ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь—март 1997. Том 4, № 1, 79–87

УДК 519.854.3

О ТОЧНОСТИ ОДНОГО АЛГОРИТМА РАЗБИЕНИЯ МНОЖЕСТВА*)

Ю. В. Шамардин, А. В. Пяткин

Рассматривается задача о разбиении n целых положительных чисел на два таких подмножества, чтобы сумма чисел в «больше M» подмножестве была минимальна. Для решения этой задачи известен алгоритм загрузки предметов в ранец [2], который находит приближенное решение с относительной погрешностью не более $1/7$, используя $O(n)$ шагов. В данной статье на основе частичного перебора разбиений и упомянутого выше алгоритма строится алгоритм E_m, $m = 0, 1, 2, \ldots$, который находит приближенное решение за $O(2^m n)$ шагов. Доказывается, что максимальная относительная погрешность $\eta(E_m)$ этого алгоритма удовлетворяет неравенствам

$$\frac{1}{7+3m} \leqslant \eta(E_m) \leqslant \frac{1}{6+2m}.$$

1. Определения и обозначения

Пусть дано n предметов, занумерованных числами из множества $N = \{1, 2, \ldots, n\}$ и имеющих целые положительные веса $w_1, w_2, \ldots, w_n$. Обозначим через $W(S)$ суммарный вес предметов из подмножества $S \subset N$. Пусть множество N разбито на два подмножества N_1 и N_2. *Весом* такого разбиения назовем величину

$$F(N_1, N_2) = \max\{W(N_1), W(N_2)\}.$$

Требуется найти такое разбиение множества N на подмножества N_1^*, N_2^*, что

$$F(N_1^*, N_2^*) = \min_{N_1, N_2} F(N_1, N_2). \tag{1}$$

Веса предметов считаем упорядоченными по невозрастанию, т. е.

$$w_1 \geqslant w_2 \geqslant \ldots \geqslant w_n. \tag{2}$$

*) Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 96–01–01582).

Вектор весов $I = (w_1, \dots, w_n)$ назовем *входом* задачи (1). Пусть A — некоторый алгоритм приближенного решения задачи (1), находящий разбиение $N_1 \cup N_2 = N$ при входе I. Обозначив через $F^*(I)$ значение минимума (1), введем погрешности

$$h(A, I) = (F(N_1, N_2) - F^*(I))/F^*(I),$$
$$\eta(A) = \sup_I h(A, I),$$

характеризующие точность алгоритма A при входе I и на всем множестве входов.

Одним из направлений исследования дискретных экстремальных задач, в том числе и (1), является оценка максимальной погрешности алгоритмов. В [1] показано, что алгоритм (обозначим его через K), предложенный Н. Кармаркаром и Р. Карпом для задачи (1) и названный разностным методом, имеет наихудшую погрешность $\eta(K) = 1/6$. В [2] рассмотрен алгоритм D, основанный на загрузке предметов в ранец вместимости $W(N)/2$, и доказано, что этот алгоритм имеет несколько лучшую точность, чем K, а именно $\eta(D) = 1/7$. Цель настоящей работы — оценка точности алгоритма E_m, $m = 0, 1, 2, \dots$, построенного на частичном переборе разбиений и алгоритме D.

Возьмем список предметов $R = \{i_1, \dots, i_r\} \subset N$, где $i_1 < i_2 < \dots < i_r$, и обозначим через $D(R)$ процедуру, которая строит разбиение $N_1 \cup N_2 = N$ путем загрузки предметов из списка R в ранец вместимости $W(N)/2$. Эта процедура включает следующие шаги.

1. Полагаем $L(R) = \varnothing$, $a = b = W(N)/2$, $j = 0$, $t = 1$; здесь $L(R)$ — формируемое подмножество предметов из R, помещенных в ранец, a — остаточная вместимость ранца, b — превышение объема ранца при загрузке в него предмета j, t — порядковый номер очередного предмета.

2. Если $w_{i_t} \leqslant a$, то полагаем $L(R) = L(R) \cup \{i_t\}$, $a = a - w_{i_t}$. Если $w_{i_t} > a$ и $w_{i_t} - a \leqslant b$, то полагаем $j = i_t$, $b = w_{i_t} - a$.

3. Полагаем $t = t + 1$ и при $t \leqslant r$ повторяем шаг 2.

4. Если $a \leqslant b$, то полагаем $N_1 = L(R)$. В противном случае

$$N_1 = (L(R) \setminus \{j + 1, \dots, n\}) \cup \{j\}.$$

Второе подмножество $N_2 = N \setminus N_1$. Вес построенного разбиения

$$F(N_1, N_2) = \min\{a, b\} + W(N)/2.$$

В [2] изучался алгоритм $D = D(N)$. Мы рассматриваем обобщающий алгоритм E_m, который, решая задачу (1), использует процедуры вида $D(S \cup \{m + 1, \dots, n\})$ и выбирает разбиение $N_1 \cup N_2 = N$ с минимальным весом $F(N_1, N_2)$ по всем подмножествам $S \subset \{1, \dots, m\} \cap N$. Здесь целое число $m \geqslant 0$ — параметр алгоритма, при этом $E_0 = D(N)$.

Будем обозначать через $\langle i, j\rangle$ набор предметов $\{i, i+1, \ldots, j\}$ и через $W(i,j)$ их суммарный вес. Множества предметов

$$H(S) = S \cap L(S \cup \langle m+1, n\rangle),$$
$$T(S) = \langle m+1, n\rangle \cap L(S \cup \langle m+1, n\rangle)$$

назовем *головной* и *хвостовой* частями загрузки ранца. Если $N_1^* \cup N_2^* = N$ — наилучшее разбиение, то обозначаем

$$S_k^* = \langle 1, m\rangle \cap N_k^*, \quad k = 1, 2.$$

Нетрудно видеть, что если $h(E_m, I) > 0$, то $H(S_k^*) = S_k^*$, $k = 1, 2$.

2. Оценка точности алгоритма E_m

Следующие две леммы выясняют необходимые ниже свойства алгоритма E_m.

Лемма 1. *Если $h(E_m, I) > 0$, то число предметов $n \geqslant m+5$.*

Доказательство. Положим $d = W(N) - F^*(I)$ и заметим, что

$$W(N_k^*) \in \{d,\ F^*(I)\}, \quad k = 1, 2. \tag{3}$$

При $n \leqslant m$ алгоритм E_m строит точное решение задачи (1). Поэтому $n \geqslant m+1$. Пусть $m+1 \in N_1^*$. Из условия леммы и включения (3) следует, что $H(S_1^*) = S_1^*$ и $m+1 \in T(S_1^*)$, т. е.

$$W(S_1^*) + w_{m+1} < d.$$

Поэтому $N_1^* \neq S_1^* \cup \{m+1\}$ и найдется предмет $i \in N_1^*$, $i > m+1$, такой, что

$$W(S_1^*) + w_{m+1} + w_i \leqslant d.$$

Отсюда вытекает, что должен быть предмет $j \in \langle m+2, i\rangle \cap T(S_1^*)$ и, следовательно,

$$W(S_1^*) + w_{m+1} + w_j < d.$$

Если $N_1^* = S_1^* \cup \{m+1, i\}$, то с учетом последнего неравенства и соотношения $w_j \geqslant w_i$ получаем неравенство $W(N_1^*) < d$, что невозможно. Следовательно, в множестве N_1^* есть по крайней мере еще один предмет и $|N_1^*| \geqslant |S_1^*| + 3$.

Аналогично показывается, что $|N_2^*| \geqslant |S_2^*| + 2$. Поэтому получаем

$$n = |N_1^*| + |N_2^*| \geqslant |S_1^*| + 3 + |S_2^*| + 2 = m + 5.$$

Лемма 1 доказана.

Лемма 2. *Пусть вход* $I_n = (w_1, \ldots, w_n)$ *удовлетворяет условиям*

$$h(E_m, I_n) > u > 0, \quad w_n \leqslant 2(1+u)F^*(I_n) - W(N).$$

Тогда при некотором $j < n$ *для входа* $I_j = (w_1, \ldots, w_j)$ *справедливы неравенства*

$$h(E_m, I_j) \geqslant h(E_m, I_n), \tag{4}$$

$$w_j > 2(1+u)F^*(I_j) - W(1, j). \tag{5}$$

Доказательство. Из условий леммы следует, что $n \in T(S)$ при любом подмножестве $S \subset \langle 1, m \rangle$. Поэтому алгоритм E_m строит по входам I_n и I_{n-1} разбиения одного и того же веса F. С учетом очевидного неравенства $F^*(I_n) \geqslant F^*(I_{n-1})$ получаем

$$h(E_m, I_{n-1}) = \frac{F}{F^*(I_{n-1})} - 1 \geqslant \frac{F}{F^*(I_n)} - 1 = h(E_m, I_n).$$

Поэтому (4) справедливо при $j = n - 1$. Если (5) не выполняется, то повторяем редукционный шаг, переходя от I_{n-1} к I_{n-2}, и так далее. Этот процесс должен завершиться выполнением (5) при некотором $j \geqslant m + 5$, что следует из леммы 1. Лемма 2 доказана.

Основным результатом работы является следующая

Теорема. *Максимальная погрешность* $\eta(E_m)$ *удовлетворяет неравенствам*

$$\frac{1}{7+3m} \leqslant \eta(E_m) \leqslant \frac{1}{6+2m}. \tag{6}$$

Доказательство. Рассмотрим вход $I = (w_1, \ldots, w_n)$ с числом предметов $n = 2m + 6$ и весами $w_i = 3$ при $i \leqslant 2m + 2$ и $w_i = 2$ при $i \geqslant 2m + 3$. Алгоритм E_m строит разбиение множества N на подмножества $N_1 = \langle 1, m+2 \rangle$, $N_2 = N \setminus N_1$ с весом $F(N_1, N_2) = 8 + 3m$, в то время как наилучшее разбиение

$$N_1^* = \langle 1, m+1 \rangle \cup \{n-1, n\}, \quad N_2^* = N \setminus N_1^*$$

имеет вес $F^*(I) = 7 + 3m$. Следовательно,

$$\eta(E_m) \geqslant h(E_m, I) = \frac{F(N_1, N_2)}{F^*(I)} - 1 = \frac{1}{7+3m},$$

т. е. нижняя оценка (6) верна.

Перейдем к доказательству верхней оценки (6). Допустим, что она не выполняется. Тогда найдется вход $I = (w_1, \ldots, w_n)$ такой, что

$$h(E_m, I) > 1/(6 + 2m). \tag{7}$$

Для сокращения записи положим $u = 1/(6+2m)$ и

$$v = 2(1+u)F^*(I) - W(N).$$

В силу лемм 1 и 2 можно считать, что

$$n \geqslant m+5, \quad w_n > v. \tag{8}$$

Величина v больше или равна $uW(N)$, поскольку $F^*(I) \geqslant W(N)/2$.

Оставшаяся часть доказательства представлена тремя леммами, в которых выясняются свойства входа I при выполнении условий (7) и (8).

Лемма 3. (a) *При любом подмножестве $S \subset \langle 1, m\rangle$ хвостовая часть $T(S)$ загрузки ранца либо пуста, либо имеет вид*

$$T(S) = \langle m+1, q\rangle, \tag{9}$$

где $q \geqslant m+1$.

(b) *Справедливы неравенства*

$$n \geqslant m + |T(S_1^*)| + |T(S_2^*)| + 2, \tag{10}$$

$$|T(S_k^*)| \geqslant 2, \quad k = 1, 2, \tag{11}$$

и, следовательно, $n \geqslant m+6$.

Доказательство. Предположение $w_{m+1} > 2v$ ведет к противоречию:

$$W(N) \geqslant (m+1)w_{m+1} + (n-m-1)w_n > (m+1)2v + (m+5-m-1)v$$
$$= (6+2m)v \geqslant (6+2m)uW(N) = W(N).$$

Поэтому

$$w_{m+1} \leqslant 2v. \tag{12}$$

(a) Допустим, что $T(S) \neq \varnothing$, но (9) не выполняется. Это означает, что найдутся предметы i и j, $m+1 \leqslant i < j$, такие, что $i \notin T(S)$, $j \in T(S)$. Тогда

$$w_i > w_j + v > 2v,$$

что противоречит (12). Следовательно, (9) верно.

(b) Положим $t_k = |T(S_k^*)|$, $k = 1, 2$. Если $|N_k^*| \leqslant |S_k^*| + t_k$, то с учетом первого утверждения леммы и условий (2) получаем оценку

$$W(N_k^*) \leqslant W(S_k^*) + W(T(S_k^*)) < (1+u)F^*(I) - v,$$

невозможную для наилучшего разбиения. Следовательно, выполняются неравенства

$$|N_k^*| \geqslant |S_k^*| + t_k + 1, \quad k = 1, 2,$$

из которых сразу следует (10).

Будем считать, что $t_1 \leqslant t_2$. Случай $t_1 = 0$, как легко видеть, невозможен. Пусть $t_1 = 1$. Покажем, что это предположение ведет к оценке $w_{m+1} > 2v$, противоречащей условию (12). В силу (9) имеем $T(S_1^*) = \{m+1\}$, $T(S_2^*) = \langle m+1, m+t_2\rangle$, а из (10) следует, что

$$m + t_2 \leqslant n - t_1 - 2 = n - 3. \tag{13}$$

Поэтому $n \notin T(S_2^*)$ и тем более $n \notin T(S_1^*)$. Следовательно,

$$W(S_1^*) + w_{m+1} + w_n > (1+u)F^*(I), \tag{14}$$

$$W(S_2^*) + W(m+1, m+t_2) + w_n > (1+u)F^*(I). \tag{15}$$

Вычтем из обеих частей неравенства (15) величину $W(N)$. После очевидных преобразований получаем

$$W(S_1^*) + W(m+t_2+1, n-1) < (1+u)F^*(I) - v.$$

Вычитая это неравенство из (14) и учитывая (13), приходим к оценке

$$w_{m+1} > v + W(m+t_2+1, n-1) - w_n \geqslant v + w_{n-2} + w_{n-1} - w_n > 2v,$$

противоречащей неравенству (12). Следовательно, $t_1 \geqslant 2$ и (11) верно. Лемма 3 доказана.

Лемма 4. *Пусть подмножество $S \subset \langle 1, m\rangle$ и предмет $i \in \langle 1, m\rangle$ таковы, что*

$$H(S) = S, \quad 2 \leqslant |T(S)| \leqslant n - m - 2, \quad w_i \leqslant 2v.$$

Тогда

1) *если $i \notin S$, то*

$$H(S \cup \{i\}) = S \cup \{i\}, \tag{16}$$

$$|T(S \cup \{i\})| = |T(S)| - 1; \tag{17}$$

2) *если $i \in S$, то*

$$H(S \setminus \{i\}) = S \setminus \{i\}, \tag{18}$$

$$|T(S \setminus \{i\})| = |T(S)| + 1. \tag{19}$$

Доказательство. 1. Положим $p = |T(S)|$ и $q = |T(S \cup \{i\})|$. По лемме 3 имеем $T(S) = \langle m+1, m+p\rangle$ и $T(S \cup \{i\}) = \langle m+1, m+q\rangle$. Так как $p \geqslant 2$ и $w_i \leqslant 2v < w_{m+1} + w_{m+2}$, то $W(S \cup \{i\}) \leqslant W(N)/2$ и (16) верно.

Очевидно, $p \geqslant q+1$. Допустим, что $p \geqslant q+2$. Тогда из разности неравенств

$$W(S \cup \{i\}) + W(m+1, m+q) + w_n > (1+u)F^*(I),$$
$$W(S) + W(m+1, m+p) < (1+u)F^*(I) - v$$

получаем

$$w_i > v + W(m+q+1, m+p) - w_n \geqslant v + w_{m+p-1} + w_{m+p} - w_n > 2v,$$

что противоречит условию $w_i \leqslant 2v$. Следовательно, $p = q+1$ и (17) верно.

2. Равенство (18) сразу следует из условия $H(S) = S$. Положим $p = |T(S)|$ и $q = |T(S \setminus \{i\})|$. По лемме 3 имеем $T(S) = \langle m+1, m+p\rangle$ и $T(S \setminus \{i\}) = \langle m+1, m+q\rangle$. Очевидно, $q \geqslant p+1$. Допустим, что $q \geqslant p+2$. Тогда из разности неравенств

$$W(S) + W(m+1, m+p) + w_n > (1+u)F^*(I),$$
$$W(S \setminus \{i\}) + W(m+1, m+q) < (1+u)F^*(I) - v$$

получаем

$$w_i > v + W(m+p+1, m+q) - w_n \geqslant v + w_{m+q-1} + w_{m+q} - w_n > 2v,$$

что противоречит условию $w_i \leqslant 2v$. Следовательно, $q = p+1$ и (19) верно. Лемма 4 доказана.

Пусть $p \geqslant 0$ — произвольное целое число.

Лемма 5. *Если $m \geqslant p$, то число предметов*

$$n \geqslant m+p+6. \tag{20}$$

Доказательство проведем индукцией по параметру p. При $p=0$ оценка (20) содержится в лемме 3. Пусть утверждение верно при значениях параметра $0, 1, \ldots p-1$. Рассмотрим случай, когда $m \geqslant p \geqslant 1$. В силу индуктивного предположения имеем $n \geqslant m+p+5$. Допустим, что $n = m+p+5$.

Рассмотрим предмет $q = m-p+1$. Предположение $w_q > 2v$ приводит к противоречию, поскольку

$$W(N) \geqslant qw_q + (n-q)w_n > 2qv + (n-q)v = (n+q)v$$
$$= (m+p+5+m-p+1)v = (6+2m)v \geqslant (6+2m)uW(N) = W(N).$$

Следовательно,

$$w_q \leqslant 2v. \tag{21}$$

Покажем, что из равенства $n = m + p + 5$ следует неравенство $w_q > 2v$, противоречащее (21).

Положим $t_k = |T(S_k^*)|$, $k = 1, 2$, и будем считать, что $t_1 \leqslant t_2$. Тогда из (11) следует, что $t_1 \geqslant 2$, а согласно (10) имеем

$$t_1 + t_2 \leqslant n - m - 2 = p + 3. \tag{22}$$

Введем множества $P_k = S_k^* \cap \langle q, m \rangle$, $k = 1, 2$. Возможны следующие случаи: 1) $|P_2| \geqslant t_1 - 1$; 2) $|P_2| < t_1 - 1$.

Случай 1. Пусть $|P_2| \geqslant t_1 - 1$. Выберем любое подмножество $Q \subset P_2$ с числом элементов $t_1 - 1$. Будем последовательно брать по одному элементу из Q, присоединять их к множеству S_1^* и вычитать из S_2^*. В результате построим две совокупности предметов $S_1 = S_1^* \cup Q$, $S_2 = S_2^* \setminus Q$, а применение леммы 4 (выполнение ее условий гарантируется неравенствами (21), (22) и $t_1 \geqslant 2$) на каждом шаге описанного процесса приводит к соотношениям

$$H(S_k) = S_k, \quad k = 1, 2, \quad |T(S_1)| = 1, \quad |T(S_2)| = t_1 + t_2 - 1.$$

В силу леммы 3 имеем $T(S_1) = \{m + 1\}$, $T(S_2) = \langle m + 1, m + t_1 + t_2 - 1 \rangle$ и, поскольку из (22)

$$m + t_1 + t_2 - 1 \leqslant n - 3,$$

предмет n не принадлежит $T(S_2)$, тем более $n \notin T(S_1)$. Поэтому справедливы неравенства

$$W(S_1) + w_{m+1} + w_n > (1 + u)F^*(I), \tag{23}$$

$$W(S_2) + W(T(S_2)) + w_n > (1 + u)F^*(I). \tag{24}$$

Вычтем из обеих частей неравенства (24) величину $W(N)$. После очевидных преобразований получаем

$$W(S_1) + W(m + t_1 + t_2\,, n - 1) < (1 + u)F^*(I) - v.$$

Вычитая это неравенство из (23), приходим к оценке

$$w_{m+1} > v + W(m + t_1 + t_2, n - 1) - w_n \geqslant v + w_{n-2} + w_{n-1} - w_n > 2v,$$

противоречащей неравенству (21). Следовательно, случай $|P_2| \geqslant t_1 - 1$ невозможен.

Случай 2. Пусть $|P_2| < t_1 - 1$. Учитывая неравенство $t_1 \leqslant p + 3 - t_2$, получаем

$$|P_1| = p - |P_2| > p - t_1 + 1 \geqslant t_2 - 2,$$

т. е. $|P_1| \geqslant t_2 - 1$. Выберем любое подмножество $Q \subset P_1$ с $t_2 - 1$ элементами и образуем множества $S_1 = S_1^* \setminus Q$, $S_2 = S_2^* \cup Q$. Те же рассуждения, что и выше, приводят к противоречию. Поэтому исходное предположение $n = m + p + 5$ неверно и должно быть $n \geqslant m + p + 6$. Таким образом, индукционный шаг и, следовательно, лемма 5 доказаны.

Применяя лемму 5 при $p = m$, получаем $n \geqslant 2m + 6$, что ведет к противоречию:

$$W(N) \geqslant nw_n > nv \geqslant (2m+6)uW(N) = W(N).$$

Следовательно, не существует входа I, удовлетворяющего условию (7). Поэтому справедлива верхняя оценка (6). Теорема доказана полностью.

Нижнюю оценку (6) можно усилить при $m = 2$. Рассмотрим вход $I = (w_1, \dots, w_7)$ с весами $w_1 = 7$, $w_2 = w_3 = 4$, $w_4 = 3$, $w_5 = w_6 = w_7 = 2$. Алгоритм E_2 строит разбиение множества $N = \langle 1, 7 \rangle$ на подмножества $N_1 = \{1, 2\}$, $N_2 = \langle 3, 7 \rangle$ с весом $F(N_1, N_2) = 13$, в то время как наилучшее разбиение $N_1^* = \{1, 4, 5\}$, $N_2^* = N \setminus N_1^*$ имеет вес $F^*(I) = 12$. Следовательно, $\eta(E_2) \geqslant 1/12$.

В заключение выскажем гипотезу, что $\eta(E_2) = 1/12$ и $\eta(E_m) = 1/(7 + 3m)$ при остальных m.

ЛИТЕРАТУРА

1. **Fischetti M., Martello S.** Worst-case analysis of the differencing method for the partition problem // Math. Programming. 1987. V. 37, N 1. P. 117–120.
2. **Lai T. C.** Worst-case analysis of greedy algorithms for the unbounded knapsack, subset-sum and partition problems // Oper. Res. Lett. 1993. V. 14, N 4. P. 215–220.

Адрес авторов:

Институт математики
им. С. Л. Соболева СО РАН,
Университетский пр., 4,
630090 Новосибирск,
Россия

Статья поступила
3 декабря 1996 г.