ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Апрель—июнь 2000. Серия 1. Том 7, № 2, 3–11

УДК 519.854

ОБ ОДНОМ АЛГОРИТМЕ НАХОЖДЕНИЯ МИНИМАЛЬНОГО ОСТОВА С ОГРАНИЧЕННЫМ СНИЗУ ДИАМЕТРОМ*)

Э. Х. Гимади, А. И. Сердюков

Рассматривается задача нахождения остовного дерева минимального веса с ограниченным снизу диаметром в полном n-вершинном взвешенном неориентированном графе. В общем случае эта задача NP-трудна. Предлагается алгоритм решения задачи с временно́й сложностью $O(n^2)$. В предположении, что веса ребер назначаются случайно, независимо и равновероятно из отрезка (a_n, b_n) при $a_n > 0$ (в непрерывном случае) или из множества $\{1, \dots, r_n\}$ (в дискретном случае), проводится вероятностный анализ предложенного алгоритма. Отдельно рассматриваются входы задачи, когда $b_n/a_n \leqslant n/\psi_n$ (соответственно при $r_n \leqslant n/\psi_n$) и нижняя оценка диаметра остовного дерева не превышает $d_n = n - n/\psi_n$ (где $\psi_n = o(n)$ и $\psi_n \to \infty$ при $n \to \infty$). В этом случае алгоритм с вероятностью, большей $1 - \exp(-0,25n/\psi_n)$, имеет относительную погрешность, не превышающую величины $\ln \psi_n/\psi_n$, т. е. является асимптотически точным.

Пусть задан полный неориентированный граф $G_n = (V, E)$ с множеством вершин $V = \{1, \dots, n\}$, в котором каждому ребру $e = (i, j) \in E$ приписана стоимость (вес) w_{ij}, $1 \leqslant i < j \leqslant n$.

Введем следующие обозначения:

$\mathscr{D}_n = \{D_n\}$ — множество всех остовных деревьев в G_n;

$w(G')$ — вес подграфа G' (сумма весов его ребер) графа G_n;

$d(D_n)$ — диаметр дерева D_n (длина максимальной относительно числа ребер цепи в D_n);

$\{d_n\}, n = 1, 2, \dots,$ — некоторая последовательность натуральных чисел;

$\widetilde{\mathscr{D}}_n = \{D_n \in \mathscr{D}_n \mid d(D_n) \geqslant d_n\}$.

Задача заключается в нахождении в графе G_n остовного дерева D_n минимального веса в множестве $\widetilde{\mathscr{D}}_n$.

*) Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (коды проектов 99–01–00601 и 97–01–00890).

Заметим, что эта задача NP-трудна при $d_n = n-1$ как классическая задача нахождении минимальной гамильтоновой цепи в G_n [3].

В качестве следствия отсюда можно получить

Утверждение 1. *При $d_n = O(n^c)$, где $0 < c < 1$ и c не зависит от n, задача NP-трудна.*

Доказательство. Пусть $\overline{G}_{\bar n} = (\overline{V}, \overline{E})$ — полный граф, где $\bar n = \lfloor n^{1/c} \rfloor$, $V \subset \overline{V} = \{1, \dots, \bar n\}$, $E \subset \overline{E}$. Теперь для каждого $i \in V$ определим весовую функцию $\overline{w}^{(i)}$, заданную на $\overline{E}$, где

$$\overline{w}_e^{(i)} = \begin{cases} w_e, & \text{если } e \in E, \\ 0, & \text{если } e = (i,k),\ k > n, \\ M & \text{в остальных случаях.} \end{cases}$$

Здесь M — достаточно большое число (большее суммы весов ребер в множестве E). Для каждого $i \in V$ рассмотрим исходную задачу на графе $\overline{G}_{\bar n}$ с весовой функцией $\overline{w}^{(i)}$ при $d_{\bar n} = n = O(\bar n^c)$. Полиномиальная разрешимость этой задачи при любом $i \in V$ влечет полиномиальную разрешимость задачи отыскания минимальной гамильтоновой цепи в графе G_n. Противоречие с NP-трудностью последней задачи доказывает утверждение 1.

Утверждение 2. *Если d_n — константа, то исходная задача на графе G_n полиномиально разрешима.*

Справедливость этого утверждения вытекает из следующей процедуры.

1. Перебираются все простые цепи, состоящие ровно из d_n ребер, и в $\mathscr{D}_n$ для каждой цепи находится минимальное дерево, содержащее данную цепь.

2. Среди найденных деревьев выбирается дерево с наименьшим весом.

Очевидно, что эти операции можно проделать за полиномиальное время (с использованием, например, алгоритма Прима [6] при поиске минимальных деревьев в п. 1). Утверждение 2 доказано.

Далее для определенности все построения и рассуждения будем вести в предположении непрерывной модели данных, зафиксировав ограничение снизу на диаметр остова величиной, равной d. Пусть $W_n = \{(w_{ij})\}$ — класс симметричных квадратных матриц порядка n, элементы которых (т. е. веса ребер графа G_n) выбираются случайно и независимо из отрезка (a_n, b_n), $a_n > 0$, с непрерывной равномерной функцией распределения. Через a_n, b_n обозначены функции от n. Полученные в статье результаты для непрерывного случая сохраняются также и для дискретной модели данных, когда элементы w_{ij} выбираются случайно, независимо и равновероятно из множества $\{1, \dots, r_n\}$, где $r_n \to \infty$ при $n \to \infty$.

В работах [4, 2] было показано, что при $r_n \leqslant n/(\psi_n \ln n)$ (в дискретном случае) и $b_n/a_n \leqslant n/(\psi_n \ln n)$ (в непрерывном случае) алгоритм «иди в ближайший город» находит асимптотически точное решение задачи коммивояжера на классе весовых матриц W_n за время $O(n^2)$. Здесь $\psi_n = o(n)$, $\psi_n \to \infty$ при $n \to \infty$. Следовательно, в этих случаях рассматриваемая задача решается асимптотически точно для любого $d_n < n$. Для дискретной модели известен более сильный результат [7]. А именно, при $r_n = O(n/\ln n)$ почти всегда можно найти гамильтонову цепь, состоящую из ребер единичного веса за полиномиальное время, т. е. рассматриваемая задача в дискретном случае при $r_n = O(n/\ln n)$ также почти всегда полиномиально разрешима для любого $d_n < n$.

Для нахождения приближенного решения поставленной задачи предлагается алгоритм $\widetilde{A}$ и проводится его вероятностный анализ при более слабом условии $b_n/a_n \leqslant n/\psi_n$ (в непрерывном случае) и $r_n \leqslant n/\psi_n$ (в дискретном случае).

Опишем алгоритм $\widetilde{A}$, состоящий из двух этапов.

Этап 1. Начиная с произвольной вершины i_0 в G_n, строится цепь $C = \{(i_0, i_1), (i_1, i_2), \dots, (i_{d-1}, i_d)\}$ за d шагов. На шаге m, $1 \leqslant m \leqslant d$, к текущей цепи $\{(i_0, i_1), (i_1, i_2), \dots, (i_{m-2}, i_{m-1})\}$ добавляется ребро (i_{m-1}, i_m), $i_m \notin \{i_0, \dots, i_{m-1}\}$, с наименьшим весом.

Длина ребра, добавленного на m-м шаге, будет обозначаться через $l_m = w_{i_{m-1} i_m}$, $m = 1, \dots, d$.

Если
$$w(C) \leqslant da_n + (b_n - a_n)(1 + h_{n,d}), \tag{1}$$
где
$$h_{n,d} = \sum_{k=n-d}^{n-1} \frac{1}{k+1}, \tag{2}$$
то выполняется этап 2.

Если (1) не выполняется, то алгоритм $\widetilde{A}$ прекращает свою работу констатацией факта неудачной работы (происходит несрабатывание алгоритма).

Этап 2. Находится минимальное остовное дерево $\widetilde{D_n} \in \widetilde{\mathscr{D}}_n$, содержащее цепь $C = \{(i_0, i_1), (i_1, i_2), \dots, (i_{d-1}, i_d)\}$, построенную на этапе 1. Дерево $\widetilde{D_n}$ принимается в качестве приближенного решения исходной задачи. Такое дерево может быть найдено с использованием, например, алгоритма Прима для задачи о минимальном остовном дереве [6], имеющего временную сложность $O(n^2)$.

Описание алгоритма $\widetilde{A}$ закончено.

Замечание 1. Алгоритм $\widetilde{A}$ реализуется с временной сложностью $O(n^2)$.

Теперь установим ряд вспомогательных утверждений.

Лемма 1. *Пусть w_0 — вес минимального остовного дерева из $\mathscr{D}_n$, т. е.*

$$w_0 = \min\{w(D_n) \mid D_n \in \mathscr{D}_n\}.$$

Тогда вес $w(\widetilde{D_n})$ остовного дерева $\widetilde{D_n}$, построенного алгоритмом $\widetilde{A}$, удовлетворяет неравенству

$$w(\widetilde{D_n}) \leqslant w_0 + w(C) - da_n.$$

Доказательство. Очевидно, что $\widetilde{D_n}$ является минимальным остовным деревом из $\mathscr{D}_n$ в задаче с модифицированной функцией w'_{ij} весов ребер в графе G_n, где

$$w'_{ij} = \begin{cases} w_{ij}, & \text{если } (i,j) \notin C, \\ a_n, & \text{если } (i,j) \in C. \end{cases}$$

Следовательно, справедливо равенство

$$\min\Big\{\sum_{(i,j)\in D_n} w'_{ij} \mid D_n \in \mathscr{D}_n\Big\} = \sum_{(i,j)\in\widetilde{D_n}} w'_{ij}$$
$$= \sum_{(i,j)\in C} w'_{ij} + w(\widetilde{D_n}) - w(C) = da_n + w(\widetilde{D_n}) - w(C). \quad (3)$$

С другой стороны, так как $w'_{ij} \leqslant w_{ij}$, имеем

$$\min\Big\{\sum_{(i,j)\in D_n} w'_{ij} \mid D_n \in \mathscr{D}_n\Big\} \leqslant \min\Big\{\sum_{(i,j)\in D_n} w_{ij} \mid D_n \in \mathscr{D}_n\Big\}$$
$$= \min\{w(D_n) \mid D_n \in \mathscr{D}_n\} = w_0. \quad (4)$$

Из соотношений (3)–(4) следует утверждение леммы 1.

Обозначим через X_k случайную величину, равную минимальному значению среди k случайных независимых величин, равномерно распределенных на отрезке $[0,1]$. Тогда вес l_m ребра, выбранного на m-м шаге первого этапа алгоритма $\widetilde{A}$, можно записать как $l_m = a_n + (b_n - a_n)X_{n-m}$, $m = 1, \dots, d$, а суммарный вес $w(C)$ цепи C в виде

$$w(C) = da_n + (b_n - a_n)\sum_{k=n-d}^{n-1} X_k. \quad (5)$$

Лемма 2. *Пусть* $\xi_k = X_k - \mathbf{E}X_k$, $k = n-d, \dots, n-1$ *и* $S = \sum\limits_{k=n-d}^{n-1} \xi_k$. *Тогда*

$$\mathbf{Pr}\{S > 1\} \leqslant e^{-0,25(n-d)}. \tag{6}$$

Доказательство. Сначала покажем, что при $T = 0,5(n-d)$ и $g_k = 2/(k+1)^2$, $k = n-d, \dots, n-1$, случайные величины ξ_k удовлетворяют условиям теоремы Петрова [5, с. 81] о вероятностях больших уклонений:

$$\mathbf{E}e^{t\xi_k} \leqslant e^{0,5 g_k t^2}$$

для всех $k = n-d, \dots, n-1$ и $0 \leqslant t \leqslant T$. Для этого нам понадобится ряд вспомогательных фактов.

Из вида функции распределения $F_{X_k}(x) = 1 - (1-x)^k$ случайной величины X_k следует

Факт 1. $\mathbf{E}X_k = 1/(k+1)$, $k = n-d, \dots, n-1$.

Факт 2. При каждом целом $k \geqslant 1$ справедливо равенство

$$\mathbf{E}e^{tX_k} = k \sum_{i=0}^{\infty} \frac{t^i}{k(k+1)\cdots(k+i)}.$$

Доказательство индукцией по k. При $k = 1$ имеем

$$\mathbf{E}e^{tX_1} = \int\limits_0^1 e^{tx} dx = \frac{e^t - 1}{t} = \sum_{i=0}^{\infty} \frac{t^i}{1 \cdot 2 \cdots (1+i)}.$$

С учетом рекуррентного соотношения

$$\mathbf{E}e^{tX_{k+1}} = \frac{k+1}{t}(\mathbf{E}e^{tX_k} - 1),$$

следующего из цепочки равенств

$$\mathbf{E}e^{tX_{k+1}} = \int\limits_0^1 e^{tx} dF_{X_{k+1}}(x) = (k+1) \int\limits_0^1 e^{tx}(1-x)^k dx$$
$$= (k+1)\left(k \int\limits_0^1 e^{tx}(1-x)^{k-1} dx - 1\right),$$

для каждого $k \geqslant 1$ из индуктивного предположения вытекает справедливость факта 2:

$$\mathbf{E}e^{tX_{k+1}} = \frac{k+1}{t}(\mathbf{E}e^{tX_k} - 1) = \frac{k+1}{t} \sum_{i=1}^{\infty} \frac{t^i}{(k+1)\cdots(k+i)}$$
$$= (k+1) \sum_{i=0}^{\infty} \frac{t^i}{(k+1)\cdots(k+1+i)}.$$

Факт 3. Для любого $k = n - d, \dots, n-1$ и $0 \leqslant t \leqslant T$ справедливо неравенство $t/(k+1) \leqslant 0,5$.

Факт 4. При любом α, $0 \leqslant \alpha \leqslant 0,5$, выполнено неравенство

$$\frac{e^{-\alpha}}{1-\alpha} \leqslant 1 + \alpha^2.$$

Это неравенство следует из оценки

$$e^{-\alpha} \leqslant 1 - \alpha + \frac{\alpha^2}{2} \leqslant (1-\alpha) + (1-\alpha)\alpha^2 = (1-\alpha)(1+\alpha^2).$$

Продолжим доказательство леммы. С учетом фактов 1 и 2 имеем

$$\mathbf{E}e^{t\xi_k} = \mathbf{E}e^{t(X_k - \mathbf{E}X_k)} = e^{-t/(k+1)}\, k \sum_{i=0}^{\infty} \frac{t^i}{k(k+1)\cdots(k+i)}$$

$$\leqslant e^{-t/(k+1)} \sum_{i=0}^{\infty} \left(\frac{t}{k+1}\right)^i = \frac{e^{-t/(k+1)}}{1 - t/(k+1)}.$$

Окончательно с использованием фактов 3 и 4 получаем требуемое условие на $\mathbf{E}e^{t\xi_k}$:

$$\mathbf{E}e^{t\xi_k} \leqslant 1 + \left(\frac{t}{k+1}\right)^2 \leqslant \exp\left(\frac{t}{k+1}\right)^2 = e^{0,5 g_k t^2}.$$

Далее, обозначив через U сумму всех g_k, $n - d \leqslant k < n$, имеем

$$U = \sum_{k=n-d}^{n-1} \frac{2}{(k+1)^2} \leqslant 2 \sum_{k=n-d}^{n-1} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \frac{2}{n-d} - \frac{2}{n} = \frac{2d}{n(n-d)}.$$

Поскольку в рассматриваемом случае $UT \leqslant d/n \leqslant 1$, можно воспользоваться неравенством из [5, с. 81] в форме

$$\mathbf{Pr}\{S > x\} \leqslant e^{-0,5Tx}.$$

Отсюда получаем

$$\mathbf{Pr}\{S > 1\} \leqslant e^{-0,25(n-d)}.$$

Лемма 2 доказана.

Далее воспользуемся следующими определениями из [1].

Определение 1. Говорят, что алгоритм A для некоторой задачи имеет оценки относительной погрешности ε_n и вероятности несрабатывания δ_n, если

$$\mathbf{Pr}\{f_A > (1 + \varepsilon_n) f^*\} \leqslant \delta_n,$$

где f_A и f^* — значения целевой функции для приближенного решения, полученного алгоритмом A, и точного решения соответственно.

Определение 2. Алгоритм A для решения некоторой задачи называется *асимптотически точным*, если существуют последовательности ε_n и δ_n, стремящиеся к нулю при $n \to \infty$.

Теорема 1. *Для матриц из класса W_n алгоритм $\widetilde{A}$ за время $O(n^2)$ находит приближенное решение $\widetilde{D_n}$ с оценками относительной погрешности решения*

$$\varepsilon_n = \frac{b_n - a_n}{na_n}(h_{n,d} + 1) \tag{7}$$

и вероятности несрабатывания

$$\delta_n = e^{-0,25(n-d)}. \tag{8}$$

Доказательство. Воспользовавшись тем, что математическое ожидание суммы случайных величин X_k, $n-d \leqslant k < n$, имеет вид

$$\mathbf{E}\Big(\sum_{k=n-d}^{n-1} X_k\Big) = \sum_{k=n-d}^{n-1} \mathbf{E}X_k = \sum_{k=n-d}^{n-1} \frac{1}{k+1} = h_{n,d},$$

$f_{\widetilde{A}} = w(\widetilde{D_n})$, $na_n \leqslant w_0 \leqslant f^*$, а также (5) и леммами 1 и 2, убеждаемся, что для алгоритма $\widetilde{A}$ оценки относительной погрешности (7) и вероятности несрабатывания (8) удовлетворяют определению 1 алгоритма с оценками ε_n и δ_n:

$$\begin{aligned}\mathbf{Pr}\{f_{\widetilde{A}} > f^*(1+\varepsilon_n)\} &\leqslant \mathbf{Pr}\{w_0 + w(C) - da_n > f^*(1+\varepsilon_n)\} \\ &\leqslant \mathbf{Pr}\{w(C) - da_n > \varepsilon_n f^*\} \leqslant \mathbf{Pr}\{w(C) - da_n > (b_n - a_n)(h_{n,d}+1)\} \\ &= \mathbf{Pr}\Big\{\sum_{k=n-d}^{n-1} X_k > (h_{n,d}+1)\Big\} = \mathbf{Pr}\{S > 1\} \leqslant e^{-0,25(n-d)} = \delta_n.\end{aligned}$$

Согласно замечанию 1 временная сложность алгоритма $\widetilde{A}$ равна $O(n^2)$. Теорема 1 доказана.

Наконец, с использованием определений 1, 2 и теоремы 1 получаем следующее утверждение об условиях асимптотической точности алгоритма $\widetilde{A}$.

Теорема 2. *В полном неориентированном n-вершинном графе с матрицей весов из класса W_n алгоритм $\widetilde{A}$ позволяет находить асимптотически точное решение задачи, если выполняются следующие условия:*

$$b_n/a_n \leqslant n/\psi_n, \tag{9}$$

$$d \leqslant d_n = n - n/\psi_n, \tag{10}$$

где $\psi_n = o(n)$ и $\psi_n \to \infty$ при $n \to \infty$.

Доказательство. Покажем, что алгоритм $\widetilde{A}$ с оценками (7) и (8) является асимптотически точным.

Для функции (2) с учетом (10) получаем

$$h_{n,d} \leqslant \ln \frac{n}{n-d} \leqslant \ln \frac{n}{n-d_n} \leqslant \ln \psi_n. \tag{11}$$

Используя условие (9) и соотношение (11), убеждаемся в сходимости к нулю при $n \to \infty$ последовательности значений относительной погрешности вида (7):

$$\varepsilon_n = \frac{b_n - a_n}{na_n}(h_{n,d} + 1) \leqslant \frac{\ln \psi_n + 1}{\psi_n} \to 0.$$

С учетом (10) требуемую сходимость получаем и для вероятности несрабатывания вида (8)

$$\delta_n = e^{-0,25(n-d)} \leqslant e^{-0,25(n-d_n)} \leqslant e^{-0,25n/\psi_n} \to 0 \text{ при } n \to \infty.$$

Теорема 2 доказана.

Следствие. *В условиях теоремы 2 алгоритм $\widetilde{A}$ за время $O(n^2)$ находит асимптотически точное решение двухкритериальной задачи, которая состоит в отыскании в графе G_n остовного дерева D_n с минимальным весом и максимальным диаметром.*

Наконец, заметим, что результаты, приведенные для непрерывной модели, справедливы и для дискретного случая с той лишь разницей, что в формуле (9) выражение b_n/a_n следует заменить величиной r_n.

ЛИТЕРАТУРА

1. **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.

2. **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.

3. **Гэри М., Джонсон Д.** Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982.

4. **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.

5. **Петров В. В.** Предельные теоремы для сумм независимых случайных величин. М.: Наука, 1987.

6. **Прим Р. К.** Кратчайшие связывающие сети и некоторые обобщения // Кибернетический сборник. М.: Мир, 1961. Вып. 2. С. 95–107.

7. **Karp R. M.** The probabilistic analysis of some combinatorial search algorithms // Algorithms and Complexity: New Directions and Recent Results. New York: Acad. Press, 1976. P. 1–19.

Адрес авторов:

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

Статья поступила
27 мая 1999 г.,
переработанный вариант —
13 декабря 1999 г.