УДК 004.2.004.8

В.Г. Зайцев, М.В. Плахотний, Є.І. Цибаєв

Національний технічний університет України «Київський політехнічний інститут», Київ

АЛГОРИТМ ДОСЛІДЖЕННЯ ХАРАКТЕРИСТИК БАГАТОПРІОРИТЕТНИХ МОДЕЛЕЙ ПРОГРАМ

Запропоновано метод і алгоритм дослідження часу виконання програм у багатопріоритетних обчислювальних системах. Метод базується на використанні класичних моделей СМО «загибелі та розмноження» шляхом багатократної зміни параметрів з урахуванням залежності окремих моделей по каналу «загибелі».

Ключові слова: модель виконання програми, час виконання програми, марковські моделі, пріоритетне обслуговування

Предложен метод и алгоритм исследования времени выполнения программ в приоритетных вычислительных системах. Метод базируется на использовании классических моделей СМО «гибели и размножения» путем многократной смены параметров с учетом зависимости отдельных моделей по каналу «гибели».

Ключевые слова: модель выполнения программы, время выполнения программы, марковские модели, приоритетное обслуживание

Proposed the method and algorithm of the research runtime programs in priority computing systems. The method is based on the use of classical models SMO «destruction and reproduction» by repeated change of the parameters of the dependence on certain models via «death».

Key words: model of program execution, the execution of the program, Markov models, priority servic

Вступ

Ефективність програмного забезпечення (ПЗ) спеціалізованих комп'ютерних систем (СКС) залежить від багатьох факторів: ефективності системи диспетчеризації; організації вирішення основних функціональних задач у реальному часі; визначення оптимальної системи пріоритетів при виконанні додатків та переривань. Виконання цих завдань тісно пов'язане з побудовою і дослідженням моделей масового обслуговування СМО, де під заявками на обслуговування розуміють завдання, а під часом обслуговування – час їх виконання. Відповідно можна вважати, послідовності завдань – потоками заявок, а інтенсивності виконання завдань – інтенсивностями обслуговування. Відповідно можна говорити про середні інтенсивності потоку завдань і закони їх розподілу, а також про середні показники часу виконання завдань і відповідні закони розподілу. Однією із специфічних особливостей СКС є те, що у більшості випадків перелік функцій СКС є незмінними за весь час їх експлуатації. Це дозволяє вважати, що перелік завдань відомий наперед та відомі характеристики потоків завдань, що вирішуються, і середній час їх виконання. Ще однією особливістю є те, що всі вимоги на виконання завдань мають задовольнятися. Отже, тут розглядаються тільки скінченні моделі.

Постановка завдання

Будемо вважати, що виконання функцій СКС забезпечуються певною кількістю завдань, виконання яких підпорядковано певним пріоритетам. Характеристики потоків заявок і час їх виконання попередньо відомі. Слід визначити середній час виконання програми, включаючи перебування у черзі, залежно від наданого пріоритету. Для цього необхідно розробити відповідні моделі СМО і запропонувати алгоритми їх дослідження.

Огляд існуючих рішень

Проблемі безпріоритетного і пріоритетного обслуговування для систем СМО із скінченними джерелами заявок присвячено дуже багато досліджень. Серед них слід визначити фундаментальні роботи [1-3]. При пріоритетному обслуговуванні розглядаються системи СМО з абсолютними, відносними та змішаними пріоритетами.

У роботі [2] зроблено дуже суттєвий висновок: при дисциплінах з абсолютним пріоритетом наявність вимог з більш низьким пріоритетом ніяк не впливає на процес обслуговування вимог з більш високим пріоритетом.

Дослідження пріоритетних СМО будується на визначенні циклу обслуговування [1] j-го класу для процесів з k класами, під яким розуміють проміжок часу між надходженням вимоги на обслуговування та моментом, коли обслуговуючий пристрій звільняється для того, щоб прийняти на обслуговування чергову вимогу того ж класу. Слід зазначити, що отримані результати дослідження у аналітичному вигляді, як стверджують самі автори, дуже складні у використанні і мають скоріше науковий, ніж практичний інтерес. Тому доцільно спробувати отримати вирішення проблеми не у вигляді аналітичних формул, а у вигляді алгоритму.

Математична модель СМО

Нехай у системі СМО задіяно n задач. Кожна i-та задача у довільний момент може виставити запит на обслуговування. Інтенсивність такого потоку вимог тоді дорівнює λ_i. Якщо обслуговуючий пристрій (процесор) вільний, він надається задачі. На її обслуговування (розв'язання) витрачається середній час $v_i = \dfrac{1}{\mu_i}$. Якщо на момент виникнення вимоги процесор зайнятий, задача стає у чергу на виконання, на що витрачається додатковий час. Загальний потік від задач можна розглядати, як суперпозицію або об'єднання вимог. Будемо вважати, що ці джерела заявок взаємонезалежні. Стаціонарна поведінка такого потоку може визначатися об'єднаною функцією розподілу величини інтервалу часу до моменту виникнення чергового запиту. Визначимо цю функцію через $F(t)$, а через $F_i\ (t)$ – відповідну функцію для i-го джерела із n взаємонезалежних джерел. Тоді,

$$1 - F(t) = \prod_{i=1}^{n} [1 - F_i(t)] dt \qquad (1)$$

Позначимо

$$H_i(t) = 1 - F_i(t); \;\; H(t) = \prod_{i=1}^{n} H_i(t). \qquad (2)$$

Функції $A'(t)$, $H(t)$ *та* $H_i(t)$ є доповнюючими функціями розподілу.

$F(t) = H(t) = 0$ *при* $t < 0$. Якщо усі функції $H_i(t)$ неперервні, то, зважаючи на (2) можна записати, що

$$\frac{H'(\mathrm{t})}{H(\mathrm{t})} = \frac{H_1'(\mathrm{t})}{H_1(\mathrm{t})} + \ \dots\ + \frac{H'_n(\mathrm{t})}{H_n(\mathrm{t})}. \qquad (3)$$

Як показано у [4], якщо проміжки часу між вимогами мають довільну, одну і ту ж функцію розподілу з єдиним обмеженням, що вона має неперервну похідну $A'(t)$, то

$$F(t) = \lambda \int_0^t 1 - A(n)\,\mathrm{dn}\,, \qquad (4)$$

де λ - математичне сподівання кількості вимог, що надходять у стаціонарному стані. Ця величина вказує середню інтенсивність вхідного рекурентного потоку.

З урахуванням (3) і (4) для $t = 0$, отримуємо

$$\lambda = \lambda_1 + \lambda_2 + \dots \lambda_n\,. \qquad (5)$$

Якщо припустити, що усі джерела мають пуасонівський розподіл, тобто

$$H_i(t) = e^{-\lambda_i t}, \quad t > 0, \qquad (6)$$

то з урахуванням (2),

$$H(t) = 1 - F(t) = \ e^{-\lambda t}, \qquad \lambda = \sum_{i=1}^{n} \lambda_i \qquad (7)$$

Виходячи з результату (7) і те, що всі джерела є скінченними, модель СМО процесу обслуговування заявок при $\lambda = \lambda_1 = \lambda_2 = \dots = \lambda_n$ може бути визначена як класична модель СМО «загибелі та розмноження» [3], граф випадкового процесу в якій показано на рис.1.

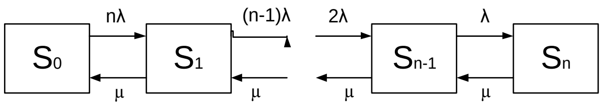

Рис. 1. Граф моделі СМО загибелі та розмноження, де S_i стани випадкового процесу СМО

Ця модель СМО добре вивчена [1; 3]. За умови, що потоки заявок мають пуасонівський розподіл, а час обслуговування – показниковий розподіл (марковська модель) усі стаціонарні характеристики моделі можуть бути визначені через граничні імовірності станів P_i, як [3]:

$$\left.\begin{aligned}
&P_0 = \ (1 + n\rho + n(n-1)\rho^2 + \dots \ n(n-1)\dots\ 1\rho^n)^{-1},\\
&P_1 = n\rho P_0,\\
&P_2 = \ n(n-1)\rho^2 P_0,\\
&\dots\dots\dots\\
&P_n = n(n-1)\dots 1\rho^n P_0, \qquad \textit{де}\ \rho = \frac{\lambda}{\mu}.
\end{aligned}\right\} \qquad (8)$$

Оскільки припущення, про показниковий характер закону обслуговування досить умовне, виникає питання про залежність величини P_0 від виду закону розподілу. Таке дослідження зробимо шляхом порівняння результатів для марковської моделі і немарковської – із постійним часом обслуговування. Дослідження немарковської моделі виконаємо методом вкладених марковських ланцюгів. Побудуємо вкладений марковський ланцюг і розглянемо його у моменти регенерації $t,\ t+\delta,\ t+2\delta\ ,\dots,t+i\delta$,

де t може бути будь-яким, зокрема $t=\ 0,\ a\ \ \delta = t\ об. = \dfrac{1}{\mu} = const.$ Якщо у деякий момент часу T у системі перебуває j вимог на обслуговування, то до моменту $T+\delta$, одна вимога буде обслужена і кількість вимог у системі стане $j-1+m_{n-j}$,

де m_{n-j} – число надходжень від $n-j$ джерел за інтервал $\{T,\ T+\delta\}$.

У результаті отримуємо матрицю (таблиця) для марковських переходів, що описує стан системи в моменти часу, розділені інтервалами тривалістю δ . Позначимо k_j^i – імовірність надходження j вимог від i джерел за час δ, що для $\delta =$const, визначається так:

$$k_j^i = C_i^j\left(1-e^{-\rho}\right)^j e^{-(i-j)\rho},\ \text{де}\ \rho=\frac{\lambda}{\mu} \qquad (9)$$

Таблиця

Матриця марковських переходів

	0	1	2	…	n-1	n
0	k_0^n	k_1^n	k_2^n	…	k_{n-1}^n	k_n^n
1	k_0^{n-1}	k_1^{n-1}	k_2^{n-1}	…	k_{n-1}^{n-1}	0
2	0	k_0^{n-2}	k_1^{n-2}	…	k_{n-2}^{n-2}	0
3	0	0	k_0^{n-3}	…	k_{n-3}^{n-3}	0
…	0	0	0	…		
n-1	0	0	0	…	k_1^1	0
n	0	0	0	…	k_0^0	0

Відповідно до наведеної матриці, запишемо рівняння для визначення граничних імовірностей станів:

$$P_j = \mathrm{k}_j^n P_0 + \sum_{i=1}^{j+1} k_{j-i+1}^{n-i} P_i\ ;\quad j=\overline{0,\ n-1}\ ;\ \sum_{i=0}^{n} P_i = 1\ ; \qquad (10)$$

Рівняння (10) можна розв’язати безпосередньо або за допомогою рекурентної формули:

$$P_{j+1} = (\mathrm{k}_0^{n-j-1})^{-1}\left\{P_j - k_j^n P_0 - \sum_{i=1}^{j} k_{j-i+1}^{n-1} P_i\right\}. \qquad (11)$$

Імовірність P_0 у формулі (11) може бути знайдена із таких міркувань. Нехай задана деяка величина:

$$Z_0 = qP_0\ ;\qquad q \neq 0. \qquad (12)$$

Тоді, позначивши $Z_i = qP_i$ та використовуючи співвідношення (11), (12), обчислимо:

$$Z_{j+1} = qP_{j+1};\qquad j=\overline{0,\ \mathrm{n}\text{-}1}. \qquad (13)$$

Для визначення q скористаємося співвідношенням:

$$\sum_{i=0}^{n} Z_i = \sum_{i=0}^{n} qP_i = q\sum_{i=0}^{n} P_i = q, \qquad (14)$$

звідки:

$$P_i = Z_i\left(\sum_{i=0}^{n} Z_i\right)^{-1}. \qquad (15)$$

Визначимо середній час обслуговування запитів T , включаючи перебування у черзі:

$$T = \frac{1}{\mu}P_1 + \frac{2}{\mu}P_2 + \ \dots\ + \frac{n}{\mu}P_n; \qquad (16)$$

з іншого боку, середня кількість запитів, що потребують обслуговування, дорівнює:

$$\omega =\ 1P_1 +\ 2P_2 +\ \dots\ + nP_n \qquad (17)$$

або $$T = \frac{1}{\mu}\omega.$$

Як відомо, у [3] отримано інший вираз для ω:

$$\omega = n - \frac{1-P_0}{\rho} \qquad (18)$$

Тобто,

$$T = \frac{1}{\mu}\left(n - \frac{1-P_0}{\rho}\right). \qquad (19)$$

Експериментальне порівняння розрахунку величини P_0 для марковської моделі СМО і немарковської, при $\mu =$const, показало, що для n у межах $2 \div 100$ та ρ від 0,01 до 0,50 максимальна розбіжність у результатах не перевищує 5%, що підтверджено також в [1].

Таким чином, за основу приймемо марковську модель «загибелі та розмноження». У загальному випадку ПЗ СКС буде відповідати декілька потоків заявок на обслуговування, кожен з яких може мати свій пріоритет. Перенумеруємо такі потоки і будемо вважати, що рівень пріоритету буде відповідати його номеру. Нехай всього буде m рівнів пріоритету. Тоді можна вважати, що загальна модель буде складатися з m моделей СМО «загибелі та розмноження», між якими існують певні залежності. Для з'ясування характеру таких залежностей розглянемо випадкову функцію $x(t)$, що має m складових [5]:

$$x(t) = f\,(x_1(t), x_2(t), \dots, x_m(t)).$$

Тобто, для кожного моменту часу t, можна розглядати стан випадкової вектор-функції у m-мірному просторі і відповідно граф випадкового процесу-$G\{x(t)\}$. Такий граф буде мати m вершин. Орієнтовані ребра графу будуть визначати зв'язок, що існує між складовими $x(t)$. Усі імовірнісні характеристики $x_i(t)$ визначаються трьома групами параметрів: числом можливих станів n_i+1; середніми інтенсивностями потоків розмноження λ_i; середніми інтенсивностями потоків загибелі μ_i. Кожний із процесів $x_i(t)$ будемо вважати марковським. Розглянемо дві випадкові складові: $x_i(t)$ та $x_j(t)$. Вони можуть бути незалежними або мати певні залежності. У випадку незалежності, граф процессу $x(t)$ не буде мати ребер. Залежність між процесами може бути чотирьох типів і відповідно мати або ні, орієнтовані ребра.

Позначимо такі ребра:

a) $R\left[x_i(t);\ x_j(t)\right]=0$, якщо залежності немає;

b) $R\left[x_i(t);\ x_j(t)\right]=1;\ x_j(t)$ залежить від $x_i(t)$;

c) $R\left[x_j(t);\ x_i(t)\right]=1;\ x_i(t)$ залежить від $x_j(t)$;

d) $\left.\begin{array}{l} R\left[x_i(t);\ x_j(t)\right]=1 \\ R[x_j(t);x_i(t)]=1 \end{array}\right\}$ взаємозалежність між $x_i(t)$ та $x_j(t)$.

На рис. 2 показано можливий вигляд графу $G\{x(t)\}$.

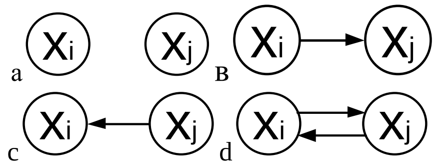

Рис. 2. Види залежності графу G

Граф, який показано на рис. 2, може бути розмічений. Біля ребер можуть бути записані параметри, за якими виникає залежність між процесами. Щодо складових $x_i(t)$ – моделей «загибелі та розмноження» у нашому випадку, залежність може бути по двох параметрах – загибелі та розмноження (по параметрах λ_i та μ_i). Будемо у подальшому вважати, що в нашому випадку взята абсолютна система пріоритетів. Тоді можна зробити такі припущення: кожний з процесів $x_i(t)$ є марковським; середній час обслуговування заявок i-го пріоритету не залежить від наявності заявок, пріоритет яких $j>i$; фактична середня інтенсивність обслуговування заявок *i*-го пріоритету залежить від наявності заявок, пріоритет яких $j<i$; потоки заявок на обслуговування для всіх рівнів пріоритету незалежні між собою. З урахуванням цих припущень, граф $G\{x(t)\}$ буде мати вигляд, як на рис. 3.

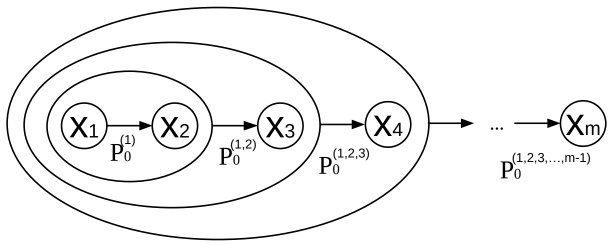

Рис. 3. Загальний вигляд графу G $\{x(t)\}$

Граф $G\{x(t)\}$ ілюструє той факт, що інтенсивність обслуговування заявок пріоритету *i* залежить від обслуговування сумісної черги заявок, пріоритет яких $j<i$. Тобто, найбільш пріоритетні заявки 1-го рівня не залежать від заявок, пріоритет яких $j>1$, а заявки 2-го рівня – тільки від заявок 1-го рівня. Розглянемо, як можна врахувати цю залежність. Заявки 2-го рівня пріоритету можуть обиратися на обслуговування тільки у тому випадку, коли заявки 1-го рівня відсутні. Їх відсутність дорівнює $P_0^{(1)}$, де $P_0^{(1)}$ гранична імовірність відсутності заявок у моделі «загибелі та розмноження» з параметрами $n_1,\ \lambda_1,\ \mu_1$.

Позначимо фактичну середню інтенсивність обслуговування заявок 2-го рівня пріоритету, що знаходиться у черзі, через μ_2^*. Вона буде дорівнювати μ_2 з імовірністю $P_0^{(1)}$ і нулю з імовірністю $1-P_0^{(1)}$. Тоді, μ_2^* буде дорівнювати:

$$\mu_2^* = 0(1-P_0^{(1)}) + P_0^{(1)}\mu_2 = P_0^{(1)}\mu_2. \qquad (20)$$

Тобто, значення $P_0^{(2)}$ можна обчислити за наведеними формулами (8), використовуючи параметри n_2, λ_2, μ_2^*. Заявки 3-го рівня пріоритету можуть обиратися на обслуговування тільки у тому випадку, коли відсутні заявки 1-го і 2-го рівнів пріоритетів. Позначимо відповідну імовірність через $P_0^{(1,2)}$. Її можна обчислити, якщо розглянути сумісну модель СМО для заявок 1-го і 2-го рівнів пріоритету. Таку об'єднану модель також можна представити, як на рис. 1, при відповідній заміні параметрів $(n,\ \lambda,\ \mu)$:

$$n = n_1 + n_2, \qquad (21)$$
$$\lambda = \frac{n_1\lambda_1 + n_2\lambda_2}{n_1+n_2}; \quad \mu = \frac{1}{\nu}; \quad \nu = \frac{n_1\nu_1 + n_2\nu_2}{n_1+n_2},$$

де ν_1, ν_2 - середній час виконання програм 1-го та 2-го рівнів пріоритету.

Відповідно для *i*–го рівня пріоритету, об'єднані моделі для обчислення

$P_0^{(1,2,3,\dots,i-1)}$ слід використати параметри:

$$n = n_1 + n_2 + \ldots + n_{i-1},$$
$$\lambda = \frac{n_1\lambda_1 + \ldots + n_{i-1}\lambda_{i-1}}{n_1+n_2+\ldots+n_{i-1}},\ \mu = \frac{1}{\nu}, \qquad (22)$$
$$\nu = \frac{n_1\nu_1 + n_2\nu_2 + \ldots + n_{i-1}\nu_{i-1}}{n_1+n_2+\ldots+n_{i-1}},\ \text{де}\ \ i = \overline{1,m}.$$

Якщо позначити через T_i середній час обслуговування заявок (розв'язання задач) *i*-го пріоритету, то їх можна обчислити за допомогою (19) моделі СМО «загибелі та розмноження» (див.рис.1), відповідно заміняючи:

$$T = T_i\ ,\ \ n = n_i\ ,\ \ \lambda = \lambda_i,\ \ \mu = \mu_i P_0^{(1,2,\dots,i-1)},\ \ \rho = \frac{\lambda}{\mu}.$$

Висновки

Запропоновано метод обчислення середнього часу розв'язання задач у багатопріоритетних спеціалізованих комп'ютерних системах. Метод базується на циклічному використанні класичної моделі СМО «загибелі та розмноження» шляхом багатократної заміни параметрів. Експериментальні дослідження алгоритму, розробленого на базі запропонованого методу, показали його високу ефективність у порівнянні з відомими обчисленнями за результатами аналітичних формул [1].

Список літератури

1. *Джейсул Н. Очереди с пріоритетами / Н. Джейсул. – М.: Мир, 1973. – 280 с.*
2. *Вагнер Г. Основы исследования операцій / Г. Вагнер. – [том 3]. – М.: Мир, 1973.– 504 с.*
3. *Вентцель Е.С. Исследование операцій / Е.С. Вентцель. – М.: Мир, 1982. – 386 с.*
4. *Риордан Дж. Вероятностные системы обслуживания / Дж. Риордан. – М.: Связь, 1966. – 184 с.*
5. *Тараканов К.В., Аналитические методы исследования систем / К.В. Тараканов, Л.А. Овчаров, А.М. Тырышкин. – М.: Советское радио, 1974 – 240 с.*

Стаття надійшла до редколегії 25.04.2013

Рецензент: д-р техн. наук, проф. С.В. Цюцюра, Київський національний університет будівництва і архітектури, Київ