ISSN 0002-306X. Изв. НАН РА и ГИУА. Сер. ТН. 2004. Т. LVII, № 1.

УДК 517.518

АВТОМАТИЗАЦИЯ И СИСТЕМЫ УПРАВЛЕНИЯ

Т.А. ГРИГОРЯН

ОБ ОЦЕНКЕ ОБЪЕМА РАБОТЫ ИНТЕРНЕТ-СЕРВИСА ПРИ ЗАПРОСАХ К БАЗЕ ДАННЫХ

Объем работы модуля интернет-сервиса при запросах к базе данных оценивается в рамках модели GI | G | 1 | ∞ очереди. Для применения известных результатов модели найдена скорость сходимости накопленной работы к ее стационарной величине.

Ключевые слова: интернет-сервис, модель GI | G | 1 | ∞, скорость сходимости.

1. Интернет-сервис обслуживания запросов к многомерной базе данных – важная проблема в сфере управления и интерпретации многомерных данных [1-3].

Нами разработан и реализован модуль регистрации и обслуживания запросов к многомерной базе данных через интернет по заказу компании “Synergy International Systems” (8100 Boone Boulevard, Suite 230, Vienna, VA, 22182, USA).

Модуль получает запрос на языке XML [4] и помещает его в очередь. Осуществляется пакетная обработка запросов по дисциплине FIFO (first in-first out) одним процессором “Intel Pentium IV”, который формирует ответ с помощью языка MDX [3].

Основной элемент реализации заключается в оценке объема накопленной и невыполненной работы в моменты поступления запросов при обработке каждого пакета дисциплиной разделения времени процессора.

Поскольку накопленная и невыполненная работа инвариантна в классе консервативных дисциплин [5], то ее можно рассчитывать в рамках модели $GI \mid G \mid 1 \mid \infty$. Анализ стационарной работы осуществлен в теории очередей [6].

Однако для использования известных формул модели $GI \mid G \mid 1 \mid \infty$ необходимо определить скорость сходимости к стационарной работе, что и является целью настоящей работы.

2. Объект исследования. В одноканальную систему обслуживания с ожиданием в случайные моменты времени $\{t_n\}$, где $0 \le t_1 \le t_2 \le \ldots$, поступают одиночные вызовы. Вызовы пронумерованы в порядке поступления числами 1,2,…. Время обслуживания n-го вызова обозначим υ_n, и пусть $u_n = t_n - t_{n-1}$, $t_0 = 0$. В момент $t = 0$ модель свободна от вызовов.

Последовательности $\{u_n\}$ и $\{\upsilon_n\}$ неотрицательных случайных величин (СВ) определены на одном и том же вероятностном пространстве $(\Omega, \Im, P)$.

Последовательности $\{u_n\}$ и $\{\upsilon_n\}$ независимы друг от друга и образуют последовательности независимых одинаково распределенных (НОР) СВ с функциями распределения (ФР) A(t) и B(t) на [0, +() соответственно.

Такую модель обозначают G I□G□1□∞ (см., напр. [6]).

Предполагаем A(+0)=0 и B(+0)=0.

Первое условие влечет

$$P(0 < t_1 < t_2 < \ldots) = 1,$$

где P- знак вероятности.

Второе условие означает, что вероятность “мгновенного” обслуживания равна нулю.

Далее предполагаем

$$0 < \alpha_1 = Mu_1 < +\infty\ ,\quad 0 < \beta_1 = M\upsilon_1 < +\infty\ ,$$

где M - знак математического ожидания.

С помощью α_1 и β_1 определяется загрузка модели $\rho_1 = \beta_1 / \alpha_1$.

Важнейшей характеристикой модели для приложений является накопленная и невыполненная работа (суммарное накопленное и невыполненное время обслуживания) $w(t)$ в момент времени t.

Эта характеристика инвариантна в классе консервативных дисциплин обслуживания. Напомним, что дисциплина обслуживания консервативна, если при ее использовании внутри модели работа не создается и не исчезает, а лишь привносится в модель извне поступлением вызовов (см. [5]).

Известно, что при $\rho_1 < 1$ и $t \to +\infty$

$$w(t) \Rightarrow w^*, \tag{1}$$

где $\Rightarrow$ - знак слабой сходимости и $W^*(x) = P(w^* < x)$, $x \geq 0$ - собственная ФР, т.е $W^*(+\infty) = 1$. Вид ФР W* известен в литературе (см., напр. [6, 7]).

Величина w^* называется стационарной работой. Для того чтобы использовать формулы, определяющие ФР W* или ее моменты, необходимо иметь оценки скорости сходимости в (1).

Настоящая работа посвящена этой проблематике.

3. Основной результат. Пусть

$$Mu_1^{\ 2} < +\infty\ ,\quad M\upsilon_1^{\ 4} < +\infty\ . \tag{2}$$

Из (2) следует конечность моментов $\alpha_k = Mu_1^{\ k}$ при $k \leq 2$ и $\beta_m = M\upsilon_1^{\ m}$ при $m \leq 4$.

Обозначим

$$A = \frac{m_4 + 12m_2^{\ 2}}{\alpha_1^{\ 2}(1-\rho_1)^4}\ , \tag{3}$$

где моменты

$$m_k = M(\upsilon_1 - \beta_1)^k\ ,\ \text{k=1-4},$$

определяются с помощью моментов β_k :

$$m_1 = m_3 = 0, \quad m_2 = \beta_2 - \beta_1^2, \quad m_4 = \beta_4 - 4\beta_3\beta_1 + 6\beta_2\beta_1^2 - 3\beta_1^4.$$

Обозначим через $w_y(t)$, $t \ge 0, \ \ y \ge 0$ накопленную и невыполненную работу в момент времени *t* при условии $w_y(0) = y$.

Для введенной выше $w(t)$ имеем $w(t) = w_y(t)$ ($_{y=0}$ при всех $t \ge 0$.

Известно, что и для $w_y(t)$ с любым неслучайным y>0 справедливо соотношение (1): при $\rho_1 < 1$ и $t \to +\infty$

$$w_y(t) \Rightarrow w^* \quad \text{(см. [6]).}$$

<u>Теорема.</u> Пусть $\rho_1 < 1$ и выполнены условия (2). Тогда при $x \ge \dfrac{\rho_1\alpha_2}{\alpha_1}$ и t >0 имеем

$$0 \le P(w(t) < x) - P(w^* < x) \le A/t, \tag{4}$$

а при y >0, $x - y \ge \dfrac{\rho_1\alpha_2}{\alpha_1}$ и t >0 :

$$-A\left(\frac{x - y - \dfrac{\rho_1\alpha_2}{\alpha_1}}{1-\rho_1} + t\right)^{-2} \le P(w_y(t) < x) - P(w(t) < x) \le 0. \tag{5}$$

Здесь константа A>0 может быть выбрана из (3).

С накопленной и невыполненной к моменту t работой $w(t)$ тесно связана другая характеристика модели G I□G□1□∞ - суммарный <u>простой</u> прибора за промежуток времени [0,t], обозначаемый через I(t).

При $\rho_1 < 1$ и $t \to +\infty$ имеем

$$I(t) \Rightarrow I^*,$$

где $I^*(x) = P(I^* < x), \ \ x \ge 0$ – собственная ФР (см. [8]).

Заменим условия (2) на

$$Mu_1^4 < +\infty, \quad M\upsilon_1^2 < +\infty. \tag{6}$$

Из (6) следует конечность моментов α_k при $k \le 4$ и β_m при $m \le 2$.

Обозначим

$$B = \frac{n_4 + 12n_2^2}{\beta_1^2(\rho_1 - 1)^4}\rho_1^4, \tag{7}$$

где моменты

$$n_k = M(u_1 - \alpha_1)^k, \ \text{k=1-4},$$

определяются с помощью моментов α_k :

$$n_1 = n_3 = 0, \ n_2 = \alpha_2 - \alpha_1^2, \ n_4 = \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4.$$

<u>Следствие.</u> Пусть $\rho_1 > 1$ и выполнены условия (6). Тогда при $x \ge \frac{\beta_2}{\rho_1\beta_1}$ и $t > 0$ имеем

$$0 \le P(I(t) < x) - P(I^* < x) \le \frac{B}{t}, \qquad (8)$$

где константа $B > 0$ может быть выбрана из (7).

4. <u>Доказательство теоремы.</u> Пусть H(t) - функция восстановления процесса $\{t_n\}$ (см. [9]).

При $t \ge 0$ обозначим через N(t) случайное число восстановлений за [0,t] процесса восстановления $\{t_n\}$, т.е. N(t) есть случайное число поступивших за [0,t] в модель вызовов. Тогда H(t)=MN(t), $t \ge 0$.

Далее обозначим через b(t), $t \ge 0$ суммарное время обслуживания поступивших за [0,t] в модель вызовов.

Для S(t)=b(t)-t, $t \ge 0$ справедливо представление

$$S(t) = \upsilon_1 + \upsilon_2 + \ldots + \upsilon_{N(t)} - t, \qquad (9)$$

где N(t) не зависит от последовательности $\{\upsilon_n\}$ НОР СВ.

При t >0 докажем неравенство

$$M(S(t) + t - \beta_1 H(t))^4 \le (m_4 + 12m_2^{\,2})\left(\frac{t}{\alpha_1} + \frac{\alpha_2}{\alpha_1^{\,2}}\right)^2. \qquad (10)$$

Пусть $\{\xi_n\}$- последовательность НОР СВ, где $\xi_n = \upsilon_n - \beta_1$, $n \ge 1$.

Из представления (9) при t >0 следует

$$M(S(t) + t - \beta_1 H(t))^4 = M\left(\xi_1 + \xi_2 + \ldots + \xi_{N(t)}\right)^4,$$

где, очевидно, N(t) не зависит от $\{\xi_n\}$. Следовательно, при t >0 имеем

$$\begin{aligned} M(S(t) + t - \beta_1 H(t))^4 &= \sum_{n\ge1} P(N(t) = n) M(\xi_1 + \xi_2 + \ldots + \xi_n)^4 = \\ &= m_4 \sum_{n\ge1} nP(N(t) = n) + 6m_2^{\,2} \sum_{n\ge1} n(n-1)P(N(t) = n) = \qquad (11) \\ &= (m_4 + 6m_2^{\,2})H(t) + 6m_2^{\,2} M(N(t))^2. \end{aligned}$$

Для H(t), t >0 имеет место неравенство Лордена [10]:

$$\frac{t}{\alpha_1} \le H(t) \le \frac{t}{\alpha_1} + \frac{\alpha_2}{\alpha_1^{\,2}}.$$

При t >0 с помощью H(t) оценивается $M(N(t))^2$:

$$M(N(t))^2 = 2H(t) * H(t) - H(t) \le 2(H(t))^2 - H(t),$$

где * - знак свертки.

Применяя неравенство Лордена к правой части (11), с учетом полученных неравенств при t >0 выводим

$$M(S(t)+t-\beta_1 H(t))^4 \le m_4 H(t) + 12 m_2^{\,2}(H(t))^2 \le$$
$$\le (m_4 + 12 m_2^{\,2})\left(\frac{t}{\alpha_1} + \frac{\alpha_2}{\alpha_1^{\,2}}\right)^2. \qquad (12)$$

При выводе (12) также принято во внимание неравенство моментов

$$\alpha_2 \ge \alpha_1^{\,2}$$

и вытекающее из него неравенство

$$\frac{t}{\alpha_1} + \frac{\alpha_2}{\alpha_1^{\,2}} \le \left(\frac{t}{\alpha_1} + \frac{\alpha_2}{\alpha_1^{\,2}}\right)^2, \ t > 0.$$

Неравенство (10) (см.(12)) доказано.

Обозначим

$$\bar{S}(t) = \sup_{0 \le u \le t} S(u), \ t > 0, \ \bar{S} = \sup_{t \ge 0} S(t)$$

и при $x > 0$ и $t > 0$ запишем равенства

$$0 \le P(w(t) < x) - P(w^* < x) = P(\bar{S} \ge x) - P(\bar{S}(t) \ge x) =$$
$$= P(\bar{S}(t) < x, \bar{S} \ge x). \qquad (13)$$

Равенство (13) является следствием известных формул модели GI G 1 ∞ (см. [7]) :

$$P(w(t) < x) = P(\bar{S}(t) < x), \ \ P(w^* < x) = P(\bar{S} < x), \ \ x \ge 0. \qquad (14)$$

Далее по неравенству Чебышева при $x > 0$, $t > 0$ имеем

$$P(\bar{S}(t) < x, \bar{S} \ge x) = \int_t^\infty P(S(u) \ge x)du \le \int_t^\infty \frac{M(S(u) + u - \beta_1 H(u))^4}{(x + u - \beta_1 H(u))^4} du. \qquad (15)$$

Используя оценку (10) и неравенство Лордена, при $x \ge \frac{\rho_1 \alpha_2}{\alpha_1^{\,2}}$ и $t > 0$ из (15) получаем

$$P(\bar{S}(t) < x, \bar{S} \ge x) \le (m_4 + 12 m_2^{\,2}) \int_t^\infty \frac{\left(\frac{u}{\alpha_1} + \frac{\alpha_2}{\alpha_1^{\,2}}\right)^2 du}{\left(x - \rho_1 \frac{\alpha_2}{\alpha_1} + u(1-\rho_1)\right)^4} = A \int_t^\infty \frac{\left(u + \frac{\alpha_2}{\alpha_1}\right)^2 du}{\left(\frac{x - \frac{\rho_1 \alpha_2}{\alpha_1}}{1-\rho_1} + u\right)^4} =$$

$$=A\int\limits_{t+\frac{\alpha_2}{\alpha_1}}^{\infty}\frac{y^2dy}{\left(\frac{x-\frac{\alpha_2}{\alpha_1}}{1-\rho_1}+y\right)^4}<A\int\limits_{t+\frac{\alpha_2}{\alpha_1}}^{\infty}\frac{dy}{\left(\frac{x-\frac{\alpha_2}{\alpha_1}}{1-\rho_1}+y\right)^2}=A\left(\frac{x-\frac{\rho_1\alpha_2}{\alpha_1}}{1-\rho_1}+t\right)^{-1},$$

откуда, в силу (13), следует оценка (4). При этом в качестве константы A>0 использована константа, определяемая равенством (3).

Теперь при x >0, y >0, t >0 приступим к оценке разности

$$P(w_y(t)<x)-P(w(t)<x).$$

Имеем

$$\begin{aligned}&0\geq P(w_y(t)<x)-P(w(t)<x)=P(w(t)\geq x)-P(w_y(t)\geq x)=\\&=-P(S(t)-\inf_{0\leq u\leq t}S(u)<x,S(t)+y\geq x)\geq -P(x>S(t)\geq x-y)\geq\\&\geq -P(S(t)\geq x-y).\end{aligned}\tag{16}$$

Неравенства (16) являются следствием (14). Поэтому при t >0, x >0, y >0 имеем

$$P(w(t)<x,w_y(t)\geq x)=P(S(t)-\inf_{0\leq u\leq t}S(u)<x,S(t)\geq x-y).$$

Как и выше, при t >0, x >0, y >0 и дополнительном условии $x-y\geq(\rho_1\alpha_2)/\alpha_1$ имеем

$$P(S(t)\geq x-y)\leq\frac{M(S(t)+t-\beta_1H(t))^4}{(x+t-\beta_1H(t))^4}\leq$$

$$\leq A\frac{\left(t+\frac{\alpha_2}{\alpha_1}\right)^2}{\left(\frac{x-y-\frac{\rho_1\alpha_2}{\alpha_1}}{1-\rho_1}+t\right)^4}<A\left(\frac{x-y-\frac{\rho_1\alpha_2}{\alpha_1}}{1-\rho_1}+t\right)^{-2},$$

откуда из (16) следует (5).

Теорема доказана.

5. <u>Доказательство следствия.</u> Если поменять местами последовательности $\{u_n\}$ и $\{\upsilon_n\}$, то получим новую модель GI □G□1□∞, загрузка которой из-за смены местами α_1 и β_1 равна $\beta_1/\alpha_1=1/\rho_1<1$, если $\rho_1>1$.

В силу известных формул модели GI □G□1□∞ (см. [8]) в новой модели GI □G□1□∞ ФР накопленной и невыполненной работы к моменту *t* и ее стационарной

ФР совпадают с ФР $P(I(t) < x)$ и $I^*(x)$ соответственно в первоначальной модели GI □G□1□∞.

Поэтому при замене условий (2) на условия (6) в случае $\rho_1 > 1$ оценка (8) следует из оценки (4).

Константа B в (8), в силу (3)-(4), может быть выбрана как в (7).

Следствие доказано.

СПИСОК ЛИТЕРАТУРЫ

1. **Thomsen E.** OLAP Solutions, John Wiley&Sons, Inc.–2002. - 660 с.
2. **Snell J., Tidwell D., Kulchenko P**. Programming Web Services with SOAP, O'Reilly & Associates, Inc. -2002. -245 с.
3. **Spofford G.** MDX Solutions, John Wiley & Sons, Inc. -2001. -432 с.
4. **Harold E. R.**, Means W. S. XML in a Nutshell, O'Reilly & Associates, Inc. -2002. -634 с.
5. **Клейнрок Л.** Вычислительные системы с очередями.-М.: Мир, 1979. -600 с.
6. **Гнеденко Б. В., Коваленко И. Н.** Введение в теорию массового обслуживания.-М.: Наука, 1987. -336 с.
7. **Боровков А. А.** Вероятностные процессы в теории массового обслуживания.-М.: Наука, 1973. -367 с.
8. **Prahbu N. U.** Stochastic Storage Processes.-Springer-Verlag, New-York-Heidelberg-Berlin, 15. - 1980. -140 с.
9. **Кокс Д. Р., Смит В. Л.** Теория восстановления.-М.: Сов. Радио, 1967. -299 с.
10. **Lorden F.** On excess over the boundary. - Ann. Math. Statistics. -1979. -41, -С. 520-527.

ЕГУ. Материал поступил в редакцию 14.10.2003.

Տ.Ա. ԳՐԻԳՈՐՅԱՆ

ԻՆՏԵՐՆԵՏ ԾԱՌԱՅՈՒԹՅԱՆ ԱՇԽԱՏԱՆՔԻ ԾԱՎԱԼԻ ԳՆԱՀԱՏՈՒՄԸ ՏՎՅԱԼՆԵՐԻ ՀԵՆՔԻՆ ՀԱՐՑՈՒՄՆԵՐԻ ԴԵՊՔՈՒՄ

Ինտերնետ–սերվիս մոդուլի աշխատանքի ծավալը տվյալների հենքին հարցումների դեպքում գնահատվում է GI | G | 1 | ∞ հերթի մոդելի շրջանակներում: Մոդելի հայտնի արդյունքների օգտագործման համար գտնված է կուտակված աշխատանքի՝ իր ստացիոնար արժեքին զուգամիտության արագությունը:

T.A. GRIGORYAN

ON ESTIMATION OF INTERNET-SERVICE WORK-VOLUME IN DATABASE REQUESTS

The work-volume of internet-service module in case of requests to the GI | G | 1 | ∞ model of queue is estimated. To use the known requests of the model, the rate of convergence of accumulated work to its stationary value is found.