Э. А. ДАНИЕЛЯН, П. Т. ХОСТИКЯН

КАТЕГОРИЙНЫЕ ВО ВРЕМЕНИ ПОЛУЧЕРЕДУЮЩИЕСЯ ПРИОРИТЕТЫ

1. Введение

При прохождении программ разных типов на однопроцессорной ЭВМ часто требуется организовать такую дисциплину счета, для которой времена реакции на программы удовлетворяют заранее заданным соотношениям. При фиксированных законах поступлений и счета программ единственный путь построения математических моделей, в которых разработчик на этапе проектирования ЭВМ имеет возможность выбора, например, отношений средних времен реакции на программы разных типов, заключается в наделении программам приоритетов, зависящих от времени. Такие модели называются моделями с динамическими приоритетами [1].

В настоящей работе предлагается и изучается одна такая динамическая модель—модель с категорийными во времени получередующимися приоритетами.

В одноканальную систему массового обслуживания с ожиданием поступают независимые пуассоновские потоки 1-вызовов, . . ., r-вызовов с параметрами $a_1>0$, . . ., $a_r>0$ соответственно. Длительности обслуживания всех вызовов независимы в совокупности, не зависят от процесса поступления и для i-вызовов ($i=\overline{1, r}$) имеют функцию распределения (ф. р.) $B_i(t)$, $B_i(+0)=0$.

Объединенный поток i-вызовов, . . ., j-вызовов ($i\leq j$) называем потоком $\overline{i, j}$-вызовов. Чем меньше индекс вызова, тем выше его приоритет.

В момент $t=0$ система свободна от вызовов, а прибор незаперт.

Опишем дисциплину категорийных во времени получередующихся приоритетов.

Временная ось разделена на непересекающиеся промежутки фиксированной длины T: $[0, T]$, $[T, 2T]$, . . ., называемые квантами. Вызовы, поступающие на некотором кванте, не обладают преимуществом перед вызовами, поступающими в систему на предыдущих квантах. Среди вызовов разных потоков, поступающих на одном и том же кванте, принята дисциплина получередующихся приоритетов.

Разъясним суть дисциплины получередующихся приоритетов, для чего индуктивным способом зададим понятие j-промежутка $(j=\overline{1, r})$.

а) Длительность обслуживания j-вызова, поступающего в свободную систему, образует j-промежуток.

б) Внутри периода занятости системы после завершения j-промежутка при наличии в очереди n_i $(i=\overline{1, r})$ i-вызовов, в случае $n_1= \ldots =n_{k-1}=0$, $n_k>0$ $(k=\overline{1, r})$ из очереди на обслуживание выбирается k-вызов и начинается k-промежуток. Данный k-промежуток состоит из n_k независимых длительностей обслуживания k-вызовов.

Вызовы одного и того же приоритетного класса, поступающие на одном и том же кванте, обслуживаются в соответствии с дисциплинами „первым пришел — первым обслужен" (FIFO), либо „последним пришел — первым обслужен" (LIFO). В дальнейшем, упоминание дисциплин FIFO и LIFO относится к вызовам одного и того же приоритетного класса, поступающих за один и тот же квант.

Категорийные во времени относительные и абсолютные приоритеты изучены в работе [2]. В настоящей работе используется метод анализа, предложенный в [2]. Поэтому доказательства полученных ниже результатов даются схематично.

Отметим, что рассматриваемая динамическая модель при $T\to 0$ переходит в модель бесприоритетную, а при $T\to\infty$ — в модель с дисциплиной получередующихся приоритетов.

Укажем на возможности применения данной динамической дисциплины при проектировании однопроцессорной ЭВМ с практически неограниченной памятью.

Пусть программы подразделяются на два типа. Обозначим через w_k $(k=1, 2)$ стационарные средние времен реакции на программы k-го типа. Очевидно, что $w_k=w_k(T)$. Требуется, чтобы отношение $w_1(T)/w_2(T)$ равнялось заданному числу c (задача 1). Если задача 1 имеет смысл, т. е. пользуясь внутренними ресурсами системы, именно, варьируя T, можно достичь искомого равенства, то данная динамическая дисциплина себя оправдывает.

Естественным образом возникает следующая оптимизационная задача 2, позволяющая с точки зрения заданного отношения выяснить возможности этой динамической дисциплины.

Пусть величины $w_k(T)$ $(k=1, 2)$ вычислены. Интуитивно понятно, что эти величины служат решениями некоторой системы линейных уравнений. К данной системе линейных уравнений добавляется критерий оптимизации

а) $\to\min(w_1(T)/w_2(T))$, либо б) $\to\max(w_1(T)/w_2(T))$.

Получилась простейшая задача линейного программирования.

Если T' и T'' суть значения T, при которых оптимизируются критерии а) и б), то мы находим границы a и b приведенных отношений.

Теперь, если $c\in[a, b]$, то задача 1 имеет смысл.

§ 2. Постановка задачи и основные результаты

Обозначим через $\overline{w}_k(t)$ $(k=\overline{1,\ r})$ условное виртуальное время ожидания k-вызова в момент t при дисциплине FIFO при условии прекращения с момента t доступа вызовов в систему, а через $w_k^1(t)$ $(w_k^2(t))$ — виртуальное время ожидания k-вызова в момент t при дисциплине FIFO (LIFO).

При $t=nT+\tau$ $(0\leqslant\tau<T)$ удобно произвести переобозначение

$$\overline{w}_k(t)=\overline{w}_{kn}(\tau),\quad w_k^l(t)=w_{kn}^l(\tau)$$

и обозначить

$$\overline{\omega}_{kn}(s,\ \tau)=Me^{-s\overline{w}_k(t)},\ \omega_{kn}^l(s,\ \tau)=Me^{-sw_k^l}(t),$$

где M—знак математического ожидания. Задача заключается в вычислении введенных функций, которые при $t=nT+\tau$ зависят от величины времени, отстающегося от момента t до момента завершения текущего кванта, т. е. зависят от величины $T'=T-\tau$.

Действительно, если в момент $t=nT+\tau$ прибор занят, то при дисциплине FIFO некоторые $\overline{1,\ k-1}$-вызовы, а при дисциплине LIFO, возможно, еше и некоторые k-вызовы, поступающие в систему за промежуток времени $[nT+\tau,\ (n+1)T]$, в силу принятой дисциплины, могут обслужиться раньше виртуального в момент t k-вызова. Следовательно, длительности обслуживания таковых вызовов включаются в виртуальное время ожидания в момент t k-вызова и функции $\omega_{kn}^l(s,\ \tau)$ зависят от T'. Исключение составляет функция $\omega_{1n}^1(s,\ \tau)$, которая не зависит от T'.

При фиксированных n, T, и τ $(t=nT+\tau)$ будут вычислены функции

$$\omega_{kn}^l(s,\ \mu,\ \tau)=\int_0^\infty e^{-\mu T'}\ \omega_{kn}^l(s,\ \tau)dT',$$

обращения которых по μ дают в точке $T'=T-\tau$ искомые функции $\omega_{kn}^l(s,\ \tau)$.

Пусть $\omega(s,\ t)$ есть преобразование Лапласа—Стильтеса от ф. р. виртуального времени ожидания в системе $M|G|1|\infty$ в момент t, в которой интенсивность поступления есть σ_r, а ф. р. длительности обслуживания вызовов есть $B_{(r)}(t)$. Здесь

$$\sigma_k=a_1+\ .\ .\ .\ +a_k,\quad \sigma_k B_{(k)}(t)=\sum_{l=1}^k a_l B_l(t).$$

Нетрудно видеть, что при $t=nT$ в рассматриваемой динамической модели

$$\overline{\omega}_k(s,\ nT)=\overline{\omega}_{kn}(s,\ 0)\equiv\omega(s,\ nT),$$

чем и объясняется интерес к функции $\omega(s, t)$.

Ниже устанавливается связь между функциями $\omega_{k+1-in}(s, \tau)$ $(i=1, 2)$ и $\omega^i_{kn}(s, \tau)$. Для $t=nT+\tau$ при дисциплине получередующегося приоритета с начальной задержкой прибора $w(nT)$ условное виртуальное время ожидания k-вызова в момент τ в случае дисциплины FIFO при условии прекращения доступа вызовов в систему с момента τ совпадает с $w_{kn}(\tau)$ для рассматриваемой динамической модели.

Итак, вкратце обрисован путь анализа функций $\omega^i_{kn}(s, \tau)$.

Введем обозначения

$$\sigma_k = \sum_{i=1}^{k} a_i(1-\beta_i(s)), \quad \beta_k(s) = \int_0^{\infty} e^{-st} dB_k(t),$$

$$\xi_k = s - \sigma_k, \; r_k = \mu + \sigma_k, \; v_k = s + a_k - a_k\beta_k(s)$$

$$\mu_{k+1}(s) = s + \sigma_k - \sigma_k \pi_k(s),$$

где $\pi_k(s)$ есть преобразование Лапласа-Стильтеса от ф. р. промежутка времени, начинающегося с поступления в свободную систему $\overline{1, k}$-вызова и завершающегося освобождением системы от $\overline{1, k}$-вызовов. Функция $\pi_k(s)$ найдена Г. П. Климовым в (3).

Лемма. Имеет место следующее равенство $(k=\overline{1, r};\ n \geqslant 0;\ \zeta \geqslant 0,\ s \geqslant 0;\ 0 \leqslant \tau < T)$:

$$\overline{\omega}^*_{kn} = \int_0^{\infty} e^{-\zeta\tau} \overline{\omega}_{kn}(s, \tau) d\tau = (\zeta - \xi_k)^{-1} \Big\{ \omega(s, nT) -$$

$$- s \frac{\omega(\mu_{r+1}(\zeta), nT)}{\mu_{r+1}(\zeta)} - \sum_{j=k+1}^{r} [\Phi_j(\zeta; 1) - \Phi_j(\zeta; \beta_j(s))] \Big\}, \qquad (1)$$

где функция $\Phi_j(\zeta; z)$ для дисциплины получередующегося приоритета определяются в (4).

Теорема. Для функций $\omega^i_{kn}(s, \tau)$ $(k=\overline{1, r};\ i=1, 2)$ при $t=nT+\tau$ $(0 \leqslant \tau < T)$ имеют место формулы $(s>0;\ \mu>0)$:

$$\omega^i_{kn}(s, \mu, \tau) = \frac{1}{r_{k-2+i}} (\overline{\omega}_{k+1-in}(s, \tau) -$$

$$- \overline{\omega}_{k+1-in}(\mu_k(\mu + g_{ki}), \tau)) + \frac{1}{\mu} \overline{\omega}_{k+1-in}(\mu + \mu_k(g_{ki}), \tau), \qquad (2)$$

где

$$g_{ki} = \begin{cases} s, & \text{при } i=1 \\ v_k, & \text{при } i=2. \end{cases}$$

3. Приложение

Ниже приводятся наброски доказательств основных результатов.

Доказательство леммы. Для дисциплины категорийных во времени получередующихся приоритетов устанавливается справедливость равенства

$$\omega(s, nT)e^{-\tau_k\tau} = e^{-s\tau}\overline{\omega}_{kn}(s, \tau) + s\int_0^{\tau} e^{-\tau_k(\tau-u)} \cdot e^{-su} \cdot P(u)\, du +$$

$$+ \sum_{j=k+1}^{r} \sum_{i\geq 1} (1-\beta_j^i(s)) \int_0^{\tau} e^{-\tau_k(\tau-u)} e^{-su} d_u P_j(i, u), \tag{3}$$

где в случае соответствующей данной модели дисциплины получередующегося приоритета: $P(u)du$—вероятность свободного и незапертого состояния прибора в промежутке $[u, u+du)$, $dP_j(i, u)$ ($j=\overline{1, r}$; $i \geq 1$)—вероятность того, что в промежутке $[u, u+du)$ начался j-промежуток, инициированный i-штук j-вызовами.

В случае вышеупомянутой дисциплины получередующихся приоритетов в начальный момент прибор случайное время $w(nT)$ заперт для обслуживания.

Формулы (3), с учетом связи $P_j(i, u)$, $P(u)$ с функциями $\Phi_j(\zeta, u)$ (см. [4]), приводятся к виду (1).

Доказательство теоремы. Доказательство формул (2) при $i=1$ слово в слово повторяет аналогичное доказательство в (2) в модели с категорийными во времени относительными и абсолютными приоритетами.

Перейдем к подсчету $\omega_{kn}^2(s, \tau)$. Пусть $\pi_k(y)$—промежуток времени, начинающийся с задержки $y>0$, в начале которого отсутствуют вызовы, и кончающийся первым после задержки y моментом освобождения системы от $\overline{1, k}$-вызовов. Возможны следующие три случая

1. $\omega_{k-1,n}(\tau) \geq T'$. В этом случае

$$\int_0^{\infty} e^{-\mu T'} M\{e^{-sw_{kn}^2(\tau)};\ \overline{\omega}_{k-1n}(\tau) \geq T'\} dT' =$$

$$= \int_0^{\infty} e^{-(\mu+\tau_k)T'} \int_{T'}^{\infty} e^{-sy} d_y P\{\overline{w}_{k-1n}(\tau) < y\} = \tag{4}$$

$$= \frac{1}{r_k}(\overline{\omega}_{k-1n}(s, \tau) - \overline{\omega}_{k-1n}(s+r_k, \tau)).$$

2. $\pi_{k-1}(\overline{w}_{k-1n}(\tau)) < T'$. Тогда

$$\int_0^{\infty} e^{-\mu T'} M\{e^{-sw_{kn}^2(\tau)};\ \pi_{k-1}(\overline{w}_{k-1n}(\tau)) < T'\} d(T' =$$

$$= \int_0^{\infty} e^{-\mu T'} dT' \int_{y=0}^{T'} d_y P\{\overline{w}_{k-1n}(\tau) < y\} \int_{z=y}^{T'} e^{-v_k z} d_z P\{\pi_{k-1}(y) < z\} =$$

$$= \mu^{-1}\omega_{k-1n}(\mu+\mu_k)(v_k), \tau). \tag{5}$$

3. Последний случай описывается неравенствами

$$w_{k-1n}(\tau) < T', \quad \pi_k w_{k-1n}(\tau)) \geqslant T'.$$

Нетрудно видеть, что в данном случае

$$\int_0^\infty e^{-\mu T'} M\{e^{-s w^2_{kn}(\tau)}; \overline{w}_{k-1n} T', \pi_{k-1}(\overline{w}_{k-1n}(\tau)) \geqslant T'\} dT' =$$

$$= \frac{1}{r_k}(\beta_{(k-1)}(s) - \beta_{(k-1)}(s+r_k)) \int_0^\infty \varphi_y^{(k+1)}(\mu+v_k; \beta_{(k-1)}(s))\, dP\{\overline{w}_{k-1n}(\tau) < y\} =$$

$$= \frac{1}{r_k}(\overline{\omega}_{k-1n}(s+r_k, \tau) - \overline{\omega}_{k-1n}(\mu_k(\mu+v_k), \tau)), \tag{6}$$

где функция $\varphi_y^{(k-1)}(\mu, z)$ введена и вычислена в (2), а $\beta_{(k-1)}(s)$ есть преобразование Лапласа — Стильтеса от $B_{(k-1)}(t)$.

Объединяя формулы (4)—(6), получаем (2) при $i=2$.

Է. Ա. ԴԱՆԻԵԼՅԱՆ, Գ. Տ. ԽՈՍՏԻԿՅԱՆ

ԿԱՏԵԳՈՐԻԿ ԸՍՏ ԺԱՄԱՆԱԿԻ ԿԻՍԱՀԵՐԹԱԳԱՅՈՂ ՆԱԽԱՊԱՏՎՈՒԹՅՈՒՆՆԵՐ

Հաճախ մեկ պրոցեսորով հաշվողական մեքենայում տարբեր տիպի ծրագրերի հաշվումը պահանջվում է այնպես կազմակերպել, որ ծրագրերի սպասման ժամանակները բավարարեն նախապես տրված առնչություններին։ Եթե պահանջների մուտքի և սպասարկման օրենքները ֆիքսված են, ապա այնպիսի մաթեմատիկական մոդելի կառուցման միակ ճանապարհը, որի սահմաններում հաշվողական մեքենայի նախագծողը ունի, օրինակ, միջին սպասման ժամանակների հարաբերություն, ընտրության հնարավորություն, կայանում է ծրագրերին ժամանակից կախված նախապատվություն տալու մեջ. այդ տիպի մոդելները կոչվում են դինամիկ նախապատվությամբ մոդելներ։

Աշխատանքում հետազոտվում է մի դինամիկ նախապատվությամբ մաթեմատիկական մոդել, որը կառուցված է կիսահերթագայող նախապատվությամբ դիսցիպլինայի հիման վրա։

ЛИТЕРАТУРА

1. *Л. Клейнрок.* Вычислительные системы с очередями. М., 1979.
2. *Э. А. Даниелян.* Время ожидания в модели с категорийными во времени приоритетами, «Кибернетика», Киев, 1980, № 6.
3. *Г. П. Климов.* Стохастические системы обслуживания, М., 1966.
4. *Э. А. Даниелян.* О классе дисциплин относительных динамических приоритетов в системе $\vec{M}_r|\vec{G}_r|1|\infty$. «Известия АН Уз. ССР», серия физ.-мат. наук, 1980, № 6.