^* = (5,3,4,2,1)$ (рис. 1c). с фитнесом $f(\sigma^*) = 247$.

	1	2	3	4	5
1	0	16	11	15	7
2	21	0	14	15	9
3	26	23	0	26	12
4	22	22	11	0	13
5	30	28	25	24	0

(a) $e = (1, 2, 3, 4, 5)$
 $f(e) = 138.$

	2	3	1	4	5
2	0	14	21	15	9
3	23	0	26	26	12
1	16	11	0	15	7
4	22	11	22	0	13
5	28	25	30	24	0

(b) $\sigma = (2, 3, 1, 4, 5)$
 $f(\sigma) = 158.$

	5	3	4	2	1
5	0	25	24	28	30
3	12	0	26	23	26
4	13	11	0	22	22
2	9	14	15	0	21
1	7	11	15	16	0

(c) $\sigma^* = (5, 3, 4, 2, 1)$
 $f(\sigma^*) = 247.$

Рис. 1 – Три различных решения для матрицы 5×5 : (a) исходная матрица; (b) не-оптимальное решение; (c) оптимальное решение.

1.2. Альтернативные постановки

В социальных науках LOP формализуется как задача агрегации k экспертных парных сравнений $\{C^{(1)}, \dots, C^{(k)}\}$ в консенсусное ранжирование π . Целевая функция минимизирует расхождение с экспертами:

$$\min_{\pi} \sum_{m=1}^k \sum_{i < j} |I[\pi(i) < \pi(j)] - C_{ij}^{(m)}|,$$

где $C_{ij}^{(m)} \in [0,1]$ — степень предпочтения i перед j у эксперта m . В комбинаторной оптимизации эквивалентная формулировка — задача о минимальном Feedback arc set (FAS) для турниров [12].

$$\min_{S \subseteq E} \sum_{(i,j) \in S} B_{ij} \text{ такой, что } G \setminus S \text{ — ациклический граф,}$$

что тождественно (1) ввиду равенства $f(\pi) = (\sum_{i,j} B_{ij}) - \text{FAS}(B)$.

В сернации артефактов задача сводится к LOP при построении матрицы B из мер сходства $s(i, j)$ [13]:

$$B_{ij} = s(i, j) - \frac{1}{n-2} \sum_{k \neq i, j} s(i, k).$$

1.3. Модификации задачи

Взвешенная LOP с r приоритетными ограничениями формализуется как:

$$\max_{\pi} f(\pi) \text{ такой, что } \pi(a_q) < \pi(b_q), \quad \forall q = 1, \dots, r,$$

где (a_q, b_q) — пары элементов с фиксированным порядком. Для input-output таблиц с m секторами обобщённая блочная LOP требует разбиения матрицы на $m \times m$ блоков и поиска перестановок π_1, π_2 :

$$\max_{\pi_1, \pi_2} \sum_{p=1}^m \sum_{q=p+1}^m \sum_{i \in G_{\pi_1(p)}} \sum_{j \in G_{\pi_2(q)}} B_{ij},$$

где G_k — группы объектов. Частичная LOP для матрицы с пропусками M_{ij} = использует индикатор известных элементов:

$$\max_{\pi} \sum_{i < j} B_{\pi(i)\pi(j)} \cdot I[M_{\pi(i)\pi(j)} \neq \emptyset].$$

Версия с регуляризацией в машинном обучении включает штраф за отклонение от априорного порядка π_0 :

$$\max_{\pi} f(\pi) - \lambda \cdot d(\pi, \pi_0), \quad \lambda > 0,$$

где $d(\cdot)$ — расстояние Кендалла.

1.4. Связанные задачи

Двумерное обобщение (double ordering problem) для бинарной матрицы D :

$$\max_{\pi, \sigma} \sum_{\pi(i) < \pi(j)} \sum_{\sigma(k) < \sigma(l)} D_{ik} D_{jl}.$$

Транспортная модификация с n пунктами и стоимостями c_{ij} :

$$\max_{\pi} \sum_{i=1}^{n-1} c_{\pi(i), \pi(i+1)}.$$

Теоретико-игровая постановка с N агентами, где целевая функция агента k зависит от его предпочтений P_k :

$$\forall k: \max_{\pi} \sum_{i < j} I[\pi(i) < \pi(j)] \cdot P_k(i, j).$$

Указанные вариации расширяют применимость LOP в междисциплинарных исследованиях, сохраняя ядро исходной постановки — оптимизацию линейного порядка на основе попарных сравнений.

2. Анализ проблемы линейного упорядочения

В ходе анализа проблемы было выявлено, что для любой перестановки индексов σ в матрице B размером $n \times n$ справедливо следующее:

- С каждым индексом $\sigma_i = k, i = 1, \dots, n$ связано $2(n - 1)$ записей из матрицы B : $(n - 1)$ из строки k и $(n - 1)$ из столбца k .
- Набор связанных записей каждого индекса $\sigma_i = k, i = 1, \dots, n$ может быть сгруппирован в пары. Например, каждой записи из строки $k, B_{k\sigma_j}$, соответствует точка из столбца $k, B_{\sigma_j k}$, симметрично расположенная относительно главной диагонали.
- Все пары записей, связанные с индексом σ_i , остаются связанными с ним даже после перестановок.
- Каждая запись $B_{\sigma_i \sigma_j}$ связана с двумя индексами, σ_i и σ_j .
- Для каждой пары $\{B_{\sigma_i \sigma_j}, B_{\sigma_j \sigma_i}\}$ одна из записей всегда расположена выше главной диагонали, а другая ниже.

Утверждения проиллюстрированы примером на рисунке 2. В этом примере представлены два различных решения: $e = (1, 2, 3, 4, 5)$ на рис. 2а и $\sigma = (1, 3, 2, 4, 5)$ на рис. 2б. На обоих рисунках записи, связанные с индексом 2, выделены жирным шрифтом. Мы видим, что, несмотря на разное упорядочивание, в обоих решениях набор записей, связанных с индексом 2, одинаков, то есть $(21, 14, 15, 9, 16, 23, 22, 28)$. Несмотря на то, что позиция индекса 2 в e и σ различна ($e_2 = 2$ и $\sigma_3 = 2$), попарное отношение связанных с ним записей остается неизменным (см. обведенные индексы на рис. 2).

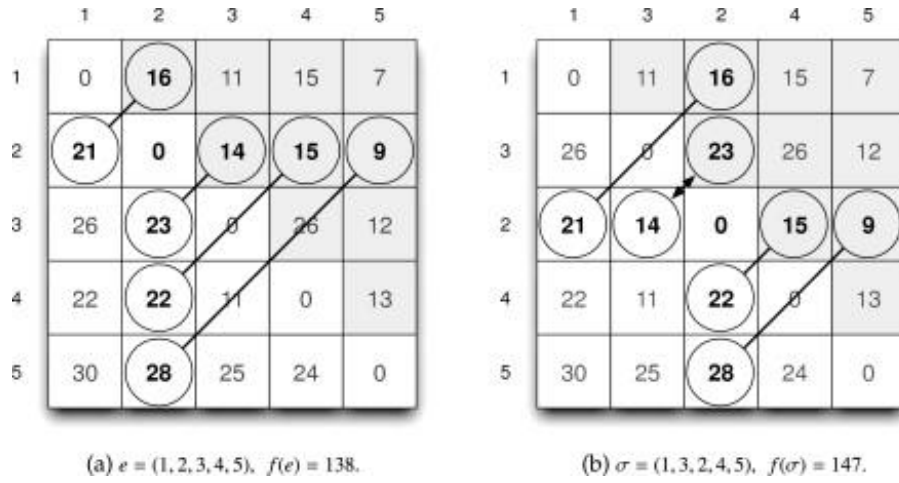


Рис. 2 – Два различных решения для экземпляра $n = 5$. Обведенные записи, соединенные ребрами, обозначают пары записей, связанные с индексом 2: (а) исходное состояние; (б) после перестановки.

Проверяя расположение связанных записей индекса 2 в σ , мы замечаем, что пара (14, 23) поменяла свои позиции в σ , 14 теперь ниже главной диагонали, а 23 выше неё. Введём новое понятие: вклад индекса в фитнес-функцию.

Когда индекс $k = 1, \dots, n$ занимает позицию i в σ , то есть $\sigma_i = k$, вклад индекса k в функцию определяется суммой записей столбца k в строках $\sigma_1, \dots, \sigma_{i-1}$ и суммой записей строки k в столбцах $\sigma_{i+1}, \dots, \sigma_n$. Иными словами, предыдущие $i - 1$ индексы $\sigma_1, \dots, \sigma_{i-1}$ и последующие $n - i$ индексы $\sigma_{i+1}, \dots, \sigma_n$ определяют вклад индекса k в фитнес-функцию. Формальное описание представлено формулой 2.

$$c(\sigma, i) = \sum_{j=1}^{i-1} B_{\sigma_j \sigma_i} + \sum_{j=i+1}^n B_{\sigma_i \sigma_j} \quad (2)$$

Вернемся к примеру на рисунке 2, благодаря обмену позициями в пару (14, 23) вклад индекса 2 изменился с 54 ($16 + 14 + 15 + 9$) в e (рис. 2а) до 63 ($16 + 23 + 15 + 9$) в σ (рис. 2б). В случае индекса 3 его вклад также увеличился, поскольку пара (14, 23) связана с обоими индексами, 2 и 3. И наоборот, в случае индексов 1, 4 и 5 их вклад не меняется от e к σ .

Если мы внимательно посмотрим на уравнение (2), то поймем, что вклад индекса $\sigma_i = k$ на самом деле не определяется конкретным упорядочиванием индексов в предыдущей и последующей позициях i , но их группировкой в этих двух наборах позиций. Как показано в примере 2, вклад индексов 1, 4 и 5 не меняется от e к σ ,

поскольку группировка остальных индексов в предыдущем и последующем наборах позиций, связанных с индексами 1, 4 и 5, была одинаковой.

Следовательно, при заданном решении σ вклад индекса $\sigma_i, i \in 1, \dots, n$, в фитнес-функцию $f(\sigma, i)$ не зависит от порядка предыдущих индексов $\sigma_1, \dots, \sigma_{i-1}$ и от порядка последующих индексов $\sigma_{i+1}, \dots, \sigma_n$.

Пример из рисунка 3 иллюстрирует, как вклад индекса 3 не зависит от упорядочивания предыдущего и последующего наборов индексов. На рис. 3а вклад индекса 3 равен 63, как результат суммы $(11 + 14 + 26 + 12)$. Если мы проверим вклад индекса 3 в σ (см. рис. 3б), то увидим, что он также равен 63, несмотря на то что индексы $\{1, 2\}$ и $\{4, 5\}$ поменялись местами.

	1	2	3	4	5
1	0	16	11	15	7
2	21	0	14	15	9
3	26	23	0	26	12
4	22	22	11	0	13
5	30	28	25	24	0

(а) $e = (1, 2, 3, 4, 5), c(e, 3) = 63.$

	2	1	3	5	4
2	0	21	14	9	15
1	16	0	11	7	15
3	23	26	0	12	26
5	28	30	25	0	24
4	22	22	11	13	0

(б) $\sigma = (2, 1, 3, 5, 4), c(\sigma, 3) = 63.$

Рис. 3 – Эффект от замены индексов в позициях 1,2 и 4,5 для индекса 3: (а) исходное состояние; (б) после перестановки.

Как было сказано выше, вклад индекса k не зависит от порядка следования индексов в предыдущих и последующих множествах. Но что произойдет, если индекс $\sigma_j = l$ будет перемещен из предыдущего набора индексов σ_i в последующий набор индексов? В отличие от предыдущего случая, вклад индекса σ_i будет изменён. В этот момент стоит вспомнить, что каждая пара записей $\{B_{\sigma_i \sigma_j}, B_{\sigma_j \sigma_i}\}$ в матрице связана с двумя индексами, σ_i и σ_j , а значит, любой обмен местоположением σ_i по определению влияет на вклад в функцию приспособленности σ_i и σ_j . Фактически, перемещение σ_i в позицию j влияет на вклад всех индексов, расположенных между

позициями i и j . Приведенный ниже пример (рисунок 4) иллюстрирует изменения в фитнес-функции, вызванные перемещением индекса.

На рисунке 4а и б показана матрица C в соответствии с решениями $e = (1, 2, 3, 4, 5)$ и $\sigma = (1, 3, 4, 2, 5)$. В этом примере мы анализируем последствия перемещения индекса $e_2 = 2$ в позицию 4. В результате этой модификации индексы 3 и 4 сдвигаются на одну позицию влево, что изменяет их вклад в фитнес-функцию. В частности, мы видим, что пары $\{14, 23\}$ и $\{15, 22\}$, связанные с индексами 2-3 и 2-4, поменялись местами. Поэтому вклад индекса 3 меняется с 63 ($11 + 14 + 26 + 12$) на 72 ($11 + 26 + 23 + 12$). Аналогично, вклад индекса 4 меняется с 69 ($15 + 15 + 26 + 13$) до 76 ($15 + 26 + 22 + 13$). Что касается индекса 2, то его вклад также меняется с 54 ($16 + 14 + 15 + 9$) до 70 ($16 + 23 + 22 + 9$). Обратите внимание, что вариация фитнес-вклада индекса 2 равна сумме вариаций индексов 3 и 4.

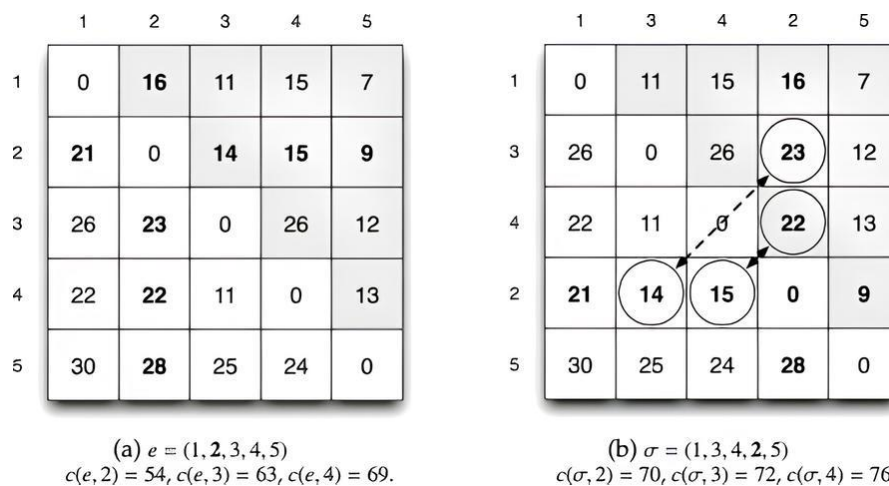


Рис. 4 – Иллюстрация влияния перемещения индекса 2 (из позиции 2 в позицию 4) на вклад индексов 2, 3 и 4 в фитнес-функцию. Числа, выделенные жирным шрифтом, обозначают записи, связанные с индексом 2. Обведенные кружком пары записей выделяют замененные записи: (а) исходное состояние; (б) после перестановки.

3. Окрестности вставок и локальная оптимальность

Среди широкого разнообразия алгоритмов для решения задачи LOP широко распространены алгоритмы, в основе которых частично или полностью лежат процедуры локального поиска. Хорошо известно, что алгоритмы локального поиска начинают со стартового или базового решения и итеративно пытаются заменить текущее решение лучшим, используя для этого систему окрестности. Среди различных систем окрестностей, предложенных для задачи LOP лучший результат показала система окрестностей вставки (Insert Neighborhood System, INS). По этой причине INS будет рассмотрен детально.

Далее даются базовые определения в контексте INS и локальной оптимальности.

3.1. Окрестность вставок

Два решения σ и σ' являются соседями в INS (N_I), если σ' получена перестановкой двух элементов i и j сдвигом остальных элементов, что представлено формулой 3.

$$\sigma' \in N_I(\sigma) \Leftrightarrow \exists i, j \in \{1, \dots, n\}, i \neq j, \text{ такие, что:}$$

$$\sigma' = \begin{cases} \begin{cases} (\sigma'_z = \sigma_z, z < i \wedge z > j) \\ (\sigma'_z = \sigma_{z+1}, i \leq z < j) \wedge \\ \sigma'_j = \sigma_i \end{cases} & \text{когда } i < j \\ \text{или} \\ \begin{cases} (\sigma'_z = \sigma_z, z < i \wedge z > j) \wedge \\ (\sigma'_z = \sigma_{z-1}, i < z \leq j) \wedge \\ \sigma'_i = \sigma_j \end{cases} & \text{когда } i > j \end{cases} \quad (3)$$

При выполнении операции вставки (перемещении индекса i в позицию j) некоторые элементы в верхнем треугольнике меняются местами с соответствующими элементами в нижнем треугольнике. Рассмотрим два различных сценария операции вставки и выделим конкретные элементы, которые обмениваются:

Случай $i > j$:

- Элементы $\{b_{\sigma_j \sigma_i}, b_{\sigma_{j+1} \sigma_i}, \dots, b_{\sigma_{i-1} \sigma_i}\}$ из верхнего треугольника (ячейки со светлой штриховкой) перемещаются в позиции $\{(j+1, j), (j+2, j), \dots, (i, j)\}$ в нижнем треугольнике

- Элементы $\{b_{\sigma_j \sigma_i}, b_{\sigma_i \sigma_{j+1}}, \dots, b_{\sigma_i \sigma_{i-1}}\}$ из нижнего треугольника (ячейки с темной штриховкой) перемещаются в позиции $\{(j, j+1), (j, j+2), \dots, (j, i)\}$ в верхнем треугольнике

Целевое значение соседней перестановки $\sigma', f(\sigma')$, вычисляется как сумма целевого значения $f(\sigma)$ и разницы обмениваемых элементов:

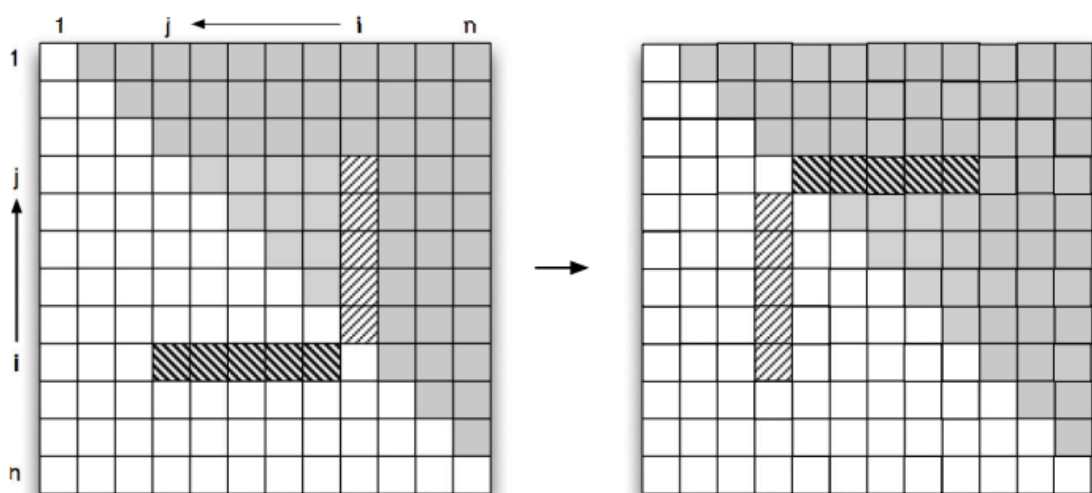
$$f(\sigma') = f(\sigma) + \sum_{z=j}^{i-1} (b_{\sigma_i \sigma_z} - b_{\sigma_z \sigma_i})$$

Случай $i < j$:

- Элементы $\{b_{\sigma_i \sigma_{i+1}}, b_{\sigma_i \sigma_{i+2}}, \dots, b_{\sigma_i \sigma_j}\}$ из верхнего треугольника (ячейки со светлой штриховкой) перемещаются в позиции $\{(j, i), (j, i+1), \dots, (j, j-1)\}$ в нижнем треугольнике
- Элементы $\{b_{\sigma_{i+1} \sigma_i}, b_{\sigma_{i+2} \sigma_i}, \dots, b_{\sigma_j \sigma_i}\}$ из нижнего треугольника (ячейки с темной штриховкой) перемещаются в позиции $\{(i, j), (i+1, j), \dots, (j-1, j)\}$ в верхнем треугольнике

Аналогично предыдущему случаю, целевое значение вычисляется как:

$$f(\sigma') = f(\sigma) + \sum_{z=i+1}^j (b_{\sigma_z \sigma_i} - b_{\sigma_i \sigma_z})$$



(a)

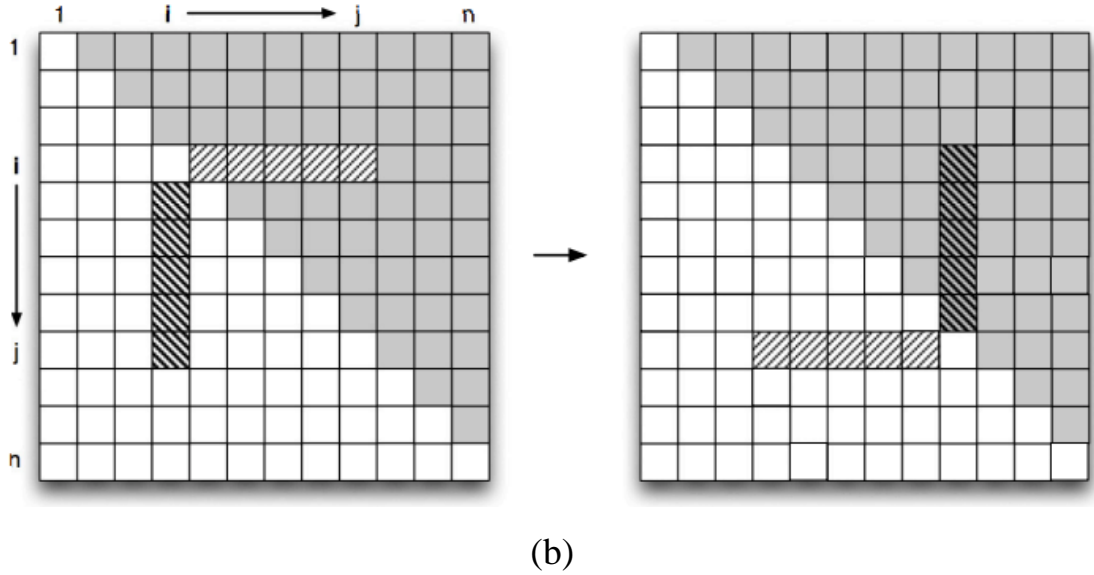


Рис. 5 – Обмен элементами при перемещении индекса. При перемещении индекса i в позицию j происходит обмен элементов матрицы. Рассматриваются два различных сценария: $i > j$ и $i < j$: (a) $i > j$. До и после операции вставки; (b) $i < j$. До и после операции вставки.

Как показано на рисунке 5, пары элементов в ячейках с различной штриховкой (темной и светлой) являются единственными элементами, которые обмениваются в результате операции вставки. Остальные элементы остаются на той же стороне главной диагонали, где и находились до операции. Каждый индекс i имеет связанные пары элементов $\{b_{i1}, b_{1i}\}, \dots, \{b_{in}, b_{ni}\}$. Вводится вектор разностей $(b_{i1} - b_{1i}, \dots, b_{in} - b_{ni})$, где каждое значение вектора описывает изменение целевой функции при обмене соответствующей пары элементов вследствие перемещения индекса i . Далее мы покажем, как этот вектор разностей для каждого индекса играет ключевую роль в определении того, приводит ли позиция индекса к локально оптимальным решениям.

3.2. Локальная оптимальность

Решение σ^* является локальным оптимумом для окрестности вставок, если все соседние решения σ имеют меньшее значение целевой функции:

$$\forall \sigma \in N_I(\sigma^*), f(\sigma^*) \geq f(\sigma)$$

Таким образом, в системе окрестности вставок решение считается локально оптимальным тогда и только тогда, когда среди всех возможных операций вставки не существует перемещения, которое улучшает текущее решение.

Согласно формуле 6, для данного решения σ и соседнего решения σ' , полученного перемещением индекса k в новую позицию, значение $f(\sigma')$ может быть вычислено путём:

- Перерасчёта вклада индекса k в новой позиции
- Добавления полученной вариации к $f(\sigma)$

Следовательно, решение является локальным оптимумом в окрестности вставок, если для любого индекса $i = 1, \dots, n$ не существует операции вставки, которая увеличивает вклад i в целевую функцию.

Для локально оптимального решения σ^* в окрестности вставок выполняется:

- Для каждого индекса $\sigma_i^*, i = 1, \dots, n$, все частичные суммы разностей между соответствующими элементами, расположенными перед i , неотрицательны:

$$\sum_{j=z}^{i-1} (b_{\sigma_j^*, \sigma_i^*} - b_{\sigma_i^*, \sigma_j^*}) \geq 0, z = i-1, \dots, 1 \quad (4)$$

- Все частичные суммы разностей для элементов, расположенных после i , неположительны:

$$\sum_{j=i+1}^z (b_{\sigma_j^*, \sigma_i^*} - b_{\sigma_i^*, \sigma_j^*}) \leq 0, z = i+1, \dots, n \quad (5)$$

4. Ограниченная окрестность вставок

В предыдущем разделе мы описали свойства (формулы 4 и 5), которым должны удовлетворять индексы решения σ , чтобы σ являлось локально оптимальным решением.

Следующий логичный шаг заключается в идентификации конкретных позиций, где индексы порождают локально оптимальные решения. Однако детальный анализ показывает, что такой подход является NP-трудной задачей, поскольку для каждого индекса в каждой позиции, требуется проверить $(n - 1)!$ перестановок.

Тем не менее, существуют позиции, в которых индексы не порождают локальный оптимум независимо от порядка остальных индексов. В отличие от предыдущего случая, обнаружение таких позиций является тривиальной задачей.

На основе вектора разностей, ассоциированного с индексами, в данном разделе мы анализируем основания для исключения позиций, в которых индексы не могут породить локально оптимальные решения. В результате анализа предлагается ограниченная окрестность вставок, исключающая операции перемещения индексов в позиции, не генерирующие локальные оптимумы. Для иллюстрации процесса идентификации таких позиций рассмотрим тривиальные граничные случаи, когда индекс k находится на первой или последней позиции:

- Для исключения первой позиции ($\sigma_1 = k$) необходимо показать, что ни одна перестановка вектора разностей не удовлетворяет уравнению (5). Достаточным условием является невыполнение (5) при $z = n$, так как сумма:

$$\sum_{j=1}^n (b_{\sigma_j k} - b_{k \sigma_j})$$

не зависит от порядка индексов $\{\sigma_2, \dots, \sigma_n\}$.

- Аналогично, для исключения последней позиции ($\sigma_n = k$) достаточно проверить невыполнение уравнения (4) при $z = 1$.

Для произвольной позиции i ($\sigma_i = k$) необходимо проверить отсутствие разбиения индексов на группы $\{1, \dots, i - 1\}$ и $\{i + 1, \dots, n\}$, при котором одновременно выполняются условия (4) и (5).

В связи с этим предлагается простой алгоритм, который начинается с сортировки вектора разностей, ассоциированного с индексом k , в порядке убывания. Поскольку наша цель - исключить неподходящие позиции, второй шаг заключается в проверке, удовлетворяет ли наиболее благоприятное разбиение вектора разностей уравнениям (4) и (5) при $\sigma_i = k$.

Процедура распределения значений включает:

- Размещение наибольшего значения вектора на позиции $i - 1$
- Второго по величине значения - на позиции $i - 2$
- И так далее для остальных значений

Для наименьших значений используется обратный порядок:

- Наименьшее значение размещается на позиции $i + 1$
- Следующее по величине - на позиции $i + 2$

Обозначим через σ' решение, полученное при таком упорядочивании вектора разностей. Тогда индекс k не порождает локально оптимальное решение в позиции i , если выполняется условие, представленное формулой 6.

$$\sum_{z=1}^{i-1} (b_{\sigma'_z k} - b_{k \sigma'_z}) < 0 \text{ или } \sum_{z=i+1}^n (b_{\sigma'_z k} - b_{k \sigma'_z}) > 0 \quad (6)$$

Продолжая данный подход для всех индексов в решении σ и всех возможных позиций, мы вычисляем бинарную матрицу ограничений R [14], представленную на рисунке 6, где нулевые элементы соответствуют позициям, в которых индексы не порождают локально оптимальные решения.

	1	2	3	4	5
1	0	16	11	15	7
2	21	0	14	15	9
3	26	23	0	26	12
4	22	22	11	0	13
5	30	28	25	24	0

Исходная матрица

	1	2	3	4	5
1	0	0	0	0	1
2	0	0	0	1	1
3	1	1	0	0	0
4	0	1	1	1	1
5	1	0	0	0	0

Матрица ограничений R

Рис. 6 – Матрица ограничений R .

Хотя на первый взгляд может показаться, что использование ограниченной окрестности не обязательно даст преимущество перед классическим жадным поиском (поскольку "плохие" решения иногда ведут к более перспективным областям пространства поиска), в случае задачи LOP обнаружена важная закономерность: операция вставки, выбираемая жадным алгоритмом (дающая максимальное улучшение), никогда не относится к запрещенным операциям согласно матрице R .

5. Методы решения задачи линейного упорядочения

LOP – это NP-полная задача, то есть мы не можем ожидать алгоритма, который решит ее за полиномиальное время. Однако LOP имеет множество практических приложений [7-9], и поэтому алгоритмы для его эффективного решения востребованы. Было предложено несколько точных и эвристических алгоритмов. Точные алгоритмы включают в себя метод границ и ветвей(B&B)[15], использующий LP-релаксацию для нижней границы, и комбинированный алгоритм Митчелла и Борчерса [16]. Современные точные алгоритмы могут решать достаточно большие задачи из определенных классов задач с числом столбцов и строк до нескольких сотен, в то время как на экземплярах из других классов гораздо меньшего размера они терпят неудачу. Независимо от типа решаемых задач, время вычислений точных алгоритмов сильно увеличивается с ростом размера матрицы.

LOP также решается с помощью ряда эвристических алгоритмов. К ним относятся жадный алгоритм Беккера [17], алгоритмы локального поиска [18], а также множество метаэвристических подходов, включая поиск с запретами [19], Scatter search [20] и алгоритмы итеративного локального поиска (ILS) [18]. В частности, подходы на основе ILS в настоящее время считаются наиболее успешными метаэвристиками.

Существующие алгоритмы обычно тестировались на множестве классов реальных и случайно сгенерированных примеров. Однако до сих пор мало известно о том, как производительность современных алгоритмов зависит от специфических характеристик различных доступных классов задач LOP, а также о том, как различия между примерами влияют на особенности их пространства поиска. Первые шаги в решении этих открытых вопросов были сделаны в работах Скьявинотто и Шутцле [18].

5.1. Метод границ и ветвей с LP-релаксацией

В данном методе задача формулируется в терминах целочисленного линейного программирования. Конкретная формулировка представлена формулой 3.

$$\max \sum_{i \neq j} W_{ij} \cdot x_{ij}, \quad (7)$$

где $x_{ij} = 1$, если элемент i предшествует j , и 0 иначе. При этом накладывается ограничение на сумму парных элементов: $x_{ij} + x_{ji} = 1$ для всех ($i \neq j$), и на транзитивность: $x_{ij} + x_{jk} - x_{ik} \leq 1$ для всех троек (i, j, k).

LP-релаксация переводит целочисленные переменные ($x_{ij} \in \{0,1\}$) в непрерывные ($0 \leq x_{ij} \leq 1$). Решение релаксированной задачи даёт верхнюю границу целевой функции. Если решение релаксации целочисленное и транзитивное, то оно оптимальное.

В иных случаях выбирается переменная (x_{ij}), значение которой в LP-решении ближе всего к 0.5 (наиболее неопределённая). Далее создаются две подзадачи или узла:

- Ветвь 1: ($x_{ij} = 1$);
- Ветвь 2: ($x_{ij} = 0$).

Подзадачи формируют список L . Для каждого узла находится решение и происходит проверка на соответствие ограничениям. Ветви нарушающие ограничения транзитивности отсекаются. Для корректных узлов вычисляется верхняя граница, если она меньше текущего лучшего целочисленного решения, то её тоже необходимо отсечь.

Алгоритм:

1. Инициализация корневого узла с LP-релаксацией

2. Пока есть неисследованные узлы:

Выбрать узел с наивысшей верхней границей и решить LP-релаксацию.

Если решение целочисленное и транзитивное, то необходимо обновить лучшее решение. Иначе:

1. Выбрать переменную для ветвления.
2. Создать подзадачи, добавив новые ограничения.
3. Добавить подзадачи в очередь узлов.

В ходе работы алгоритма могут быть исследованы все узлы. Поэтому он обладает экспоненциальной сложностью.

5.2. Комбинированный алгоритм Митчелла и Борчерса

Этот метод объединяет метод внутренней точки (Interior Point Method, IPM) и симплекс-метод в рамках алгоритма отсечений (Cutting Plane Algorithm) для решения задач линейного упорядочения (LOP)[16]. Его ключевая идея — использовать сильные стороны обоих методов:

- IPM — быстро находит приближённые решения на начальных этапах;
- Симплекс-метод — эффективно уточняет решение и генерирует отсечения.

В данном алгоритме существует три фазы:

1. Фаза внутренней точки

Нахождение решения, приближённого к оптимальному. Решается LP-релаксация LOP с помощью IPM за счёт высокой эффективности для задач большой размерности;

2. Фаза генерации отсечений

Устранение нарушений транзитивности, цикличности и нецелочисленных решений.

3. Фаза симплекс-метода

Уточнение решения после добавления отсечений. Высокая эффективность для локального поиска точных решений в суженном пространстве.

Фазы 2 и 3 повторяются, пока не будет найдено решение.

Преимущества подхода:

1. Скорость и точность

IPM быстро приближается к оптимуму. А симплекс-метод точно обрабатывает добавленные отсечения.

2. Сокращение итераций

Динамические отсечения сужают пространство решений, уменьшая число ветвей.

3. Устойчивость к размерности

IPM хорошо работает для задач большой размерности.

Комбинированный алгоритм Митчелла и Борчерса для решения LOP объединяет скорость метода внутренних точек (IPM) и точность симплекс-метода, дополненных

динамическими отсечениями. ИРМ быстро приближается к оптимуму, решая LP-релаксацию, после чего симплекс-метод уточняет решение, добавляя отсечения для устранения циклов и нарушений транзитивности. Это позволяет эффективно решать крупномасштабные задачи линейного упорядочения, сокращая число итераций и вычислительные ресурсы за счёт гибридного подхода, который балансирует между скоростью ИРМ и точностью симплекса.

5.3. Алгоритм Беккера

Конструктивная эвристика, предложенная О. Беккером [17]. На первом шаге выбирается индекс, максимизирующий стоимость, представленную формулой 8.

$$q_i = \frac{\sum_{k=1}^n c_{ik}}{\sum_{k=1}^n c_{ki}} i = 1 \dots n \quad (8)$$

Он помещается на первую позицию в перестановке. Затем этот индекс вместе с соответствующими столбцом и строкой удаляется, и из полученной подматрицы. Вычисляются новые значения q_i для оставшихся индексов. Эти шаги повторяются до тех пор, пока список индексов не станет пустым, что приводит к вычислительным затратам $O(n^3)$. А простая вариация этого алгоритма - вычислить значения q_i только один раз в начале работы алгоритма, отсортировать эти значения в неубывающем порядке, чтобы получить перестановку индексов. При использовании этого варианта решение может быть вычислено за $O(n^2)$.

5.4. Жадный алгоритм последовательной вставки

Greedy Construction — это одна из базовых эвристик для решения задачи линейного упорядочивания (LOP). Алгоритм строит допустимое решение, последовательно добавляя элементы в порядке, который локально максимизирует целевую функцию на каждом шаге. Этот метод прост в реализации, работает быстро и часто даёт хорошие начальные решения для последующего улучшения другими методами.

Шаги алгоритма:

1. Инициализация

Создаётся пустой порядок: `order = []`. Все элементы считаются недобавленными.

2. Последовательное добавление элементов

Для каждого недобавленного элемента i вычисляется его вклад (delta), если добавить его в текущую последнюю позицию. Вклад delta— это сумма весов дуг от i ко всем уже добавленным элементам. Выбираем элемент с максимальным delta и добавляем его в order.

3. Завершение

Повторяем шаг 2, пока все элементы не будут добавлены.

Модификации:

1. Лучшая вставка (Best insertion)

На каждом шаге проверяются все возможные позиции для нового элемента, а не только конец. Улучшается качество решения, но возрастает сложность.

2. Жадный алгоритм с предпросмотром (Look-Ahead Greedy)

На каждом шаге оценивается влияние добавления элемента на следующие n шагов. Чаще всего на 2 или 3. Требуется больше вычислений, но улучшает качество решений.

3. Жадный алгоритм со случайным выбором кандидата

Вместо строго максимума выбирается случайный элемент из топ- k кандидатов. Позволяет получать разные решения при последовательных запусках.

4. Взвешенный жадный алгоритм

Вводятся дополнительные веса для элементов. Например, на основе степеней в графе.

5. Обратный алгоритм

Порядок строится с конца. Сначала выбирается последний элемент, потом предпоследний и т. д.

Жадный алгоритм является фундаментальной эвристикой для решения LOP, сочетающей простоту реализации с достаточно высокой эффективностью. Этот метод особенно ценен при необходимости быстрого получения субоптимального решения или в качестве стартовой точки для более сложных алгоритмов. Его ключевое преимущество заключается в интуитивно понятном механизме пошагового построения решения, однако качество конечного результата существенно зависит от структуры

входных данных – в некоторых случаях жадный выбор может приводить к попаданию в выраженные локальные оптимумы. Модификации алгоритма, такие как Best Insertion или Look-Ahead Greedy, позволяют улучшить качество решения ценой увеличения вычислительной сложности, что делает их применение оправданным для задач средней размерности. В практических реализациях данный метод часто комбинируют с техниками локального поиска, что позволяет нивелировать его основной недостаток – отсутствие механизмов выхода из локальных оптимумов, сохраняя при этом вычислительную эффективность базового подхода.

5.5. Попарное жадное упорядочивание

Попарное жадное упорядочивание представляет собой эвристический подход к решению задачи линейного упорядочивания (LOP), основанный на последовательном построении решения через анализ и выбор оптимальных пар элементов. Этот метод отличается от классического жадного алгоритма тем, что на начальном этапе рассматривает пары элементов, что позволяет лучше учитывать их взаимное влияние на целевую функцию. Подход особенно эффективен в случаях, когда важны парные взаимодействия между элементами.

Шаги алгоритма:

1. Инициализация

Создаётся пустой частичный порядок.

2. Построение начальной пары

Для всех возможных пар элементов вычисляется сумма их взаимных весов ($W[i][j] + W[j][i]$). Выбирается пара с максимальной суммой. Она добавляется в частичный порядок

3. Последовательное расширение порядка

Для каждого не добавленного элемента определяется оптимальная позиция в существующем порядке путём проверки всех возможных мест вставки. Вычисляется изменение целевой функции для каждой потенциальной вставки. Элемент вставляется в позицию, максимизирующую целевую функцию.

4. Критерий остановки

Если остались необработанные элементы, то шаг 3. Иначе – конец.

Вычислительная сложность попарного жадного упорядочивания зависит от количества элементов и способа реализации. На этапе выбора начальной пары требуется $O(n^2)$ операций для оценки всех пар. Затем, для каждого из оставшихся элементов выполняется проверка $O(k)$ возможных позиций вставки (где k - текущая длина порядка), что в худшем случае приводит к общей сложности $O(n^3)$. Однако на практике можно оптимизировать процесс, сокращая количество проверяемых позиций или используя эвристики для предварительного отбора кандидатов.

Для улучшения базового алгоритма можно применить следующие модификации:

1. Ограниченный просмотр позиций

Проверка не всех возможных позиций вставки, а только наиболее перспективных на основе дополнительных критериев.

2. Ранжирование кандидатов

Предварительная сортировка элементов по определенному показателю (например, по сумме исходящих весов) для определения порядка их добавления.

3. Гибридный подход

Комбинация с другими эвристиками, например, использование локального поиска после построения начального решения для его дальнейшего улучшения.

Попарное жадное упорядочивание предлагает сбалансированный подход к решению LOP, сочетая относительную простоту реализации с учетом парных взаимодействий между элементами. Хотя метод не гарантирует нахождения глобального оптимума, он демонстрирует хорошую эффективность на практике, особенно при работе с задачами, где важную роль играют парные отношения. Основное преимущество алгоритма заключается в его способности строить качественные начальные решения, которые могут служить основой для более сложных методов оптимизации. Ключевыми направлениями для улучшения являются оптимизация вычислительной сложности и разработка адаптивных стратегий выбора пар и позиций вставки, позволяющих повысить качество итогового решения без существенного увеличения времени работы.

5.6. Деструктивные эвристики

Решают модифицированную задачу LOP, когда доступно исключение элементов из упорядочивания. Деструктивные эвристики представляют собой класс алгоритмов для задачи линейного упорядочивания (LOP), основанных на принципе последовательного удаления элементов из полного порядка. В отличие от конструктивных методов, которые строят решение "снизу вверх", деструктивные подходы начинают с полного множества элементов и итеративно устраняют наименее значимые компоненты. Такие методы особенно эффективны, когда можно определить относительно слабые связи между элементами, удаление которых минимально ухудшает качество решения.

Шаги алгоритма

1. Инициализация

Начинаем с полного порядка (например, случайной перестановки всех элементов).

2. Оценка элементов

Для каждого элемента в текущем порядке вычисляем, насколько ухудшится целевая функция при его удалении.

3. Удаление элемента

Выбираем и удаляем элемент, исключение которого приводит к наименьшему ухудшению целевой функции.

4. Повторение

Продолжаем процесс, пока не останется минимально допустимый порядок (например, два элемента).

5. Реконструкция

На основе оставшихся элементов строим итоговое упорядочивание, добавляя ранее удаленные элементы в оптимальные позиции.

Модификации

1. Адаптивное удаление

Динамическая корректировка критериев удаления в зависимости от текущего состояния решения

2. Частичная реконструкция

Периодическое перестроение части порядка вместо полного удаления элементов

3. Взвешенное удаление

Учет дополнительных характеристик элементов (степеней в графе, статистических показателей) при выборе кандидатов на удаление

4. Гибридный подход

Комбинация с конструктивными методами после достижения определенного уровня декомпозиции

Деструктивные эвристики предлагают перспективный альтернативный подход к решению LOP, особенно эффективный в случаях, когда качественное начальное упорядочивание может быть получено путем исключения слабых связей. Основное преимущество этих методов заключается в их способности выявлять и устранять наименее значимые компоненты решения, что часто приводит к более устойчивым конфигурациям по сравнению с чисто конструктивными подходами. Хотя вычислительная сложность может быть выше, чем у простых жадных алгоритмов, применение оптимизаций и гибридных стратегий делает эти методы практичными для различных классов задач. Наибольший потенциал деструктивные эвристики демонстрируют при работе с задачами, где имеется выраженная иерархия значимости элементов или четко различимые кластеры связей.

5.7. Комбинированные конструктивно-деструктивные методы

Комбинированные конструктивно-деструктивные методы представляют собой мощный класс эвристик для задачи линейного упорядочивания (LOP), объединяющий преимущества обоих подходов. Эти алгоритмы циклически чередуют фазы построения и разрушения решения, что позволяет преодолевать ограничения чисто конструктивных или деструктивных методов. Особенностью таких подходов является их способность постепенно улучшать решение, избегая преждевременной сходимости к локальным оптимумам.

Шаги алгоритма:

1. Фаза конструктивного построения:

Применяется жадный алгоритм для создания начального решения, последовательно добавляя элементы по выбранному критерию.

2. Оценка устойчивости:

Анализируется полученное решение, выделяя хорошо и плохо размещенные элементы.

3. Фаза частичного разрушения:

Удаляется определенная доля наименее устойчивых элементов (обычно 20-40%).

4. Фаза реконструкции:

Повторно добавляются удаленные элементы, используя усовершенствованные критерии вставки.

5. Итеративное повторение:

Циклически выполняются фазы 2-4 до достижения критерия остановки.

Модификации:

1. Адаптивное разрушение

Автоматическая регулировка доли удаляемых элементов на основе динамики улучшений

2. Селективная реконструкция

Приоритетное рассмотрение определенных классов элементов при повторной вставке

3. Гибридные критерии

Комбинация различных метрик для оценки качества размещения элементов

4. Многоуровневая стратегия

Применение разных методов на различных стадиях процесса оптимизации

Комбинированные конструктивно-деструктивные методы демонстрируют высокую эффективность при решении сложных экземпляров LOP, превосходя по качеству решения чисто конструктивные или деструктивные подходы. Их основное преимущество заключается в способности последовательно улучшать решение, избегая при этом застревания в локальных оптимумах. Хотя эти методы требуют больше вычислительных ресурсов, чем простые эвристики, их применение оправдано в случаях, когда необходимо получить высококачественное решение. Наиболее перспективными направлениями развития являются адаптивные стратегии управления параметрами алгоритма и интеллектуальные методы выбора элементов для разрушения

и реконструкции, позволяющие оптимизировать баланс между качеством решения и вычислительными затратами.

5.8. Иерархическое упорядочивание

Иерархическое упорядочивание представляет собой специализированный графовый метод решения задачи линейного упорядочивания (LOP), основанный на принципе рекурсивной декомпозиции. Этот подход особенно эффективен для задач с выраженной кластерной структурой, где элементы естественным образом группируются в иерархические уровни. Метод имитирует стратегию "разделяй и властвуй", сначала упорядочивая макроуровни системы, а затем детализируя порядок внутри них.

Шаги алгоритма:

1. Кластеризация графа

Исходный взвешенный ориентированный граф разбивается на кластеры с использованием метрики связности.

2. Построение метаграфа

Создаётся граф более высокого уровня, где вершины представляют кластеры, а веса - агрегированные связи между ними.

3. Рекурсивное упорядочивание

Применяется жадный алгоритм для упорядочивания метаграфа, затем рекурсивно повторяется процесс для каждого кластера.

4. Согласование уровней

Корректируется порядок на каждом уровне для минимизации межкластерных конфликтов.

5. Финальная сборка

Объединяются упорядоченные кластеры в итоговую линейную последовательность.

Модификации:

1. Адаптивная кластеризация

Динамический выбор уровня детализации кластеров на основе локальных характеристик графа

2. Многоуровневое согласование

Итеративная оптимизация порядка между соседними уровнями иерархии

3. Гибридная стратегия

Комбинация с локальными методами поиска для постобработки кластерных решений

4. Взвешенная агрегация

Учет дополнительных метрик при построении метаграфа для лучшего сохранения структуры исходной задачи

Иерархическое упорядочивание предлагает принципиально новый взгляд на решение LOP, особенно эффективный для крупномасштабных задач с внутренней модульной структурой. Метод превосходит традиционные подходы в случаях, когда элементы естественным образом образуют иерархические группы, демонстрируя при этом хорошую масштабируемость. Ключевым преимуществом является способность алгоритма работать с разными уровнями абстракции, что позволяет находить компромисс между качеством решения и вычислительными затратами. Наиболее перспективными направлениями развития остаются адаптивные стратегии кластеризации и методы интеллектуального согласования уровней иерархии, которые могут значительно улучшить качество итоговых решений для сложных графовых структур.

5.9. Алгоритм локального поиска

Алгоритм локального поиска является одним из базовых методов решения LOP[18]. Этот алгоритм начинает работу с произвольного решения и последовательно улучшает его, исследуя окрестность текущей перестановки с помощью операций вставки. На каждом шаге вычисляется изменение целевой функции, если находится лучшее решение, оно принимается в качестве нового текущего состояния. Процесс продолжается до тех пор, пока в окрестности не останется улучшающих перемещений, что соответствует достижению локального оптимума.

Шаги алгоритма:

1. Инициализация начального решения

Начальная перестановка, полученная случайно или с помощью алгоритма

2. Вычисление целевой функции

Для текущей перестановки σ вычисляется сумма весов:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n W[\sigma(i), \sigma(j)]$$

3. Определение окрестности

Соседние решения генерируются с помощью операции вставки элемента на новую позицию. Для элемента на позиции i рассматриваются все возможные позиции $j \neq i$.

4. Вычисление изменения целевой функции (дельта)

При перемещении элемента $x = \sigma(i)$ в позицию j :

Если $j < i$:

$$\Delta = \sum_{k=i+1}^j (W[x, \sigma(k)] - W[\sigma(k), x]) \quad (9)$$

Если $j > i$:

$$\Delta = \sum_{k=i+1}^j (W[\sigma(k), x] - W[x, \sigma(k)]) \quad (10)$$

5. Поиск улучшения

Для каждого элемента x находится позиция j , дающая максимальное Δ . После этого выбирается перемещение с наибольшим дельта $\Delta > 0$.

6. Обновление решения

Переместите элемент на новую позицию, обновите перестановку σ и сумму S .

7. Критерий остановки

Если улучшений нет (все $\Delta \leq 0$), алгоритм завершается. Текущее решение — локальный оптимум.

Алгоритм локального поиска для LOP демонстрирует высокую эффективность на небольших и средних размерностях задачи, однако его главным ограничением является зависимость от начального решения и риск застревания в локальных

оптимумах. Для улучшения результатов часто применяются модификации, такие как мультистартовый поиск или гибридизация с метаэвристиками (например, табу-поиском). Несмотря на простоту, этот метод остается популярным инструментом для приближенного решения LOP благодаря своей прозрачности и гибкости настройки.

5.10. Алгоритм поиска с запретами

Табу-поиск (Tabu Search, TS) — это метаэвристический алгоритм, расширяющий идеи локального поиска за счёт механизмов избегания циклов и выхода из локальных оптимумов [19]. В отличие от стандартного локального поиска, TS использует список табу, запрещающий возврат к недавно посещённым решениям, и стратегии диверсификации, что позволяет исследовать новые области пространства решений.

Шаги алгоритма:

1. Инициализация

Задаётся начальное решение случайным образом или через алгоритм. Инициализируется пустой список табу, который хранит запрещённые перестановки. Определяется критерий останова через максимально допустимое количество итераций.

2. Генерация окрестности

Для текущей перестановки σ рассматриваются все возможные вставки элемента x на новую позицию $j \neq x$. Вычисляется изменение целевой функции ΔS по формулам (9) и (10).

3. Выбор лучшего допустимого хода

Среди всех соседей выбирается тот, что даёт максимальное положительное улучшение ΔS . Если данный ход в списке табу, то он игнорируется и выбирается следующий по убыванию приращения ΔS сосед. Выполняется ход. Если обнаружено застревание в локальном минимуме, то применяются стратегии диверсификации (например, пертурбация или рестарт из случайной позиции).

4. Обновление списка табу

Запрещается обратное перемещение элемента на несколько итераций (например, 7-10). Список табу обновляется по принципу FIFO.

5. Критерий остановки

Алгоритм завершается при достижении лимита итераций или отсутствии решений.

Самой дорогостоящей по времени является операция генерации окрестности $O(n^2)$. За счёт ограничения на количество итераций общая сложность алгоритма в худшем случае будет ограничена:

$$O(k \cdot n^2),$$

где k — максимальное число итераций.

Табу-поиск демонстрирует высокую эффективность в решении LOP, особенно для больших матриц, благодаря способности избегать застревания в локальных оптимумах. Однако его производительность сильно зависит от настройки параметров (длины списка табу, стратегии диверсификации). В сравнении с локальным поиском, TS требует больше вычислений, но обеспечивает лучшее качество решений, что делает его предпочтительным выбором для сложных экземпляров LOP. Для дальнейшего ускорения можно комбинировать TS с другими метаэвристиками (например, реактивным поиском).

5.11. Алгоритм поиска с рассеиванием

Scatter search (SS) — это популяционная метаэвристика, которая комбинирует решения из элитного множества для генерации новых перспективных кандидатов [20]. В LOP SS эффективен благодаря способности сохранять разнообразие решений и комбинировать их лучшие фрагменты. В отличие от генетических алгоритмов, SS использует детерминированные методы выбора и комбинирования решений, что делает его менее зависимым от случайности и более стабильным для задач с выраженной структурой, такой как LOP.

Шаги алгоритма:

1. Инициализация популяции

Генерация начального множества решений для включения их в популяцию P .
Размер популяции обычно небольшой (5-20 решений).

2. Улучшение решений

Каждое решение из популяции локально оптимизируется (например, алгоритмом локального поиска). Как итог, имеется набор элитных решений E .

3. Построение Reference Set (RS)

Выбор b лучших по целевой функции решений из E для включения в RS. Добавление d разнообразных решений (например, с максимальным расстоянием до других) в RS.

4. Генерация новых решений (Subset Generation)

Комбинирование новых решений из пар в RS с помощью операторов линейной комбинации, плавной трансформации решений и других.

5. Улучшение новых решений

Применение локального поиска к сгенерированным решениям.

6. Обновление Reference Set

Замена худших решений в RS на новые улучшенные кандидаты. Критерием может выступать, как целевая функция или разнообразие (функция Хэмминга), так и их взвешенная сумма с динамически меняющимися весами.

7. Критерий остановки

Достигнуто максимальное число итераций без улучшений или достигнут лимит времени.

Scatter Search демонстрирует высокую эффективность для LOP, особенно в случаях, где важно сочетание интенсификации (глубокий поиск вокруг элитных решений) и диверсификации (за счёт включения разнообразных кандидатов). Его сила — в способности систематически комбинировать решения, а не полагаться на случайные мутации. Однако алгоритм требует тщательной настройки

5.12. Итеративный локальный поиск

Итеративный локальный поиск (Iterated Local Search, ILS) — это метаэвристика, объединяющая локальный поиск с механизмами выхода из локальных оптимумов [18]. В задаче линейного упорядочивания (LOP) ILS эффективен благодаря чередованию фаз интенсификации (глубокий поиск вокруг текущего решения) и диверсификации (возмущение для исследования новых областей). Алгоритм особенно

полезен для задач средней и большой размерности, где классический локальный поиск застревает в субоптимальных решениях.

Шаги алгоритма:

1. Генерация начального решения

Создание исходной перестановки σ_0 .

2. Локальный поиск

Применения локального поиска к σ_0 . Результат – локальный оптимум σ^* .

3. Perturbation (Возмущение)

Контролируемая модификация σ^* , чтобы выйти из его окрестности. Для LOP могут быть поменяны местами случайные пары элементов. Сдвиг блоков.

Важно, чтобы сила возмущения была минимальной, но достаточной для выхода из локального оптимума.

4. Критерий принятия решения

Новое решение σ' принимается, если оно лучше текущего ($S(\sigma') > S(\sigma_{best})$).

В вероятностных версиях может приниматься и решение хуже, где эта вероятность вычисляется по специальному закону.

5. Завершение

Достигнуто максимальное число итераций или отсутствие улучшений в течение t итераций (например, $t = 50$).

Итеративный локальный поиск — мощный метод для решения LOP, особенно для матриц размерности $n \geq 100$. Он легко адаптируется под разные алгоритмы пертурбаций, сочетает глубокий поиск и глобальное исследование, устойчив за счёт выхода из локальных минимумов.

Таким образом, выбор метода зависит от требований к точности, доступных ресурсов и специфики данных. Современные исследования LOP акцентируют развитие гибридных подходов, сочетающих скорость эвристик и точность математических моделей.

5.13. Локальный поиск на ограниченной окрестности вставки

Restricted Insert Neighbourhood Local Search (RIN-LS) — это модификация классического локального поиска, разработанная специально для задач большой размерности [14]. В контексте LOP алгоритм сочетает мощь LS с оптимизированной стратегией исследования окрестности, где вместо полного перебора всех возможных вставок рассматривается только их ограниченное подмножество. Этот подход особенно эффективен для матриц с большой размерностью, где традиционные методы становятся вычислительно непрактичными.

Шаги алгоритма:

1. Инициализация

Генерация начального решения σ_0

2. Локальный поиск с ограниченной окрестностью

Для каждого элемента рассматриваются только k ближайших позиций ($k \ll n$, типично $k = \log_2 n$). Используются префиксные суммы для вычисления изменения целевой функции за $O(1)$ на операцию

3. Критерии принятия решения

Детерминированное принятие только улучшающих решений. Или ограниченный откат к предыдущим решениям при длительном застое

4. Остановка

По количеству итераций.

Предложенный алгоритм RIN-LS демонстрирует значительное превосходство над традиционными методами решения LOP для задач большой размерности.

5.14. Меметический алгоритм

Меметический алгоритм (Memetic Algorithm, MA) представляет собой гибридный подход, комбинирующий генетический алгоритм с локальной оптимизацией [18]. В контексте задачи линейного упорядочения (LOP) данный метод демонстрирует высокую эффективность благодаря способности сочетать глобальный поиск в пространстве решений с локальным улучшением особей. Особенно результативен МА

при работе с задачами большой размерности, где традиционные точные методы становятся неприменимыми.

Шаги алгоритма:

1. Инициализация популяции

2. Оценка приспособленности

$$S(\pi_i) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n W[\pi_i(i), \pi_i(j)]$$

3. Селекция

Турнирный отбор или метод рулетки

4. Рекомбинация

Ordred Crossover (OX), Partially Mapped Crossover (PMX)

5. Мутация

Инверсия или транспозиция с вероятностью $p_m \approx 0.1$

6. Локальный поиск

7. Обновление популяции

Меметический алгоритм демонстрирует превосходные результаты при решении LOP, особенно для задач большой размерности. Ключевые преимущества:

- Эффективное сочетание глобального и локального поиска
- Гибкость настройки под конкретные условия задачи

5.15. Алгоритм великого потопа

Алгоритм "Великого потопа" (Great Deluge Algorithm, GDA) представляет собой метаэвристический подход для решения задач оптимизации, вдохновленный природным процессом адаптации к изменяющимся условиям [21]. В контексте LOP GDA демонстрирует особую эффективность благодаря способности избегать локальных оптимумов за счет динамического управления уровнем принятия решений ("уровнем воды"). Рисунок 7 иллюстрирует основные принципы работы алгоритма.

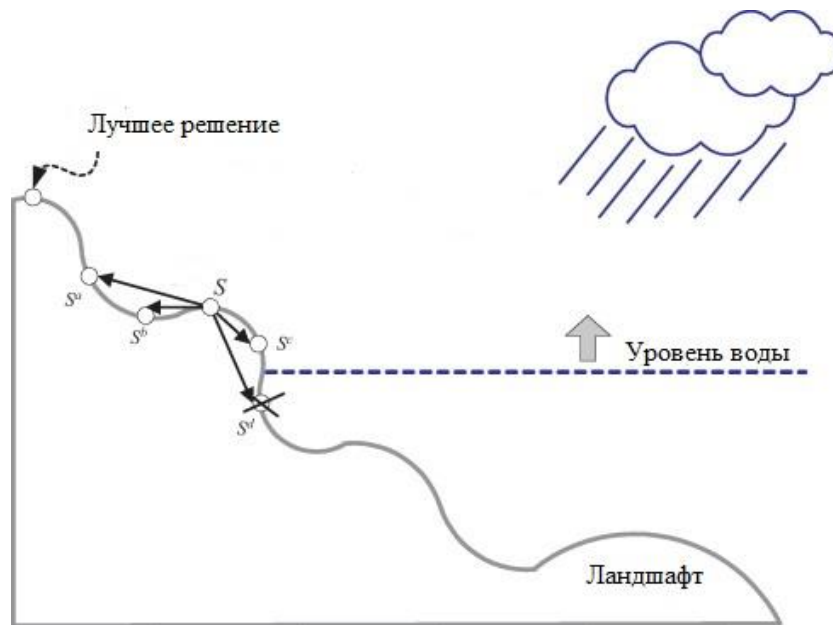


Рис. 7 – Принцип работы алгоритма великого потопа

iwl – коэффициент при начальном уровне воды, $dryratio$ - коэффициент испарения.

Шаги алгоритма:

1. Генерация начального решения s

2. Инициализация параметров

$$waterlevel = iwl \times cost(s)$$

3. Случайный выбор вершины v для вставки

4. Вставка v в найденную позицию

Получение нового решения s'

5. Если $cost(s') > waterlevel$:

1. Обновление решения: $s = s'$

2. Расчет величины испарения: $filling = dryratio \times waterlevel$

3. Увеличение уровня воды: $waterlevel += filling$

6. Если пройдено n итераций без улучшений:

$$\text{Сброс уровня воды } waterlevel = iwl \times cost(s)$$

Алгоритм GDA демонстрирует высокую эффективность при решении сложных задач LOP с многочисленными локальными оптимумами. Его ключевые преимущества включают: адаптивный механизм управления уровнем принятия решений через динамическое изменение $waterlevel$, и устойчивость к застреванию благодаря

периодическому сбросу параметров. Наибольшую эффективность алгоритм проявляет при работе с задачами повышенной размерности, обеспечивая оптимальный баланс между качеством получаемых решений и вычислительными затратами, что делает его предпочтительным выбором для сложных практических приложений.

6. Применения

LOP направлена на поиск оптимальной последовательности элементов, максимизирующей сумму весов над главной диагональю матрицы. Эта задача имеет широкое применение в различных областях, где требуется ранжирование объектов с учётом их взаимного влияния, приоритетов или структурных зависимостей. Универсальность LOP делает её инструментом для анализа данных, оптимизации процессов и принятия решений в междисциплинарных исследованиях.

В экономике LOP используется для анализа межотраслевых взаимодействий [4]. Например, матрица весов может отражать объёмы поставок между секторами экономики. Оптимальное упорядочивание позволяет определить последовательность отраслей, максимизирующую совокупный экономический эффект. В финансах LOP применяется для ранжирования активов по их волатильности или корреляции, что помогает строить сбалансированные инвестиционные портфели.

Социология

В социологических исследованиях LOP помогает ранжировать социальные факторы по степени влияния на определённый показатель (например, уровень образования на доход). В психологии задача используется для анализа результатов тестов, где важно определить порядок вопросов или реакций, минимизирующий когнитивную нагрузку. Например, упорядочивание стимулов в эксперименте для выявления скрытых поведенческих паттернов.

Биоинформатика

В биоинформатике LOP применяется для анализа данных генной экспрессии [9]. Матрица весов может отражать силу взаимодействия между генами, а оптимальное упорядочивание помогает выявить цепочки регуляторных процессов. В генетике LOP используется для построения филогенетических деревьев, где важно определить последовательность мутаций, объясняющую эволюционное развитие видов.

Информационные технологии

В IT LOP востребована в задачах ранжирования поисковой выдачи, рекомендательных систем и обработки естественного языка. Например, матрица весов может кодировать релевантность документов запросу, а оптимальный порядок элементов улучшает точность выдачи. В машинном обучении LOP используется для обучения моделей ранжирования (Learning to Rank), где алгоритмы учатся предсказывать порядок объектов, максимизирующий пользовательскую удовлетворённость.

Логистика

В логистике LOP помогает оптимизировать маршруты доставки, где матрица весов отражает время или стоимость перемещения между точками. Упорядочивание пунктов позволяет минимизировать общие затраты. В управлении цепями поставок LOP используется для расстановки приоритетов между этапами производства, чтобы сократить простои и повысить эффективность.

Археология

В археологии задача линейного упорядочения (LOP) играет ключевую роль в решении проблем стратиграфии — определения хронологической последовательности культурных слоёв и артефактов [3]. Например, матрица Харриса, используемая для визуализации стратиграфических отношений, кодирует информацию о взаимном расположении слоёв: если слой А перекрывает слой В, это указывает, что А моложе В. Однако при наличии множества пересекающихся слоёв и фрагментарных данных ручной анализ становится крайне трудоёмким. LOP позволяет автоматизировать этот процесс, преобразуя матрицу Харриса в матрицу весов, где элемент $W[i, j]$ отражает частоту случаев, когда слой i должен предшествовать слою j . Решение LOP даёт оптимальную последовательность слоёв, максимизирующую согласованность стратиграфических данных. Это особенно важно при анализе крупных раскопок, таких как многослойные поселения или некрополи, где точное датирование слоёв критично для реконструкции исторических процессов. Алгоритмы на основе LOP (например, метаэвристики) успешно применяются для обработки таких данных, сокращая время анализа и минимизируя субъективные ошибки, что способствует более объективному восстановлению хронологии археологических памятников.

LOP служит мощным инструментом для решения задач, требующих учёта парных взаимодействий и структурных зависимостей. Её применение охватывает экономику, социологию, биоинформатику, IT и логистику, демонстрируя междисциплинарную ценность. С развитием алгоритмов оптимизации и роста объёмов данных роль LOP будет только возрастать, особенно в областях, где критически важен анализ сложных систем и принятие решений на основе множества параметров.

7. Практическая часть

Задача линейного упорядочивания (LOP) сохраняет высокую востребованность в анализе экономических систем, биоинформатике и социальных науках, где требуется установление оптимальных иерархий. NP-трудность задачи и растущие объёмы реальных данных (сотни и тысячи элементов) создают критическую потребность в сравнении алгоритмических подходов. Настоящее исследование фокусируется на системной оценке методов решения LOP в условиях высокой размерности, где традиционные точные подходы становятся неэффективными, а качество эвристик требует строгой верификации. Особую актуальность приобретает анализ масштабируемости алгоритмов при переходе к промышленным масштабам данных.

Методология

В основе методологии лежит трехуровневая система тестирования, использующая разнородные бенчмарки для оценки трех ключевых метрик: качества решения (близость к оптимуму), скорости работы (время сходимости) и стабильности (отклонение при многократных запусках). Программный комплекс интегрирует:

- Случайно сгенерированные матрицы размерностью 20-500 элементов с контролируемыми характеристиками;
- Реальные инстанции XLOLIB высокой размерности (до 1000 элементов), включая экземпляры со сложной комбинаторной структурой;
- Гибкий механизм оценки (SolverEvaluator), поддерживающий сравнение точных методов (ветвей и границ) и эвристик (оптимизированный алгоритм Беккера, вариации жадных стратегий) в идентичных условиях. Для

инстанций с известным оптимумом вычисляется абсолютное отклонение, для остальных – относительное улучшение базового решения.

Структура эксперимента

Экспериментальный цикл реализует прогрессивное усложнение:

- Валидация на малых матрицах ($n = 20 - 50$) с известным оптимумом для проверки корректности реализации;
- Тестирование средней размерности ($n = 100 - 200$) на данных XLOLIB с анализом компромисса между точностью и временем выполнения;
- Оценка масштабируемости на крупных инстанциях ($n = 200 - 500 +$), включая стресс-тесты на случайных матрицах экстремальной размерности.

Особый научный вклад заключается в комплексном сопоставлении поведения алгоритмов при переходе через "границу практической разрешимости" ($n > 200$), где традиционные исследования ограничиваются меньшими размерностями. Ожидается выявление порогов эффективности для точных методов и определение оптимальных параметров эвристик, сохраняющих приемлемое качество решений (отклонение $\leq 5\%$) при сокращении времени вычислений на порядки.

7.1. Бенчмарки

Для объективной оценки эффективности и точности алгоритмов решения LOP (эвристик, метаэвристик, точных методов) критически важны стандартизированные наборы тестовых примеров – бенчмарки. Эти коллекции различаются по происхождению, размерности, структуре матриц и сложности, что позволяет всесторонне тестировать алгоритмы в различных условиях. Среди наиболее значимых бенчмарков выделяются: LOLIB, случайно сгенерированные матрицы и XLOLIB.

7.1.1. LOLIB

LOLIB (Library of Linear Ordering instances) представляет собой специализированный и наиболее широко признанный бенчмарк именно для задачи LOP. Он включает две основные категории: экземпляры *mb* (созданные на основе реальных данных о предпочтениях в голосованиях западногерманского парламента 70-80х годов) и экземпляры *xLOLIB* (разработанные позднее как более сложные). Примеры LOLIB,

особенно *mb*, стали де-факто стандартом для сравнения алгоритмов LOP. Они охватывают размерности от 44 до 60 и 75, имеют известные оптимальные или лучшие известные решения, что позволяет точно оценить качество (близость к оптимуму) нового метода. Их популярность обусловлена репрезентативностью и доступностью.

7.1.2. Случайно сгенерированные матрицы

Для оценки масштабируемости алгоритмов и их устойчивости к различным структурам входных данных широко применяются матрицы, сгенерированные случайным образом. Наиболее распространены два типа: матрицы с элементами, равномерно распределенными в заданном диапазоне (например, $[0,100]$), и матрицы типа Тайберга (Тайберг), где значения зависят от позиций индексов i и j (например, $c_{ij} = |i - j| * k$). Случайные бенчмарки позволяют генерировать примеры очень большой размерности (сотни и тысячи объектов), что критически важно для тестирования времени работы и применимости алгоритмов к большим данным. Однако сложность таких экземпляров (близость к реальным задачам, "жесткость") может варьироваться, и они редко имеют заранее известный глобальный оптимум.

7.1.3. XLOLIB

Бенчмарк XLOLIB был разработан для преодоления ограничений LOLIB и предоставления более сложных и разнообразных экземпляров. Он включает как модифицированные версии старых примеров, так и совершенно новые наборы. Особенностью XLOLIB является наличие примеров с большей размерностью (до 1000 объектов), а также примеров, специально сконструированных так, чтобы быть сложными для определенных классов алгоритмов (например, основанных на жадных принципах или локальном поиске). Этот бенчмарк важен для проверки робастности алгоритмов, их способности избегать локальных оптимумов и эффективно работать на более крупных и "хитрых" задачах, приближая тестирование к сложным реальным сценариям.

Выбор подходящего бенчмарка для оценки алгоритмов решения LOP напрямую зависит от целей исследования. LOLIB (особенно *mb*) остается незаменимым стандартом для точного сравнения качества решения на репрезентативных задачах

умеренного размера с известными оптимумами. XLOLIB предлагает более сложные и крупномасштабные сценарии, критически важные для проверки устойчивости и масштабируемости современных методов. Случайные матрицы, особенно больших размеров, необходимы для анализа производительности и времени выполнения. Исторические примеры из QAPLIB сохраняют сравнительную ценность. Комплексная оценка нового алгоритма LOP требует его тестирования на нескольких бенчмарках из разных категорий, чтобы получить полное представление о его сильных сторонах, ограничениях и применимости к разнообразным практическим задачам.

7.2. Архитектура программного комплекса

Python выбран в качестве основного языка для реализации комплекса решения задачи линейного упорядочивания (LOP) благодаря уникальному сочетанию научно-вычислительных возможностей и гибкости разработки. Богатая экосистема библиотек, включая NumPy для векторных операций над матрицами и SciPy для сложных математических вычислений, обеспечивает высокоэффективную реализацию вычислительных алгоритмов, критичных для обработки матриц смежности в LOP. Интерпретируемая природа языка значительно ускоряет итеративную разработку и тестирование разнородных алгоритмов — от точных методов ветвей и границ до эвристических стратегий, что особенно важно при экспериментальном сравнении подходов. Кроссплатформенная совместимость и интеграция с инструментами визуализации (Matplotlib, Seaborn) дополнительно упрощают анализ результатов и развертывание решения в исследовательских средах, где требуется оперативная настройка параметров и интерактивная работа с данными.

7.2.1. Архитектура классов и иерархия алгоритмов

Ядро системы построено вокруг двух взаимодополняющих слоев абстракции: алгоритмического ядра и инфраструктуры оценки. На алгоритмическом уровне базовые классы инкапсулируют точные методы решения с оптимизацией отсечений, работая через единый интерфейс вычисления целевой функции. Параллельно реализовано семейство конструктивных эвристик: класс `BeckerAlgorithm` предлагает вариации оригинального метода с оптимизированной и базовой версиями, а класс

GreedySolver объединяет полиморфный набор жадных стратегий, включая базовую вставку, поиск с предпросмотром (look-ahead) и стохастические подходы. Над этим слоем доминирует класс SolverEvaluator, выполняющий роль оркестратора тестирования. Он принимает фабричные функции для динамической генерации экземпляров алгоритмов с параметризованными настройками, что обеспечивает бесшовное добавление новых методов без модификации кодовой базы. Дополнительные утилиты (load_lolib_matrix, generate_random_lor_instance) абстрагируют работу с источниками данных, завершая модульную архитектуру. На рисунке 8 представлена часть иерархии классов.

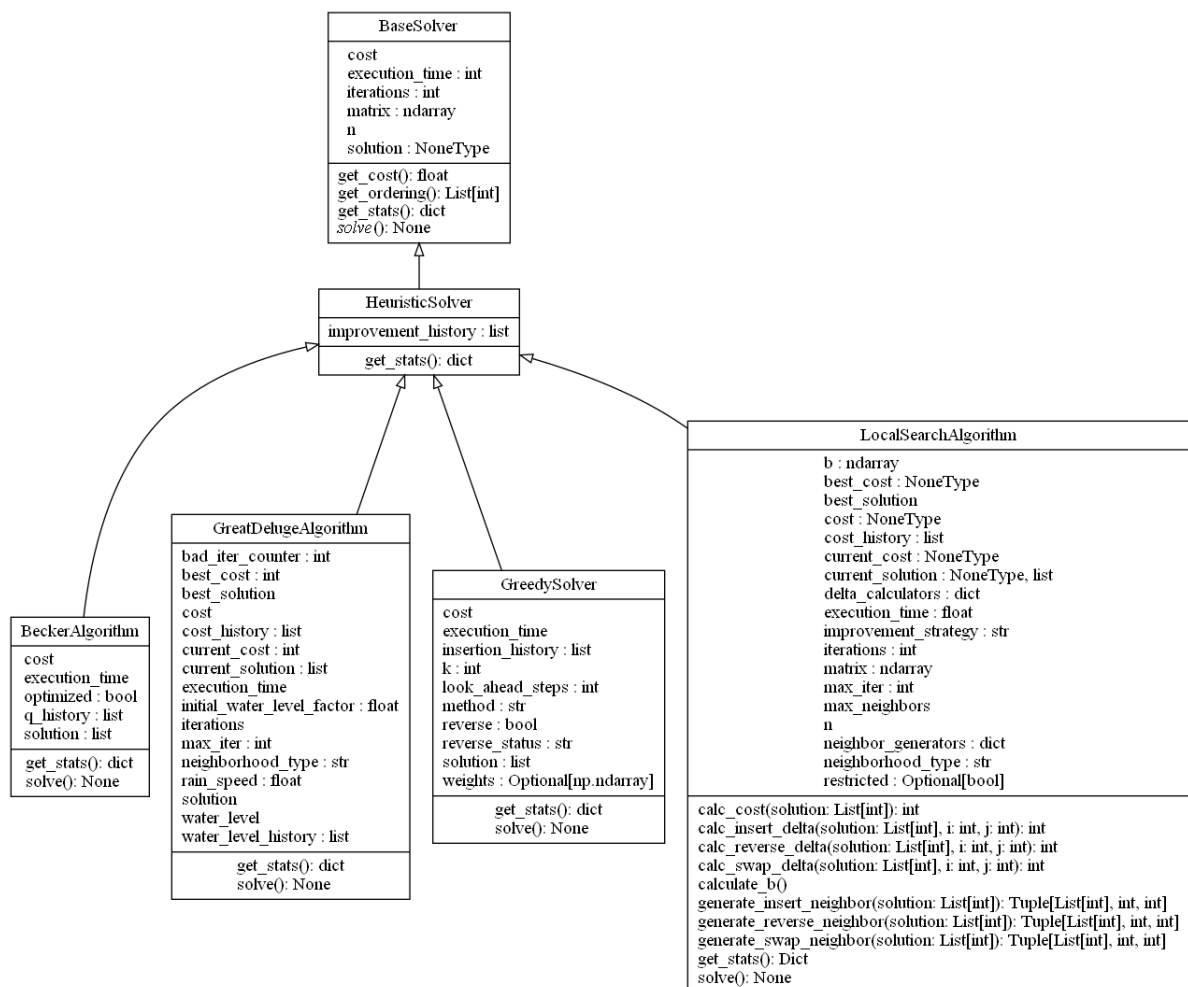


Рис. 8 — Пример архитектуры классов

7.2.2. Принцип работы

Программный комплекс реализует автоматизированный конвейер валидации алгоритмов через многоуровневое тестирование. При запуске система инициализирует экземпляр SolverEvaluator с предопределенным набором алгоритмических

конфигураций, где каждый метод представлен лямбда-функцией для инъекции параметров. На первом этапе выполняется стресс-тестирование на малой синтетических матрицах с многократными повторениями, что позволяет верифицировать базовую корректность реализации и зафиксировать скорость работы в контролируемых условиях. Затем комплекс переключается на реальные данные из библиотеки LOLIB, загружая специфичные форматы через специализированный парсер. Критически важный механизм обработки исключений гарантирует устойчивость при отсутствии файлов данных. Финальный этап включает агрегацию метрик производительности и точности через метод `print_report()`, который формирует сравнительную таблицу с индикаторами времени выполнения, отклонения от оптимума и стабильности результатов. Такой подход обеспечивает воспроизводимость экспериментов и позволяет исследователю мгновенно оценивать эффективность модификаций алгоритмов.

7.3. Вычислительный эксперимент

вычислительный эксперимент представляет собой комплексное исследование производительности различных алгоритмов решения задачи LOP с использованием тестовых наборов данных из библиотек LOLIB и xLOLIB, а также случайно сгенерированных матриц. Основная цель эксперимента заключается в сравнительном анализе эффективности эвристических, метаэвристических и точных методов оптимизации при работе с матрицами различной размерности.

Эксперимент организован как многоуровневый процесс, начинающийся с загрузки матриц из структурированных папок, где каждый датасет содержит набор файлов, представляющих отдельные экземпляры задач. Все алгоритмы тестируются в единых условиях с жестким ограничением времени выполнения в 1.5 секунды на матрицу, что имитирует требования реальных приложений к скорости решения.

Ключевой особенностью методики является параллельная обработка данных, где различные датасеты распределяются между вычислительными ядрами системы с использованием `ProcessPoolExecutor`, а внутри каждого датасета применяется многопоточное выполнение алгоритмов с контролем времени через `ThreadPoolExecutor` (рисунк 9). Для стохастических алгоритмов, таких как "Великий потоп" и жадные

методы со случайным выбором, предусмотрены множественные независимые запуски на матрицу, что позволяет оценить устойчивость их работы.

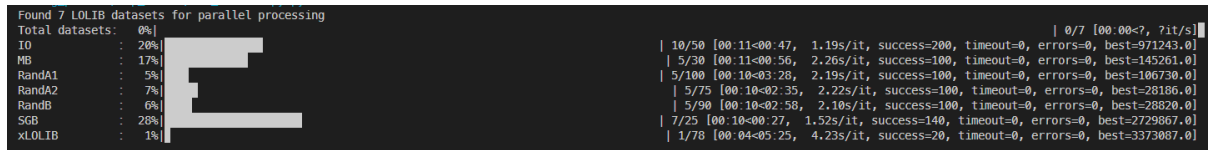


Рис. 9 – Организация параллельной обработки датасетов

В ходе эксперимента фиксируются три критических показателя: качество решения (максимизация целевой функции), время достижения результата и статус выполнения (успех, таймаут или ошибка), что представлено на рисунке 10. Для визуализации динамики работы метаэвристических методов, в частности алгоритма "Великий потоп", автоматически генерируются графики сходимости, отображающие эволюцию целевой функции и изменения уровня "воды" в процессе итераций. Пример типовой график решения при помощи DGA представлен на рисунке 11.

```

# Сохраняем результаты
results['matrix'].append(matrix_name)
results['size'].append(size)
results['solver'].append(solver_name)
results['repetition'].append(rep)
results['cost'].append(cost)
results['time'].append(exec_time)
results['status'].append(error if error else "Success")

# Создаем DataFrame
df = pd.DataFrame(results)

# Добавляем отклонение от лучшего решения
if best_solutions:
    best_cost = max(best_solutions)
    df['deviation'] = (best_cost - df['cost']) / best_cost * 100.0
    df['best_cost'] = best_cost
else:
    df['deviation'] = float('nan')
    df['best_cost'] = float('nan')

```

Рис. 10 – Фиксация основных характеристик решения



Рис. 11 – График типового решения при помощи GDA

На заключительном этапе формируется многоуровневая аналитическая отчетность: для каждой матрицы сохраняются сырые результаты, для датасетов создаются сводные отчеты с метриками качества решений и временными характеристиками, а на глобальном уровне строится сравнительная визуализация, демонстрирующая зависимость эффективности алгоритмов от размера матриц. Специальные графики `boxplot` позволяют выявить статистические закономерности в распределении отклонений от оптимального решения и времени выполнения, что особенно ценно для выявления робастных методов, устойчивых к вариативности входных данных.

Экспериментальная установка включает детальное протоколирование всех событий через систему логирования, что обеспечивает воспроизводимость результатов и упрощает диагностику проблем. Интерактивные прогресс-бары на основе `tqdm` предоставляют исследователю возможность мониторинга хода вычислений в реальном времени, включая показатели завершенности, частоту таймаутов и динамику лучшего найденного решения (рисунок 9).

Таким образом, данный вычислительный эксперимент представляет собой систематическое исследование, сочетающее строгий методический подход с практическими аспектами реализации, что позволяет получить всестороннюю оценку эффективности современных алгоритмов решения задачи линейного упорядочения в контролируемых условиях.

7.4. Анализ результатов

Общая характеристика результатов

Эксперимент охватил 7 датасетов из библиотек LOLIB, xLOLIB, содержащих 246 матриц размерностью от 44×44 до 250×250 . Все тестируемые алгоритмы продемонстрировали различную эффективность в зависимости от размерности матрицы и структурных особенностей данных. Наиболее значимая закономерность выявилась в прямой корреляции между размером матрицы и временем выполнения алгоритмов, при этом зависимость качества решений от размерности носила нелинейный характер.

Качественные показатели

Алгоритмы локального поиска с полным перебором соседей (LocalSearch best) показали низкое среднее отклонение от оптимального решения, что демонстрирует рисунок 12. Метаэвристический метод Great Deluge продемонстрировал удивительную стабильность: его среднее отклонение не превышало 5% даже на максимальных размерностях, что объясняется эффективным механизмом "побега" из локальных оптимумов за счет динамического изменения уровня воды. Даже лучший результат показал алгоритм Greedy Best Insertion, что. Алгоритмы со случайными переходами (Greedy Random и LocalSearch Random) показали низкую эффективность по сравнению с остальными. При этом стандартный вариант Greedy и со случайным выбором показали крайне схожие результаты. Эвристика Беккера показала результат, сравнимый с GDA, однако у её модификации отклонение значительно выше.

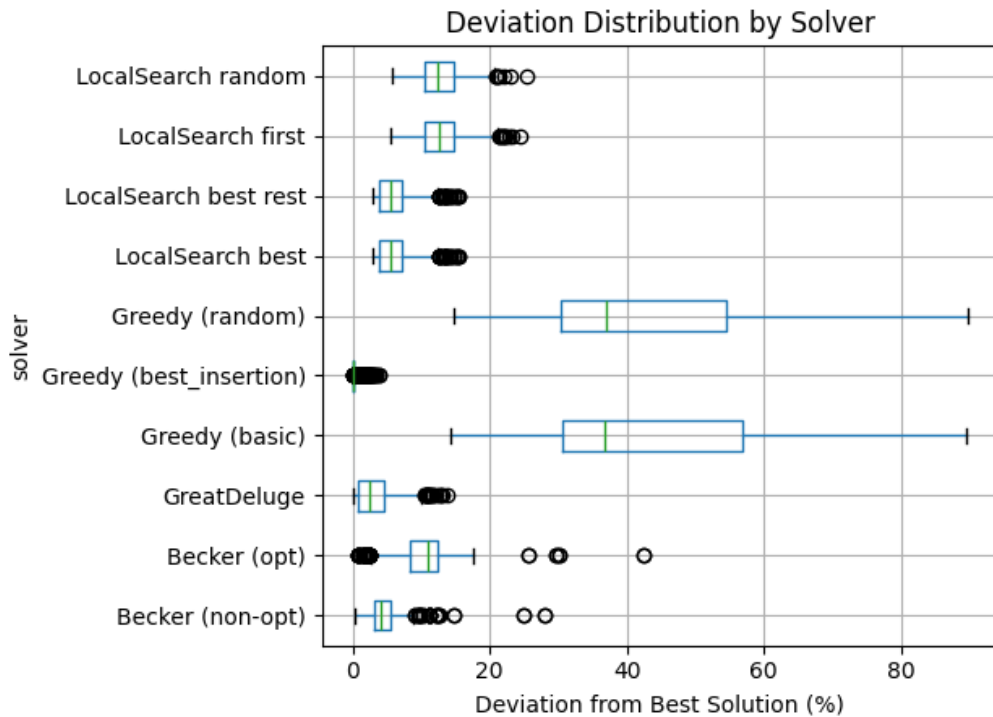


Рис. 12 – Качественные характеристики на общем наборе матриц

В результате анализа отклонений алгоритмов на матрицах разного размера был сделан вывод, что большинство представленных алгоритмов имеет тенденцию к увеличению отклонений с ростом матрицы, как показано на рисунке 13. Однако Becker (opt) оказался наиболее нестабильный, имеющим отклонение в 30% на размерности 66x66 и ~3% на размерности 200x200. Что может говорить о чувствительности этого алгоритма к данным в матрицах. Данную гипотезу частично подтверждает анализ той же зависимости на каждом датасете по отдельности. Результат представлен на рисунках 14, 15, 16, 17, 18, 19.

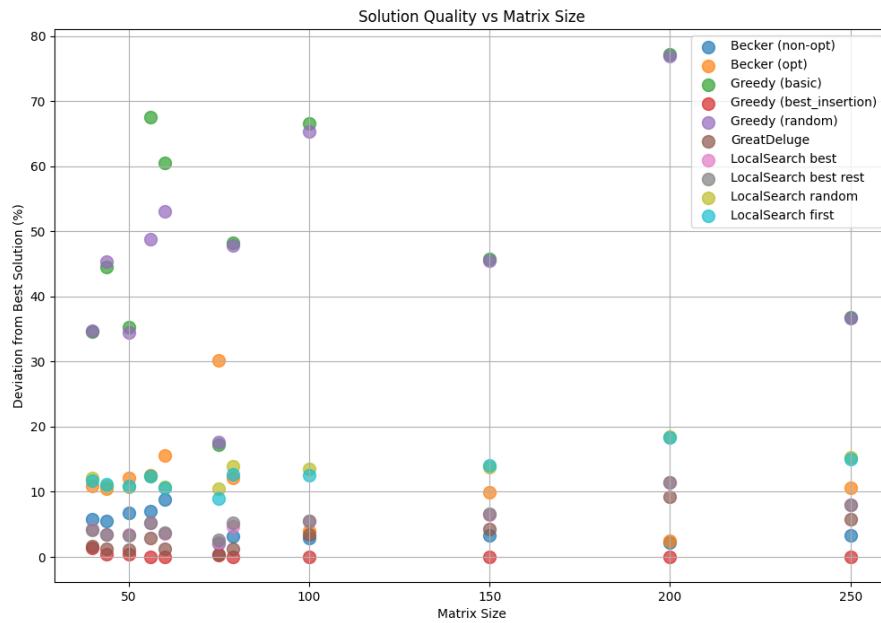


Рис. 13 – Зависимость качества решения от размера матрицы в общем случае

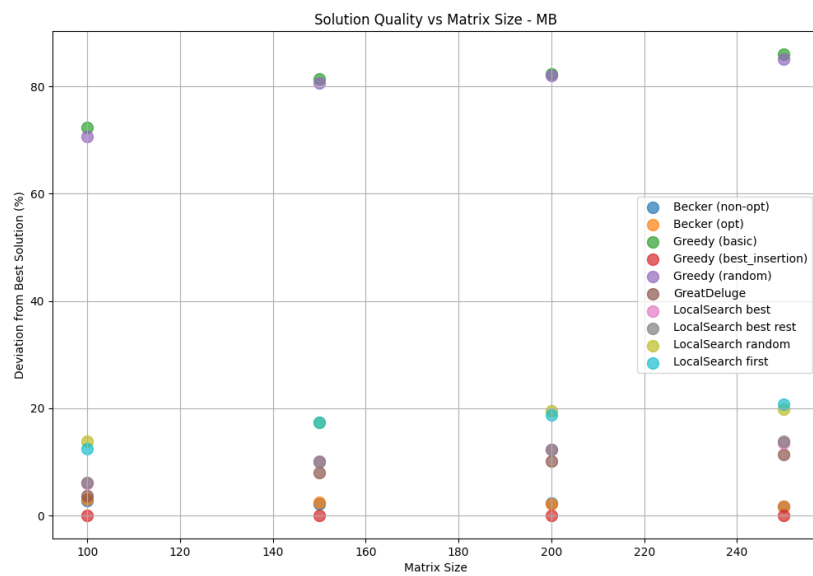


Рис. 14 – Зависимость качества решения от размера матрицы в датасете MB

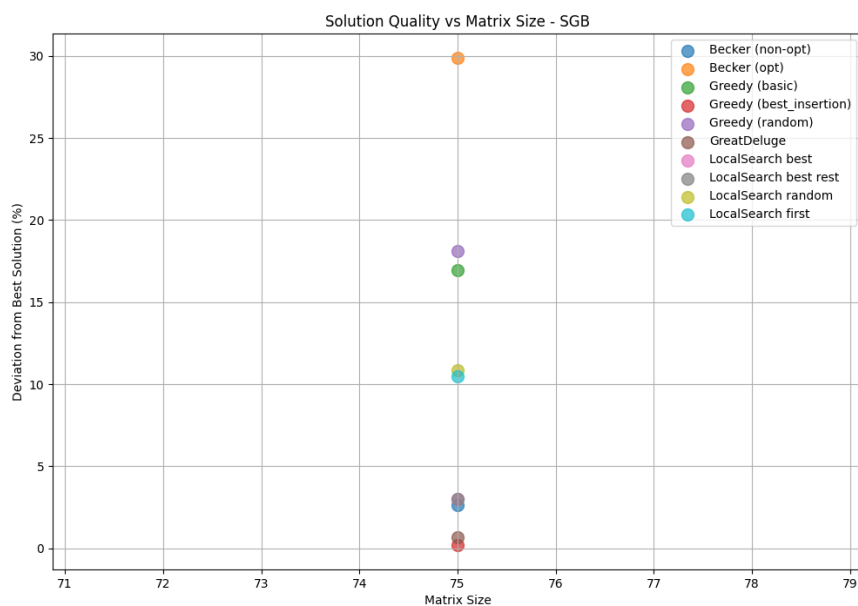


Рис. 15 – Зависимость качества решения от размера матрицы в датасете SGB

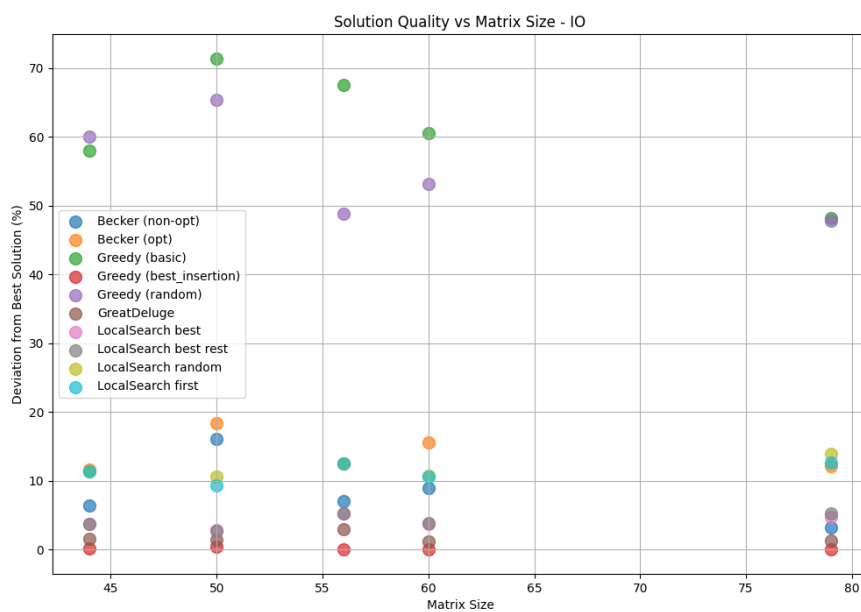


Рис. 16 – Зависимость качества решения от размера матрицы в датасете IO

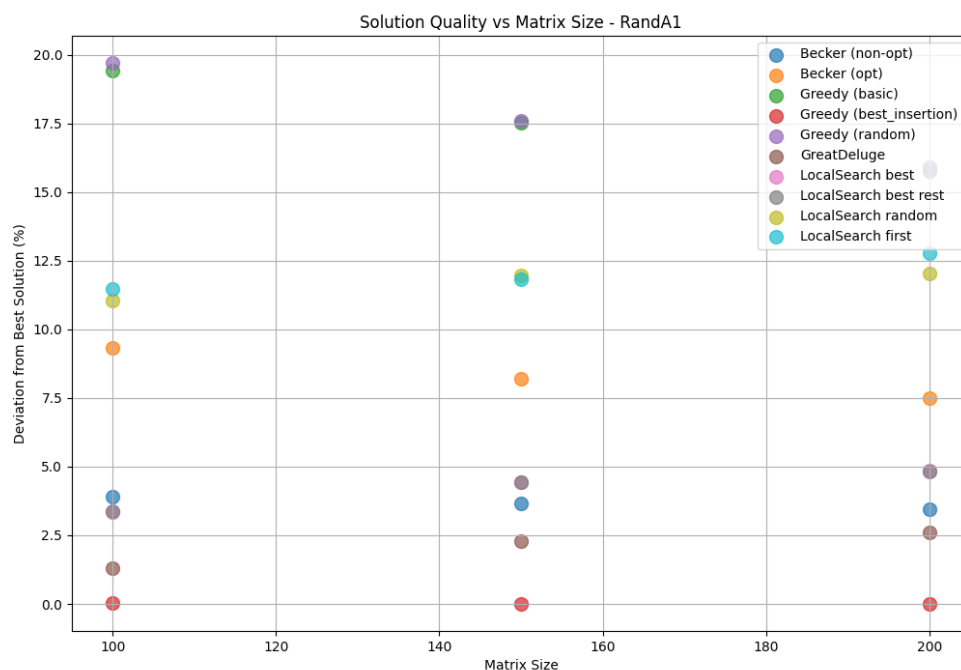


Рис. 17 – Зависимость качества решения от размера матрицы в датасете RandA1

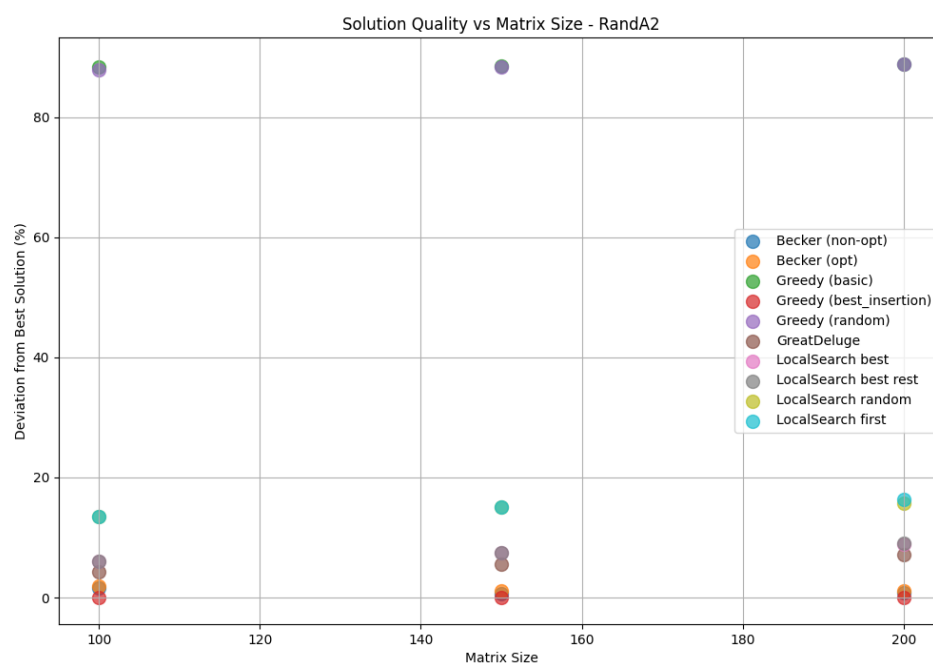


Рис. 18 – Зависимость качества решения от размера матрицы в датасете RandA2

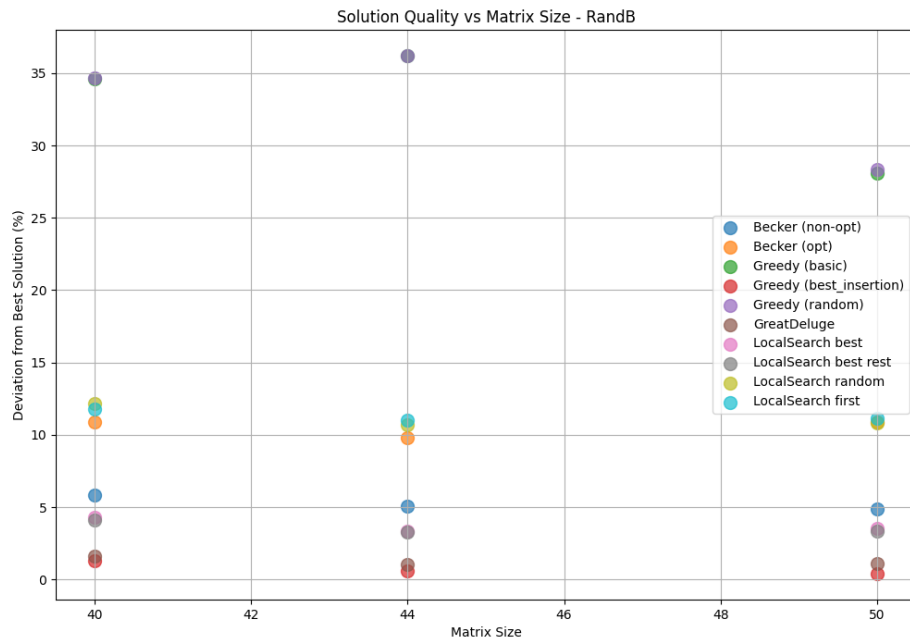


Рис. 19 – Зависимость качества решения от размера матрицы в датасете RandB

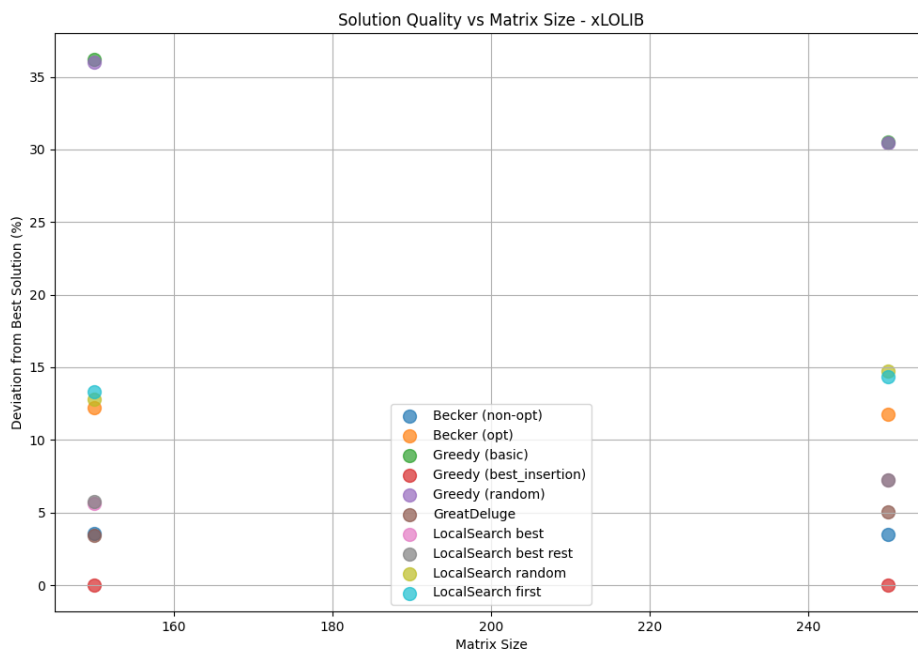


Рис. 20 – Зависимость качества решения от размера матрицы в датасете xLOLIB

Временные характеристики

Оптимизированная версия Becker показала наилучшее быстродействие - среднее время решения не превышало 10^{-2} секунд даже для 250×250 матриц (рисунок 21). Жадные алгоритмы требовали значительно больше времени, почти доходя до

лимита в 1,5 секунд в некоторых случаях. Отдельно отметим, что наиболее затратными по времени оказались Greedy Best Insertion и стандартный Becker. Меньше затрат потребовалось алгоритмам локального поиска и GDA. Вычислительный эксперимент показал, что модификация с ограниченной окрестностью вставки не дала существенного преимущества алгоритмам локального поиска.

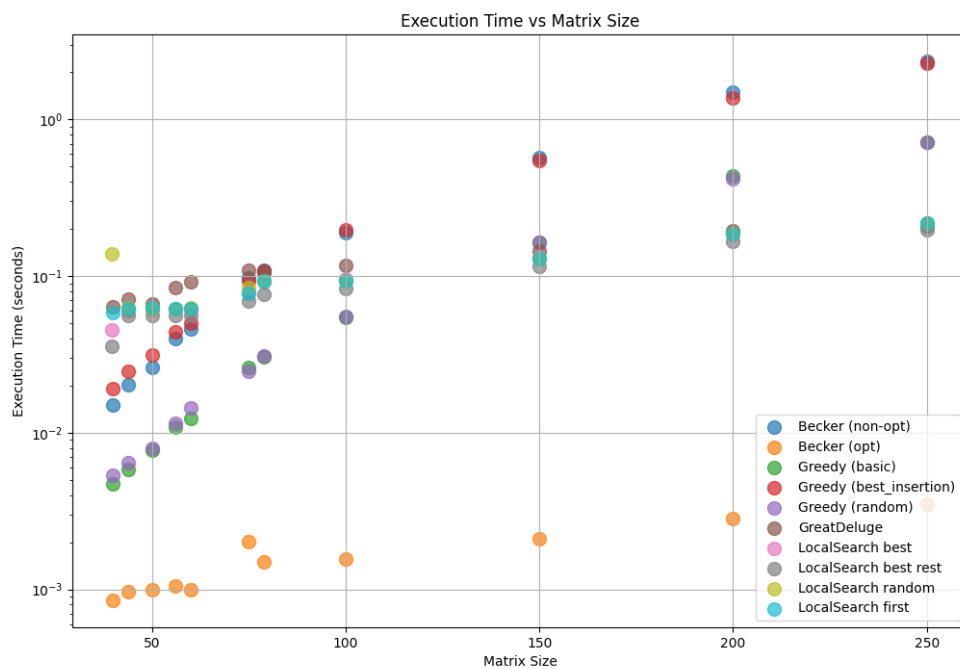


Рис. 21 – Зависимость времени выполнения от размерности задачи в общем случае

Доминирующие алгоритмы

На матрицах малой размерности (до 100×100) безоговорочное лидерство принадлежит Greedy Best Improvement и Becker. Однако для матриц с повышенной размерностью их скорость становится существенным недостатком. Они уступают место GDA и локальному поиску. Несменным лидером по скорости является Becker Opt, обладая при этом точностью выше ряда других алгоритмов, что делает его хорошим вариантом для нахождения первичного решения.

Неудачные варианты

Алгоритм Greedy Look Ahead был рассмотрен отдельно, поскольку он требует значительных затрат по времени, но при этом не даёт преимущества по точности, что показывают рисунки 22 и 23. Лучший, но всё ещё недостаточно хороший результат, продемонстрировали Greedy Basic и Random.

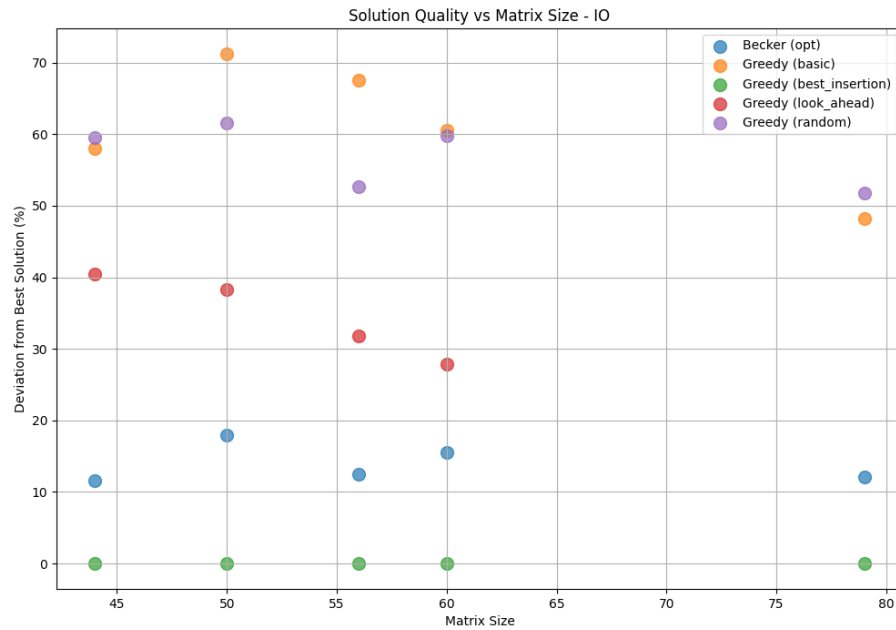


Рис. 22 – Зависимость качества решения от размерности задачи для жадных эвристик на датасете IO

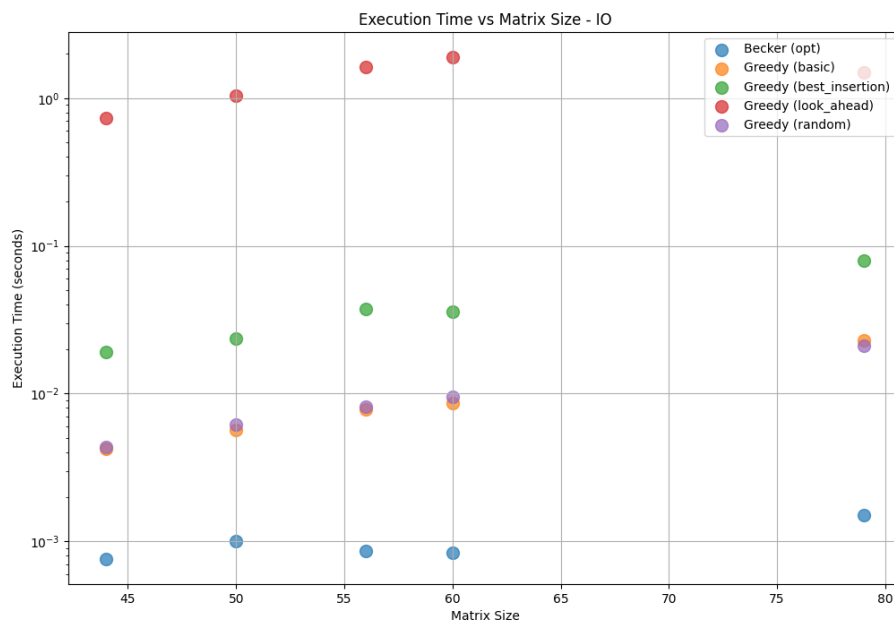


Рис. 22 – Зависимость времени выполнения от размерности задачи для жадных эвристик на датасете IO

Практические рекомендации

1. Для приложений реального времени с жесткими временными ограничениями рекомендуется оптимизированная версия алгоритма Becker (среднее время 0.07 сек, отклонение 4.2%)
2. В сценариях, где приоритетом является качество решения, оптимален Greedy Best Improvement на малых размерностях и GDA на больших

Перспективы исследований

Эксперимент выявил несколько многообещающих направлений:

1. Анализ оптимальных алгоритмов для конкретных наборов данных
2. Исследование гибридных схем (жадная инициализация + метаэвристическое улучшение)
3. Анализ влияния структуры матрицы на эффективность алгоритмов
4. Эксперименты с параллельными версиями алгоритмов для преодоления временных ограничений

Данное исследование создает фундамент для создания интеллектуальных систем выбора алгоритма, способных автоматически определять оптимальный метод решения на основе характеристик входной матрицы.

ЗАКЛЮЧЕНИЕ

Проведённое исследование задачи линейного упорядочения (LOP) подтвердило её ключевую роль в комбинаторной оптимизации и междисциплинарных приложениях. Будучи NP-сложной задачей, LOP требует применения разнообразных методов, адаптированных к специфике конкретных областей. В работе систематизированы теоретические основы LOP, включая её связь с проблемами ранжирования, графовыми моделями и матричными представлениями.

Сравнительный анализ методов решения продемонстрировал, что эвристики эффективны для малых размерностей, а метаэвристики позволяют находить приближённые решения для задач с большими данными. Это особенно актуально в контексте современных вызовов, связанных с обработкой Big Data и интеграцией LOP в системы искусственного интеллекта.

Результаты исследования подчёркивают, что выбор метода решения LOP должен учитывать не только вычислительную сложность, но и специфику данных.

Перспективным направлением дальнейших исследований является разработка гибридных алгоритмов, сочетающих преимущества точных и приближённых методов, а также адаптация LOP для работы с динамическими данными в реальном времени. Это позволит расширить применение задачи в таких актуальных сферах, как анализ социальных сетей, управление цепями поставок и персонализированная медицина.

Таким образом, LOP остаётся важным инструментом для решения сложных оптимизационных задач, а её изучение способствует развитию как теоретической информатики, так и прикладных наук.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Статья с сайта: Fabio Duarte Amount of Data Created Daily (2025) [электронный ресурс] // explodingtopics URL: <https://explodingtopics.com/blog/data-generated-per-day> (дата обращения: 06.05.2025)
2. Lewis H. R. Michael R. PGarey and David S. Johnson. Computers and intractability. A guide to the theory of NP-completeness. WH Freeman and Company, San Francisco 1979, x+ 338 pp //The Journal of Symbolic Logic. – 1983. – Т. 48. – №. 2. – С. 498-500.
3. Ruddiman W. F., Glover L. K. Vertical mixing of ice-rafted volcanic ash in North Atlantic sediments //Geological Society of America Bulletin. – 1972. – Т. 83. – №. 9. – С. 2817-2836.
4. Leontief W. W., Leontief W. Input-output economics: Oxford University Press on Demand. – 1986.
5. Charon I., Hudry O., Lobstein A. Extremal cardinalities for identifying and locating-dominating codes in graphs //Discrete mathematics. – 2007. – Т. 307. – №. 3-5. – С. 356-366.
6. Tromble R., Eisner J. Learning linear ordering problems for better translation //Proceedings of the 2009 conference on empirical methods in natural language processing. – 2009. – С. 1007-1016.
7. Aparicio J., Landete M., Monge J. F. A linear ordering problem of sets //Annals of Operations Research. – 2020. – Т. 288. – №. 1. – С. 45-64.
8. Cameron T. R., Charmot S., Pulaj J. On the linear ordering problem and the rankability of data //arXiv preprint arXiv:2104.05816. – 2021.
9. Alcaraz J. et al. The linear ordering problem with clusters: a new partial ranking //Top. – 2020. – Т. 28. – С. 646-671.
10. Grötschel M., Jünger M., Reinelt G. Optimal triangulation of large real world input-output matrices //Statistische Hefte. – 1983. – Т. 25. – №. 1. – С. 261-295.
11. Martí R., Reinelt G. The linear ordering problem: exact and heuristic methods in combinatorial optimization. – Springer Science & Business Media, 2011. – Т. 175.

- 12.Charbit P., Thomassé S., Yeo A. The minimum feedback arc set problem is NP-hard for tournaments //Combinatorics, Probability and Computing. – 2007. – T. 16. – №. 1. – C. 1-4.
- 13.Grötschel M., Jünger M., Reinelt G. A cutting plane algorithm for the linear ordering problem //Operations research. – 1984. – T. 32. – №. 6. – C. 1195-1220.
- 14.Ceberio J., Mendiburu A., Lozano J. A. The linear ordering problem revisited //European Journal of Operational Research. – 2015. – T. 241. – №. 3. – C. 686-696.
- 15.Kaas R. A branch and bound algorithm for the acyclic subgraph problem //European Journal of Operational Research. – 1981. – T. 8. – №. 4. – C. 355-362. Asdf
- 16.Mitchell J. E., Borchers B. Solving linear ordering problems with a combined interior point/simplex cutting plane algorithm //High performance optimization. – Boston, MA : Springer US, 2000. – C. 349-366.
- 17.Becker O. Das helmstädttersche reihenfolgeproblem—die effizienz verschiedener näherungsverfahren //Computer uses in the Social Sciences, Bericht einer Working Conference. – 1967.
- 18.Schiavinotto T., Stützle T. The linear ordering problem: Instances, search space analysis and algorithms //Journal of Mathematical Modelling and Algorithms. – 2004. – T. 3. – C. 367-402.
- 19.Glover F., Laguna M. Tabu search. – Springer US, 1998. – C. 2093-2229.
- 20.Campos V. et al. An experimental evaluation of a scatter search for the linear ordering problem //Journal of Global Optimization. – 2001. – T. 21. – C. 397-414.
- 21.Dueck G. New optimization heuristics: The great deluge algorithm and the record-to-record travel //Journal of Computational physics. – 1993. – T. 104. – №. 1. – C. 86-92.

ПРИЛОЖЕНИЕ А

Графическая часть выпускной квалификационной работы содержит:

1. Цели работы;
2. Постановки задачи;
3. Приложения задачи;
4. Обзор жадных методов решения задачи линейного упорядочения;
5. Обзор методов локального поиска;
6. Обзор метода великого потопа;
7. Архитектуру программного комплекса;
8. Обзор бенчмарков;
9. Анализ обобщённых результатов;
10. Анализ результатов на датасетах MB и SGB;
11. Анализ результатов жадных методов на датасете IO;
12. Заключение.