Управляемые системы, ИМ СО РАН, 1993 г., вып. 31

УДК 519.8

АЛГОРИТМ С ОЦЕНКОЙ ДЛЯ ЗАДАЧИ МИНИМИЗАЦИИ МАКСИМАЛЬНОГО ЗАПАЗДЫВАНИЯ НА ОДНОРОДНЫХ МАШИНАХ

А. В. Кононов

В статье рассматривается задача составления расписания для независимых работ на параллельных машинах. Машины имеют разные скорости. Для каждой работы известны директивные сроки начала и завершения обслуживания. Требуется найти такой порядок запуска работ и их распределение по машинам, чтобы минимизировать максимальное запаздывание.

Впервые задачи такого типа рассматривались Джексоном [1]. Доказательство сильной NP-полноты задачи для одной машины приведено в [2]. Обзор результатов в данной области можно найти в [3].

В работе изучается приближенный алгоритм решения задачи, который при загрузке машин отдает предпочтение работе с наиболее ранним директивным сроком завершения (*EDD-earliest due date*). Алгоритм имеет трудоемкость $O(n\cdot\log n+m\cdot\log m)$, n - число работ, m - количество машин. Предлагается оценка абсолютного отклонения полученного решения от оптимального

$$L_{max}-L^*_{max} \le \left(\max_{i=1..n} P_i / \min_{l=1..m} S_l\right)\left(1+(m-1)\cdot \min_{l=1..m} S_l/S\right),$$

где

S_l - скорость машины, $S = \sum_{l=1..m} S_l$,

P_i - нормативное время выполнения i-й работы, $i=1,\dots,n$. Аналогичный результат при $S_l=1$, $l=1,\dots,m$ содержится в работе [4].

1. Постановка задачи

Пусть $I=\{1,\dots,n\}$ - множество независимых работ, выполнение которых возможно на любой из m параллельных машин. Для каждой работы $i\in I$ известны:

$r_i \ge 0$ - время поступления работы на обслуживание;

$P_i > 0$ - время на выполнение работы;

$d_i \ge 0$ - директивный срок завершения обработки.

Каждая работа i начинает обслуживаться не ранее момента времени r_i и выполняется только на одной машине. Прерывания в обслуживании работ не разрешаются. Каждая машина j из множества $M=\{1,\dots,m\}$ имеет свою ско-

рость $S_j \geq 0$ и может обслуживать в любой момент времени не более одной работы. Если работа i выполняется на машине j, то время ее обслуживания есть $p_{ij}=p_i/S_j$. Пусть t_{ij} - начало выполнения работы i на машине j. Величину $L_i=t_{ij}+p_i/S_j-d_i$ назовем запаздыванием работы i, а величину $L_{max}=\max\limits_{i\in I} L_i$ - максимальным запаздыванием. Под расписанием будем понимать множество $L=\{t_{ij}:\ i\in I,\ j\in M\}$ времен начала обслуживания, удовлетворяющих приведенным выше условиям. Задача состоит в нахождении расписания L^*, для которого величина L_{max} принимает наименьшее значение.

Пусть $q_i=D-d_i$, где $D \geq \max\limits_{i\in I} d_i$. Величину q_i принято называть остаточным временем обслуживания или хвостом [2]. Обозначим $f_i=t_{ij}+p_i/S_j+q_i$ - время полного обслуживания работы i, тогда исходная задача может быть записана в следующем виде:

$$\max_{i\in I} f_i \longrightarrow \min_{(t_{ij})} .$$

2. Алгоритм EDD.

Рассмотрим следующий приближенный полиномиальный алгоритм решения исходной задачи. Выберем из множества незагруженных машин самую быструю и назначим на нее готовую к выполнению работу с наибольшей величиной q_i. Пусть t - текущее время, $\tau(j)$ - время освобождения j-й машины.

Алгоритм:

0. Упорядочим множество I по невозрастанию величин q_i, множество M - по невозрастанию величин S_j и будем считать, что $q_1\geq q_2\geq\ldots\geq q_n$, $S_1\geq S_2\geq\ldots\geq S_m$.
1. Полагаем $t=\min r_i$, $\tau(j)=0$ для всех $j\in M$, $I^0=I$.
2. Находим минимальный номер i, для которого $r_i\leq t$, $i\in I^0$.
3. Находим минимальный номер j, для которого $\tau(j)\leq t$.
4. Полагаем $t_{ij}=t$, $I^0=I^0\setminus\{i\}$, $\tau(j)=t+p_i/S_j$.
5. Если $I^0=\emptyset$, то конец, в противном случае полагаем $t=\min \tau(j)$ и возвращаемся на 2.

Трудоемкость алгоритма определяется сложностью упорядочения множеств I, M и оценивается величиной $O(n\cdot\log n+m\cdot\log m)$.

3. Нижняя граница.

Пусть $J\subseteq I$ произвольное подмножество I, и $f^*=\min\limits_{(t_{ij})} \max\limits_{i\in I} f_i$.

Лемма 1. Справедлива оценка

$$f^*\geq G(J)=\min_{i\in J} r_i+\sum_{i\in J} p_i/S+\min_{i\in J} q_i. \qquad (1)$$

Доказательство. В оптимальном расписании каждая ра-

бота стартует не раньше $\min_{i\in I} r_i$ и заканчивается не позже $f^*-\min_{i\in J} q_i$. Т.е. все работы из J производятся в промежуток времени $[\min_{i\in J} r_i;\ f^*-\min_{i\in J} q_i]$. Время выполнения всех работ из J не превышает величины $\sum_{i\in J} P_i/S$, т. е. $\sum_{i\in J} P_i/S \le f^*-\min_{i\in J} r_i - \min_{i\in J} q_i$. Откуда получаем требуемое.■

Пусть $P(I)$ - множество подмножеств I. Величина $G=\max_{J\in P(I)} G(J)$ также является нижней оценкой для f^*.

Лемма 2. Величина $G=\max_{J\in P(I)} G(J)$ может быть найдена за $O(n\cdot \log n)$ операций.

Доказательство. Рассмотрим задачу с прерываниями на одной машине и набором параметров $(r_i,P_i/S,q_i)$ для каждой работы $i\in I$. Требуется минимизировать время полного обслуживания всех работ:

$$f=\min_{\bar t_i}\max_{i\in I}(\bar t_i+q_i), \tag{2}$$

где $\bar t_i$ - время завершения работы i. Оптимальное решение этой задачи находится с трудоемкостью $O(n\cdot \log n)$ при помощи следующего алгоритма [5]. Начиная с момента времени $t=\min_{i\in I} r_i$, машина обслуживает из множества работ, готовых к выполнению, работу с наибольшим q_i. Если во время выполнения некоторой работы в очередь приходит работа j с $q_i<q_j$, то обслуживание i-й работы прерывается, и на обслуживание поступает работа j.

Покажем, что $G=f$. Заметим, что для этой задачи G является нижней границей. Т. е. $G\le f$. Убедимся теперь, что $G\ge f$. Пусть i_0-работа, для которой $f=\bar t_{i_0}+q_{i_0}$. Обозначим через t последний момент времени в промежутке $[0,\bar t_{i_0}]$, когда либо обслуживалась деталь j такая, что $q_j<q_{i_0}$, либо машина простаивала без работы. Пусть J-множество работ, обслуживаемых в промежутке $[t,\bar t_{i_0}]$. Если $j\in J$, то $r(j)\ge t$, $q(j)\ge q(i_0)$. Кроме того, $J\ne\emptyset$, так как $i_0\in J$. Тогда

$$G(J)=\min_{i\in J} r_i+\sum_{i\in J} P_i/S+\min_{i\in J} q_i \ge t+\sum_{i\in J} P_i/S+\min_{i\in J} q_i=\bar t_{i_0}+q_{i_0}=f.$$

Лемма 2 доказана.■

4. Основное утверждение.

В этом разделе приводится неулучшаемая оценка качества алгоритма EDD. Пусть f-значение целевой функции, полученное алгоритмом EDD.

Теорема 1. Для погрешности $\delta=f-f^*$ справедлива оценка

$$\delta\le(\max_{i\in I} P_i/\min_{l\in M} S_l)\cdot(1+(m-1)\cdot\min_{l\in M} S_l/S). \tag{3}$$

Д о к а з а т е л ь с т в о. Пусть I_0 - работа, на которой достигается значение f, т. е.

$$t_{i_0}+P_{i_0}/S_{k_0}+q_{i_0}=f, \qquad (4)$$

где k_0 - машина, обслуживающая работу I_0. Введем два специальных множества.

$J=\{j\in I:\ t_j<t_{i_0},\ q_j<q_{i_0}\}$ - множество работ J, которые начали обслуживаться до начала I_0 и имели остаточное время обслуживания меньше q_{i_0}.

$W=\{w:\ 0<w<t_{i_0},\ S(w)<m\}$ - множество моментов времени, когда есть незагруженные машины, $S(w)$ - количество машин, находящихся в работе в момент w.

Если $J\neq\emptyset$, то $t_{j_0}=\max\limits_{j\in J} t_j$ - последний момент времени, когда начато обслуживание работы с $q_i<q_{i_0}$. Если $W\neq\emptyset$, то $w_0=\sup\limits_{w\in W} w$.

Обозначим $I(\alpha,\beta)=\{I\in I:\ r_i\geq\alpha, q_i\geq\beta\}$, $G(\alpha,\beta)=G(I(\alpha,\beta))$. Так как $I(\alpha,\beta)\subseteq I$, то $G(\alpha,\beta)$ является нижней границей для f^*. Разберем три случая, возникающие после построения расписания алгоритмом EDD.

Случай 1. J пусто и W пусто.

Пусть Ω - множество работ, стартовавших строго до работы I_0, т. е. $\Omega=\{w\in I:\ t_w<t_{i_0}\}$. Так как $J=\emptyset$, то $\Omega\subset I(0,q_{i_0})$. Очевидно, что $I_0\in I(0,q_{i_0})$. Кроме того, из условия $W=\emptyset$ следует, что

$$\sum_{w\in\Omega} P_w \geq S\, t_{i_0}\ . \qquad (5)$$

Используя неравенства (1) и (5), можно записать:

$$G(0,q_{i_0}) \geq \Big(\sum_{w\in\Omega} P_w+P_{i_0}\Big)/S+q_{i_0} \geq t_{i_0}+P_{i_0}/S+q_{i_0}\ . \qquad (6)$$

В итоге получаем

$$f-f^* \leq P_{i_0}/S_{k_0}-P_{i_0}/S \leq (\max_{i\in I} P_i/\min_{l\in M} S_l)(1-\min_{l\in M} S_l/S), \qquad (7)$$

и оценка, приведенная в теореме, справедлива.

Случай 2. W не пусто и либо J пусто, либо $w_0\geq t_{j_0}$.

Пусть Ω - множество работ, стартующих между моментами времени w_0 и t_{i_0}, т. е. $\Omega=\{w\in I:\ w_0<t_w<t_{i_0}\}$. Тогда для любого $w\in\Omega$ имеем $q_w\geq q_{i_0}$ и $\Omega\subseteq I(w_0,q_{i_0})$. Кроме того $I_0\in I(w_0,q_{i_0})\setminus\Omega$. Отсюда получаем:

$$G(w_0,q_{i_0}) \geq w_0+\sum_{w\in\Omega} P_w/S+P_{i_0}/S+q_{i_0}\ . \qquad (8)$$

Рассмотрим множество $H=\{h\in I:\ t_h<w_0\leq t_h+P_h\}$ работ, начавшихся до момента w_0 и закончившихся после него. Количество работ множества H не больше, чем $m-1$. Работы из объединения $H\cup\Omega$ выполняются в промежутке $[w_0,t_{i_0}]$ всеми машинами без простоев и, следовательно,

$$\left(\sum_{h\in H} P_h + \sum_{w\in\Omega} P_w\right)/S \ge t_{i_0} - W_0 \Rightarrow \tag{9}$$

$$\sum_{w\in\Omega} P_w/S \ge t_{i_0} - W_0 - \sum_{h\in H} P_h/S. \tag{10}$$

Из (8) и (10) находим:

$$f^* \ge G(W_0, q_{i_0}) \ge t_{i_0} + P_{i_0}/S + q_{i_0} - \sum_{h\in H} P_h/S. \tag{11}$$

Окончательно из(4) и (11) получаем:

$$f - f^* \le P_{i_0}/S_{k_0} - P_{i_0}/S + \sum_{h\in H} P_h/S \le$$

$$\le \max_{i\in I} P_i \cdot (1/\min_{l\in M} S_l - 1/S) + (m-1)\cdot \max_{i\in I} P_i/S \le$$

$$\le (\max_{i\in I} P_i/\min_{l\in M} S_l)\cdot(1+(m-2)\min_{l\in M} S_l/S), \tag{12}$$

и оценка ,приведенная в теореме, справедлива.

Случай 3. W пусто или $W_0 \le t_{j_0}$. Пусть $\Omega = \{w\in I:\ t_{j_0} < t_w < t_{i_0}\}$ - множество работ, стартующих между моментами времени t_{j_0} и t_{i_0}. Если $w\in\Omega$, то $q_w \ge q_{i_0}$ и, следовательно, $\Omega \subseteq I(t_{j_0}, q_{i_0})$. По-прежнему, $i_0 \in I(t_{j_0}, q_{i_0}) \setminus \Omega$. Имеем

$$G(t_{j_0}, q_{i_0}) \ge t_{j_0} + \sum_{w\in\Omega} P_w/S + P_{i_0}/S + q_{i_0}\ . \tag{13}$$

Рассмотрим множество работ, начавшихся до момента t_{j_0} и закончившихся после него: $H = \{h\in I:\ t_h < t_{j_0} \le t_h + P_h\}$. Работ из множества H не больше m. Работы из множества $H \cup \Omega$ выполняются всеми машинами без простоев в промежутке $[t_{j_0}, t_{i_0}]$ и, следовательно,

$$\left(\sum_{h\in H} P_h + \sum_{w\in\Omega} P_w\right)/S \ge t_{i_0} - t_{j_0} \Rightarrow \tag{14}$$

$$\sum_{w\in\Omega} P_w/S \ge t_{i_0} - t_{j_0} - \sum_{h\in H} P_h/S. \tag{15}$$

Из (13) и (15) находим:

$$f^* \ge G(t_{j_0}, q_{i_0}) \ge t_{i_0} + P_{i_0}/S + q_{i_0} - \sum_{h\in H} P_h/S\ . \tag{16}$$

Для разности $f - f^*$ получаем оценку:

$$f - f^* \le P_{i_0}/S_{k_0} - P_{i_0}/S + \sum_{h\in H} P_h/S \le$$

$$\le \max_{i\in I} P_i \cdot (1/\min_{l\in M} S_l - 1/S) + m\cdot \max_{i\in I} P_i/S \le$$

$$\le (\max_{i\in I} P_i/\min_{l\in M} S_l)(1+(m-1)\min_{l\in M} S_l/S). \tag{17}$$

Теорема доказана.∎

Замечание. Пусть $\min_{l\in M} S_l = 1$. Так как $(m-1)\min_{l\in M} S_l/S \le 1$, то $f - f^* \le 2\cdot \max_{i\in I} P_i$, что соответствует оценке Карлье для идентичных машин [4].

Т е о р е м а 2. Оценка ,приведенная в теореме 1, наилучшая из возможных.

Д о к а з а т е л ь с т в о. Рассмотрим пример, в котором имеется M машин. Первая со скоростью K, назовем ее быстрой, остальные - медленные - со скоростями равными 1. Требуется обслужить $N=KZ+M+1$ работ, где $Z=(K-1)(K+M-1)L$, а L - любое натуральное число. Первые M работ, далее именуемых M-работами, имеют параметры:

$$r_1=r_2=\ldots=r_M=0,\quad p_1=p_2=\ldots=p_M=K-\varepsilon,\quad q_1=q_2=\ldots=q_M=0.$$

Для $(M+1)$-й работы, которую будем называть критической, исходные данные имеют вид:

$$r_{M+1}=\varepsilon,\quad p_{M+1}=K,\quad q_{M+1}=q-\varepsilon,$$

а для последних KZ работ, которые назовем Z-работами, исходные данные таковы:

$$r_{M+2}=r_{M+3}=\ldots=r_{KZ+M+1}=\varepsilon,\quad p_{M+2}=p_{M+3}=\ldots p_{KZ+M+1}=(K-1)/Z,$$

$$q_{M+2}=q_{M+3}=\ldots q_{KZ+M+1}=q,\quad K-1/KZ>\varepsilon>0,\quad q>K>1.$$

Алгоритм EDD получит следующее расписание выполнения работ (рис. 1). Первыми обслуживаются M-работы. Все Z-работы выполняются на быстрой машине после завершения ею обслуживания одной из M-работ. Критическая работа будет выполнена на одной из медленных машин после завершения ими обслуживания M-работ. В результате $f=2K+q-2\varepsilon$.

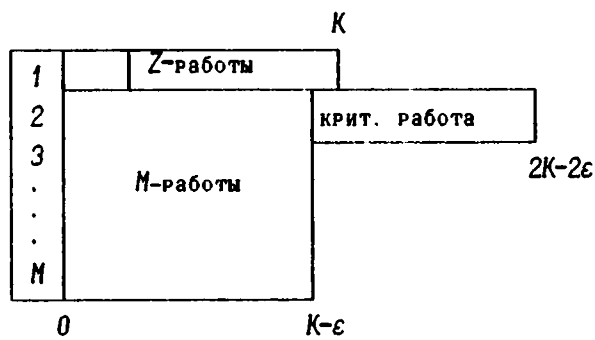

Рис. 1

Оптимальное решение задачи изображено на рис. 2. Критическая работа с момента времени ε выполняется на быстрой машине, Z-работы расписываются по машинам следующим образом: на медленных машинах выполняются с момента ε подряд $K^2L(M-1)$ работ, а на быстрой машине $KZ-K^2L(M-1)$ работ после выполнения критической. После завершения обслуживания Z-работ выполняются M-работы.

Критическая работа завершится в момент времени $1+\varepsilon$. Все Z-работы закончатся в момент $K^2/(K+M-1)+\varepsilon$, а M-работы будут выполнены в момент $K^2/(K+m-1)+K$. В результате получаем $f^*= K^2/(K+M-1)+q+\varepsilon$.

Для разности $f-f^*$ имеем

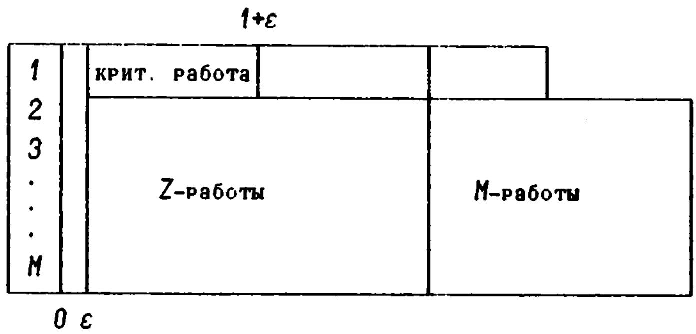

Рис. 2

$$f-f^*=2K-K^2/(K+M-1)-3\varepsilon=K(K+2M-2)/(K+M-1)-3\varepsilon \ .$$

Согласно теореме 1, эта разность не превышает величины

$$K(1+(M-1)/(K+M-1))=K(K+2M-2)/(K+M-1) \ .$$

Так как ε может быть сколь угодно мало, получаем нужную оценку. ■

Поступила в ред.-изд. отдел
6 ноября 1991 г.

Литература

1. Jackson J.R. Scheduling a production line to minimize maximum tardiness // Research 43 California, Los Angeles. - 1955. - P. 10-20.

2. Lenstra J.K., Rinnooy Kan A.H.G., Brucker P. Complexity of machine scheduling problems // Ann. Discrete. Math. - 1977.- V. 1. - P.343-362.

3. Lawler E.L., Lenstra J.K., Rinnooy Kan A.H.G., Shmoys D.B. Sequencing and scheduling: Algoritms complexity // Report BS-R 8909, Department of Operations Research, Statistics and System Theory. Amsterdam - 1989. - 70 p.

4. Carlier J. Scheduling jobs with release dates and tails on identical machines to minimize the makespan // Eur. J. Oper. Res. - 1987. - V. 29. - P. 298-306.

5. Horn W. Some simple scheduling algoritms // Naval. Res. Log. Q. - 1974 - V. 21. - P.177-185.