УДК 681.3.076.4

ОПТИМАЛЬНОЕ РАЗМЕЩЕНИЕ ИНФОРМАЦИОННЫХ МАССИВОВ В ПАМЯТИ НА МАГНИТНЫХ ЛЕНТАХ ДЛЯ СЛУЧАЯ ДВУНАПРАВЛЕННОГО ПОИСКА

В. Н. БУРКОВ, В. Б. СОКОЛОВ

(Москва)

Рассматриваются задачи оптимального размещения информации в памяти на магнитных лентах при различных стратегиях управления лентами и различных критериях оптимизации.

В крупных информационно-поисковых системах и системах обработки данных, имеющих основную память на магнитных лентах (МЛ), при жестких ограничениях на время и эффективность получения решения возникает задача оптимального размещения информационных массивов (ИМ) в памяти. Общие постановки этой задачи рассмотрены в работе [1]. Там же даны некоторые частные решения для случая однонаправленного поиска.

В настоящей работе рассматриваются возможные подходы к решению задачи об оптимальном размещении ИМ в памяти на МЛ для случая двунаправленного поиска и однонаправленного считывания. Как и в работе [1], задача о размещении ИМ решается для двух стратегий управления движением МЛ с возвратом в некоторое начальное положение и без возврата.

ЗАДАЧА 1. СТРАТЕГИЯ С ВОЗВРАТОМ

1. Постановка задачи

Пусть память состоит из m блоков МЛ, каждая из которых имеет длину L. Управление МЛ выполнено таким образом, что после каждого обращения МЛ возвращается в некоторое начальное положение. Поиск нужных ИМ может вестись при движении ленты в любом направлении, а считывание — при движении только в одном строго определенном направлении.

Пусть имеется n информационных массивов (ИМ), для каждого из которых известна длина y_i ($i = 1, 2, ..., n$) и вероятность обращения P_i (обращения предполагаются независимыми). Допустим, что все $y_i = y$ и число L кратно y. Требуется так разместить n ИМ по m блокам МЛ, чтобы среднее время обращения ко всем ИМ, $\overline{T}_{\text{обр}}$, было минимальным при следующих условиях:

$$\sum_{i=1}^{n} y_i = ny \leqslant mL, \tag{1}$$

$$y > 0, \tag{2}$$

$$0 \leqslant P_i \leqslant 1, \tag{3}$$

$$\sum_{i} P_i = 1. \tag{4}$$

2. Решение задачи 1

Получим вначале решение задачи для случая $m = 1$, а потом обобщим его на случай произвольного m.

Среднее время обращения можно представить в виде суммы средних времен поиска $\overline{T}_п$, считывания $T_с$ и возврата $T_в$:

$$\overline{T}_{обр} = \overline{T}_п + \overline{T}_с + \overline{T}_в. \quad (5)$$

Пусть скорости поиска, считывания и возврата равны $V_п$, $V_с$ и $V_в$ соответственно. Обозначим через i_k ($k = 1, 2, ..., l$) номер массива, расположенного справа от считывающей головки на k-м месте; i_{l+j} ($j = 1, 2, \ldots, n-l$) — номер массива, расположенного слева от считывающей головки на j-м месте. Тогда для одной ленты будем иметь:

Правая часть ленты

$$T_п = \frac{1}{V_п}(y_0 P_{i_1} + yP_{i_2} + 2yP_{i_3} + \ldots + (l-1)\,yP_{i_l}) = \frac{y}{V_п}\sum_{k=1}^{l-1} P_{i_{k+1}} k, \quad (6)$$

где

$$y_0 = 0, \qquad \overline{T}_с = \frac{y}{V_с}\sum_{k=1}^{l} P_{i_k}, \quad (7)$$

$$\overline{T}_в = \frac{y}{V_в}(P_{i_1} + P_{i_2} + \ldots + lP_{i_l}) = \frac{y}{V_в}\sum_{k=1}^{l} kP_{i_k}, \quad (8)$$

$$\overline{T}_{обр_п} = \frac{y}{V_п}\sum_{k=1}^{l-1} kP_{i_{k+1}} + \frac{y}{V_с}\sum_{k=1}^{l} P_{i_k} + \frac{y}{V_в}\sum_{k=1}^{l} kP_{i_k}. \quad (9)$$

Левая часть ленты

$$\overline{T}_п = \frac{1}{V_п}(yP_{i_{l+1}} + 2yP_{i_{l+2}} + \ldots + (n-l)\,yP_{i_п} = \frac{y}{V_п}\sum_{j=1}^{n-l} jP_{i_{l+j}}, \quad (10)$$

$$\overline{T}_с = \frac{y}{V_с}\sum_{j=1}^{n-l} P_{i_{l+j}}, \quad (11)$$

$$\overline{T}_в = \frac{y}{V_в}\sum_{j=1}^{n-l-1} jP_{i_{l+j}}, \quad (12)$$

$$\overline{T}_{обр\ л} = \frac{y}{V_п}\sum_{j=1}^{n-l} jP_{i_{l+j}} + \frac{y}{V_с}\sum_{j=1}^{n-l} P_{i_{l+j}} + \frac{y}{V_в}\sum_{j=1}^{n-l-1} jP_{i_{l+j+1}}. \quad (13)$$

Полное среднее время обращения равно

$$\overline{T}_{обр} = \overline{T}_{обр_п} + \overline{T}_{обр\ л} = y\sum_{k=1}^{l} P_{i_k}\left[\frac{k-1}{V_п} + \frac{k}{V_в} + \frac{1}{V_с}\right] +$$

$$+ y\sum_{j=1}^{n-l} P_{i_{l+j}}\left[\frac{j}{V_п} + \frac{1}{V_с} + \frac{j-1}{V_в}\right]. \quad (14)$$

Если $V_в = V_п$, то правую и левую части МЛ можно рассматривать как две независимые ленты.

Пусть массивы упорядочены так, что $P_1 \geqslant P_2 \geqslant P_3 \geqslant \ldots \geqslant P_n$. Тогда в соответствии с результатом, полученным в [1], в оптимальном решении массивы с P_{2i-1} и P_{2i} должны располагаться на i-х местах (один массив справа от считывающей головки, а другой — слева). В общем случае $V_в > V_п$ и такое размещение может не быть оптимальным. Если при $V_в = V_п$ безразлично, какой из массивов (с P_{2i-1} или P_{2i}) размещается справа, то при $V_в > V_п$ это уже не так. Справа от считывающей головки следует располагать массив с большим значением вероятности обращения, т. е. P_{2i-1}. Таким образом, все массивы с нечетными номерами располагаются справа от считывающей головки в порядке убывания P_i, все массивы с четными номерами располагаются слева от считывающей головки также в порядке убывания P_i. Оценим отклонение от оптимальности в таком решении при $V_в \neq V_п$. Очевидно, ошибка максимальна при $\sum_{k=1}^{l} P_{i_k} = \sum_{j=1}^{n-l} P_{i_{l+j}} = 1/2$, так как в противном случае на левую часть ленты всегда можно поместить массивы с меньшим значением ΣP_{i_k}. С другой стороны, наибольшее значение величины

$$\frac{y}{V_п}\sum_{k=1}^{l} P_{i_k} + \frac{y}{V_в}\sum_{j=1}^{n-l} P_{i_{l+j}} \tag{15}$$

достигается при $l = n$ и равно $y / V_п$. Таким образом, ошибка не превышает

$$\frac{y}{2V_п} - \frac{y}{2V_в} = \frac{y(V_в - V_п)}{2V_п V_в}. \tag{16}$$

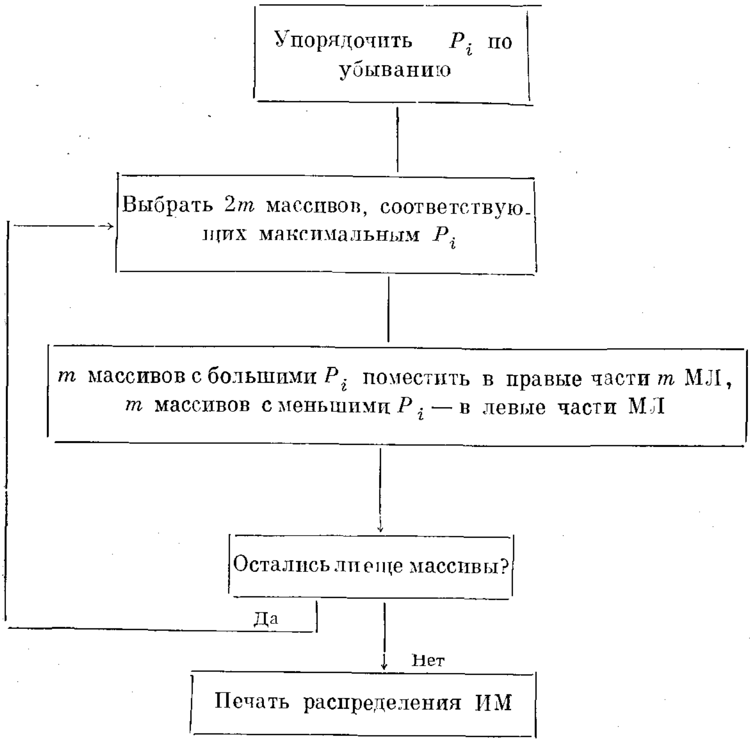

Рис. 1

Следовательно, даже при $V_в \gg V_п$ ошибка по абсолютной величине не больше

$$\varepsilon = y / 2V_п, \tag{17}$$

а значит, очень мала.

В случае m лент среднее время обращения $\overline{T}_{\text{обр}}$ будет равно

$$\overline{T}_{\text{обр}} = y \sum_{s=1}^{m} \left\{ \sum_{k=1}^{l_s} P_{i_k^s} \left[\frac{k-1}{V_{\text{п}}} + \frac{1}{V_{\text{c}}} + \frac{k}{V_{\text{в}}} \right] + \right.$$
$$\left. + \sum P_{i_{l+j}^s} \left[\frac{j}{V_{\text{п}}} + \frac{1}{V_{\text{c}}} + \frac{j-1}{V_{\text{в}}} \right] \right\}, \tag{18}$$

где i_k^s — номер ИМ, находящегося на k-м месте s-й ленты. Для минимизации $\overline{T}_{\text{обр}\, m}$ можно предложить алгоритм (рис. 1), который в случае $V_{\text{п}} = V_{\text{в}}$ дает оптимальное решение, а в случае $V_{\text{п}} < V_{\text{в}}$ решение будет отличаться от оптимального не более чем на $\varepsilon = y / 2V_{\text{п}}$, независимо от числа лент.

ЗАДАЧА 2

Перейдем теперь к рассмотрению следующей задачи. Пусть условия остаются те же, что и в задаче 1, только длины массивов y_i могут быть любыми. В качестве критерия будем использовать не среднее время считывания любого массива, а максимальное взвешенное время считывания ИМ, т. е. произведение времени считывания ИМ на вероятность обращения к нему. Тогда для одной ленты получим:

Правая часть МЛ

$$\overline{T}_{i_k} = P_{i_k} \left[\frac{1}{V_{\text{п}}} \sum_{j<k} y_{i_j} + \frac{y_{i_k}}{V_{\text{c}}} + \frac{1}{V_{\text{в}}} \sum_{j \leqslant k} y_{i_j} \right], \tag{19}$$

где $\overline{T}_{i_k}$ — время считывания ИМ, находящегося на k-м месте, P_{i_k} — вероятность обращения к ИМ на k-м месте, y_{i_j} — длина массива на j-м месте.

Левая часть МЛ

$$\overline{T}_{i_k} = P_{i_k} \left[\frac{1}{V_{\text{п}}} \left(\sum_{j \leqslant k} y_{i_j} \right) + \frac{y_{i_j}}{V_{\text{c}}} + \frac{1}{V_{\text{в}}} \sum_{j<k} y_{i_j} \right]. \tag{20}$$

Из (19) и (20) следует, что обе части МЛ можно рассматривать как независимые ленты с однонаправленным поиском и положением возврата в начальной точке. Ошибка при этом будет не больше

$$\varepsilon = \frac{V_{\text{в}} - V_{\text{п}}}{V_{\text{п}} V_{\text{в}}} \max_i P_i y_i. \tag{21}$$

Таким образом, решим следующую задачу:

$$\max_k P_{i_k} \left[\frac{1}{V_{\text{п}}} \sum_{j<k} y_{i_j} + \frac{y_{i_k}}{V_{\text{c}}} + \frac{1}{V_{\text{в}}} \sum_{j \leqslant k} y_{i_j} \right] \to \min. \tag{22}$$

Предположим, что известно максимальное время поиска ИМ. Обозначим его M. Тогда

$$P_{i_k} \left[\frac{1}{V_{\text{п}}} \sum_{j<k} y_{i_j} + \frac{y_{i_i}}{V_{\text{c}}} + \frac{1}{V_{\text{в}}} \sum_{j \leqslant k} y_{i_j} \right] \leqslant M. \tag{23}$$

Запишем (23) в несколько иной форме:

$$\left(\frac{1}{V_{\text{п}}} + \frac{1}{V_{\text{в}}} \right) \sum_{1}^{k-1} y_{i_j} \leqslant \frac{M}{P_{i_k}} - \left(\frac{1}{V_{\text{c}}} + \frac{1}{V_{\text{в}}} \right) y_{i_k}. \tag{24}$$

Допустим, что номера ИМ совпадают с номерами мест на МЛ. Тогда, вводя обозначения:

$$x_i = \left(\frac{1}{V_k} + \frac{1}{V_\text{в}}\right) y_i, \tag{25}$$

$$z_i = \left(\frac{1}{V_\text{с}} + \frac{1}{V_\text{в}}\right) y_i, \tag{26}$$

получим

$$\sum_{1}^{k-1} x_i \leqslant \frac{M}{P_k} - z_k \tag{27}$$

или

$$\sum_{1}^{k} x_i \leqslant \frac{M}{P_k} - z_k + x_k. \tag{28}$$

Наконец, обозначив

$$\frac{M}{P_k} - z_k + x_k = \delta_k, \tag{29}$$

запишем (28) в виде

$$\sum_{1}^{k} x_i \leqslant \delta_k. \tag{30}$$

В левой части (30) стоит сумма длин ИМ, расположенных на первых k местах МЛ, умноженная на постоянный коэффициент. Ясно, что с ростом k левая часть (30) монотонно растет. Очевидно, для того чтобы обеспечить максимально равномерное выполнение неравенства (30) нужно, чтобы правая его часть также монотонно росла. Действительно, если при каком-то размещении ИМ окажется, что $\delta_s > \delta_{s+1}$, то можно записать:

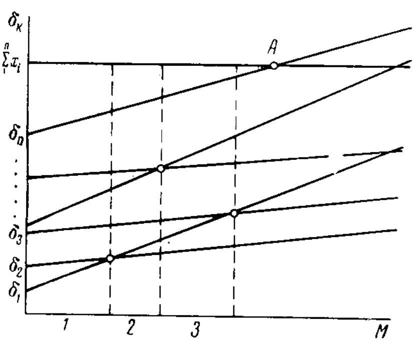

Рис. 2

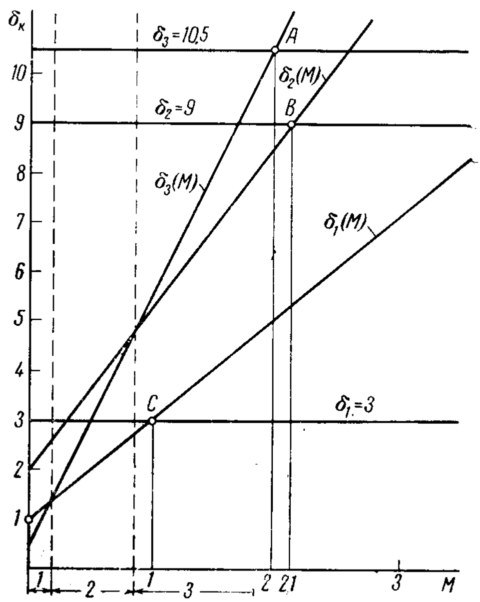

Рис. 3

$$\sum_{1}^{s-1} x_i + x_s \leqslant \delta_s, \tag{31}$$

$$\sum_{1}^{s-1} x_i + x_s + x_{s+1} \leqslant \delta_{s+1}. \tag{32}$$

Если поменять местами ИМ на s-м и $(s+1)$-м местах, то получим:

$$\sum_{1}^{s-1} x_i + x_{s+1} \leqslant \delta_{s+1}, \tag{33}$$

$$\sum_{1}^{s-1} x_i + x_{s+1} + x_s \leqslant \delta_{s+1} < \delta_s, \tag{34}$$

т. е. если $\delta_s > \delta_{s+1}$, то из (31), (32) тем более следует (33), (34). Следовательно, можно утверждать, что оптимальное упорядочение ИМ при минимаксном критерии есть упорядочение по возрастанию δ_k.

Рассмотрим теперь процедуру, которая позволяет найти оптимальное упорядочение для конкретных значений y_k, P_k $(k = 1, 2, ..., n)$.

Построим графики функций $\delta_k = \delta_k(M)$. Очевидно, эти функции представляют собой прямые, определенные для $M \geqslant 0$ (рис. 2). В плоскости $\delta_k M$ можно выделить области, для которых упорядочение δ_k не меняется (области *1*, *2*, *3* и т. д. на рис. 2).

В соответствии с предыдущим утверждением искомое решение должно находиться в одной из таких областей. Поиск можно начинать с самой первой области и двигаться к последней в порядке их расположения. Можно несколько сократить количество просматриваемых областей.

Действительно, нам известна $\sum_{i=1}^{n} y_i$ и, следовательно, $\sum_{1}^{n} x_i$, поэтому, проведя прямую $\delta_k = \sum_{1}^{n} x_i$ и найдя первую точку пересечения ее с прямыми $\delta_k = \delta_k(M)$, тем самым определим первую область, с которой следует начинать поиск. Поиск решения будем проводить следующим образом. Допустим, что ИМ упорядочены в соответствии с порядком δ_k в рассматриваемой области. Вычислим все Σx_i при принятом порядке ИМ и проверим выполнение неравенства (30) для всех δ_k. Если хотя бы в одном случае неравенство (30) не выполняется, перейдем в соседнюю область и повторим всю рассмотренную процедуру, если же в проверяемой области неравенство (30) выполняется для всех δ_k, то это означает, что требуемое упорядочение ИМ найдено. Остается только определить численное значение максимального взвешенного времени обращения. Это несложно сделать, проведя прямые Σx_i для всех индексов i и найдя точки их пересечения с соответствующими по порядку в выбранной области прямыми $\delta_k = \delta_k(M)$. Максимальная абсцисса точек пересечения равна искомому значению максимального взвешенного времени обращения.

Таблица 1

i	1	2	3
P_i	0,5	0,3	0,2
y_i	2	4	1

Для иллюстрации предложенной процедуры оптимального размещения рассмотрим простой пример.

Пример

Пусть на МЛ с параметрами $V_п = 1$, $V_в = V_с = 2$ нужно поместить три ИМ, для которых известны P_i и y_i (см. табл. 1). При этом требуется, чтобы $\max_k T_{обр\,k} \to \min$.

Пользуясь формулой (20), найдем времена обращения ко всем массивам для различных порядков расположения массивов на МЛ. Запишем их в табл. 2.

Из табл. 2 видно, что наилучшим в смысле выбранного критерия является порядок 1—2—3.

Таблица 2

Место ИМ	$T_{\text{обр}}$ при порядке ИМ					
	1—2—3	1—3—2	2—1—3	2—3—1	3—1—2	3—2—1
1	1,0	1,0	1,2	1,2	0,2	0,2
2	2,1	0,8	4,0	1,4	1,75	1,65
3	2,0	2,55	2,0	4,75	2,55	4,75

Проверим теперь предложенное правило. Для этого построим семейство прямых $\delta_k = \delta_k(M)$ (рис. 3) и выделим на нем области одинакового упорядочения ИМ (области *1, 2, 3* на рис. 3). Вычислим $\sum\limits_1^n x_i$. В данном случае

$$\sum_1^n x_i = \sum_1^3 x_i = \left(\frac{1}{V_{\text{п}}} + \frac{1}{V_{\text{в}}}\right) \sum_1^3 y_i = (1 + 0{,}5)(2 + 4 + 1) = 10{,}5.$$

На графике рис. 3 строим прямую $\delta = 10{,}5$. Первая точка пересечения этой прямой с прямой $\delta_3 = \delta_3(M)$, точка A, оказывается в третьей области. Упорядочиваем ИМ по порядку δ_k в этой области. Видим, что получаемый при этом порядок 1—2—3 совпадает с оптимальным, полученным вначале методом полного перебора всех вариантов по табл. 2. Вычислим теперь $\sum\limits_1^k x_i$ для выбранного порядка ($k = 1, 2$):

$$\sum_1^1 x_i = (1 + 0{,}5)2 = 3, \qquad \sum_1^1 x_i = (1 + 0{,}5)(2 + 4) = 9$$

и построим прямые $\delta_1 = 3$ и $\delta_2 = 9$. Находим точки пересечения прямых $\delta = \delta_1$, $\delta = \delta_2$ и $\delta = \delta_3$ с прямыми $\delta = \delta_1(M)$, $\delta = \delta_2(M)$ и $\delta = \delta_3(M)$ соответственно. Абсциссы этих точек дают следующие значения для взвешенных времен обращения к каждому ИМ: $T_1 = 1$, $T_2 = 2,1$, $T_3 = 2$. Видим, что максимальное время равно $T_{\text{обр. макс}} = 2,1$, т. е. совпадает с $T_{\text{обр. макс}}$ первого столбца табл. 2. Таким образом, мы еще раз убедились, что предложенная процедура дает строго оптимальное решение задачи о размещении ИМ на одной ленте.

В случае m лент при минимаксном критерии ситуация осложняется, так как здесь неизвестна заранее Σx_i для одной ленты. Здесь можно предложить следующий алгоритм размещения. Построим снова семейство прямых $\delta_k = \delta_k(M)$ $(k = 1, 2, \ldots, n)$ и выделим области постоянного порядка δ_k. Будем искать решение в каждой из выделенных областей, просматривая их либо последовательно, либо методом дихотомии. Примем следующую процедуру поиска решения внутри каждой области: разобьем все множество δ_k на группы с количеством элементов m в каждой (в последней группе число элементов может быть меньше m). Массивы, соответствующие первой группе δ_k, поместим на первые места m лент; массивы, соответствующие второй группе δ_k — на вторые места в той же последовательности по m, что и в первом случае, и т. д. В результате получим упорядочение ИМ, которое будет оптимальным по каждой ленте. Если при этом удовлетворяется неравенство (30), то можно считать, что получено локально-оптимальное решение. Это решение можно попытаться улучшить, переставляя ИМ с ленты с большим минимаксом на ленту с меньшим при сохранении порядка по δ_k.

Если внутри рассматриваемой зоны неравенство (30) не выполняется, то нужно перейти к оценке следующей области и повторять весь рассмотренный процесс до получения ближайшей к началу области, где неравество (30) выполняется. С учетом того, что количество областей конечно, предлагаемый алгоритм сходится за конечное число шагов. Можно предложить еще один алгоритм оптимального размещения ИМ для m лент, который во многих случаях дает решение лучше предыдущего. Построим снова прямые $\delta_k = \delta_k(M)$ и выделим области постоянного порядка δ_k. Опять будем искать решение в каждой из областей, начиная с первой, следующим образом: для рассматриваемой области найдем все $\sum\limits_{1}^{k} x_i$ для всех индексов k. Проверим неравенство (30) для всех δ_k, начиная с минимального. Если неравенство выполняется, то соответствующий ИМ ставится на отведенное для него по порядку δ_k место на первой ленте и проверяется неравенство для следующего δ_k. Если оно и для него выполняется, то новый ИМ заносится на следующее по порядку место первой МЛ и т. д. до заполнения первой ленты. Затем, если и при последующих проверках всюду выполняются неравенства (30), точно таким же образом заполняются все остальные ленты. В случае, если для какого-то δ_k неравенство не выполняется, ищется лента, на которой неравенство (30) выполняется для данного δ_k при минимальной разности $\delta_k - x_k - \sum\limits_{i \in Q} x_i$, где Q — множество массивов на данной ленте. Массив, соответствующий рассматриваемому δ_k, помещается следующим по порядку на найденную ленту. После этого проверяется следующее по порядку δ_k и производится заполнение первой из незаполненных МЛ и т. д. В случае, если внутри рассматриваемой области удается найти минимальное M и такое распределение ИМ по МЛ, при котором неравенство (30) выполняется для всех $\delta_k(M)$, это распределение и есть искомое решение. Если решение не находится в рассматриваемой области, то проверяется соседняя область и т. д. Весь процесс повторяется до получения решения.

Рассмотрим в заключение одну из задач размещения ИМ для стратегии 2.

ЗАДАЧА 3. СТРАТЕГИЯ УПРАВЛЕНИЯ МЛ БЕЗ ВОЗВРАТА

Пусть теперь управление МЛ выполнено так, что лента после каждого обращения остается в том месте, в котором она оказалась в результате обращения.

Допустим, что снова требуется разместить n ИМ по m МЛ таким образом, чтобы среднее время обращения ко всем ИМ было минимальным.

Для массивов снова известны их длины y_i, вероятности обращений к ним P_i и, кроме того, попарные вероятности P_{ij}, т. е. вероятности обращения к j-м ИМ после обращения к i-м $(i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n)$.

Ограничимся пока решением задачи 3 для случая $m = 1$ и $y_i = y$. Легко показать, что в этом случае среднее время обращения равно

$$\overline{T}_{\text{обр}} = \frac{y}{V_{\text{п}}} \sum_{k=1}^{n} P_{i_k} \left[\sum_{s=k+1}^{n} P_{i_k i_s}(s-k-1) + \sum_{s=1}^{k} P_{i_k i_s}(k-s+1) \right] + \\ + \frac{y}{V_{\text{c}}} \sum_{1}^{n} P_{i_k}, \tag{35}$$

где i_k — номер ИМ, находящегося на k-м месте.

Взаимная зависимость обращений к ИМ сильно затрудняет задачу минимизации формы (35) в смысле получения глобального оптимума. Однако, как будет показано ниже, существует условие, позволяющее получить довольно простой алгоритм поиска локально-оптимальных решений.

Рассмотрим любое фиксированное упорядочение ИМ и упорядочение, отличающееся от исходного расположением любых двух соседних массивов, например l-го и $(l+1)$-го. Пусть на этих местах стояли массивы с номерами i и j. Найдем разность средних времен обращений, соответствующих двум рассматриваемым порядкам ИМ:

$$\Delta T_{ij} = \frac{y}{V_{\text{п}}}\Bigg[\sum_{k=1}^{i-1}(P_k P_{ki} + P_i P_{ik}) + \sum_{i+2}^{n}(P_k P_{k(i+1)} + P_{(i+1)} P_{(i+1)k}) +$$

$$+ 2P_i P_{i+1} - \sum_{1}^{i-1}(P_k P_{i+1} + P_{i+1} P_k) -$$

$$- \sum_{i+2}^{n}(P_k P_{ki} + P_i P_{ik}) - 2P_{i+1} P_{(i+1)i}\Bigg]. \qquad (36)$$

Здесь со знаком плюс записаны члены, учитывающие проигрыш от перестановки l-го и $(l+1)$-го ИМ, а со знаком минус — члены, учитывающие выигрыш.

Обозначим:

$$q_{ij} = P_i P_{ij}, \qquad (37)$$

$$S_{ki} = q_{ki} + q_{ik}. \qquad (38)$$

Тогда с учетом того, что

$$\sum_{1}^{n} q_{(i+1)k} = P_{(i+1)}; \qquad (39)$$

и

$$\sum_{i+2}^{n} S_{(i+1)k} = 2P_{(i+1)} - \sum_{1}^{i+1} S_{(i+1)k}, \qquad (40)$$

для разности средних времен получим с точностью до постоянного множителя

$$\Delta T_{ij} = \sum_{1}^{i-1}(S_{ki} - S_{k(i+1)}) + (P_{(i+1)} - P_i) +$$

$$+ (q_{i(i+1)} - q_{(i+1)i}) + (q_{ii} - q_{(i+1)(i+1)}) \qquad (41)$$

или, обозначив

$$\sigma_{ij} = (P_{i+1} - P_i) + (q_{i(i+1)} - q_{(i+1)i}) + (q_{ii} - q_{(i+1)(i+1)}), \qquad (42)$$

окончательно будем иметь

$$\Delta T_{ij} = \sum_{k=1}^{i-1}(S_{ki} - S_{k(i+1)}) + \sigma_{ij}. \qquad (43)$$

Легко видеть, что выражение (43) не зависит от порядка ИМ, стоящих до и после выбранной пары. Следовательно, его можно использовать для оценки эффекта перестановки любых двух соседних ИМ. Очевидно, что порядок ИМ в рассматриваемой паре должен меняться на обратный всякий раз, когда выполняется условие отрицательности ΔT_{ij}. Таким образом, необходимое условие оптимальности порядка ИМ можно записать в виде

$$\Delta T_{i_l i_{l+1}} > 0 \qquad (44)$$

$(l = 1, 2, \ldots, n-1$; i_l — номер ИМ на l-м месте$)$.

Условие (44) позволяет сразу получить тривиальный алгоритм поиска допустимого решения. Действительно, можно взять любое начальное упорядочение и перестановкой соседних ИМ до выполнения условия (44) получить допустимое решение. Однако легко убедиться, что таких решений может быть довольно много, поэтому условие (44) не является достаточным условием оптимальности. Это и понятно. Оно сформулировано только для перестановок внутри пар соседних ИМ и никак не учитывает всех остальных возможных перестановок.

Процедура получения локально-оптимального решения из случайного исходного упорядочения ИМ может оказаться довольно трудоемкой. Можно предложить алгоритм, в этом смысле лучший. Вычислим на основе исходных данных матрицы $\|q_{ij}\|$, $\|S_{ij}\|$ и $\|\sigma_{ij}\|$.

Так как для первой пары любого упорядочения первый член правой части выражения (43) равен нулю, то в качестве первой пары можно выбрать по матрице $\|\sigma_{ij}\|$ любые ИМ, для которых $\Delta T_{ij} > 0$.

Пусть, например, выбраны ИМ k и l. Найдем ΔT_{lj}, где j пробегает все значения, кроме l и k. Поместим после l любой из оставшихся ИМ, для которого $\Delta T_{lj} > 0$. Пусть это массив с номером r. Вычислим ΔT_{rj}, где j принимает все значения, кроме k, l и r, и повторим всю рассмотренную процедуру до размещения последнего ИМ. Если окажется, что для последнего ИМ условие $\Delta T_{ij} > 0$ выполняется, то решение получено и процедура заканчивается; в противном случае делается шаг назад и выбирается другая допустимая пара.

В заключение рассмотрим пример, иллюстрирующий предложенный алгоритм. Пусть задано пять ИМ с вероятностями обращения $P_1 = 0{,}24$, $P_2 = 0{,}28$, $P_3 = 0{,}20$, $P_4 = 0{,}16$ и $P_5 = 0{,}12$ и найдена матрица

$$\|q_{ij}\| = \begin{Vmatrix} 0{,}10 & 0{,}06 & 0{,}02 & 0{,}02 & 0{,}04 \\ 0{,}02 & 0{,}12 & 0{,}04 & 0{,}08 & 0{,}02 \\ 0{,}04 & 0{,}06 & 0{,}06 & 0{,}02 & 0{,}02 \\ 0{,}02 & 0{,}02 & 0{,}06 & 0{,}02 & 0{,}04 \\ 0{,}06 & 0{,}02 & 0{,}02 & 0{,}02 & 0 \end{Vmatrix}.$$

Пользуясь (38) и (42), находим матрицы $\|S_{ij}\|$ и $\|\sigma_{ij}\|$:

$$\|S_{ij}\| = \begin{Vmatrix} 0{,}20 & 0{,}08 & 0{,}06 & 0{,}04 & 0{,}10 \\ 0{,}08 & 0{,}24 & 0{,}10 & 0{,}10 & 0{,}04 \\ 0{,}06 & 0{,}10 & 0{,}12 & 0{,}08 & 0{,}04 \\ 0{,}04 & 0{,}10 & 0{,}08 & 0{,}04 & 0{,}06 \\ 0{,}10 & 0{,}04 & 0{,}04 & 0{,}06 & 0 \end{Vmatrix},$$

$$\|\sigma_{ij}\| = \begin{Vmatrix} 0 & 0{,}06 & -0{,}02 & 0 & -0{,}04 \\ -0{,}06 & 0 & -0{,}04 & 0{,}04 & -0{,}04 \\ 0{,}02 & 0{,}04 & 0 & -0{,}04 & -0{,}02 \\ 0 & -0{,}04 & 0{,}04 & 0 & 0 \\ 0{,}04 & 0{,}04 & 0{,}02 & 0 & 0 \end{Vmatrix}.$$

По матрице $\|\sigma_{ij}\|$ выбираем первую пару ИМ. Здесь возможен следующий выбор: 1—2, 2—4, 3—1, 3—2, 4—3, 5—1, 5—2, 5—3. Возьмем для большей устойчивости решения пару 1—2 с максимальным значением σ_{ij} ($\sigma_{12} = 0{,}06$). Находим далее ΔT_{2j} ($j = 3, 4, 5$):

$$\begin{aligned} \Delta T_{23} &= \sigma_{23} + S_{12} - S_{13} = -0{,}04 + 0{,}08 - 0{,}06 = -0{,}02, \\ \Delta T_{24} &= \sigma_{24} + S_{12} - S_{14} = 0{,}04 + 0{,}08 - 0{,}04 = 0{,}08, \\ \Delta T_{25} &= \sigma_{25} + S_{12} - S_{15} = -0{,}04 + 0{,}08 - 0{,}10 = -0{,}06. \end{aligned}$$

Получили единственное положительное значение ΔT, ΔT_{24}. Значит, следующим по порядку за ИМ с номером 2 должен стоять ИМ с номером 4. Фиксируем место четвертого ИМ. Находим ΔT_{4j} ($j = 3; 5$):

$$\begin{aligned} \Delta T_{4j} &= \sigma_{4j} + (S_{14} - S_{1j}) + (S_{24} - S_{2j}), \\ S_{14} + S_{24} &= 0{,}14, \\ \Delta T_{43} &= 0{,}14 + 0{,}04 - 0{,}16 = 0{,}02, \\ \Delta T_{45} &= 0{,}14 + 0 - 0{,}14 = 0. \end{aligned}$$

Ставим за четвертым ИМ массив с номером 3. Тогда остается последний, пятый, ИМ, который автоматически помещается после третьего ИМ. Проверяем условие $\Delta T_{35} > 0$:

$$\begin{aligned} \Delta T_{35} &= \sigma_{35} + (S_{13} + S_{23} + S_{43}) - (S_{15} + S_{25} + S_{45}) = \\ &= -0{,}02 + 0{,}24 - 0{,}20 = 0{,}02 > 0. \end{aligned}$$

Условие выполняется, следовательно, решение найдено. Порядок ИМ при найденном решении 1—2—4—3—5.

Заключение

Решение задач оптимального размещения ИМ в памяти на магнитных лентах представляет большой практический интерес.

Действительно, это позволяет уменьшить время и стоимость решения основных задач и повысить эффективность использования аппаратуры вычислительной системы.

С другой стороны, при этом проявляется еще один интересный эффект. Оптимизация размещения ИМ уменьшает средний рабочий путь МЛ и, следовательно, облегчает режим их работы. В результате повышается надежность и долговечность памяти.

В настоящей работе рассмотрены алгоритмы решения трех задач оптимального размещения ИМ в памяти на МЛ для случая двунаправленного поиска. Задачи 1 и 3 решены при ограничениях $y_i = y = \mathrm{const}$ (в задаче 1) и $m = 1$ (в задаче 3). Представляет интерес обобщение решения на случай произвольных длин ИМ и произвольного числа блоков МЛ.

Поступила в редакцию
15 мая 1968 г.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. Соколов В. Б., Бурков В. Н. Оптимальное размещение информации в памяти большого объема. Доклад на конференции «Проблемы создания больших информационных вычислительных систем и обработки информации на ЭВМ». Киев, 27—30 мая 1968 г.

OPTIMAL INFORMATION AUOCATION IN THE MAGNETIC TAPE MEMORY FOR BIDIRECTIONAL SEARCH

V. N. BURKOV, V. B. SOKOLOV

The problems of optimal information auocation in the magnetic tape memory for different strategies of tape control and different optimization criteria are considered.
