ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь—июнь 2001. Серия 2. Том 8, № 1, 22–39

УДК 519.8

О НЕКОТОРЫХ РЕЗУЛЬТАТАХ ДЛЯ ЗАДАЧИ КОММИВОЯЖЕРА НА МАКСИМУМ*)

Э. Х. Гимади, $\boxed{\text{А. И. Сердюков}}$

Приводится краткий обзор исследований приближенных алгоритмов с оценками для задачи коммивояжера на максимум (MAX TSP). Основное внимание обращается на специальные классы задачи MAX TSP: симметрическая, полуметрическая, метрическая (симметрическая и полуметрическая), полиэдральная и евклидова. Для задачи со случайными входами представлены результаты вероятностного анализа двух алгоритмов, использующих эвристики «Иди в удаленный город» и «Склейка контуров».

Введение

В статье приводится краткий обзор исследований по разработке приближенных алгоритмов с оценками для задачи коммивояжера на максимум (MAX TSP).**) Основной акцент сделан на работах авторов настоящей статьи.***)

Задача MAX TSP заключается в следующем. Дана произвольная матрица (c_{ij}) порядка n с неотрицательными вещественными элементами. Требуется найти такую перестановку $(i_1, i_2, \ldots, i_n)$ множества $\{1, 2, \ldots, n\}$, что сумма $c_{i_1 i_2} + c_{i_2 i_3} + \ldots + c_{i_{n-1} i_n} + c_{i_n i_1}$ принимает максимальное значение.

Из работы [27] (S. Kosaraju, J. Park, C. Stein) известны приложения MAX TSP к анализу последовательностей ДНК и сжатию данных.

---

*) Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект 99-01-00601) и Федеральной целевой программы «Интеграция» (проект 274).

**) Представлен в качестве пленарного доклада на Третьей Международной конференции «Дискретный анализ и исследование операций» (DAOR'2000) (Новосибирск, 26 июня — 1 июля 2000 г.) [23].

***) Предполагается, что работы других авторов будут более подробно отражены в готовящейся книге «The traveling salesman problem and its variation» (edited by G. Gutin and A. Punnen), глава «The maximum traveling salesman problem» (Александр Барвинок совместно с авторами данного обзора).

Будем говорить, что алгоритм имеет оценку точности ρ, если отношение длины найденного гамильтонова тура к длине оптимального маршрута не меньше ρ. Рассматриваемая задача MAX SNP-трудна [17, 29], т. е. из существования полиномиальной аппроксимационной схемы (PTAS) для ее решения следует, что $P = NP$. Это обстоятельство стимулировало исследования приближенных полиномиальных алгоритмов с оценками точности их решения: 1/2 (M. Fisher, G. Nemhauser, L. Wolsey [20]), 4/7 (М. М. Ковалев, В. М. Котов [9]), 38/63 (S. Kosaraju, J. Park, C. Stein [27]).

Ниже основное внимание обращается на специальные классы задачи MAX TSP: симметрическая, полуметрическая, метрическая (т. е. симметрическая и полуметрическая), полиэдральная и евклидова. Для задачи со случайными входами представлены результаты вероятностного анализа двух алгоритмов, использующих эвристики «Иди в самый удаленный город» и «Склейка контуров».

Подмножество $E' \subset E$ ребер неориентированного графа $G = (V, E)$ называется *паросочетанием* (*совершенным паросочетанием*), если каждая вершина $v \in V$ инцидентна не более (соответственно ровно) одному ребру. Подмножество $E' \subset E$ ребер неориентированного графа G называется 2-*фактором*, если каждая вершина $v \in V$ инцидентна ровно двум ребрам в E'. Подграф $G' = (V, A')$ ориентированного графа, $A' \subset A$, называется *контурным покрытием*, если каждая вершина подграфа G' имеет полустепени захода и исхода, равные 1. Для отыскания каждого указанного подграфа (паросочетания, 2-фактора, контурного покрытия) максимального веса известны алгоритмы с временной сложностью $O(n^3)$ [6, 21, 22, 24, 28].

1. Полуметрическая и метрическая MAX TSP

Пусть $G = (V, A)$ — полный ориентированный граф с n вершинами (без петель) с заданной весовой функцией c на его дугах, которая удовлетворяет неравенству треугольника

$$c(v_1, v_2) + c(v_2, v_3) \geqslant c(v_1, v_3)$$

для любых трех вершин v_1, v_2, v_3 из V. В случае выполнения неравенства треугольника MAX TSP называется *полуметрической*. Полуметрическую MAX TSP называем *метрической*, если $c(v_1, v_2) = c(v_2, v_1)$ для каждой пары $v_1, v_2 \in V$.

Приведем результат, следующий из работ [10] и [15].

Утверждение 1. *Пусть в графе G построено контурное покрытие*

$$C = \{C_1, C_2, \ldots, C_\mu\}.$$

Положим $\nu = \nu(\mathrm{C}) = \min\{\nu_i \mid 1 \leqslant i \leqslant \mu\}$, где $\nu_i = |C_i|$ — число вершин в C_i. Тогда за время $O(\nu n)$ в графе G могут быть построены два гамильтоновых контура H_1 и H_2 с весами

$$c(H_1) \geqslant \left(1 - \frac{1}{2\nu}\right)c(C) \quad \text{(А. В. Косточка, А. И. Сердюков [10])}, \tag{1}$$

$$c(H_2) \geqslant \left(1 - \frac{1}{n}\right)^{\mu-1} c(C) \quad \text{(А. И. Сердюков [15])}. \tag{2}$$

Легко видеть, что оценка (2) лучше оценки (1) при

$$\nu \leqslant \frac{0.5}{1-(1-1/n)^{\mu-1}},$$

что заведомо выполняется, например, при $n \geqslant 2\,\nu\,\mu$.

Полуметрическая задача нахождения максимального контурного покрытия полиномиально разрешима при $\nu = 2$ (например, с помощью алгоритма решения задачи об оптимальном назначении [6]), а метрическая задача — при $\nu = 3$ (например, с использованием алгоритма отыскания 2-фактора [28]). В обоих случаях временная сложность одинакова и равна $O(n^3)$. Отсюда и из (1) имеем

Следствие 1. *Полуметрическая* MAX TSP *может быть решена за время $O(n^3)$ с оценкой точности* 3/4.

Следствие 2. *Метрическая* MAX TSP *может быть решена за время $O(n^3)$ с оценкой точности* 5/6.

М. М. Ковалевым и В. М. Котовым [8] для метрического случая предложен алгоритм с оценкой точности $(\nu+2)/(\nu+3)$, которая совпадает с оценкой (1) при $\nu = 3$ и хуже при $\nu > 3$.

2. Симметрическая MAX TSP

Задача MAX TSP на неориентированном графе $G = (V_n, E)$ называется *симметрической*. В этом случае $c(v_1, v_2) = c(v_2, v_1)$ для любых вершин $v_1, v_2 \in V_n$.

Для симметрической MAX TSP построены полиномиальные алгоритмы с оценками точности 2/3 (M. Fisher, G. Nemhauser, L. Wolsey [20]) и 13/18 (М. Ковалев, В. Котов [7]). R. Hassin и S. Rubinstein в [25] предложили полиномиальный алгоритм с оценкой точности 5/7. Наилучший по точности алгоритм решения симметрической MAX TSP предложил А. И. Сердюков [13].

Утверждение 2 [13]. *Симметрическая* MAX TSP *решается за время $O(n^3)$ с оценкой точности* 3/4.

Для получения данного приближения используются решения двух полиномиально разрешимых задач: задачи отыскания в графе G паросочетания максимального веса и задачи отыскания 2-фактора максимального веса.

Совокупность ребер в решениях обеих задач представима в виде объединения двух частичных туров. (*Частичным туром* называем множество попарно непересекающихся путей в графе.) Частичный тур максимального веса дополняется ребрами до гамильтонова цикла. В основе алгоритма решения задачи лежит процедура P (для случая четного n), а также ее модификация $\widetilde{P}$, используемая в случае нечетного n.

2.1. Процедура P и свойство сохранения частичного тура

Начальная фаза.

— В G отыскиваются максимальные по весу 2-фактор $C = C(G) = C_1 \cup \ldots \cup C_r$ и паросочетание $M = M(G)$.

— Устанавливаются множества $T_1 = C$, $T_2 = M$.

Общий шаг (обработка очередного цикла C_i, $1 \leqslant i \leqslant r$).

— Фиксируются произвольная вершина $v' \in C_i$ и инцидентные ей ребра $e_1, e_2 \in C_i$.

— Из T_1 удаляется, а к T_2 добавляется то ребро $e' \in \{e_1, e_2\}$, которое сохраняет T_2 частичным туром.

Заключительный шаг. Когда все циклы обработаны, полагается

$$T = \begin{cases} T_1, & \text{если } W(T_1) \geqslant W(T_2), \\ T_2 & \text{в противном случае.} \end{cases}$$

Лемма 1. *Результатом работы процедуры P является частичный тур T с длиной*

$$W(T) \geqslant \frac{1}{2}\,(W(C) + W(M)).$$

Доказательство. Покажем, что по окончанию процедуры получены частичные туры T_1 и T_2. Относительно T_1 это очевидно, так как на каждом шаге из очередного цикла удаляется одно ребро. Покажем это для T_2 (по индукции). В начале имеется частичный тур $T_2 = M$. Пусть перед шагом i частичный тур T_2 представлен совокупностью непересекающихся максимальных (по включению) цепей, внутренние вершины которых принадлежат циклам $C_1, \ldots, C_{i-1}$. Очевидно, что вершина $v' \in C_i$, зафиксированная на шаге i, является концом некоторой цепи частичного тура. Обозначим через v'' другой конец этой цепи. Из двух ребер e_1, e_2, инцидентных вершине v', всегда можно выбрать ребро

e', не инцидентное вершине v''. Добавив его к частичному туру T_2, снова получим частичный тур.

Отсюда следует, что на последнем шаге из двух частичных туров T_1 и T_2 выбирается максимальный по весу частичный тур T такой, что

$$W(T) \geqslant \frac{1}{2}\,(W(C) + W(M)).$$

2.2. Модифицированная процедура $\widetilde{P}$

Процедура $\widetilde{P}$ используется в случае нечетного n. Ее отличие от P состоит в выполнении начальной фазы.

C и M находятся как и в P, но теперь имеется вершина v^0, не инцидентная никакому ребру из M. Пусть $v^0 \in C_1$.

Определим ребро $e_0 = (v^0, v') \notin C$ максимального веса $c(e_0)$. Обозначим через $e_1, e_2 \in C_1$ ребра, инцидентные вершине v^0.

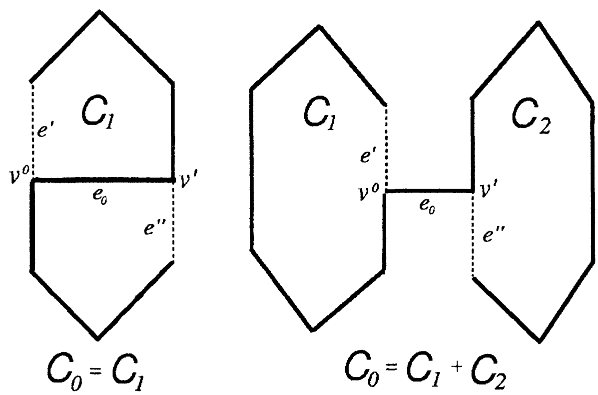

Рис. 1

Пусть (рис. 1)

$$C_0 = \begin{cases} C_1, & \text{если } v' \in C_1, \\ C_1 \cup C_2, & \text{если } v' \in C_2. \end{cases}$$

В $C_0 \setminus M$ выберем ребро $e' \in \{e_1, e_2\}$ и ребро e'', инцидентное вершине v'. В случае $C_0 = C_1$ дополнительно потребуем, чтобы e' было несмежно с e''.

Результатом начальной фазы являются два множества

$$T_1 = e_0 \cup C \setminus \{e', e''\}, \quad T_2 = M \cup \{e', e''\}.$$

Далее выполняются пункты 2 и 3 процедуры P, но общий шаг начинается с $i = 2$ (при $C_0 = C_1$) или с $i = 3$ (при $C_0 = C_1 \cup C_2$).

Замечание. Нетрудно видеть, что в начальной фазе процедуры множество T_2 является частичным туром, а в конце процедуры $\widetilde{P}$ оба множества T_1 и T_2 суть частичные туры, из которых выбирается тур $\widetilde{T}$ максимального веса.

2.3. Описание алгоритма решения симметричной MAX TSP

Этап 1. В случае четного n с помощью процедуры P в графе G находится частичный тур T_0 и осуществляется переход к этапу 5, в противном случае выполняется следующий этап.

Этап 2. С помощью процедуры P находится частичный тур $T(\overline{G})$ в графе $\overline{G} = (V \setminus v^0, \overline{E})$, ребра которого имеют веса

$$\overline{c}(e) = \begin{cases} c(e), & \text{если } e \neq \overline{e}, \\ c(e_1) + c(e_2) & \text{в противном случае,} \end{cases}$$

где ребра $e_1, e_2, \overline{e}$ образуют цикл.

Полагается

$$\overline{T} = \begin{cases} T(\overline{G}) \cup \{e_1, e_2\} \setminus \overline{e}, & \text{если } \overline{e} \in T(\overline{G}), \\ T(\overline{G}) & \text{в противном случае.} \end{cases}$$

Этап 3. С помощью модифицированной процедуры $\widetilde{P}$ в графе G строится частичный тур $\widetilde{T}$.

Этап 4. Полагается

$$T_0 = \begin{cases} \overline{T}, & \text{если } W(\overline{T}) \geqslant W(\widetilde{T}), \\ \widetilde{T} & \text{в противном случае.} \end{cases}$$

Этап 5. Частичный тур T_0 достраивается до гамильтонова цикла H.

Описание алгоритма закончено.

2.4. Обоснование оценки 3/4

Обозначим через H^* гамильтонов цикл максимального веса в графе G и через e^* ребро минимального веса в цикле H^*. При четном n с учетом леммы 1 имеем

$$W(H) \geqslant W(T_0) \geqslant \frac{1}{2}\,(W(S) + W(M)) \geqslant \frac{3}{4}\,W(H^*).$$

При нечетном n возможны два случая.

Случай 1. $e_1 \in H^*$ и $e_2 \in H^*$. В этом случае $\overline{W}(\overline{H}^*) = W(H^*)$, где $\overline{H}^*$ — гамильтонов цикл максимального веса в графе $\overline{G}$, содержащем ребро $\overline{e}$. Имеем

$$W(H) \geqslant W(\overline{T}) \geqslant \frac{1}{2}(\overline{W}(\overline{C}) + \overline{W}(\overline{M})) \geqslant \frac{3}{4}\overline{W}(\overline{H}^*) = \frac{3}{4}W(H^*),$$

где $\overline{C} = C(\overline{G})$ и $\overline{M} = M(\overline{G})$.

Случай 2. По крайней мере одно из ребер e_1, e_2 не принадлежит циклу H^*. В этом случае с учетом неравенств

$$W(C) \geqslant W(H^*) \quad \text{и} \quad W(M) \geqslant \frac{1}{2}\,(W(H^*) - c(e^*))$$

имеем

$$W(H) \geqslant W(\widetilde{T}) \geqslant \frac{1}{2}\,(W(C) + W(M) + c(e_0))$$
$$\geqslant \frac{1}{2}\left(W(H^*) + \frac{1}{2}\,(W(H^*) - c(e^*)) + c(e_0)\right) \geqslant \frac{3}{4}\,W(H^*) + \frac{1}{4}(c(e_0) - c(e^*)).$$

Остается убедиться, что $c(e_0) \geqslant c(e^*)$. Действительно, в случае $e_0 \in H^*$ последнее неравенство очевидно. При $e_0 \notin H^*$ в рассматриваемом случае имеется по крайней мере одно ребро $e' \notin C$, инцидентное вершине v^0 и принадлежащее циклу H^*. В силу определения ребра e^0 имеем $c(e^0) \geqslant c(e') \geqslant c(e^*)$. Утверждение 2 доказано.

С использованием техники алгоритмов [13] и [25] R. Hassin и S. Rubinstein в [26] предложили полиномиальный рандомизированный алгоритм с ожидаемой оценкой точности ρ для любого $\rho < 25/33$ (что чуть-чуть лучше $3/4$).

3. Полиэдральная MAX TSP

Пусть в евклидовом пространстве $\mathbf{R}^k$ со стандартным скалярным произведением $\langle\cdot,\cdot\rangle$ заданы векторы $a_1, \ldots, a_r$. Для определения функции расстояния между точками в пространстве введем единичный шар (содержащий начало координат внутри)

$$B = \{x \in \mathbf{R}^k \mid \langle a_i, x\rangle \leqslant 1,\ i = 1, \ldots, r\}.$$

Тогда для упорядоченной пары произвольных точек $(x, y) \in \mathbf{R}^k \times \mathbf{R}^k$ определим расстояние $c(x, y) = f(y - x)$, где $f : \mathbf{R}^k \to \mathbf{R}^+$ есть функционал Минковского от множества B, т. е. $f(x) = \inf\{\lambda \in \mathbf{R} \mid x \in \lambda B,\ \lambda > 0\}$.

Рассмотрим n точек $x^1, \ldots, x^n$ из $\mathbf{R}^k$ таких, что расстояние между x^i и x^j равно $c_{ij} = f(x^j - x^i), 1 \leqslant i, j \leqslant n$. Тогда задачу коммивояжера на максимум для заданного множества точек назовем *полиэдральной* MAX TSP. Через m обозначим число фасет шара B.

До недавнего времени сложностной статус полиэдральной MAX TSP оставался неопределенным. В последнее время группой исследователей (S. Fekete, A. Barvinok, D. Johnson, G. Woeginger, R. Woodroofe) установлены следующие результаты для полиэдральной MAX TSP с центрально-симметрическим (относительно начала координат) и ограниченным шаром B (см., например, [17–19]).

1. Задача NP-трудна в $\mathbf{R}^k$, $k \geqslant 3$, если число фасет единичного шара B является частью ее входа (S. Fekete [19]).

2. A. Barvinok [17] представил полную полиномиальную аппроксимационную схему при фиксированном m. Этот результат позднее был перекрыт в совместной работе [18] (A. Barvinok, D. Johnson, G. Woeginger, R. Woodroofe), в которой предложен точный алгоритм с временной сложностью $O(n^{m-2} \log n)$.

3. В случае норм L_1 («манхеттенская норма») и L_∞ на плоскости $\mathbf{R}^2$ S. Fekete [19] построил точный алгоритм с временной сложностью $O(n)$.

Результаты остаются справедливыми и без требования центральной симметричности и ограниченности множества B.

А. И. Сердюков использовал технику, отличную от техники указанных выше авторов.

Утверждение 3 (А. И. Сердюков [16]). *Пусть G — полный ориентированный граф с множеством вершин $V = \{x^1, \ldots, x^n\}$, расположенных в заданных точках пространства $\mathbf{R}^k$, $m' = \lfloor m/2 \rfloor$ и $C = \{C_1, C_2, \ldots, C_\mu\}$ — максимальное контурное покрытие в графе G, где $\mu > m'$. Тогда за время $O(n^3)$ можно построить контурное покрытие $C' = \{C'_1, C'_2, \ldots, C'_{m'}\}$ в графе G такое, что $c(C') = c(C)$.*

Отсюда и из оценки (2) имеем

Следствие 3. *Для полиэдральной* MAX TSP *в $\mathbf{R}^k$ при фиксированном k может быть построен алгоритм с временной сложностью $O(n^3)$ и оценкой точности, не меньшей чем $1 - (\lfloor m/2 \rfloor - 1)/n$.*

Очевидно, этот алгоритм асимптотически точен при $m = o(n)$.

Для получения указанного приближения используется точное решение (полиномиально разрешимой) задачи отыскания максимального контурного покрытия. На каждом шаге алгоритма (всего $\mu - 1$ шагов) из текущего контурного покрытия (начиная с $C = \{C_1, C_2, \ldots, C_\mu\}$) определенным образом выбираются два контура, которые затем склеиваются в один. При этом, по существу, учитываются свойства функции расстояния $c(x, y)$. Результатом работы алгоритма получается гамильтонов контур с указанной оценкой точности.

Следствие 4. *При $m = 3$ полиэдральная* MAX TSP *в $\mathbf{R}^2$ разрешима за время $O(n^3)$.*

Детали доказательства утверждения 3 и следствий 3 и 4 изложены А. И. Сердюковым в работах [15, 16].

В то время как полиэдральная MAX TSP задача NP-трудна (если число фасет m является частью входа), имеет место

Следствие 5. *При $m = O(\ln n)$ для полиэдральной* MAX TSP *существует полиномиальная аппроксимационная схема* (PTAS) *с временной сложностью $n^{O(1/\varepsilon)}$.*

4. Евклидова MAX TSP

Пусть в евклидовом пространстве $\mathbf{R}^k$ со стандартным скалярным произведением задана евклидова метрика

$$c(x,y) = \left((y_1 - x_1)^2 + \ldots (y_k - x_k)^2\right)^{1/2}, \quad x, y \in \mathbf{R}^k.$$

Пусть в $\mathbf{R}^k$ заданы n точек $x^1, \ldots, x^n$ с указанным попарным расстоянием $c_{ij} = c(x^i, x^j)$, $1 \leqslant i, j \leqslant n$. Тогда задачу коммивояжера на максимум для заданного множества точек назовем *евклидовой* MAX TSP.

Вопрос о сложностном статусе евклидовой MAX TSP до недавнего времени был открыт (и остается открытым при $k = 2$). NP-трудность евклидовой MAX TSP в $\mathbf{R}^k$ при $k \geqslant 3$ установлена S. Fekete [19].

А. И. Сердюков [14] представил асимптотически точный алгоритм решения задачи коммивояжера на максимум в k-мерном евклидовом пространстве $\mathbf{R}^k$.

Утверждение 4 (А. И. Сердюков [14]). *Евклидова* MAX TSP *решается асимптотически точно за время $O(n^3)$ с оценкой точности $1 - \gamma_k n^{-2/(k+2)}$, где константа γ_k зависит только от размерности пространства.*

4.1. Описание алгоритма Сердюкова

Отправной точкой в работе [14] является построение максимального взвешенного паросочетания M в $\mathbf{R}^k$, время отыскания которого и определяет трудоемкость $O(n^3)$ алгоритма в целом.

Найденное паросочетание M представляется в виде совокупности $\{I_1, \ldots, I_m\}$ прямолинейных интервалов (отрезков), $m = \lfloor n/2 \rfloor$. Решающую роль в обосновании асимптотической точности алгоритма сыграл факт существования среди большого числа отрезков в $\mathbf{R}^k$ такой пары отрезков, что угол между ними достаточно мал (стремится к нулю с ростом числа отрезков).

После упорядочения интервалов по убыванию весов последние t интервалов объявляются легкими, остальные — тяжелыми (число t определяется специальным образом).

Очевидно, что суммарный вес тяжелых интервалов оценивается снизу величиной

$$W_1 = W(M)\Bigl(1 - \frac{t}{m}\Bigr). \tag{3}$$

Подход в [14] основывается на следующих геометрических результатах. Для любой пары точек $x, y \in \mathbf{R}^k$ обозначим через $w(x,y) = ||x - y||$.

Лемма 2. *Для любых двух интервалов $I_j = (x_j, y_j)$, $I_k = (x_i, y_i)$ в $\mathbf{R}^k$ справедливы неравенства*

$$w(x_j,x_i) + w(y_j,y_i) + w(x_j,y_i) + w(y_j,x_i) \geqslant 2\max\{(x_j,y_j), w(x_i,y_i)\};$$
$$\max\{w(x_j,x_i) + w(y_j,y_i), w(x_j,y_i) + w(y_j,x_i)\} \geqslant \max\{w(x_j,y_j), w(x_i,y_i)\};$$
$$\max\{w(x_j,x_i)+w(y_j,y_i), w(x_j,y_i)+w(y_j,x_i)\} \geqslant \cos\frac{\alpha}{2}(w(x_j,y_j)+w(x_i,y_i)),$$

где $\alpha \leqslant \pi/2$ — угол между интервалами I_j, I_i.

Лемма 3. *Пусть в евклидовом пространстве $\mathbf{R}^k$ фиксированной размерности k задано произвольное множество из t прямолинейных отрезков. Тогда наименьший угол между двумя отрезками из этого множества ограничен сверху константой $\alpha(k,t)$ такой, что $\alpha(k,t) \to 0$ при $t \to \infty$. При этом*

$$\sin^2\frac{\alpha(k,t)}{2} \leqslant \frac{\gamma_k}{t^{2/(k-1)}}, \tag{4}$$

где константа γ_k не зависит от числа отрезков.

Доказательство. Путем параллельного переноса, сжатия или растяжения добьемся, чтобы каждый рассматриваемый отрезок являлся диаметром единичного шара в $\mathbf{R}^k$. Обозначив через $V_k(r)$ объем k-мерного шара радиуса r в $\mathbf{R}^k$, а через $S^k(r)$ объем его $(k-1)$-мерной поверхности, имеем

$$2t \leqslant \frac{S_{k-1}(1)}{V_{k-1}(\sin\frac{\alpha(k,t)}{2})} = \frac{S_{k-1}(1)}{\sin^{k-1}\frac{\alpha(k,t)}{2}\, V_{k-1}(1)},$$

откуда получаем формулу для константы $\gamma_k = \left(\frac{S_{k-1}(1)}{2V_{k-1}(1)}\right)^{2/(k-1)}$ и неравенство (4).

Определение. Назовем *α-цепью* такую упорядоченную последовательность тяжелых интервалов, что угол между любыми двумя соседними интервалами не превышает числа α.

Лемма 4. *Паросочетание $M = \{I_1, \ldots, I_m\}$ можно представить в виде совокупности $(t-1)$ $\alpha(k,t)$-цепей, перемежаемых t легкими интервалами* (рис. 2, где тяжелые интервалы выделены толстыми отрезками, а легкие — тонкими).

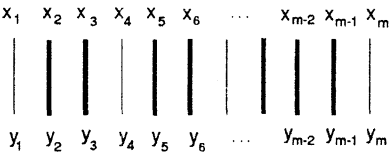

Рис. 2

Доказательство. Путем дублирования ребер из M получается максимальное взвешенное цикловое покрытие S заданных в $\mathbf{R}^k$ точек (вершин неориентированного графа) двухвершинными циклами. Покрытие S состоит из подмножества S_1 двухвершинных тяжелых циклов и подмножества S_2 двухвершинных легких циклов.

Далее S_1 за $m-2t-1$ шагов модифицируется в $\widetilde{S}_1$ методом склеивания циклов. На каждом шаге в текущем (модифицированном) множестве $\widetilde{S}_1$ (в начале $\widetilde{S}_1 = S_1$) выбирается такая система из t тяжелых интервалов, что в каждом цикле содержится не более одного такого интервала. В выбранной системе ищется пара интервалов с наименьшим углом между ними. После этого циклы, содержащие эти интервалы, склеиваются в один. Существование такой пары гарантируется леммой 3.

По окончании $m-2t-1$ шагов процедуры склейки получаем покрытие, состоящее из множеств $\widetilde{S}_1$ и S_2. Но теперь множество $\widetilde{S}_1$ представлено не двухвершинными тяжелыми циклами, как вначале, а $t-1$ модифицированными циклами, каждый из которых соответствует некоторой $\alpha(k,t)$-цепи (состоящей, как отмечено, из тяжелых интервалов).

Требуемая по лемме совокупность получается, если между каждой парой легких интервалов паросочетания M поместить по одной $\alpha(k,t)$-цепи. Лемма 4 доказана.

Заметим,что временная сложность такого представления не превышает $O(nt^2)$.

Далее полученное покрытие используется для построение мультиграфа $\overline{G}$. Этот мультиграф содержит все ребра $I_j = (x_j, y_j)$, $j = 1, \dots, m$, максимального взвешенного паросочетания с кратностью 4, а также дополнительные ребра (изображенные на рис. 3 пунктиром) с кратностью 2, за исключением 8 дополнительных ребер (с кратностью 1), смежных с легкими интервалами I_1 или I_m. На рис. 3 ребра мультиграфа $\overline{G}$ изображены вместе с кратностью их вхождения.

Согласно лемме 2 вершины мультиграфа можно занумеровать так, что для любой пары соседних тяжелых ребер (x_j, y_j) и (x_{j+1}, y_{j+1}), $1 \leqslant j \leqslant m$, справедливо неравенство

$$w(x_j, x_{j+1}) + w(y_j, y_{j+1}) \geqslant \cos\frac{\alpha(d,l)}{2}(w(x_j, y_j) + w(x_{j+1}, y_{j+1})). \qquad (5)$$

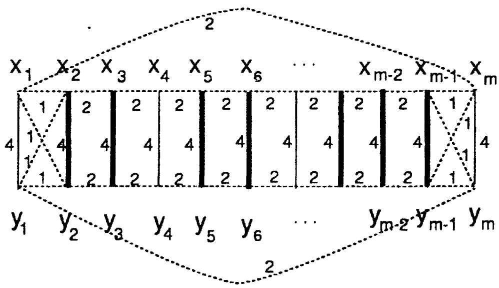

Рис. 3

Если же одно из соседних ребер пары $(I_j,\ I_{j+1})$ легкое (кроме ребер $I_1,\ I_m$), то

$$w(x_j, x_{j+1}) + w(y_j, y_{j+1}) \geqslant \max(w(x_j, y_j), w(x_{j+1}, y_{j+1})). \tag{6}$$

Мультиграф $\overline{G}$ на рис. 3 покрывается четырьмя гамильтоновыми циклами H_i, $i = 1, 2, 3, 4$. В случае нечетного m они имеют вид

$$\begin{aligned} H_1 &= (x_1, y_1, y_2, x_2, x_3, y_3, \ldots, y_{m-1}, x_{m-1}, y_m, x_m, x_1), \\ H_2 &= (y_1, x_1, y_2, x_2, x_3, y_3, \ldots, y_{m-1}, x_{m-1}, x_m, y_m, y_1), \\ H_3 &= (x_1, y_1, x_2, y_2, y_3, x_3, \ldots, x_{m-1}, y_{m-1}, y_m, x_m, x_1), \\ H_4 &= (y_1, x_1, x_2, y_2, y_3, x_3, \ldots, x_{m-1}, y_{m-1}, x_m, y_m, y_1). \end{aligned}$$

(В случае четного m последовательность вершин в циклах слегка видоизменяется.) Каждый цикл содержит все ребра максимального паросочетания M. В качестве приближенного решения среди H_i, $i = 1, 2, 3, 4$, выбирается гамильтонов цикл, имеющий наибольший вес. Этот цикл обозначим через H^0.

4.2. Доказательство асимптотической точности

Ясно, что

$$\begin{aligned} W(H^*) \geqslant W(H^0) &\geqslant \frac{1}{4}(W(H_1) + W(H_2) + W(H_3) + W(H_4)) \\ &= \frac{1}{4}\Big(4W(M) + W' + 2\sum_{j=2}^{m-2} (w(x_j, x_{j+1}) + w(y_j, y_{j+1})) + W''\Big), \end{aligned}$$

где W' и W'' — суммарные веса дополнительных ребер, смежных легким ребрам I_1 и I_m соответственно.

Из (3) и леммы 2 следует, что

$$W(H^0) \geqslant W(M) + W_1 \cos\frac{\alpha(k,t)}{2} = W(M)\Big(1 + \Big(1 - \frac{t}{m}\Big)\cos\frac{\alpha(k,t)}{2}\Big).$$

С учетом леммы 3, а также соотношений $2W(M) \geqslant W(H^*)(1-1/n)$ и $\cos\alpha = 1 - 2\sin^2\alpha/2$ получаем

$$\rho_A = \frac{W(H^0)}{W(H^*)} \geqslant \frac{1}{2}\Big(1 + \Big(1 - \frac{t}{m}\Big)\Big(1 - 2\sin^2\frac{\alpha(k,t)}{4}\Big)\Big)(1 - 1/n)$$
$$\geqslant 1 - \frac{t+1}{n} - \sin^2\frac{\alpha(k,t)}{4} \geqslant 1 - \frac{t+1}{n} - \frac{\gamma_k}{2} t^{\frac{-2}{k-1}}.$$

Окончательно, положив $t = \lfloor n^{\frac{k-1}{k+1}} \rfloor$, убеждаемся в асимптотической точности алгоритма [34]: если $n \to \infty$, то

$$\rho_A \geqslant 1 - \beta_k\ n^{\frac{-2}{k+1}} \to 1,$$

где константа β_k зависит только от размерности пространства.

Следствие 6. *Для евклидовой* MAX TSP *существует* PTAS *с временной сложностью*

$$O(n^3 + n^2\ 2^{n(k,\varepsilon)}),$$

где $n(k,\varepsilon) = (\beta_k/\varepsilon)^{(k+1)/2}$.

5. Вероятностный анализ алгоритма «Иди в самый удаленный город»

Предположим, что на входе задана матрица расстояний (c_{ij}), элементы которой являются независимыми случайными величинами, выбираемыми из сегмента $[a_n, b_n]$, где $a_n \geqslant 0$, с одинаковой функцией распределения $P(x) = \mathbf{P}r\,\{\xi < x\}$ случайной величины $\xi = (c_{ij} - a_n)/(b_n - a_n)$, $0 \leqslant \xi \leqslant 1$. Пусть $\mathscr{A}$ обозначает алгоритм «Иди в самый удаленный (непройденный) город» на соответствующем классе задач MAX TSP (временная сложность алгоритма $O(n^2)$).

Пусть K_n — вероятностное пространство входных данных задачи MAX TSP с n городами. Следуя [2], говорим, что алгоритм имеет *оценки качества работы* $(\varepsilon_n,\ \delta_n)$ на задаче из класса K_n, если вероятность нахождения решения с оценкой точности $1-\varepsilon_n$ не меньше $1-\delta_n$. Алгоритм называется *асимптотически точным*, если существуют оценки $(\varepsilon_n,\ \delta_n)$ такие, что $\varepsilon_n \to 0$ и $\delta_n \to 0$ при $n \to \infty$.

Алгоритм «Иди в самый удаленный город» можно отнести к классу «жадных» алгоритмов. Для заданной матрицы расстояний (c_{ij}) алгоритм начинает свою работу с произвольного города, затем на каждом шаге осуществляется переход в самый удаленный из оставшихся (непройденных) городов и т. д. На последнем шаге происходит возврат в исходный город.

5.1. Равномерное распределение

Используя статистику Колмогорова–Смирнова, R. Vohra [30] в случае равномерного распределения ($P(x) = x,\ 0 \leqslant x \leqslant 1$) показал, что для всякого $\delta > 0$ существует такая положительная константа α, что для решения, полученного с помощью алгоритма $\mathscr{A}$, справедливо следующее вероятностное неравенство:

$$\mathbf{P}r\left\{\frac{F_{\mathscr{A}}(I)}{F^*(I)} \geqslant 1 - \alpha\sqrt{\frac{\ln(1/\delta)}{n}}\right\} \geqslant 1 - \delta. \tag{7}$$

Положим $\alpha = \sqrt{2}$, $\delta = n^{\lambda/2}$, где λ — положительная константа. Тогда (7) может быть записано в виде

$$\mathbf{P}r\left\{\frac{F_{\mathscr{A}}(I)}{F^*(I)} \geqslant 1 - \sqrt{\frac{\lambda \ln n}{n}}\right\} \geqslant 1 - n^{\lambda/2}. \tag{8}$$

Понятно, что в этом случае алгоритм $\mathscr{A}$ асимптотически точен.

5.2. Распределения, отличные от равномерного

Далее остановимся на распределениях, отличных от равномерного. Пусть

$$\psi_n = \int\limits_0^{1-1/n} \frac{dx}{1 - P(x)}.$$

Утверждение 4 [1, 4]. *Пусть выполнено одно из следующих условий*:

a) $\psi_n = o(n)$ *и* $\psi_n \to \infty$ *при* $n \to \infty$;

b) $P(x) > x$ *при* $0 \leqslant x \leqslant 1$ *и* $\psi_n = o(n)$;

c) $P(x) \leqslant x$ *при* $0 \leqslant x \leqslant 1$.

Тогда алгоритм А является асимптотически точным.

Доказательство утверждения основано на следующем. Сначала оценивается ожидаемый вес маршрута. Затем с использованием неравенства Чебышёва показывается, что типичные значения веса близки к тривиальной верхней оценке nb_n.

Следует заметить, что в отличие от алгоритма «Иди в ближайший (непройденный) город» в задаче на минимум (Э. Х. Гимади, В. А. Перепелица [5, 11]) для асимптотической точности алгоритма «Иди в самый удаленный город» дополнительные ограничения на величину разброса коэффициентов матрицы (c_{ij}) (типа $b_n/a_n = o(n/\ln n)$) не накладываются.

5.3. Вероятности больших уклонений

Используя неравенство Петрова [12] для вероятностей больших уклонений, можно получить существенно лучшую по сравнению с [30] оценку относительной погрешности решения $\varepsilon_n = O(n^{-1}\ln(1/\delta_n))$, причем на более широком классе вероятностных распределений $P(x) \leqslant x$, $0 \leqslant x \leqslant 1$.

Утверждение 5. *Для задачи MAX TSP из класса K_n с функцией распределения $P(x) \geqslant x$, $0 \leqslant x \leqslant 1$, алгоритм $\mathscr{A}$ является асимптотически точным с оценками качества*

$$\varepsilon_n = \frac{\lambda \ln n}{n}, \quad \delta_n = n^{-\lambda/2},$$

где λ — положительная константа.

Доказательство может быть получено из следующего утверждения, вытекающего из теоремы Петрова [12, с. 81].

Лемма 5. *Пусть $\xi_{ki} \in (0,1)$ — случайные независимые равномерно распределенные величины и X_k — независимая случайная переменная, равная $\min\{\xi_{ki} \mid 1 \leqslant i \leqslant k\}$. Тогда для суммы S случайных величин $\xi_k = X_k - \mathbf{E}X_k$, $k = 1,\dots,n-1$, $\xi_n = \xi_1$, справедливо неравенство*

$$\mathbf{P}r\{S > \lambda \ln n\} \leqslant n^{-\lambda/2}.$$

6. Вероятностный анализ алгоритма «Склейка контуров»

В статье [3] предложен двухэтапный алгоритм решения задачи коммивояжера с временной сложностью $O(n^3)$. Основу первого этапа составляет рандомизированный вариант алгоритма для точного решения задачи о назначениях. На втором этапе это решение преобразуется в некоторое приближенное решение задачи MAX TSP с использованием эвристики «Склейка контуров».

В отличие от других известных работ, где входные данные задачи предполагаются независимыми случайными величинами, в работе Э. Х. Гимади, Н. И. Глебова, А. И. Сердюкова [3] вероятностный анализ алгоритма проводится для таких вероятностных распределений на множестве входов (матриц расстояний между городами), относительно которых столбцы случайной матрицы являются последовательностью симметрично зависимых случайных величин. Обосновываются оценки относительной погрешности и вероятности несрабатывания алгоритма. При некоторых дополнительных ограничениях на величину коэффициента разброса элементов матрицы расстояний устанавливается асимптотическая точность алгоритма для класса вероятностных распределений, более широкого по сравнению с рассматривавшимися ранее.

Утверждение 6 [3]. *Пусть столбцы матрицы расстояний* (c_{ij}) *образуют последовательность симметрично зависимых случайных величин. Тогда задача* MAX TSP *может быть решена с временной сложностью* $O(n^3)$ *и оценками качества*

$$\varepsilon_n = 2c^* \ln n / F^*_{AP}, \quad \delta_n = (e/n)^{0.38}, \tag{9}$$

где c^* — *максимальный элемент матрицы* (c_{ij}) *и* F^*_{AP} — *оптимальное значение функционала задачи о назначениях на максимум для матрицы* (c_{ij}).

Следствие 7. *Если* $a_n \leqslant c_{ij} \leqslant b_n$, $i \neq j$, $1 \leqslant i, j \leqslant n$; $a_n > 0$, *то решение с оценками* (9) *является асимптотически точным при*

$$\frac{b_n}{a_n} = o\Big(\frac{n}{\ln n}\Big).$$

ЛИТЕРАТУРА

1. **Гимади Э. Х.** Задача коммивояжера на максимум: условия асимптотической точности алгоритма «Иди в самый удаленный город» // Управляемые системы: Сб. науч. тр. Новосибирск: Ин-т математики СО АН СССР, 1989. Вып. 29. С. 11–15.
2. **Гимади Э. Х., Глебов Н. И., Перепелица В. А.** Алгоритмы с оценками для задач дискретной оптимизации // Проблемы кибернетики. М.: Наука, 1975. Вып 31. С. 35–42.
3. **Гимади Э. Х., Глебов Н. И., Сердюков А. И.** Алгоритм для приближенного решения задачи коммивояжера и его вероятностный анализ // Сиб. журн. исслед. операций. 1994. Т. 1, № 2. С. 8–17.
4. **Гимади Э. Х., Максишко Н. М.** Обоснование условий асимптотической точности приближенного алгоритма решения задачи коммивояжера на максимум в случае дискретного распределения // Управляемые системы: Сб. науч. тр. Новосибирск: Ин-т математики СО АН СССР, 1990. Вып. 30. С. 26–29.
5. **Гимади Э. Х., Перепелица В. А.** Асимптотически точный подход к решению задачи коммивояжера // Управляемые системы: Сб. науч. тр. Новосибирск: Ин-т математики СО АН СССР, 1974. Вып. 12. С. 35–45.
6. **Диниц Е. А., Кронрод М. А.** Один алгоритм решения задачи о назначении // Докл. АН СССР. 1969. Т. 189, № 1. С. 23–25.
7. **Ковалев М. М., Котов В. М.** Субоптимальные алгоритмы для решения задачи коммивояжера // Вестн. Белорус. ун-та. Сер. физика, математика, механика. 1982. № 1. С. 1–31.
8. **Ковалев М. М., Котов В. М.** Серия эвристик // 30 Междунар. науч. коллоквиум. Ильменау (ГДР), 1985. С. 53–56.

9. **Ковалев М. М., Котов В. М.** Оценки погрешности серий приближенных алгоритмов // Вестн. Белорус. ун-та, 1986. Вып. 3(1). С. 44–50.

10. **Косточка А. В., Сердюков А. И.** Полиномиальные алгоритмы с оценками 3/4 и 5/6 для задачи коммивояжера на максимум // Управляемые системы: Сб. науч. тр. Новосибирск: Ин-т математики СО АН СССР, 1985. Вып. 26. С. 55–59.

11. **Перепелица В. А., Гимади Э. Х.** К задаче нахождения минимального гамильтонова контура на графе со взвешенными дугами // Дискретный анализ: Сб. науч. тр. Новосибирск: Ин-т математики СО АН СССР, 1969. Вып. 15. С. 57–65.

12. **Петров В. В.** Предельные теоремы для суммы независимых случайных величин. М.: Наука, 1987.

13. **Сердюков А. И.** Алгоритм с оценкой для задачи коммивояжера на максимум // Управляемые системы: Сб. науч. тр. Новосибирск: Ин-т математики СО АН СССР, 1984. Вып. 25. С. 80–86.

14. **Сердюков А. И.** Асимптотически точный алгоритм для задачи коммивояжера на максимум в евклидовом пространстве // Управляемые системы: Сб. науч. тр. Новосибирск: Ин-т математики СО АН СССР, 1987. Вып. 27. С. 79–87.

15. **Сердюков А. И.** Полиномиальные алгоритмы с оценками точностей решений для одного класса задач коммивояжера на максимум // Комбинаторно-алгебраические методы в дискретной оптимизации. Нижний Новгород: Нижегород. гос. ун-т, 1991. С. 107–114.

16. **Сердюков А. И.** Задача коммивояжера на максимум в конечномерных вещественных пространствах // Дискрет. анализ и исслед. операций. 1995. Т. 2, № 1. С. 50–56.

17. **Barvinok A. I.** Two algorithmic results for the traveling salesman problem // Math. Oper. Res. 1996. V. 21, N 1. P. 65–84.

18. **Barvinok A. I., Johnson D. S., Woeginger G., Woodroofe R.** The maximum traveling salesman problem under polyhedral norms // Integer Programming and Combinatorial Optimization. Berlin: Springer, 1998. P. 195–201. (Lecture Notes in Comput. Sci.; V. 1412).

19. **Fekete S. P.** Simplicity and hardness of the maximum travelling salesman problem under geometric distances // Proc. of the Tenth Annual ACM–SIAM Symposium on Discrete Algorithms (Baltimore, MD, 1999). New York: ACM, 1999. P. 337–345.

20. **Fisher M. L., Nemhauser G. L., Wolsey L. A.** An analysis of approximations for finding a maximum weight Hamiltonian circuit // Oper. Res. 1979. V. 27, N 6. P. 799–809.

21. **Gabow H. N.** An efficient reduction technique for degree-constrained subgraph and bidirected network flow problems // Proc. of the 15th Annual ACM Symposium on Theory of Computing. New York: ACM, 1983. P. 448–456.
22. **Gabow H. N.** Data structures for weighted matching and nearest common ancestors with linking // Proc. of the First Annual ACM–SIAM Symposium on Discrete Algorithms. Philadelphia, PA: SIAM, 1990. P. 434–443.
23. **Gimadi E. Kh., Serdyukov A. I.** The traveling salesman problem // Материалы конф. DAOR'2000. Новосибирск: Изд-во Ин-та математики, 2000. С. 22–27.
24. **Hartvigsen D.** Extensions of matching theory: PhD Thesis. Pittsburg, PA: Carnegie Mellon Univ., 1984.
25. **Hassin R., Rubinstein S.** An approximation algorithm for the maximum traveling salesman problem // Inform. Process. Lett. 1998. V. 67, N 3. P. 125–130.
26. **Hassin R., Rubinstein S.** Better approximations for Max TSP // Inform. Process. Lett. 2000. V. 75. P. 181–186.
27. **Kosaraju S. R., Park J. K., Stein C.** Long tours and short superstrings // 35st Annual IEEE Symposium on Foundation of Computer Science. Los Alamitos, CA: Comput. Soc. Press, 1994. P. 166–177.
28. **Lovasz L., Plummer M. D.** Matching theory. Amsterdam: North-Holland Publ. Co, 1986. (Annals of Discrete Mathematics; V. 29).
29. **Papadimitriu C. H., Yannakakis M.** The traveling salesman problem with distance one and two // Math. Oper. Res. 1993. V. 18. N 1. P. 1–11.
30. **Vohra R. V.** Probabilistic analysis of the longest Hamiltonia tour problem // Networks. 1988. V. 18, N 1. C. 13–18.

Адрес автора:

Институт математики
им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4,
630090 Новосибирск, Россия.
E-mail: gimadi@math.nsc.ru

Статья поступила
8 января 2001 г.