

Оптимизация, системный анализ и исследование операций

© 2019 г. Г.Н. ЖУКОВА, канд. физ.-мат. наук (gzhukova@hse.ru)

(Национальный исследовательский университет
«Высшая школа экономики», Москва),

М.В. УЛЬЯНОВ, д-р техн. наук (muljanov@mail.ru)

(Институт проблем управления РАН им. В.А. Трапезникова, Москва
ВМК МГУ им. Ломоносова),

М.И. ФОМИЧЕВ (michan94@yandex.ru)

(Национальный исследовательский университет
«Высшая школа экономики», Москва)

КОМБИНИРОВАННЫЙ ТОЧНЫЙ АЛГОРИТМ ДЛЯ АСИММЕТРИЧНОЙ ЗАДАЧИ КОММИВОЯЖЕРА: ПОСТРОЕНИЕ И СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ ВРЕМЕННОЙ ЭФФЕКТИВНОСТИ¹

Приведены результаты сравнительного статистического анализа времени решения несимметричной задачи коммивояжера (NTSP) методом ветвей и границ (без предвычисления тура) и комбинированным методом. Комбинированный метод состоит из приближенного алгоритма Lin-Kernighan-Helsgaun, используемого для вычисления начального тура, и метода ветвей и границ. Показано, что использование приближенного решения, найденного с помощью алгоритма Lin-Kernighan-Helsgaun, позволяет существенно уменьшить время поиска точного решения задачи коммивояжера методом ветвей и границ для задач из некоторого класса. Построен прогноз времени поиска точного решения методом ветвей и границ и комбинированным алгоритмом. Вычислительный эксперимент показал, что доля задач, которые комбинированным алгоритмом были решены быстрее, чем методом ветвей и границ, растет с ростом размерности задачи.

Ключевые слова: задача коммивояжера, метод ветвей и границ, аппроксимация вероятностного распределения, квантиль вероятностного распределения, вероятностный прогноз времени.

DOI: 10.1134/S0005231019110096

1. Введение

Задача коммивояжера заключается в поиске гамильтонова цикла с минимальной суммой весов ребер в полном (ориентированном в случае несимметричной задачи) графе. Уже при относительно небольшом (порядка десятков)

¹ Исследование выполнено за счет гранта Российского научного фонда (проект № 19-01665).

порядке матрицы, описывающей задачу, для осуществления полного перебора всевозможных гамильтоновых циклов требуется слишком много времени, поэтому были предприняты многочисленные попытки усовершенствовать перебор. Одним из наиболее известных и широко применяемых алгоритмов оптимизированного перебора является метод ветвей и границ (англ. Branch and Bound Method, сокращенно *B&B*) [1–8].

В рамках проведенного исследования использовалась классическая реализация метода ветвей и границ [1], которая предполагает следующие особенности:

- дуга ветвления (k, l) выбирается так, чтобы получить большую по величине нижнюю границу подмножества множества допустимых решений, которые не содержат дугу (k, l) ;
- приведенная матрица стоимостей хранится только в корне дерева, в узлах поискового дерева решений матрицы не хранятся, матрица стоимостей отсекается при необходимости для каждого из узлов дерева решений;
- для хранения листьев поискового дерева решений используется упорядоченный список;
- процедура удаления вершин поискового дерева решений осуществляется каждый раз при нахождении нового кандидата на оптимальное решение.

Метод ветвей и границ пытались улучшить, например в [9], в том числе за счет использования в качестве начального тура гамильтонова цикла, найденного другим (например, жадным) алгоритмом [10–14]. Такой способ позволяет исключить из рассмотрения вершины дерева решений с нижней оценкой, не меньшей суммы весов дуг этого тура. Однако показано [15], что только в случае рекорда, очень близкого к оптимальному туру (т.е. со стоимостью, превышающей не более чем на 5% стоимость оптимального решения), значительное число вершин может быть исключено. Предвычисленные туры со стоимостью, значительно большей стоимости оптимального тура, не ускоряют работу метода ветвей и границ.

Ранее [16] авторами статьи исследовалась такая характеристика индивидуальной задачи коммивояжера, как сложность [17], равная числу вершин поискового дерева, построенного при решении задачи методом ветвей и границ.

Поскольку число вершин дерева решений зависит от того, использовался ли предвычисленный тур, и от качества этого тура, то будем далее называть сложностью индивидуальной задачи коммивояжера число вершин, построенных классическим методом ветвей и границ (без предвычислений). Число вершин, порожденных методом ветвей и границ с использованием предвычисленного тура, будем называть сложностью индивидуальной задачи при использовании предвычисленного тура, полученного некоторым эвристическим алгоритмом.

Отметим, что сложность индивидуальной задачи коммивояжера при использовании предвычисленного тура не будет значительно меньше сложности индивидуальной задачи коммивояжера (без предвычислений), если стоимость предвычисленного тура значительно больше стоимости оптимального тура.

2. Постановка задачи

Объектом исследования в данной статье является несимметричная задача коммивояжера. Решаются индивидуальные несимметричные задачи коммивояжера с матрицей стоимостей, элементы которой — псевдослучайные числа с равномерным распределением.

Предмет исследования — временная эффективность программной реализации комбинированного точного алгоритма, включающего в себя эвристический алгоритм для поиска начального тура и метод ветвей и границ для получения точного решения задачи коммивояжера.

Авторы ставят перед собой в рамках данной статьи следующие задачи:

1. Выбор эвристического алгоритма для получения первоначального псевдооптимального тура, обеспечивающего начиная с некоторой размерности асимметричной задачи коммивояжера сокращение среднего значения времени получения точного решения. В комбинированном алгоритме (далее $E + B\&B$) время получения точного решения представляет собой сумму времени работы эвристического алгоритма E и метода ветвей и границ (далее $B\&B$) с предвычисленным туром, т.е. надо найти такой эвристический алгоритм E и такую размерность задачи n_0 , что

$$(1) \quad \forall n > n_0 \quad \bar{t}_{E+B\&B}(n) < \bar{t}_{B\&B}(n).$$

2. Статистическое исследование экспериментально полученных значений времени работы программных реализаций алгоритмов $B\&B$ и $E + B\&B$ с целью сравнения быстродействия алгоритмов. Построение аппроксимирующих семейств распределений времени работы алгоритмов $B\&B$ и $E + B\&B$. Прогнозирование среднего и квантилей времени работы алгоритма $E + B\&B$ при помощи линейных функций, зависящих от размерности ATSP.
3. Оценка качества прогноза путем сравнения прогнозируемых значений среднего времени работы алгоритма $E + B\&B$, а также квантилей аппроксимирующего распределения с соответствующими значениями контрольной выборки.
4. Определение характерных особенностей индивидуальных задач, для которых комбинированный алгоритм $E + B\&B$ обеспечивает ощутимое сокращение времени расчета:

$$(2) \quad t_{E+B\&B}(A) < t_{B\&B}(A).$$

5. Исследование зависимости времени решения задачи коммивояжера от сложности.
6. Формулировка рекомендаций по применению комбинированного алгоритма.

3. Выбор эвристического алгоритма

Классический метод ветвей и границ не предполагает наличие какого-либо тура в момент начального запуска, поэтому сначала многократно проводится процедура ветвления, порождаются вершины поискового дерева, а исключение вершин из рассмотрения не ведется. В зависимости от особенностей

индивидуальной задачи число вершин дерева решений может быть довольно большим к моменту, когда будет найден первый тур. Использование тура, найденного заранее каким-либо приближенным методом, позволяет начать исключение вершин из рассмотрения раньше, в результате будет порождено меньшее число вершин дерева решений, что уменьшит сложность задачи и время поиска оптимального тура.

С другой стороны, на получение начального тура придется тратить время, так что общее время расчетов будет складываться из времени, затраченного на поиск начального тура, и времени работы метода ветвей и границ с найденным туром (будем называть его далее предвычисленным туром). В связи с этим будем искать настолько быстрый алгоритм поиска начального тура, чтобы суммарное время работы алгоритма для нахождения предвычисленного тура и алгоритма метода ветвей и границ с найденным предвычисленным туром было не больше времени работы классического метода ветвей и границ.

Из экспериментального исследования, проведенного одним из авторов статьи [15], следует:

- если предвычисленный тур является оптимальным, то сложность индивидуальной задачи при использовании предвычисленного тура меньше в среднем приблизительно на 35–40% ;
- если стоимость предвычисленного тура на 5% больше, чем оптимального, то сложность индивидуальной задачи при использовании предвычисленного тура не будет заметно меньше, чем сложность при решении этой задачи классическим *B&B*.

Обозначим через ε точность предвычисленного тура:

$$(3) \quad \varepsilon = \frac{p_E(A) - p_{B\&B}(A)}{p_{B\&B}(A)},$$

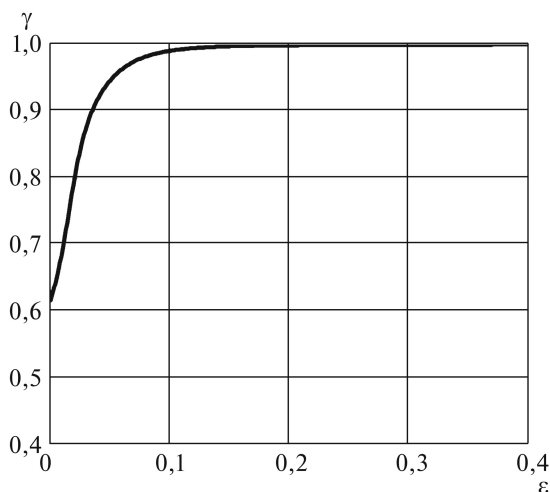
где $p_{B\&B}(A)$ и $p_E(A)$ — соответственно стоимость оптимального тура задачи коммивояжера с матрицей стоимостей A и стоимость тура, найденного приближенным алгоритмом E . Пусть γ — отношение сложности при использовании предвычисленного тура к сложности индивидуальной задачи коммивояжера:

$$(4) \quad \gamma = \frac{c_{B\&B}(A) - c_{E+B\&B}(A)}{c_{B\&B}(A)}.$$

На рисунке представлена зависимость γ от точности ε предвычисленного тура для $n = 45$ [15], для других размерностей график аналогичный.

Зависимость получена для несимметричных задач коммивояжера с элементами матрицы стоимостей, полученными генератором псевдослучайных чисел с равномерным распределением.

Таким образом, для поиска начального тура для метода ветвей и границ необходим приближенный алгоритм, быстро находящий решение, очень близкое к оптимальному. Будем искать подходящий алгоритм в классе эвристических алгоритмов.



Зависимость γ от ε .

На основе предварительного анализа литературных источников [15, 18–28] было решено для получения предвычисленного тура использовать эвристические алгоритмы, улучшающие решения. В 1971 г. S. Lin и B.W. Kernighan [29] предложили эффективный эвристический алгоритм, основанный на итерационном улучшении случайно полученного тура. Позднее, в 2000 г. Keld Helsgaun представил в [30] модифицированную реализацию алгоритма Lin-Kernighan. Алгоритм Lin-Kernighan-Helsgaun (далее LKH) находит допустимое решение, затем строит два множества дуг таких, что если все дуги первого множества удалить из найденного тура и заменить их дугами из второго множества, то результатом будет тур с меньшей суммой весов дуг. Процесс замены дуг повторяется до тех пор, пока невозможно сформировать два подходящих множества дуг. Все ограничения на множества дуг и особенности их выбора подробно описаны в [30].

Этот алгоритм достаточно часто находит не только субоптимальные, но и оптимальные решения за приемлемое время даже для задач большой размерности (порядка 10^4). Именно благодаря этому свойству, алгоритм LKH был выбран в качестве эвристического алгоритма для получения предвычисленного тура. В предлагаемом исследовании алгоритм LKH запускается только один раз, аналогично [31].

4. Статистическое исследование времени работы $B\&B$ и $E + B\&B$

Сравнение времени работы $B\&B$ без предвычисленного тура и комбинированного алгоритма вначале проведем на основе выборочных средних (табл. 1), затем сравним выборочные квантили распределений времени работы этих алгоритмов и доли задач в выборке и ее частях, которые были решены комбинированным алгоритмом быстрее, чем $B\&B$.

Вычислительный эксперимент проводился на стационарном компьютере со следующими характеристиками:

- процессор Intel i7 8700K 4700 MHz;

Таблица 1. Минимальное, максимальное и среднее время работы $B\&B$ и $E + B\&B$, мс.

n	min $B\&B$	min комб.	max $B\&B$	max комб.	среднее $B\&B$	среднее комб.	СКО $B\&B$	СКО комб.
30	227	2764	588975	589430	24070	31457	30747	30655
35	324	5287	2010030	1967021	78111	87304	108205	106739
40	445	5723	15011713	12654423	263125	270031	467713	439801
45	2144	8307	68240602	49713361	875652	848704	210841	1708932
50	4013	10533	1464202466	441979575	3046440	2634915	21328841	7631699

- оперативная память Corsair Vengeance LPX CMK32GX4M2B3466C16R DDR4 3466MHz 32 ГБ;

- материнская плата ASRock Fatal1ty Z370 Gaming K6;

- операционная система Arch with kernel version 4.14.13-1-ARCH.

Для минимизации шумов операционной системы фоновые процессы, которые не нужны для исследования, были отключены, а также отсутствует графический пользовательский интерфейс, а управление операционной системой осуществлялось посредством командной строки. Алгоритмы реализованы на языке C++ и скомпилированы в исполняемый файл с помощью компилятора gcc 7.2.1 20171224.

На данной аппаратной конфигурации был проведен вычислительный эксперимент по решению задач коммивояжера с матрицами стоимостей, имеющими порядок от 30 до 50. Для каждого значения порядка матрицы стоимостей было рассмотрено 10^4 задач коммивояжера, каждая задача была решена методом ветвей и границ и комбинированным методом. Все выборочные характеристики были вычислены при объеме выборки 10^4 .

Как видно из табл. 1, среднее время работы комбинированного алгоритма больше для задач размерностей до 40, а для размерностей более 45 комбинированный алгоритм работает несколько быстрее в среднем. Кроме того, среднеквадратическое отклонение времени работы комбинированного алгоритма меньше, чем у $B\&B$, причем с ростом размерности задач до 50 среднеквадратическое отклонение (СКО) времени работы комбинированного алгоритма становится меньше почти в 3 раза по сравнению с $B\&B$.

Также можно заметить, что максимальное время работы комбинированного алгоритма растет медленнее с ростом размерности, чем в случае $B\&B$, так что для $n = 45$ максимальное значение времени работы комбинированного алгоритма на 27% меньше, а для $n = 50$ — на 70% меньше.

В табл. 1 n — число вершин графа в задаче коммивояжера (равное порядку матрицы стоимостей), время здесь и далее в статье измеряется в миллисекундах.

Для построения прогноза квантилей и среднего времени работы алгоритмов построим аппроксимирующее семейство распределений для натурального логарифма времени работы алгоритмов. Для определения типа распределения воспользуемся квантильными коэффициентами асимметрии и эксцесса [32, 33] аналогично тому, как это было сделано для сложности индивидуальных задач коммивояжера в [16, 34].

Для выборок времени работы $B\&B$ по 10000 задач для каждой размерности были получены значения квантильного коэффициента асимметрии от $-0,037$ до $0,018$ и квантильного коэффициента эксцесса от $1,206$ до $1,262$. Для комбинированного алгоритма соответственно имеем квантильный коэффициент асимметрии от $-0,022$ до $0,079$ и квантильный коэффициент эксцесса от $1,194$ до $1,297$. Для нормального распределения независимо от значений параметров квантильный коэффициент асимметрии равен 0 , а квантильный коэффициент эксцесса — примерно $1,23$, поэтому будем далее использовать нормальное распределение в качестве аппроксимирующего распределения для натурального логарифма времени работы $B\&B$ и комбинированного алгоритма.

Параметры аппроксимирующего нормального распределения будем подбирать, строя линейную зависимость выборочных средних и выборочных среднеквадратических отклонений от размерности задач. Для диапазона размерностей от 30 до 50 с шагом 1 были получены выборки по 10000 матриц стоимостей для каждой размерности. Распределение элементов матриц было выбрано равномерным непрерывным на интервале $[0, 1]$. Каждая задача была решена $B\&B$, затем те же самые задачи в том же порядке были решены комбинированным алгоритмом. При решении задачи коммивояжера комбинированным алгоритмом приближенный алгоритм ЛКН запускался один раз, затем полученный псевдооптимальный тур использовался в качестве предвычисленного тура в $B\&B$.

По полученным выборкам были вычислены значения выборочных средних натурального логарифма времени работы обоих алгоритмов, а также среднеквадратические отклонения. Далее методом наименьших квадратов были вычислены коэффициенты сдвига и масштаба линейных функций, наилучшим образом приближающих зависимость выборочных средних и среднеквадратического отклонения от размерности задачи (для комбинированного алгоритма и $B\&B$). Полученные линейные функции были использованы для аппроксимации и экстраполяции зависимости от размерности задачи параметров нормального распределения, приближающего распределение натурального логарифма времени работы каждого алгоритма.

Обозначим $\mu_{B\&B}(n)$, $\sigma_{B\&B}(n)$, $\mu_{E+B\&B}(n)$, $\sigma_{E+B\&B}(n)$, найденные линейные зависимости выборочных средних и среднеквадратических отклонений времени работы $B\&B$ и комбинированного алгоритма соответственно:

$$(5) \quad \mu_{B\&B}(n) = 0,217n + 3,06,$$

$$(6) \quad \sigma_{B\&B}(n) = 0,015n + 0,59,$$

$$(7) \quad \mu_{E+B\&B}(n) = 0,219n + 2,96,$$

$$(8) \quad \sigma_{E+B\&B}(n) = 0,0135n + 0,65.$$

Отклонения выборочных средних (вычисленных по выборкам объема 10000 для каждой размерности задач от 30 до 50) от соответствующих значений, найденных по формулам (5)–(8), составляют менее $0,5\%$, для среднеквадратических отклонений — менее 2% .

Далее для размерностей задач от 30 до 50 были вычислены квантили нормальных распределений с параметрами, найденными по формулам (5)–(8).

Качество аппроксимации натурального логарифма времени работы каждого алгоритма нормальным распределением с параметрами (5)–(8) оценим, сравнивая вычисленные квантили нормального распределения с выборочными квантилями натуральных логарифмов времени работы алгоритмов. Используем данные о времени работы $V\&B$ и комбинированного алгоритма при решении одного набора по 10000 задач коммивояжера для каждой размерности от 30 до 50. Выборочные квантили натурального логарифма времени работы $V\&B$ уровня 25, 50 и 75% отличаются от квантилей нормального распределения с параметрами, вычисленными по (5)–(8), на 0,1–0,5%. Различие между выборочными квантилями и квантилями соответствующего нормального распределения монотонно уменьшается с ростом размерности задач для квантилей уровня от 10 до 90%.

Ввиду того, что логарифмическая функция монотонно возрастает (используется натуральный логарифм), квантили времени работы $V\&B$ (а также комбинированного алгоритма) можно найти, потенцируя соответствующие квантили натурального логарифма времени. Выборочные квантили времени работы $V\&B$ уровня 25, 50% отличаются от экспонент квантилей нормального распределения (или, что то же самое, квантилей логнормального распределения) с параметрами, вычисленными по (5)–(8) менее, чем на 7%, для квантилей уровня 75% отличие менее 4%. С ростом размерности задач до 50 отличия квантилей времени уровней 5–99% от аппроксимирующих значений уменьшаются до 1–2%.

Выборочные квантили времени работы комбинированного алгоритма при небольших значениях уровня отличаются от аппроксимирующих значений сильнее, чем у $V\&B$, особенно для небольших (до 40) размерностей задач. Так, отличие квантилей уровня 5% уменьшается от 66 до 7% с ростом размерности задачи от 30 до 50. Отличие квантилей уровней 50 и 75% от аппроксимирующих значений уменьшается соответственно от 40 и 25 до 1%. В то же время квантили уровня 90% и больше отличаются в меньшей степени от аппроксимирующих значений, в основном на 1–5% даже для размерностей от 35.

Различное поведение квантилей низких порядков у времени работы $V\&B$ и комбинированного алгоритма объясняется влиянием времени, затраченного приближенным алгоритмом ЛКН на вычисление тура. Это время не пропорционально времени, затраченному $V\&B$ на решение этой задачи, оно мало по сравнению со средним временем решения задачи этой же размерности, но от близкого к минимальному наблюдаемому времени решения задачи может отличаться несильно.

5. Построение прогноза и оценка его качества

Используем аппроксимацию натурального логарифма времени решения задачи коммивояжера нормальным распределением для построения прогноза среднего и среднеквадратического отклонения натурального логарифма времени работы $V\&B$ и алгоритма как функций размерности.

Далее будем использовать связь нормального и логнормального распределений: если случайная величина X имеет логнормальное распределение, то

Таблица 2. Сравнение прогнозируемых квантилей с выборочными

n	ω_{25} $B\&B$	ω_{25} $E + B\&B$	ω_{50} $B\&B$	ω_{50} $E + B\&B$	ω_{70} $B\&B$	ω_{70} $E + B\&B$	ω_{90} $B\&B$	ω_{90} $E + B\&B$
55	-4	-2	-8	-6	-9	-8	-8	-9
60	-2	-4	-12	-14	-11	-17	0	-13
65	-22	-28	-21	-26	-11	-27	14	16

случайная величина $\ln X$ распределена нормально с теми же параметрами μ и σ . Для удобства приведем формулы плотностей соответствующих друг другу нормального (9) и логнормального (9) распределений [35]:

$$(9) \quad f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

$$(10) \quad f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}.$$

Прогноз квантилей времени решения задачи коммивояжера $B\&B$ и комбинированного алгоритма можно получить по формулам

$$(11) \quad q_p = e^{q_p^{\text{норм}}},$$

где $q_p^{\text{норм}}$ — квантили нормального распределения с параметрами, вычисленными по (5)–(8).

Сравним прогнозируемые (по результатам статистической обработки выборок объема 10000 задач размерности от 30 до 50) значения квантилей времени работы $B\&B$ и $E + B\&B$ с соответствующими значениями контрольной выборки объема 1000 размерностей 55, 60 и 65. В табл. 2 представлены относительные отличия (в %) прогнозируемых и наблюдаемых квантилей уровня 25, 50, 70 и 90, вычисленные по формуле

$$(12) \quad \omega_p = 100 \frac{q_p^{\text{выборочн}} - q_p^{\text{норм}}}{q_p^{\text{выборочн}}}.$$

Для прогноза среднего значения натурального логарифма времени работы алгоритмов можно воспользоваться зависимостями (5) и (7), используемыми в качестве аппроксимации параметра μ нормального распределения, приближенно описывающего распределение натурального логарифма времени работы алгоритмов. Прогноз среднего времени работы алгоритмов можно получить по формулам, связывающим математическое ожидание и дисперсию логнормально распределенной случайной величины X с параметрами μ и σ соответствующей нормальной случайной величины:

$$(13) \quad EX = e^{\mu + \sigma^2/2},$$

$$(14) \quad DX = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2},$$

$$(15) \quad \sqrt{DX} = EX \sqrt{e^{\sigma^2} - 1}.$$

Таблица 3. Среднее относительное ускорение и доля ускоренных задач

n	среднее, прогноз	среднее, выборка	отн. откл., %	СКО, прогноз	СКО, выборка	отн. откл., %
натуральные логарифмы времени $B\&B$						
55	14,99	14,95	-0,24	1,41	1,41	0,24
60	16,07	16,06	-0,08	1,48	1,53	2,88
65	17,16	17,08	-0,48	1,56	1,75	10,88
натуральные логарифмы времени комбинированного алгоритма						
55	15,0	14,96	-0,29	1,39	1,35	-2,65
60	16,1	16,02	-0,47	1,46	1,42	-2,16
65	17,19	16,97	-1,29	1,52	1,57	3,1
время $B\&B$						
55	8851632	18426169	52	22402702,86	289538698,43	92
60	29210827	57601429	49	83689156,34	686238735,54	88
65	96940941	197362756	51	315309546,87	1534229312,24	79
время комбинированного алгоритма						
55	8662623	9622779	9	21134291	53524540	60
60	28510984	27846776	-2	77706642	87834992	11
65	94265716	87403339	-7	287648920	325819590	11
натуральные логарифмы времени $B\&B$, без максимумов						
55	14,99	14,94	-0,33	1,41	1,38	-2,06
60	16,07	16,05	-0,12	1,48	1,51	1,68
65	17,16	17,07	-0,53	1,56	1,74	10,11
натуральные логарифмы времени комбинированного алгоритма, без максимумов						
55	15,0	14,95	-0,36	1,39	1,33	-4,33
60	16,1	16,01	-0,5	1,46	1,41	-2,85
65	17,19	16,96	-1,32	1,52	1,56	2,46
время $B\&B$, без максимумов						
55	8851632	8579177	-3,17	22402702,86	20893391,09	-7,22
60	29210827	36285516	19,49	83689156,34	130486873,56	35,86
65	96940941	152130922	36,27	315309546,87	557043066,24	43,39
время комбинированного алгоритма, без максимумов						
55	8662623	7649809	-13	21134291	15181657	-39
60	28510984	25822374	-10	77706642	60203855	-29
65	94265716	78828873	-19	287648920	180950805	-58

Для контроля качества построенного прогноза среднего времени работы алгоритмов и среднеквадратического отклонения используем контрольные выборки объема 1000 задач размерности 55, 60 и 65. В табл. 3 приведены прогнозируемые значения среднего натурального логарифма времени (во втором столбце) и его среднеквадратического отклонения (в пятом столбце) для обоих алгоритмов, а также прогнозируемое и наблюдаемое среднее время и его среднеквадратическое отклонение. В третьем и шестом столбцах таблицы даны средние значения и среднеквадратические отклонения выборок, в четвертом и седьмом столбцах — относительные отклонения выборочных значений от прогнозируемых, вычисленные по формулам:

$$\omega = \frac{Z^* - \bar{Z}}{Z^*} \cdot 100,$$

где Z^* — прогнозируемое значение величины Z , а \bar{Z} — выборочное значение Z . В четвертом столбце приведены относительные отклонения выборочных значений от прогнозируемых для среднего, в седьмом — для среднеквадратического отклонения.

Под двойной чертой в этой таблице представлены результаты таких же расчетов, проведенных по тем же выборкам, но из каждой выборки было удалено максимальное значение времени. Как видно, среднее время не может служить надежной характеристикой временной эффективности алгоритма, поскольку оно очень сильно меняется под влиянием единичных задач, на решение которых понадобилось значительно больше времени, чем в среднем на остальные.

С одной стороны, чем больше объем выборки, тем более точно математическое ожидание наблюдаемой случайной величины может быть оценено выборочным средним. С другой — чем больше выборка, тем больше вероятность встретить в ней задачу, которая будет решаться очень долго и серьезно повлияет на величину выборочного среднего.

Резюмируя сказанное выше, отметим, что более предпочтительными характеристиками временной эффективности кажутся квантили распределения времени работы алгоритма, они гораздо менее чувствительны к отдельным выделяющимся значениям.

Полученные результаты свидетельствуют о том, что построенная аппроксимация может быть использована для грубой оценки времени решения задачи коммивояжера при размерностях больше 50. Большие отличия выборочных и теоретических квантилей при размерности 65 по сравнению с размерностями 55 и 60 могут означать, что используемый вычислительный комплекс работает на пределе своих возможностей. В этом случае для решения большого числа задач большей размерности следует использовать более производительную вычислительную систему.

6. Определение характерных особенностей индивидуальных задач

Анализируя данные об относительном ускорении решения задачи коммивояжера, вычисленном по формуле

$$(16) \quad \nu = 100 \frac{t_{B\&B} - t_{E+B\&B}}{t_{B\&B}},$$

можно заметить, что некоторые задачи решаются комбинированным алгоритмом медленнее, чем $B\&B$, но в основном это задачи, решаемые обоими алгоритмами быстро. В этом случае ускорение расчетов благодаря наличию тура незначительно, в то время как приближенный алгоритм тратит на поиск тура хоть и небольшое время, но сравнимое со временем работы $B\&B$, так что сумма времени работы приближенного алгоритма и $B\&B$ с предвычисленным туром оказывается больше времени работы $B\&B$ без предвычислений. С другой стороны, для задач, требующих много времени на нахождение точного решения, ускорение за счет использования предвычисленного тура может достигать 50%, т.е. комбинированный алгоритм работает в 2 раза быстрее на некоторых задачах.

Таблица 4. Среднее относительное ускорение и доля ускоренных задач, $e^{\mu+k\sigma}$

n	e^μ , сред. уск., %	e^μ , доля	$e^{\mu+\sigma}$, сред. уск., %	$e^{\mu+\sigma}$, доля	$e^{\mu+2\sigma}$, сред. уск., %	$e^{\mu+2\sigma}$, доля	$e^{\mu+3\sigma}$, сред. уск., %	$e^{\mu+3\sigma}$, доля
30	-28,24	6,44	-10,16	11,99	-3,53	17,8	0,87	36,36
35	-11,28	7,67	-4,11	12,83	-0,38	42,71	2,39	100,0
40	-3,58	13,62	-0,22	40,33	2,69	94,01	9,17	100,0
45	0,17	46,7	2,59	94,14	7,67	100,0	21,28	100,0
50	2,17	75,49	5,89	98,57	15,09	100,0	33,37	100,0

Для более детального анализа из выборок объема 10000 для каждой размерности задач коммивояжера от 30 до 50 были извлечены подмножества задач, для которых время решения $B\&B$ было больше, чем

$$(17) \quad e^{\mu+k\sigma}, \quad k = 0, 1, 2, 3,$$

где μ и σ — параметры нормального распределения, аппроксимирующего распределение логарифма времени решения задач коммивояжера. Эти параметры немного отличаются у $B\&B$ и комбинированного алгоритма.

Таким образом, для каждой размерности задачи коммивояжера было получено 4 подмножества задач из выборки объема 10000. Для каждого такого подмножества было вычислено относительное ускорение каждой задачи комбинированным алгоритмом по формуле (16), а также среднее в этом подмножестве ускорение и доля задач, для которых ускорение было больше 0. Результаты анализа приведены в табл. 4.

Как видно из таблицы, доля задач, решаемых быстрее комбинированным алгоритмом, растет с ростом времени работы $B\&B$. Среднее относительное ускорение варьируется от -28% до 33%. Несмотря на сравнительную малость среднего относительного ускорения, оно значительно влияет на величину среднего времени расчета в выборке из 10000 задач. Так, ускорение 0,2% наиболее долго решаемых задач размерности 50 в среднем на 33% в совокупности со значительно меньшим ускорением остальных задач приводит к уменьшению среднего времени решения для всей выборки почти на 1% по сравнению с $B\&B$ без предвычислений.

Аналогичный расчет был проведен на основе квантилей логнормального распределения с параметрами, соответствующими параметрам нормального распределения, аппроксимирующего логарифм времени работы $B\&B$. При фиксированной размерности задач коммивояжера параметры μ и σ логнормального распределения времени работы каждого алгоритма соответствуют параметрам μ и σ нормального распределения, аппроксимирующего натуральный логарифм времени работы.

Из выборок объема 10000 для каждой размерности задач коммивояжера от 30 до 50 были извлечены подмножества задач, для которых время решения $B\&B$ было больше, чем квантиль уровня $p = 50, 60, 70$ и 80 логнормального распределения с параметрами, вычисленными по формулам (5) и (6). Для каждой размерности задачи коммивояжера опять было получено 4 подмножества задач из выборки объема 10000. Среднее относительное ускорение и доля задач, для которых ускорение было больше 0, приведены в табл. 5.

Таблица 5. Среднее относительное ускорение и доля ускоренных задач, q_p

n	q_{50} , сред. уск., %	q_{50} , доля	q_{60} , сред. уск., %	q_{60} , доля	q_{70} , сред. уск., %	q_{70} , доля	q_{80} , сред. уск., %	q_{80} , доля
30	-12,36	10,63	-28,24	6,44	-22,51	7,64	-17,51	8,91
35	-4,96	11,53	-11,28	7,67	-9,06	8,23	-7,06	9,32
40	-0,72	31,75	-3,58	13,62	-2,54	16,7	-1,6	21,87
45	2,16	89,43	0,17	46,7	0,73	57,76	1,34	72,98
50	5,01	97,79	2,17	75,49	2,81	86,86	3,67	94,75

Заметим, что среднее относительное ускорение для задач размерности не более 40 отрицательное, это значит, что большой вклад в эту величину вносят задачи, которые решаются $B\&B$ быстрее, чем комбинированным алгоритмом. При больших размерностях на решение задачи коммивояжера тратится больше времени и применение комбинированного алгоритма приводит чаще к сокращению времени поиска оптимального решения, кроме того, величина относительного сокращения у задач, решаемых наиболее долго, становится больше.

7. Зависимость эффективности комбинированного алгоритма от сложности индивидуальных задач ATSP

При вычислении сложности индивидуальной задачи учитываются вершины дерева решений с нижней оценкой менее стоимости уже найденного тура. В случае классической реализации $B\&B$ (без предвычислений) до тех пор, пока не найден какой-то тур, учитываются все порожденные вершины дерева решений, а после того, как тур получен, только те, у которых нижняя оценка ниже стоимости этого тура.

Значения сложности индивидуальных задач коммивояжера фиксированной размерности (от 20 до 50) различаются на порядки, что серьезно затрудняет исследование временной эффективности алгоритмов. Так, в выборке из 100 000 задач размерности 40 встретились задачи со сложностями 77 и 598 893, причем выборочная средняя сложность составила 8145.

Поскольку задачи с небольшим значением сложности решаются программной реализацией $B\&B$ достаточно быстро, то использование предвычисленного тура не приведет к заметному уменьшению времени работы алгоритма. Более того, сумма времени, затраченного на поиск рекорда даже таким быстрым алгоритмом, как метод ЛКН, и времени работы $B\&B$ с полученным предвычисленным туром будет больше времени работы $B\&B$ без предвычисленного тура.

С другой стороны, задачи с большой сложностью требуют много времени на получение решения с помощью $B\&B$, и в этом случае достаточно быстро найденное приближенное решение ЛКН приводит к такому уменьшению времени работы $B\&B$, что сумма времени получения тура ЛКН и поиска точного решения $B\&B$ уже будет меньше времени работы $B\&B$ без предвычисленного тура.

В связи с этим каждая выборка 10000 матриц фиксированной размерности (от 35 до 45) была упорядочена по убыванию сложности, затем в упорядоченной выборке был найден номер задачи $N(C_{\max})$ такой, что для всех задач

с меньшими номерами время работы комбинированного алгоритма меньше, чем классического $V\&B$. $N(C_{\max})$ почти монотонно возрастает от 11 до 292 с ростом размерности задачи от 35 до 45. Также для каждой размерности от 35 до 45 были найдены значения сложности $C_{\text{крит}}$, такие что все задачи большей сложности из выборки решаются комбинированным алгоритмом быстрее, чем классическим $V\&B$. $C_{\text{крит}}$ растет от 40000 до 100000 с ростом размерности задачи от 35 до 45.

Кроме того, эти же выборки, содержащие по 10000 матриц для каждой размерности от 35 до 45, были упорядочены по убыванию времени работы классического $V\&B$. Число задач с наибольшим временем решения алгоритмом $V\&B$, которые были решены комбинированным алгоритмом быстрее, увеличивается от 11 до 171. Обозначим через \hat{t} время работы классического $V\&B$, такое что все задачи, решаемые $V\&B$ за время, большее \hat{t} , решаются комбинированным алгоритмом быстрее, чем классическим $V\&B$. Это время \hat{t} растет от 1100000 мс до 6400000 мс с ростом размерности задачи от 35 до 45.

Таким образом, применение комбинированного алгоритма сокращает время решения задачи коммивояжера в основном в случае задач с большим значением сложности. Эти задачи решаются классическим $V\&B$ дольше задач с небольшим значением сложности, поэтому закономерность «более сложная задача решается быстрее комбинированным алгоритмом» сохраняется в виде «задача, которая решается классическим $V\&B$ дольше, решится комбинированным алгоритмом быстрее». В отличие от этого в упорядоченной по времени работы ЛКН выборке не наблюдается аналогичной закономерности.

8. Формулировка рекомендаций по применению комбинированного алгоритма

Ввиду того, что задачи небольших размерностей (до 40–45) решаются методом ветвей и границ относительно быстро, не рекомендуется использовать комбинированный алгоритм в таких случаях. Причина в том, что уменьшение суммарного времени вычислений за счет использования предвычисленного тура будет настолько мало, что окажется меньше времени, требующегося на поиск начального тура даже таким быстрым приближенным алгоритмом, как ЛКН. В результате комбинированный алгоритм будет работать дольше, чем классический метод ветвей и границ.

В случае решения задач размерности больше 45 разумно использовать комбинированный алгоритм для решения большой серии задач коммивояжера, поскольку среднее время решения будет несколько меньше, так что общее время расчетов сократится.

Комбинированным алгоритмом стоит решать задачи с большим значением сложности, именно для таких задач ускорение расчетов за счет использования предвычисленного тура максимально.

9. Заключение

В статье на основании проведенных исследований, экспериментов с программными реализациями алгоритмов и статистического исследования полу-

ченных времен решения индивидуальных несимметричных задач коммивояжера показано что:

- алгоритм Lin-Kernighan-Helsgaun быстро находит достаточно близкий к оптимальному решению тур, позволяющий в некоторых случаях существенно уменьшить время поиска точного решения задачи коммивояжера методом ветвей и границ;
- комбинированный метод, в котором на первом этапе используется эвристический алгоритм для нахождения предвычисленного тура, а на втором этапе применяется программная реализация метода ветвей и границ с полученным туром в качестве начального, может быть построен с использованием эвристического алгоритма Lin-Kernighan-Helsgaun;
- программная реализация предложенного комбинированного алгоритма начиная с размерностей задачи больших, чем 45, показывает лучшие времена решений по оценке в среднем, чем классический метод ветвей и границ, несмотря на то, что доля задач, которые решаются быстрее, составляет менее половины;
- логнормальное распределение является удовлетворительным приближением распределения времен решения индивидуальных задач как для программных реализаций классического метода ветвей и границ, так и для предложенного комбинированного метода в диапазоне размерностей от 30 до 50;
- результаты вероятностного квантильного прогнозирования времен на основе полученного семейства логнормальных распределений достаточно хорошо совпадают с полученными экспериментальными данными для размерностей 55, 60 и 65, что позволяет использовать полученные результаты для вероятностного прогнозирования времен.

Параметры аппроксимирующего распределения были оценены по выборкам задач коммивояжера (объема 10000 для каждой размерности задачи от 30 до 50), затем методом наименьших квадратов была построена линейная зависимость этих параметров от размерности задачи. Сравнение квантилей аппроксимирующего распределения с выборочными квантилями натурального логарифма времени решения контрольных выборок объема 1000 задач коммивояжера размерностей 55, 60 и 65 показало, что прогноз удовлетворительный.

Также исследована зависимость эффективности применения комбинированного алгоритма от сложности задачи коммивояжера. Для этого было проведено дополнительное статистическое исследование выборки тех индивидуальных задач, для которых сокращение времени расчета комбинированным алгоритмом значительно. В результате показано, что в основном это такие индивидуальные задачи, сложность которых велика для классического метода ветвей и границ. На основе проведенных исследований авторы рекомендуют использовать комбинированный метод для индивидуальных несимметричных задач коммивояжера начиная с размерности порядка 45 и для индивидуальных задач меньшей размерности, для которых известно, что их сложность велика. Авторы видят дальнейшее развитие исследования в поиске таких характеристических функционалов исходных матриц индивидуальных задач, которые позволили бы определить за полиномиальное время принадлежность

задачи к классу сложных (имеющих большое значение сложности, т.е. много вершин в поисковом дереве, построенном $B\&B$).

СПИСОК ЛИТЕРАТУРЫ

1. *Little J.D.C., Murty K.G., Sweeney D.W., Karel C.* An Algorithm for the Traveling Salesman Problem // *Oper. Res.* 1963. No. 11. P. 972–989.
2. *Land A.H., Doig A.G.* An automatic method of solving discrete programming problems // *Econometrica*. 1960. V. 28 No. 3. P. 497–520.
3. *Гудман С., Хидетниеми С.* Ведение в разработку и анализ алгоритмов. М.: Мир, 1981.
4. *Korte B., Vygen J.* Combinatorial Optimization: Theory and Algorithms, Springer Publishing Company, Inc., 2007.
5. *Matai R., Mittal M.L., Singh S.* Travelling salesman problem: An overview of applications, formulations and solution approaches/Davendra, D. (ed.) Chapter 1 in *Travelling Salesman Problem: Theory and Applications*. Intech Open Access Publisher, Rijeka, 2010. P. 1–24.
6. *Slominski L.* Probabilistic analysis of combinatorial algorithms: a bibliography with selected annotations // *Computing*. 1982. V. 28. P. 257–267.
7. *Колесников А.В., Кириков И.А., Листопад С.В. и др.* Решение сложных задач коммивояжера методами функциональных гибридных интеллектуальных систем / Под ред. А.В. Колесникова. М.: ИПИ РАН, 2011.
8. *Ульянов М.В.* Ресурсно-эффективные компьютерные алгоритмы. Разработка и анализ. М.: ФИЗМАТЛИТ, 2008.
9. *Сергеев С.И.* Задача коммивояжера. Использование нелинейных разрешающих функций // *АиТ*. 2013 № 6. С. 101–120.
Sergeev S. Nonlinear Resolving Functions for the Travelling Salesman Problem // *Autom. Remote Control*. 2013. V. 74. No. 6. P. 978–994.
10. *Oliver I., Smith D., Holland J.* A Study of Permutation Crossover Operators on the Traveling Salesman Problem. J. Grefenstette (Ed.) // *Proc. 2 Int. Conf. Genet. Algorithm*, Lawrence Erlbaum Associat., Hillsdale N.J., 1987. P. 224–230.
11. *Cotta C., Aldana J., Nebro A., Troya J.* Hybridizing Genetic Algorithms with Branch and Bound Techniques for the Resolution of the TSP/ D. Pearson, N. Steele, R. Albrecht (Eds.), *Artificial Neural Nets and Genetic Algorithm 2*. Wien N.Y.: Springer-Verlag, 1995. P. 277–280.
12. *Goldberg D., Lingle R.J.* Alleles, Loci, and the Travelling Salesman Problem J. Grefenstette (Ed.) // *Proc. 1st Int. Conf. Genet. Algorithms*, Lawrence Erlbaum Associat., Hillsdale N.J., 1985. P. 154–159.
13. *Toriello A.* Optimal Toll Design: a Lower Bound Framework for the Asymmetric Traveling Salesman Problem // *Math. Programm.* 2014. V. 144. No. 1/2. P. 247–264.
14. *Cornuéjols G., Karamanov M., Li Y.* Early Estimates of the Size of Branch-and-Bound Trees // *INFORMS J. Comput.* 2006. V. 18. No. 1. P. 86–96.
15. *Фомичёв М.И.* Сравнительный анализ метаэвристических алгоритмов решения несимметричной задачи коммивояжера // *Системы управления и информационные технологии* 2017. № 3 (69). С. 88–92.
16. *Головешкин В.А., Жукова Г.Н., Ульянов М.В., Фомичев М.И.* Распределение логарифма сложности индивидуальных задач коммивояжера при фиксированной длине входа // *Современ. информ. технологии и ИТ-образование*. 2016. Т. 12. № 3. Ч. 2. С. 131–137.

17. *Knuth D.E.* Estimating the Efficiency of Backtracking Programs // *Math. Comput.* 1975. V. 29. P. 121–136.
18. *Пантелеев А.В.* Метаэвристические алгоритмы поиска глобального экстремума. М.: Изд-во МАИ-ПРИНТ, 2009.
19. *Applegate D.L., Bixby R.E., Chavatal V. and Cook W.J.K.* The Traveling Salesman Problem, Princeton, N.J.: PUP, 2006.
20. *Colorni A., Dorigo M., Maniezzo V.* Distributed Optimization by Ant Colonies. Proc. First Eur. Conf. on Artificial Life, Paris, France, F. Varela and P. Bourguine (Eds.), Elsevier Publishing, 1991. P. 134–142.
21. *Dorigo M., Gambardella L. M.* Ant colony system: a cooperative learning approach to the traveling salesman problem//*IEEE Trans. Evol. Comput.* Apr. 1997. V. 1. No. 1. P. 53–66.
22. *Bonavear E., Dorigo M.* Swarm Intelligence: from Natural to Artificial Systems, Oxford, the UK: OUP, 1999.
23. *Gamboa D., Rego C., Glover F.* Data Structures and Ejection Chains for Solving Large Scale Traveling Salesman Problems//*Eur. J. Oper. Res.* 2005. V. 160 (1). P. 154–171.
24. *Kaplan H., Lewenstein M., Shafrir N., Sviridenko M.* Approximation algorithms for asymmetric TSP by decomposing directed regular multigraphs//*J. ACM (JACM)*. 2005. V. 52. No. 4. P. 602–626.
25. *Mötnke T., Svensson O.* Removing and Adding Edges for the Traveling Salesman Problem// *J. ACM (JACM)*. 2016. V. 63. No. 1. P. 1–28.
26. *Бородин В.В., Ловецкий С.Е., Меламед И.И., Плотинский Ю.М.* Экспериментальное исследование эффективности эвристических алгоритмов решения задачи коммивояжера // *АиТ.* 1980. № 11. С. 76–84.
Borodin V.V., Lovetskij S.E., Melamed I.I., Plotinskij Yu.M. Experimental Study of the Efficiency of Heuristic Algorithms for Solution of the Traveling Salesman Problem // *Autom. Remote Control.* 1980. V. 41. No. 11. P. 1543–1550.
27. *Cormen T.H., Leiserson C.E., Rivest R.L., Stein C.* Introduction to Algorithm, 3rd ed. Cambridge. MA: MIT Press, 2009.
28. *Stutzle T., Hoos H.H.* Max-Min ant system// *Future Generat. Comput. Syst.* 2000. V. 16. P. 889–914.
29. *Lin S., Kernighan B.W.* An Effective Heuristic Algorithm for the Traveling-Salesman Problem // *Oper. Res.* 1973. V. 21. P. 498–516.
30. *Helsgaun K.* An effective implementation of the Lin–Kernighan traveling salesman heuristic // *Eur. J. Oper. Res.* Elsevier. 2000. V. 126(1). P. 106–130.
31. *Жукова Г.Н., Ульянов М.В., Фомичев М.И.* Эффективный по времени точный комбинированный алгоритм для асимметричной задачи коммивояжера. *Бизнес-информатика.* 2018. № 3 (45). С. 20–28.
32. *Moors J.J.A.* A Quantile Alternative for Kurtosis// *The Statist.* 1988. V. 37. P. 25–32.
33. *Moors J.J.A., Wagemakers R.Th.A., Coenen V.M.J., et al.* Characterizing Systems of Distributions by Quantile Measures// *Statist. Neerlandica.* 1996. V. 50. No. 3. P. 417–430.
34. *Головешкин В.А., Жукова Г.Н., Ульянов М.В., Фомичев М.И.* Вероятностный прогноз сложности индивидуальных задач коммивояжера на основе идентификации распределения сложности по экспериментальным данным // *АиТ.* 2018. № 7. С. 149–166.

Goloveshkin V.A., Zhukova G.N., Ulyanov M.V., Fomichev M.I. Probabilistic Prediction of the Complexity of Traveling Salesman Problems Based on Approximating the Complexity Distribution from Experimental Data // Autom. Remote Control. 2018. V. 79. No. 7. P. 1296–1310.

35. *Крамер Г.* Математические методы статистики. М.: Мир, 1975.

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

Поступила в редакцию 14.12.2018

После доработки 04.04.2019

Принята к публикации 25.04.2019