

Оптимизация, системный анализ и исследование операций

© 2025 г. Ю.Ф. ЛЕОНОВА (yuliya.igosheva@gmail.com),
А.В. ПАНЮКОВ, д-р физ.-мат. наук (paniukovav@susu.ru)
(Южно-Уральский государственный университет, Челябинск)

СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ КАЧЕСТВА АЛГОРИТМА СОЕДИНЕНИЯ ЦИКЛОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КОММИВОЯЖЕРА НА МИНИМУМ

Задача коммивояжера является одной из наиболее изученных в комбинаторной оптимизации, однако исследование новых подходов и улучшение существующих методов остается актуальной задачей. В данной работе проведен анализ качества алгоритма соединения циклов для задачи коммивояжера на минимум. Представлены результаты вычислительного эксперимента на пяти семействах задач, проанализированы точность и временная сложность алгоритма. Для симметричных экземпляров задачи построена регрессионная модель, описывающая зависимость оценки относительной погрешности от числа вершин. Показано, что полиномиальная модель наилучшим образом аппроксимирует полученные данные и удовлетворяет основным статистическим предположкам. Полученные результаты позволяют оценить характер роста ошибки и обосновать применимость алгоритма к экземплярам задачи коммивояжера большой размерности.

Ключевые слова: задача коммивояжера, эвристический алгоритм, асимптотическая точность алгоритма, статистическое исследование.

DOI: 10.31857/S0005231025050065, **EDN:** AXACAX

1. Введение

Задача коммивояжера является классической задачей комбинаторной оптимизации и при этом остается одной из наиболее важных и актуальных в области прикладной математики и информатики. Ее значимость проявляется во множестве практических сфер, таких как логистика, транспорт, производство, планирование маршрутов и др. Задача коммивояжера на минимум состоит в поиске гамильтонова цикла с минимальной суммой весов ребер в полном графе.

В последние десятилетия широко исследовались различные методы решения задачи коммивояжера, включая точные алгоритмы, аппроксимационные и эвристические методы [1, 2]. Однако с появлением новых вычислительных технологий и постоянным увеличением объема данных возникает необходи-

мость в разработке более эффективных и масштабируемых алгоритмов для решения этой проблемы.

Задача коммивояжера является NP-трудной [3]. Поэтому для ее решения актуальной является разработка полиномиальных приближенных алгоритмов.

В данной работе представлено эмпирическое исследование и статистический анализ точности и временной сложности эвристического алгоритма для решения задачи коммивояжера – алгоритма соединения циклов. В разделе 2 приведены характеристики оценок точности и времени работы популярных эффективных алгоритмов для решения задачи коммивояжера на минимум. В разделе 3 дается краткое описание исследуемого алгоритма. В разделе 4 описана генерация задач и проведение вычислительного эксперимента. Результаты вычислительного эксперимента на случайных входных данных представлены в разделе 5: приведены данные по точности и времени решения задач. В разделе 6 приводится статистический анализ точности алгоритма соединения циклов. В разделе 7 сформулированы основные выводы по результатам экспериментов.

2. Обзор состояния проблемы

Эвристические и приближенные алгоритмы, такие как алгоритм ближайшего соседа, жадный алгоритм, ближайшая вставка и другие, широко используются в практике благодаря их простоте и эффективности. Они часто

Таблица 1. Верхняя граница оценки точности и время работы алгоритмов

Алгоритм	Верхняя граница	Время работы
Ближайший сосед [5]	$0,5 \lceil \log_2 N + 1 \rceil$	$O(N^2)$
Ближайший сосед с двойным концом [6]		$O(N^2) \log N$
Жадный [5]		$O(N^2)$
Ближайшее сложение [6]	$2 - 2/N$	$O(N^2)$
Ближайшая вставка [6, 7]		$O(N^2) \log N$
Самая дешевая вставка [6, 7]		$O(N^2)$
Самая дальняя вставка [6, 7]		
Произвольная вставка [7]		
Вставка ближайшего сегмента		
Двойное минимальное пространственное дерево [8]		$O(N^3)$
Двойное минимальное пространственное дерево модифицированное		
Кристофидес [9]	$3/2 - 1/N$	$O(N^3)$
Кривая Мура [10]	$\log N$	$O(N \log N)$
Кривая Серпинского [11]		
2-Opt [12]	≈ 2	$O(N^2)$

демонстрируют хорошие результаты для небольших наборов данных, но могут не обеспечивать оптимальное решение для задач большой размерности.

В табл. 1 представлены априорные оценки точности и времени некоторых эвристических алгоритмов [4] для задачи коммивояжера на минимум. Верхние границы для алгоритмов рассчитаны как отношение $f(s)/f(s_O)$, где $f(s)$ – полученная длина тура, а $f(s_O)$ – оптимальная длина тура.

Приведенные алгоритмы для решения задачи коммивояжера имеют различные преимущества и недостатки в зависимости от характеристик входных данных и требуемой точности результата.

В последующих разделах будет представлен анализ точности и временной сложности алгоритма соединения циклов.

3. Алгоритм соединения циклов для решения задачи коммивояжера

Авторами статьи [13] был исследован эвристический алгоритм для решения задачи коммивояжера – алгоритм соединения циклов (Cycles Merging Algorithm – CMA). Приведем краткое описание алгоритма.

Первый шаг алгоритма состоит в нахождении в заданном графе 2-регулярного суграфа экстремального веса, т.е. покрытия этого графа циклами. Данную конструкцию принято называть 2-фактором минимального веса. Задача поиска 2-фактора минимального веса может быть сведена к задаче о назначении, для которой известны точные полиномиальные алгоритмы.

На втором шаге проверяется условие единственности цикла полученного решения. Если 2-фактор представлен единственным циклом, то этот цикл является решением задачи. Если 2-фактор представлен несколькими циклами, то перебираются попарно различные циклы r и t , в каждом цикле выбранной пары перебираются по одному ребру.

Пусть выбраны $e_{\{r \oplus t\}}^r = [v_1, v_2]$ и $e_{\{r \oplus t\}}^t = [u_1, u_2]$, для них находятся две пары сопряженных ребер ($f = [v_1, u_1], g = [v_2, u_2]$) и ($f = [v_1, u_2], g = [v_2, u_1]$), используемых для соединения циклов (рис. 1). Выполняется поиск набора

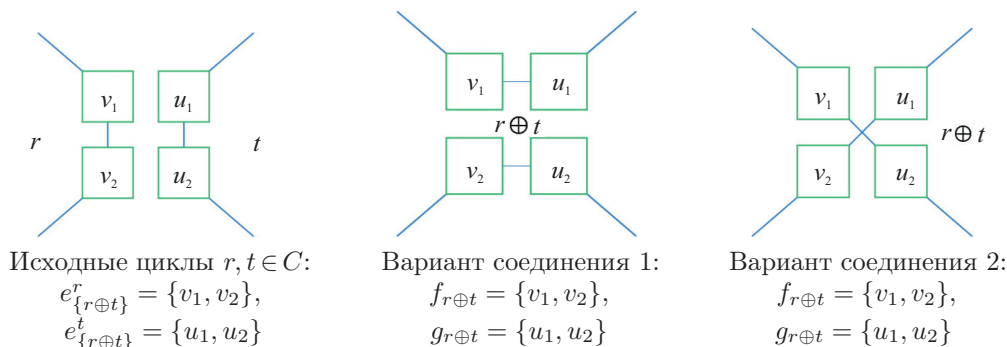


Рис. 1. Варианты соединения циклов.

$r, t, e_{\{r \oplus t\}}^r, e_{\{r \oplus t\}}^t, f, g$ из перечисленных элементов такой, что соединенный цикл имеет минимальный вес.

Затем найденная пара циклов заменяется объединенным циклом с минимальной стоимостью. Алгоритм завершает работу, когда текущий 2-фактор содержит один цикл.

Для описанного алгоритма сформулирована и доказана теорема о вычислительной сложности алгоритма.

Теорема 1. Вычислительная сложность алгоритма соединения циклов не превосходит величины $O(|V|^3)$.

Более подробное описание алгоритма соединения циклов и доказательство теоремы о вычислительной сложности приведены в [13].

Для метрической задачи коммивояжера на максимум для данного алгоритма ранее была установлена теоретическая граница точности [13], что позволяет строго оценивать его качество в этом классе задач. Однако для задачи коммивояжера на минимум такие теоретические границы отсутствуют, поэтому представляет интерес проведение эмпирического исследования точности алгоритма.

4. Описание тестовых задач

Для оценки качества алгоритма соединения циклов для задачи коммивояжера на минимум вычислительный эксперимент проводился на наборах задач из библиотеки TSPLIB [14], а также на случайно сгенерированных экземплярах с различными характеристиками матрицы стоимостей.

Оценка эффективности алгоритма проводилась на следующих семействах задачи коммивояжера:

1. Асимметричные экземпляры задачи коммивояжера из TSPLIB. TSPLIB содержит реальные и искусственно сгенерированные задачи, широко используемые в тестировании эвристических и точных алгоритмов. Асимметричность стоимости переездов между вершинами делает решение более сложным, так как нарушается свойство симметрии, присущее многим классическим методам.

2. Евклидовы экземпляры задачи коммивояжера из TSPLIB ($n \leq 3000$). Этот тип данных моделирует геометрические версии задачи, где вершины соответствуют точкам на плоскости, а стоимости переходов между ними определяются евклидовым расстоянием. Такие задачи широко встречаются в логистике, навигации и сетевом планировании.

3. Асимметричные матрицы стоимости $W = [w(i, j)]$, где $w(i, j)$ – случайные числа из $\{0, 1, 2, \dots, 10^5\}$. Равномерное распределение значений в фиксированном интервале моделирует случайные системы с независимыми весами переходов. Это позволяет протестировать алгоритм в условиях полной неопределенности структуры стоимости.

4. Асимметричные матрицы стоимости, где $w(i, j)$ – случайные числа из $\{0, 1, 2, \dots, i \times j\}$. В данном семействе диапазон возможных стоимостей увеличивается по мере роста индексов вершин, что моделирует задачи, где связь между вершинами становится дороже или сложнее по мере увеличения их номеров. Такой вид матрицы позволяет проанализировать поведение алгоритма на задачах с неравномерной сложностью связей.

5. Симметричные матрицы стоимости W , где $w(i, j)$ – случайные числа из $\{0, 1, 2, \dots, 10^5\}$ для $i < j$. В отличие от асимметричных случаев, симметричная структура ограничивает множество возможных решений, но при этом приводит к увеличению числа коротких циклов в исходном покрытии. Это затрудняет процесс объединения циклов в один маршрут и может негативно влиять на точность эвристического алгоритма.

6. Симметричные матрицы, где $w(i, j)$ – случайные числа из $\{0, 1, 2, \dots, i \times j\}$ для $i < j$. Аналогично асимметричному случаю, но с учетом симметрии весов. Такой тип задачи моделирует случаи, где удаленные вершины связаны с более высокой стоимостью переходов. Однако наличие множества коротких циклов на ранних этапах решения делает задачу вычислительно сложной и может ухудшать точность приближенного решения.

7. Экземпляры задачи на наклонной плоскости, где вес дуги определяется как

$$w(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} - \max(0, y_i - y_j) + 2 \max(0, y_j - y_i),$$

причем координаты вершин генерировались независимо и равномерно на интервале $\{0, 1, 2, \dots, 10^5\}$. Эта модель имитирует ситуацию, когда перемещение по вертикали сопряжено с дополнительными затратами, например при передвижении по горным дорогам. Разница в стоимости движения вверх и вниз делает задачу асимметричной и усложняет ее решение.

Для семейств 3–6 количество вершин n варьировалось от 100 до 3000 с шагом 100. Все результаты являются средними по 100 испытаниям каждое.

В эксперименте используются экземпляры, созданные как случайным, так и детерминированным образом. Количество и разнообразие используемых семейств позволяет проверить надежность тестируемого алгоритма.

5. Результаты вычислительного эксперимента для задачи коммивояжера на минимум

Для первого и второго семейства экземпляров задачи (асимметричные и евклидовы задачи из TSPLIB) результаты приведены в [15]. Анализ точности и вычислительной эффективности алгоритма соединения циклов для остальных семейств рассматриваются в рамках данной работы.

5.1. Анализ оценки относительной погрешности

На рис. 2 приведены оценки относительной погрешности, полученные в ходе вычислительного эксперимента для различных семейств задач коммивояжера. Элементы матрицы стоимостей этих задач были сгенерированы случайным образом в соответствии с правилами генерации, описанными в разделе 4.

Анализ результатов показывает различия в поведении оценки относительной погрешности в зависимости от семейства рассматриваемых экземпляров.

Для асимметричных экземпляров задачи коммивояжера с равномерно распределенными элементами матрицы весов из множества $\{0, 1, 2, \dots, 10^5\}$ (семейство 3) относительная погрешность уменьшается с ростом числа вершин. При $n = 100$ ее среднее значение составляет 9,15%, а при $n = 3000$ снижается до 1,97%. Аналогичная тенденция наблюдается и для семейства 4, однако величина погрешности в среднем остается ниже, уменьшаясь с 5,12% при $n = 100$ до 1,06% при $n = 3000$.

Для симметричных экземпляров задачи коммивояжера, представленных семействами 5 и 6, наблюдается иное поведение. Относительная погрешность в обоих случаях возрастает с увеличением числа вершин. Так, для семейства 5 она увеличивается с 103% при $n = 100$ до 541% при $n = 3000$. Аналогичный рост характерен и для семейства 6, с увеличением погрешности с 65 до 175% при тех же значениях n .

Для экземпляров наклонных плоскостей (семейство 7) оценка относительной погрешности остается относительно стабильной на всем интервале значений n , колеблясь в пределах 36,6–37,7%. Данный результат свидетельствует

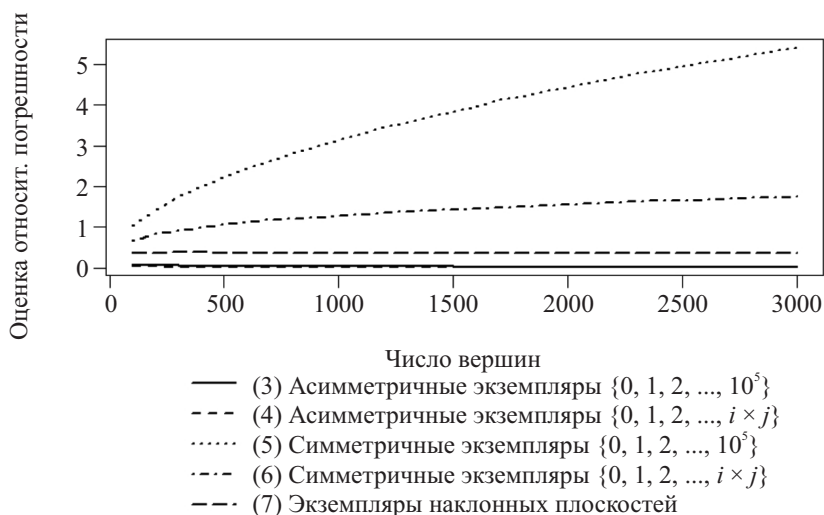


Рис. 2. Оценка относительной погрешности алгоритма соединения циклов для различных семейств задачи коммивояжера.

о меньшей зависимости точности алгоритма от размерности задачи в данном семействе экземпляров.

Таким образом, анализ показывает, что:

- для асимметричных экземпляров задачи коммивояжера оценка относительной погрешности уменьшается с увеличением числа вершин;
- для симметричных экземпляров задачи коммивояжера наблюдается обратная тенденция – погрешность увеличивается с ростом размерности задачи;
- семейство наклонных плоскостей демонстрирует стабильные значения относительной погрешности, незначительно изменяющиеся с увеличением числа вершин.

5.2. Анализ времени решения

На рис. 3 представлено среднее время работы алгоритма соединения циклов при решении одной задачи для различных семейств экземпляров задачи коммивояжера.

Анализ временных характеристик показывает различное поведение роста времени решения в зависимости от семейства рассматриваемых экземпляров.

Для асимметричных экземпляров задачи коммивояжера с равномерно распределенными элементами матрицы весов из множества $\{0, 1, 2, \dots, 10^5\}$ (семейство 3) время решения увеличивается от 0,105 с при $n = 100$ до 259,88 с при $n = 3000$. Для семейства 4 ($W = [w(i, j)]$, где $w(i, j)$ – равномерно распределенные случайные числа из $\{0, 1, 2, \dots, i \times j\}$) время решения растет быстрее, начиная с 0,14 с при $n = 100$ и достигая 3057,04 с при $n = 3000$. Это

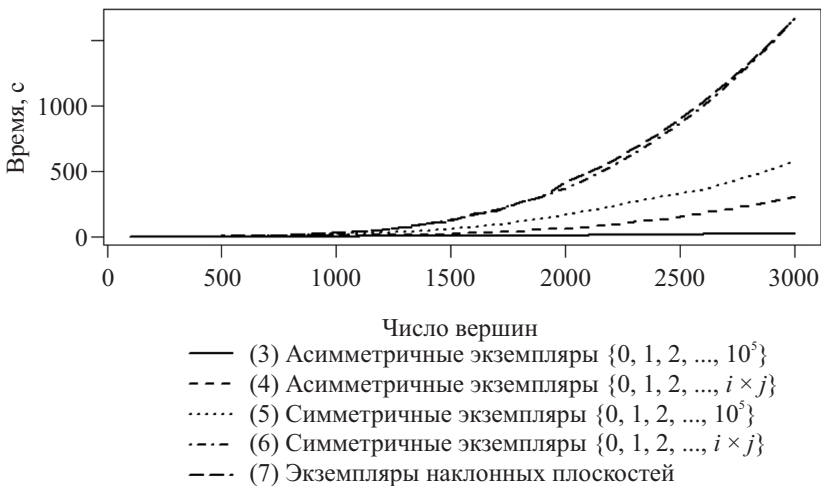


Рис. 3. Время решения одного экземпляра задачи коммивояжера алгоритмом соединения циклов для различных семейств задач.

свидетельствует о значительном влиянии структуры матрицы стоимостей на вычислительную сложность алгоритма.

Для симметричных экземпляров задачи коммивояжера наблюдается более быстрый рост времени решения. В частности, для семейства 5 ($W = [w(i, j)]$, где $w(i, j)$ – равномерно распределенные случайные числа из $\{0, 1, 2, \dots, 10^5\}$) оно возрастает с 0,114 с при $n = 100$ до 5760,3 с при $n = 3000$. Для семейства 6 ($W = [w(i, j)]$, где $w(i, j)$ – равномерно распределенные случайные числа из $\{0, 1, 2, \dots, i \times j\}$) этот рост еще более выражен: от 0,27 с до 16 669,63 с на том же интервале значений n . Таким образом, специфика структуры весов оказывает влияние не только на точность, но и на вычислительную сложность алгоритма.

Для экземпляров наклонных плоскостей (семейство 7) также наблюдается значительный рост времени решения. При $n = 100$ среднее время составляет 0,279 с, а при $n = 3000$ оно достигает 16 694,93 с, что аналогично показателям для шестого семейства задач и существенно превышает данные по всем остальным семействам задач. Это указывает на сложность обработки таких экземпляров алгоритмом соединения циклов.

В целом анализ временных характеристик показывает, что:

- асимметричные экземпляры задачи коммивояжера решаются быстрее, чем симметричные, однако структура матрицы стоимостей существенно влияет на рост времени решения;
- для симметричных экземпляров наблюдается ускоренный рост времени работы алгоритма, особенно для семейства 6;
- экземпляры наклонных плоскостей демонстрируют наибольшие значения времени решения, что может быть связано с особенностями геометрической структуры задач.

Сложность симметричной неевклидовой задачи коммивояжера для алгоритмов построения тура, начинающихся с решения задачи о назначении, обусловлена низкой точностью нижней границы, полученной на основе этого метода. В частности, для подавляющего большинства симметричных экземпляров задача о назначении приводит к цикловому покрытию, содержащему большое количество коротких циклов (табл. 2). Это значительно затрудняет процесс их объединения в единый маршрут, увеличивая как вычислительные затраты, так и итоговую относительную погрешность.

Как видно из таблицы, для асимметричных экземпляров отношение числа начальных циклов к числу вершин существенно уменьшается с увеличением размерности задачи, снижаясь с 0,0382 (100 вершин) до 0,0070 (1000 вершин). Это указывает на тенденцию к формированию более длинных циклов, что способствует более эффективному соединению маршрутов на последующих этапах работы алгоритма.

В отличие от этого для симметричных экземпляров данное отношение остается практически неизменным, колеблясь в диапазоне 0,4768–0,4920. Это свидетельствует о том, что при формировании 2-фактора для данного се-

Таблица 2. Отношение числа начальных циклов к числу вершин задачи

Число вершин	Асимметричные экземпляры	Симметричные экземпляры
100	0,0382	0,4920
200	0,0225	0,4768
300	0,0162	0,4771
400	0,0139	0,4892
500	0,0118	0,4848
600	0,0101	0,4867
700	0,0086	0,4896
800	0,0076	0,4902
900	0,0071	0,4890
1000	0,0070	0,4878

мейства задач образуется множество коротких подциклов, что затрудняет их объединение в итоговый маршрут.

Рост времени решения для симметричных экземпляров объясняется тем, что процесс соединения большого количества коротких циклов требует значительно большего количества итераций, что приводит к усложнению структуры алгоритма и росту его вычислительной сложности. Как показано на рис. 3, время работы алгоритма соединения циклов для симметричных экземпляров существенно превышает аналогичные значения для асимметричных задач.

Таким образом, для повышения точности и сокращения времени решения симметричной неевклидовой задачи коммивояжера требуется дальнейшее изучение специализированных эвристик и приближенных алгоритмов, адаптированных для работы с данным семейством задач.

6. Статистический анализ точности алгоритма соединения циклов для симметричной задачи коммивояжера

В рамках исследования точности алгоритма соединения циклов для симметричной задачи коммивояжера на минимум был проведен статистический анализ эмпирических данных, полученных для экземпляров задачи с матрицей весов, элементы которой генерируются равномерно из множества $\{0, 1, 2, \dots, i \times j\}$ (пятое семейство экземпляров задач из раздела 4). Поскольку для симметричной задачи коммивояжера на минимум отсутствуют теоретические границы точности алгоритма, для прогноза зависимости относительной погрешности от числа вершин были использованы методы регрессионного анализа.

Рассматривались три возможные регрессионные модели: степенная, полиномиальная второго порядка и логарифмическая. Эти модели были выбраны на основе наблюдений о характере роста относительной погрешности с увеличением числа вершин.

6.1. Выбор модели регрессии

Исследовательский анализ является необходимым этапом для проверки допущений [16], связанных с регрессионными моделями, и выбора наиболее подходящей модели для аппроксимации данных. В рамках этого этапа анализа были проведены визуальные проверки и вычислены статистические характеристики для каждой модели, что позволило сравнить их эффективность и точность.

Визуальные методы, такие как точечная диаграмма рассеяния и матрица точечной диаграммы, использовались для анализа взаимосвязи между числом вершин и относительной погрешностью. Также были использованы гистограммы и тесты на нормальность распределения остатков. Статистические тесты, такие как коэффициент детерминации R^2 , скорректированный R^2 , а также тесты Дарбина–Уотсона и Бреуша–Пагана, позволили оценить соответствие модели статистическим предположениям.

6.1.1. Степенная регрессия

Степенная модель имеет вид

$$(1) \quad y = ax^b.$$

После логарифмирования уравнение принимает линейный вид:

$$(2) \quad \log y = \log a + b \times \log x.$$

Для этой модели были получены следующие результаты:

- Коэффициент детерминации $R^2 = 0,9676$, скорректированный $R^2 = 0,9665$, что свидетельствует о высоком качестве аппроксимации данных.
- Значимость коэффициентов модели подтверждается p -значениями, которые меньше 2×10^{-16} .
- Тест Дарбина–Уотсона показал отсутствие значимой автокорреляции остатков ($DW = 1,9527$, $p = 0,3711$).
- Тест Бреуша–Пагана выявил наличие гетероскедастичности ($p = 1,24 \times 10^{-6}$), что указывает на нестабильность дисперсии остатков.
- Глобальный тест на соответствие предположениям линейной регрессии показал нарушение предположений ($p = 2,81 \times 10^{-7}$).

Таким образом, хотя степенная модель имеет высокий коэффициент детерминации и предоставляет хорошую аппроксимацию, выявленные проблемы с гетероскедастичностью и нарушением линейных предположений указывают на ее ограниченность.

6.1.2. Полиномиальная регрессия

Полиномиальная модель второго порядка имеет вид

$$(3) \quad \log y = a + b \times \log x + c \times \log^2 x.$$

Статистические характеристики этой модели:

- $R^2 = 0,9999$, скорректированный $R^2 = 0,9999$, что подтверждает почти идеальное соответствие данных модели.
- Все коэффициенты модели статистически значимы ($p < 0,001$).
- Тест Дарбина–Уотсона показал отсутствие автокорреляции ($DW = 2,2082$, $p = 0,5771$).
- Тест Бреуша–Пагана не выявил гетероскедастичности ($p = 0,1035$).
- Глобальный тест показал небольшие отклонения от предпосылок ($p = 0,002$).

Полиномиальная модель второго порядка показала значительно лучшее соответствие данным по сравнению со степенной моделью, устранив проблемы гетероскедастичности. Однако тест на глобальное соответствие предпосылкам линейности указывает на возможные отклонения, которые тем не менее не имеют значительного влияния на точность модели.

6.1.3. Логарифмическая регрессия

Логарифмическая модель описывается уравнением

$$(4) \quad y = a + b \times \ln x.$$

Статистические результаты:

- R^2 немного ниже, чем у полиномиальной модели, но остается высоким.
- Остатки соответствуют требованиям нормальности и гомоскедастичности.
- Глобальный тест подтверждает удовлетворительное соответствие предпосылкам.

Логарифмическая модель, как и полиномиальная модель, показала хорошее соответствие данным и предпосылкам регрессии с легко интерпретируемыми коэффициентами.

6.1.4. Выбор модели

На рис. 4 представлены графики эмпирических и модельных значений для различных регрессионных моделей.

Анализ показал, что степенная модель имеет высокий коэффициент детерминации, но нарушает предпосылки регрессии. Полиномиальная модель второго порядка значительно улучшает качество аппроксимации, устраняя гетероскедастичность. Добавление третьего порядка улучшает качество предсказаний, но может привести к избыточному усложнению. Логарифмическая модель хорошо описывает замедляющийся рост, но имеет ограничения при экстраполяции.

С учетом результатов тестов и интерпретации данных наиболее подходящими являются полиномиальная модель второго порядка и логарифмическая регрессия. Полиномиальная модель обеспечила наилучшее соответствие данным, устранив основные проблемы степенной регрессии. Логарифмическая модель хорошо описывает замедляющийся рост, но при экстраполяции

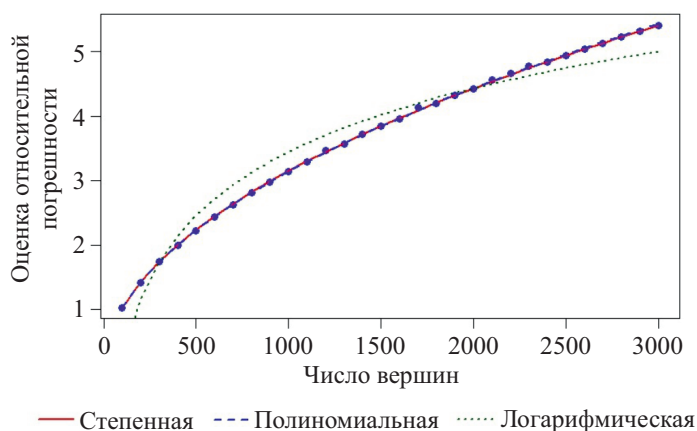


Рис. 4. Сравнение моделей регрессии для симметричных экземпляров из множества $\{0, 1, 2, \dots, 10^5\}$.

на значения, выходящие за пределы обучающего набора, она может давать смещение, так как логарифмическая зависимость предполагает бесконечный рост при больших x , что не всегда соответствует реальным данным.

В дальнейшем анализе будет использована полиномиальная модель второго порядка, как наиболее точная и удовлетворяющая ключевым статистическим требованиям (табл. 3).

Таблица 3. Глобальная проверка предпосылок полиномиальной регрессии

Тест	Значение	p -значение
Глобальная статистика	16,843	0,0021 (нарушены)
Асимметрия	3,189	0,0742 (допустимо)
Эксцесс	2,321	0,1277 (допустимо)
Функция связи	9,898	0,0017 (нарушены)
Гетероскедастичность	1,436	0,2308 (допустимо)

Полиномиальная модель в логарифмической шкале демонстрирует хорошее соответствие данным, однако небольшие отклонения от линейности могут указывать на ее ограничения при экстраполяции. В целом результаты подтверждают ее применимость, но при прогнозировании для больших значений необходимо учитывать возможное смещение.

6.2. Оценка и интерпретация модели

Построенная полиномиальная регрессионная модель описывает зависимость средней относительной погрешности от числа вершин. Включение квадратичного члена позволило учесть нелинейный характер роста погрешности. Коэффициент $b_1 = 0,4465$ при $\log n$ указывает на ускоряющийся рост

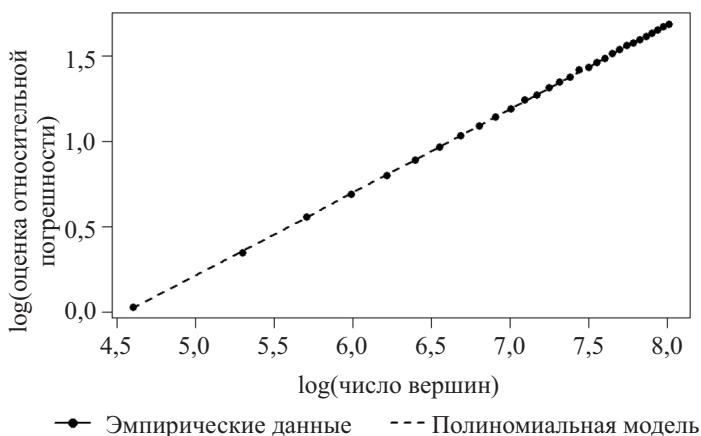


Рис. 5. Эмпирические и модельные значения логарифмически преобразованных данных для полиномиальной регрессии для симметричных экземпляров из множества $\{0, 1, 2, \dots, 10^5\}$.

ошибки при увеличении размерности задачи, а положительное значение $b_2 = 0,0035$ при $(\log n)^2$ подтверждает небольшое ускорение роста при малых значениях n . Однако данный эффект сглаживается при больших размерностях задачи, что свидетельствует о тенденции к стабилизации погрешности.

Ключевые статистические показатели модели:

- $R^2 = 0,9999$ (практически полное объяснение дисперсии).
- F -статистика: 129 100 ($p < 2 \times 10^{-16}$), что подтверждает значимость модели и высокую степень соответствия данным.
- Доверительные интервалы показывают, что все коэффициенты модели статистически значимы ($p < 0,01$).

На рис. 5 представлено графическое сравнение эмпирических данных и предсказанных значений модели.

Анализ глобального теста предпосылок регрессии показал нарушение функции связи ($p = 0,0017$). Однако отсутствие гетероскедастичности и автокорреляции остатков позволяет считать полиномиальную модель подходящей для описания зависимости, несмотря на небольшие отклонения от предположений линейности.

6.3. Прогнозирование для задач большой размерности

На основании построенной полиномиальной модели выполнена экстраполяция для прогнозирования относительной погрешности для задач размером до 10 000 вершин (рис. 6). Результаты показывают, что при больших n тенденция к росту ошибки сохраняется, однако эффект ускорения, обусловленный квадратичным членом, становится менее выраженным. Это подтверждает, что относительная погрешность стремится к стабилизации, а ее увеличение замедляется.

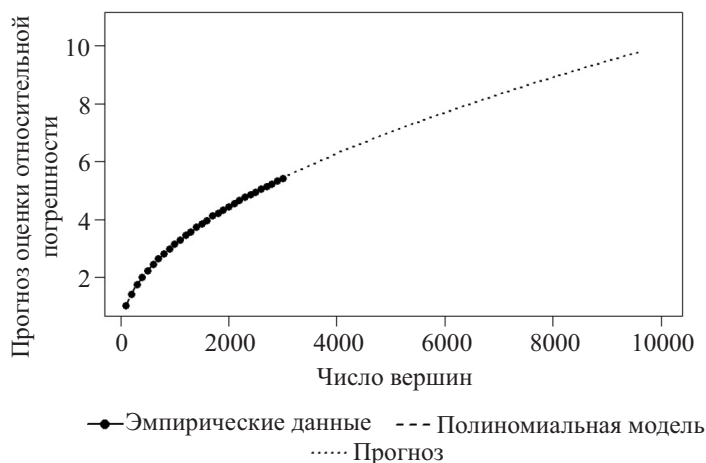


Рис. 6. Прогнозирование оценки относительной погрешности для задач большой размерности для симметричных экземпляров из множества $\{0, 1, 2, \dots, 10^5\}$.

Статистический анализ демонстрирует, что предложенная полиномиальная модель хорошо описывает зависимость относительной погрешности от размерности задачи и позволяет выполнять предсказания для больших n . Хотя модель сохраняет высокую точность аппроксимации, возможные отклонения при экстраполяции указывают на необходимость дальнейшего исследования поведения ошибки для задач экстремально больших размеров. Тем не менее полученные результаты подтверждают, что алгоритм соединения циклов демонстрирует устойчивый характер изменения погрешности и может быть применим к решению симметричной задачи коммивояжера в случае большой размерности.

7. Заключение

В статье приведены результаты эмпирического исследования точности и временной сложности алгоритма соединения циклов, а также статистического анализа точности для одного из рассматриваемых семейств симметричной задачи коммивояжера на минимум.

Эмпирический анализ выполнен на семи семействах задач, из которых в статье представлены результаты для пяти семейств случайно сгенерированных экземпляров. Исследование показало, что качество решения и вычислительная сложность алгоритма зависят от структуры матрицы стоимостей. Для асимметричных экземпляров оценка относительной погрешности снижается до 1–2% при 3000 вершинах, тогда как для симметричных экземпляров и экземпляров наклонных плоскостей наблюдается постепенный рост относительной погрешности. Время решения для всех семейств увеличивается полиномиально, при этом для симметричных экземпляров и экземпляров

наклонных плоскостей оно оказывается значительно выше, чем для асимметричных.

Проведенный анализ показывает, что относительная погрешность алгоритма соединения циклов для симметричной задачи коммивояжера демонстрирует полиномиальный рост с увеличением размерности задачи. Это подтверждается статистической обработкой экспериментальных данных, где полиномиальная регрессионная модель обеспечивает наилучшее соответствие эмпирическим наблюдениям.

Несмотря на рост погрешности, предложенный алгоритм сохраняет приемлемую точность и полиномиальное время работы, что делает его перспективным инструментом для решения экземпляров задачи коммивояжера большой размерности. Для дальнейшего повышения качества решений целесообразно исследовать гибридные подходы, сочетающие соединение циклов с дополнительными эвристиками и локальными улучшениями.

СПИСОК ЛИТЕРАТУРЫ

1. *Zhang C., Sun P.* Heuristic Methods for Solving the Traveling Salesman Problem (TSP): A Comparative Study // 2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). 2023. P. 1–6. <https://doi.org/10.1109/PIMRC56721.2023.10293957>
2. *Toaza B., Esztergár-Kiss D.* A review of metaheuristic algorithms for solving TSP-based scheduling optimization problems // Appl. Soft Comput. 2023. V. 148. 110908 p. <https://doi.org/10.1016/j.asoc.2023.110908>
3. *Reingold E.M., Nievergelt J., Deo N.* Combinatorial algorithms: theory and practice. Prentice Hall College Div, 1977. <https://doi.org/10.2307/2987917>
4. *Авдошин С.М., Береснева Е.Н.* Метрическая задача коммивояжера: экспериментальное исследование Парето-оптимальных алгоритмов // Труды Института системного программирования РАН. 2017. 29(4). С. 123–138. [https://doi.org/10.15514/ISPRAS-2017-29\(4\)-8](https://doi.org/10.15514/ISPRAS-2017-29(4)-8)
5. *Flood M.M.* The traveling-salesman problem // Oper. res. 1956. V. 4. P. 61–75. <https://doi.org/10.1287/opre.4.1.61>
6. *Rosenkrantz D., Stearns R., Lewis II P.* An analysis of several heuristics for the traveling salesman problem // SIAM J. Comput. 1977. V. 6. P. 563–581. <https://doi.org/10.1137/0206041>
7. *Hahsler M., Hornik K.* TSP – Infrastructure for the traveling salesperson problem // J. Statist. Softwar. 2007. V. 23. No. 2. <https://doi.org/10.18637/jss.v023.i02>
8. *Kay E., Christofides N.* Graph Theory: An Algorithmic Approach // Oper. Res. Quarterly. 1976. V. 27. No. 4. P. 1027. <https://doi.org/10.2307/3009122>
9. *Christofides N.* Worst-case analysis of a new heuristic for the travelling salesman problem // Graduat. School Industr. Administrat. CMU. 1976. <https://doi.org/10.1007/s43069-021-00101-z>
10. *Buchin K.* Space-filling curves. Organizing points sets: space-filling curves, delaunay tessellations of random point sets, and flow complexes. Berlin, Free University of Berlin, 2007. P. 5–30.

11. *Bartholdi J., Platzman L., Collins R., Warden W.* A minimal technology routing system for meals on wheels // *Interfaces*. 1983. V. 13. No. 3. P. 1–8.
<https://doi.org/10.1287/inte.13.3.1>
12. *Aarts E., Lenstra J.* Local search in combinatorial optimization, Princeton, New Jersey: Princeton University Press, 2003. <https://doi.org/10.1515/9780691187563>
13. *Панюков А.В., Леонова Ю.Ф.* Алгоритм соединения циклов для метрической задачи коммивояжера на максимум // *Вест. Южно-Урал. гос. ун-та. Сер. Вычисл. мат. и информатика*. 2021. Т. 10. No. 4. С. 26–36.
<https://doi.org/10.14529/cmse210402>
14. *Reinelt G.* TSPLIB—A traveling salesman problem library // *ORSA J. Comput.* 1991. V. 3. P. 376–384. <https://doi.org/10.1287/ijoc.3.4.376>.
15. *Леонова Ю.Ф.* Исследование качества алгоритма соединения циклов на примерах из библиотеки TSPLIB / Ю.Ф. Леонова // *Актуальные проблемы прикладной математики, информатики и механики : сборник трудов Международной научной конференции, Воронеж, 13–15 декабря 2021 г. Воронеж, 2022. С. 573–579.*
16. *Heins J., Bossek J., Pohl J., et. al.* A study on the effects of normalized TSP features for automated algorithm selection // *Theoretical Computer Science*. 2023. V. 940. P. 123–145. <https://doi.org/10.1016/j.tcs.2023.01.045>

Статья представлена к публикации членом редколлегии А.А. Лазаревым.

Поступила в редакцию 11.03.2024

После доработки 28.02.2025

Принята к публикации 14.03.2025