

ОБ ЭФФЕКТИВНОЙ РАБОТЕ ЭЦВМ, РЕШАЮЩЕЙ ЗАДАЧИ ПЛАНИРОВАНИЯ И ОПЕРАТИВНОГО УПРАВЛЕНИЯ

А. М. Бейлин, А. Н. Ефимов

Харьков

При эксплуатации ЭЦВМ часто приходится решать ряд задач оптимального обслуживания с точки зрения экономических критериев, заявок, поступающих на машину.

Одной из таких задач является следующая.

Из n источников на ЭЦВМ поступают простейшие потоки заявок с параметрами β_i ($i=1, 2, \dots, n$). Тогда суммарный поток также будет простейшим с параметром $\beta = \beta_1 + \beta_2 + \dots + \beta_n$. Пусть время обработки заявок всех типов подчиняется показательному закону $F(t) = 1 - e^{-\mu t}$, и при этом имеет место $\frac{\beta}{\mu} < 1$.

Кроме обработки случайных заявок (режим оперативного обслуживания), ЭЦВМ должна решать также задачи, требующие длительного времени счета, значительно превышающего μ^{-1} . Например, задачи

математического программирования большой размерности, которые для краткости будем называть «основными». В этом случае необходимо обеспечить такой порядок счета, при котором потери (за счет ожидания заявок, изменений режима работы ЭЦВМ и структуры «основной» программы) были бы минимальными.

В большинстве современных ЭЦВМ, имеющих блоки переключения программ (БПП), переход от одной программы к другой может осуществляться после любой из выполненных команд. Однако, в случае счета программ, практически полностью занимающих оперативное запоминающее устройство, это явно недопустимо, ввиду грозящей потери некоторой части информации. Поэтому наиболее естественной представляется следующая стратегия обслуживания (в классе данных стратегий и ищется оптимальная). «Основная» программа при составлении разбивается на блоки так, чтобы после каждого из них с минимальными потерями уже посчитанных результатов можно было бы перейти в оперативный режим. При накоплении в буферной памяти ЭЦВМ m заявок досчитывается текущий блок и начинается режим оперативного обслуживания. В этом режиме ЭЦВМ работает до полного отсутствия заявок в памяти, т. е. обслуживаются и те заявки, которые поступили в течение оперативного режима, затем основная программа восстанавливается и считается дальше. В течение этого времени вновь накапливаются заявки и т. д.

Так как при разбивке «основной» программы на блоки удобно иметь определенную свободу в выборе длительности счета очередного блока, то программист из многих последовательностей псевдослучайных чисел $\xi_1, \xi_2, \dots$, подчиняющихся показательному закону с параметром α , выбирает наиболее ему удобную и составляет «основную» программу так, что длительность счета i -го блока равна ξ_i . В этом случае средние затраты на обслуживание в единицу времени можно определить по формуле

$$\rho(m, \alpha) = \frac{\alpha(M[v_m] + 1) + M \left[\sum_{j=1}^i \tau_j b_j \right] + d}{M[\tau_{v_m+1} + T]}, \quad (1)$$

где α — средние затраты на компоновку одного блока за счет увеличения используемого объема памяти и неэкономичного использования машинного времени при частой разбивке «основной» программы на блоки;

b_j — штраф за единицу времени ожидания в памяти заявки j -го типа ($j=1, 2, \dots, n$);

d — средние затраты на одно восстановление «основной» программы.

Восстановление — это проверка отсутствия искажений, которые могли появиться в основной программе в течение оперативного периода, ввод уже посчитанных промежуточных результатов и констант, которые могли быть затерты в оперативном режиме из-за ограниченности оперативной памяти и т. д.

v_{m+1} — число разрешающих моментов перехода в оперативный режим до того, как он начнется;

τ_i — время ожидания начала обслуживания заявкой, пришедшей i -ой по счету;

τ_{v_m+1} — время, в течение которого накапливаются заявки;

T — время работы ЭЦВМ в оперативном режиме;

I — количество заявок, обслуженных в очередном оперативном режиме.

Необходимо определить такие m и α , при которых функция затрат $p(m, \alpha)$ принимает наименьшее значение. Следовательно, решаются как бы две взаимосвязанные задачи:

1) какое количество заявок нужно накапливать в памяти, чтобы целесообразно было переходить в оперативный режим;

2) какая должна быть структура «основной» программы для наиболее экономичного обнаружения момента перехода в оперативный режим?

Обозначим $n_m = M[v_m]$. Математическое ожидание числа заявок из простейшего потока с параметром α , пришедших за время t , равно αt . Если τ_j — интервал времени между приходом $(j-1)$ -ой и j -ой заявки-

ми, то величина $\sum_{j=1}^m \tau_j$ подчиняется эрланговскому закону распределения с функцией плотности распределения

$$u(t) = \frac{\beta (\beta t)^{m-1}}{(m-1)!} e^{-\beta t}.$$

Используя понятие рандомизации [1], получим

$$n_m = \int_0^{\infty} \alpha t u(t) dt = m \frac{\alpha}{\beta}. \quad (2)$$

Определим

$$M[\eta_{m+1}]$$

$$\eta_{m+1} = \sum_{r=1}^{m+1} \xi_r, \quad \text{где } \xi_r \text{ — интервал времени}$$

между двумя соседними разрешающими моментами перехода в оперативный режим.

Используя тождество Вальда, получим

$$M[\eta_{m+1}] = (1 + n_m) M[\xi] = \frac{1}{\alpha} + \frac{m}{\beta}.$$

Вычислим $M[T] = M\left[\sum_{i=1}^T \mu_i\right]$, где μ_i — время обслуживания i -ой заявки.

$$M\left[\sum_{i=1}^T \mu_i\right] = \sum_{l_0=m}^{\infty} p_{l_0} M\left[\sum_{i=1}^{l_0} \mu_i\right] = \frac{1}{\mu} \sum_{l_0=m}^{\infty} l_0 p_{l_0}. \quad (3)$$

Здесь p_{l_0} — вероятность того, что в оперативном режиме будет обслужено ровно l_0 заявок.

По формуле полной вероятности

$$p_{l_0} = \sum_{t=0}^{l_0-m} p_{m+t} \psi_{m+t, l_0}. \quad (4)$$

где p'_{m+1} — вероятность того, что оперативный период начнется с наличия в памяти $m+i$ заявок;
 ψ_{m+i, l_0} — вероятность того, что ЭВМ, подключенная к очереди длины $m+i$, обслужит в оперативном режиме ровно l_0 заявок.

Определим p'_{m+i} . Величина δ_i (время от момента прихода m -ой заявки до начала оперативного режима) подчинена показательному закону — $P(\delta_i < t) = 1 - e^{-\alpha t}$. Это следует из леммы [2].

Лемма. Пусть случайная величина ξ имеет экспоненциальное распределение, т. е.

$$P(\xi < t) = \begin{cases} 1 - e^{-\alpha t} & t > 0 \\ 0 & t < 0; \quad \alpha > 0. \end{cases}$$

Тогда для любого числа $\tau \geq 0$

$$P(\xi - \tau < t, \xi > \tau) = P(\xi < t)$$

p'_{m+i} можно интерпретировать как вероятность прихода i заявок за время δ_i . Еще раз используя понятие рандомизированного времени [1], получим

$$p'_{m+1} = \int_0^{\infty} \frac{(\beta t)^i}{i!} e^{-\beta t} \cdot \alpha e^{-\alpha t} dt = \frac{\alpha}{\alpha + \beta} \left(\frac{\beta}{\alpha + \beta} \right)^i. \quad (5)$$

В [3] указывается, что

$$\psi_{m, n} = \frac{m}{n} \gamma_{n-m}^{(n)},$$

где $\gamma_{n-m}^{(n)}$ — вероятность того, что ровно $n-m$ требований из простейшего потока поступят в течение n независимых длительностей обслуживания. В нашем случае

$$\begin{aligned} \psi_{m+i, l_0} &= \frac{m+i}{l_0} \gamma_{l_0-m-i}^{(l_0)} = \frac{m+i}{l_0} \int_0^{\infty} \frac{(\beta t)^{l_0-m-i}}{(l_0-m-i)!} \cdot e^{-\beta t} \frac{\mu (\mu t)^{l_0-1}}{(l_0-1)!} e^{-\mu t} dt = \\ &= \frac{m+i}{l_0} \left(\frac{\beta}{\beta + \mu} \right)^{l_0-m-i} \left(\frac{\mu}{\beta + \mu} \right)^{l_0} \binom{2l_0-m-i-1}{l_0-1}. \end{aligned} \quad (6)$$

Перейдем к определению последнего члена в формуле (1):

$$M \left[\sum_{j=1}^l \gamma_j b_j \right] = M[b_j] \sum_{l_0=m}^{\infty} p_{l_0} \sum_{i=1}^{l_0} M[\gamma_i]. \quad (7)$$

Это имеет место, поскольку затраты на задержку заявки j -го типа в единицу времени не зависят от длительности ожидания в очереди, а обслуживание заявок производится в порядке поступления.

Обозначим через π_j вероятность того, что прибывшая заявка — j -го типа. Тогда [4] для случая $n=2$

$$\pi_j = \frac{\beta_j}{\beta_1 + \beta_2} \quad (j=1, 2).$$

Пользуясь методикой [4], легко можно показать, что и в общем случае

$$\pi_j = \frac{\beta_j}{\beta_1 + \beta_2 + \dots + \beta_n}$$

и, следовательно,

$$M[b_j] = \frac{\beta_1 b_1 + \beta_2 b_2 + \dots + \beta_n b_n}{\beta_1 + \beta_2 + \dots + \beta_n}. \quad (8)$$

Для определения $M[\gamma_i]$ введем функцию $\delta(A)$

$$\delta(A) = \begin{cases} 1, & \text{если } A \geq 0 \\ 0, & \text{если } A < 0 \end{cases}$$

Тогда

$$\begin{aligned} p\{\delta(A) = 1\} &= p\{A \geq 0\}, \quad p\{\delta(A) = 0\} = p\{A < 0\} \\ M[\delta(A)] &= p\{A \geq 0\}. \end{aligned} \quad (9)$$

Тогда можно записать

$$\gamma_i = \delta(A_i) \cdot A_i,$$

где

$$A_i = \gamma_{i,m+1} + \sum_{j=0}^{i-1} \mu_j - \sum_{j=1}^i \tau_j (\mu_0 = 0, i = 1, 2, \dots). \quad (10)$$

Будем различать два случая:

1. $i \leq m$, тогда $\gamma_i = A_i$ ($\delta(A_i) = 1$ для $i \leq m$) и

$$\begin{aligned} A_i &= \left(\gamma_{i,m+1} - \sum_{j=1}^m \tau_j \right) + \sum_{j=0}^{i-1} \mu_j + \sum_{j=i+1}^m \tau_j = \delta_i + \sum_{j=1}^{i-1} \mu_j + \sum_{j=i+1}^m \tau_j. \\ M[\gamma_i] &= M[A_i] = \alpha^{-1} + (i-1)\mu^{-1} + (m-i)\beta^{-1} \end{aligned} \quad (11)$$

2. $i > m$, в этом случае (10) перепишем в виде

$$A_i = \left(\gamma_{i,m+1} - \sum_{j=1}^m \tau_j \right) + \sum_{j=0}^{i-1} \mu_j - \sum_{j=m+1}^i \tau_j = \delta_i + \sum_{j=0}^{i-1} \mu_j - \sum_{j=m+1}^i \tau_j.$$

Величины δ_i , $\sum_{j=0}^{i-1} \mu_j$ и $\sum_{j=m+1}^i \tau_j$ взаимно независимы, и их функции плотности распределения имеют вид соответственно:

$$u(t) = \alpha e^{-\alpha t}; \quad v_i(t) = \frac{\mu(\mu t)^{i-2}}{(i-2)!} e^{-\mu t}; \quad w(t) = \frac{\beta(\beta t)^{i-m-1}}{(i-m-1)!} e^{-\beta t}$$

для случая $i > m$

$$M[\gamma_i] = M[\delta(A_i)] \cdot M_{A_i > 0}[A_i]$$

$M_{A_i > 0}[A_i]$ — условное математическое ожидание при условии, что $A_i > 0$. Легко можно показать, что

$$M_{A_i > 0}[A_i] = \frac{\int_0^\infty t f_i(t) dt}{P(A_i) > 0}, \quad (12)$$

где $f_i(t)$ — функция плотности распределения случайной величины A_i для $t > 0$. Используя (9) и (12)

$$M[\gamma_i] = M[\delta(A_i)] \cdot M_{A_i > 0}[A_i] = \int_0^{\infty} t f_i(t) dt \quad (i > m),$$

но $f_i(t)$ можно получить, применяя последовательно операции нахождения функции плотности распределения для суммы независимых случайных величин, а затем для разности, т. е.

$$f_i(t) = \int_0^{\infty} b_i(t+x) w_i(x) dx, \quad \text{где } b_i(t) = \int_0^t U(t-x) v_i(x) dx.$$

После преобразований получим:

$$M[\gamma_i] = \int_0^{\infty} t f_i(t) dt = \frac{1}{\alpha} \left(\frac{\beta}{\alpha + \beta} \right)^{i-m} \left[\left(\frac{\mu}{\mu - \alpha} \right)^{i-1} - \left(\frac{\alpha}{\mu} \right)^2 \sum_{k=2}^i \left(\frac{\mu}{\mu - \alpha} \right)^{k-1} \times \right. \\ \left. \sum_{r=0}^{i-k} \binom{r+i-m-1}{r} \left(\frac{\mu}{\mu + \beta} \right)^r (i-k+r+1) \right] \quad (i > m). \quad (13)$$

Перепишем (1) в следующем виде

$$\rho(m, \alpha) = \frac{\alpha(n_m + 1) + M[b_i] \sum_{l_0=m}^{\infty} p_{l_0} \sum_{i=1}^{l_0} M[\gamma_i] + d}{(1 + n_m) \alpha^{-1} + \mu^{-1} \sum_{l_0=m}^{\infty} l_0 p_{l_0}}. \quad (14)$$

Все компоненты (14) найдены выше в явном виде. Действительно, n_m вычисляется по (2), $M[b_i]$ — по (8), p_{l_0} — по (5), (6) и (4), $M[\gamma_i]$ — по (11) или (13). Остается определить значения α и m , при которых функция затрат $\rho(m, \alpha)$ принимает наименьшее значение.

Нахождение минимума осложняется тем, что одна из переменных дискретная. Хорошо разработанных и удобных алгоритмов для отыскания экстремумов таких функций не существует [5]. Один из приемов, заключающийся в том, что дискретную переменную полагают непрерывной, а затем округляют до ближайшего целого, в нашем случае принципиально не подходит. Поэтому наиболее целесообразным является применение алгоритма «конфигураций» [5] с последующим усовершенствованием Вуда [6].

Пример. $n = 3$; $\beta_1 = 0,5$; $\beta_2 = 0,4$; $\beta_3 = 0,1$; $\frac{b_1}{a} = 40$; $\frac{b_2}{a} = 50$; $\frac{b_3}{a} = 100$;

$$\frac{d}{a} = 300; \quad \mu = 100.$$

В этом случае $\min \rho_1(m, \alpha) = \min \frac{\rho(m, \alpha)}{a}$ достигается при $m = 15$, $\alpha = 19,6$.

ЛИТЕРАТУРА

1. В. Феллер. Введение в теорию вероятностей и ее приложения, т. 2, «Мир», 1967.
2. Г. П. Климов. Стохастические системы обслуживания. «Наука», 1966.
3. Д. Кокс, У. Смит. Теория очередей. «Мир», 1966.
4. Д. Кокс, У. Смит. Теория восстановления. «Советское радио», 1967.
5. Д. Дж. Уайлд. Методы поиска экстремума. «Наука», 1967.
6. C. F. Wood. Recent developments in «Direct Search» techniques, Westinghouse Research Report, 62—159—522, R, (July, 1962).

ВЕРОЯТНОСТНЫЕ ХАРАКТЕРИСТИКИ РЕАГЕНТНОГО РЕЖИМА ПРИ ФЛОТАЦИИ УГЛЕЙ

В. И. Салыга, Н. Р. Хаджигов

Харьков

Реагентный режим является одним из наиболее эффективных каналов регулирования промышленного флотопроцесса. Реагенты вводятся в пульпу для ускорения процесса, повышения эффективности и из-

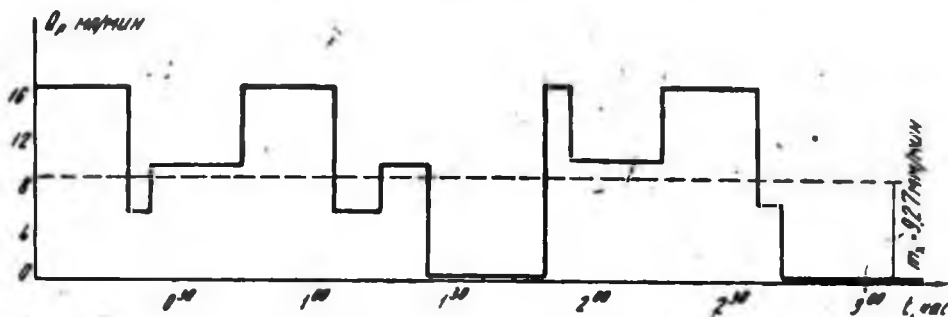


Рис. 1.

бирательности флотации. На результативный признак при этом влияют состав, сочетание, температура и порядок подачи реагентов.

Большинство из этих факторов для конкретной обогащательной фабрики являются стабильными, обусловленными технологическим процессом. Для каждой углеобогащательной фабрики характерны колебания расхода реагентов, поскольку они определяются изменением таких параметров процесса, как количество флотируемой пульпы, ее плотность, аэрированность и гранулометрический состав угля.

Большое количество случайных возмущающих воздействий приводит к тому, что расход реагентов также является случайной функцией времени.

Определение вероятностных характеристик реагентного режима позволит найти вероятное значение расхода реагентов в единицу времени, а также в сочетании с результативными признаками динамические характеристики по каналам «реагенты — зольность концентрата» и «реагенты — зольность хвостов».

В настоящей работе использовались данные реагентного режима по Горловской ЦОФ (рис. 1). В качестве собирателя на фабрике применяется синтетический спирт, а вспениватель — керосин. Подача реагентов производится в агрегат приготовления пульпы. Изменение количества реагента рассматривается как функция, принимающая слу-