

МОДЕЛИРОВАНИЕ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С КОНКУРЕНЦИЕЙ ЗА ПРИБОРЫ

А.Н. Дудин, С.А. Дудин, О.С. Дудина

*Белорусский государственный университет, пр. Независимости, 4, 220030,
г. Минск, Беларусь, dudin@bsu.by, dudins@bsu.by, dudina@bsu.by*

Рассматривается модель массового обслуживания, состоящая из двух систем, конкурирующих за приборы. Запросы двух типов поступают в соответствии с маркированным марковским входным потоком. Построен процесс изменения состояний системы. Выписан инфинитезимальный генератор данного процесса

Ключевые слова: Конкурирующие системы; маркированный марковский входной поток; многомерные цепи Маркова.

MODELING QUEUING SYSTEMS WITH COMPETITION FOR SERVERS

A.N. Dudin, S.A. Dudin, O.S. Dudina

*Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus,
dudins@bsu.by, dudina@bsu.by
Corresponding author: dudin@bsu.by*

We consider a queuing model consisting of two systems competing for servers. Two types of customers arrive according to the marked Markovian arrival process. The process of system states is analyzed. The infinitesimal generator of this process is written out.

Keywords: Competing systems; marked Markovian arrival process; multidimensional Markov chains.

Введение

Теория массового обслуживания – общепризнанный математический аппарат для оптимизации распределения ограниченных ресурсов в различных системах и сетях. Зачастую, приборы, которые не задействованы в данный момент, не простаивают, а могут быть использованы для выполнения других задач с целью получения

дополнительной прибыли. Обычно при исследовании систем с переменным числом приборов в случае необходимости подключения прибора считается, что прибор из неактивного состояния просто переключается в активное состояние. В реальности незадействованный прибор, который считается выключенным, может выполнять другую задачу, и включение его как дополнительного прибора может означать то, что прибор должен будет прекратить выполнять эту другую задачу. То есть, подключение прибора для работы в данной системе означает его выключение в другой параллельно работающей системе. В результате этого оптимальная стратегия подключения приборов может сильно отличаться от стратегии, полученной для классической системы с включением приборов.

В данной работе мы исследуем модель массового обслуживания состоящую из двух систем массового обслуживания, конкурирующих за приборы. Каждая из систем имеет свое минимальное число приборов, которое не может быть уменьшено. Остальные приборы могут работать как в первой, так и во второй системе. Первая система считается приоритетной и может забирать приборы из второй системы.

Литература, посвященная системам обслуживания с конкуренцией за приборы довольно бедна. Похожая модель была исследована в работе [1]. В отличие от нее, где предполагалось, что запросы двух типов поступают независимо друг от друга, в данной работе мы предполагаем, что поступления запросов в обе системы зависимы и описываются маркированным марковским входным потоком (ММАР). Кроме того, мы предполагаем, что запросы находящиеся в буфере, могут проявлять нетерпеливость и покидать систему без обслуживания. Данные обобщения существенно повышают адекватность модели, но и усложняют ее аналитическое исследование.

1. Математическая модель

Мы рассматриваем модель, состоящую из двух взаимодействующих многолинейных систем массового обслуживания, конкурирующих за приборы. Структура модели представлена на рисунке.

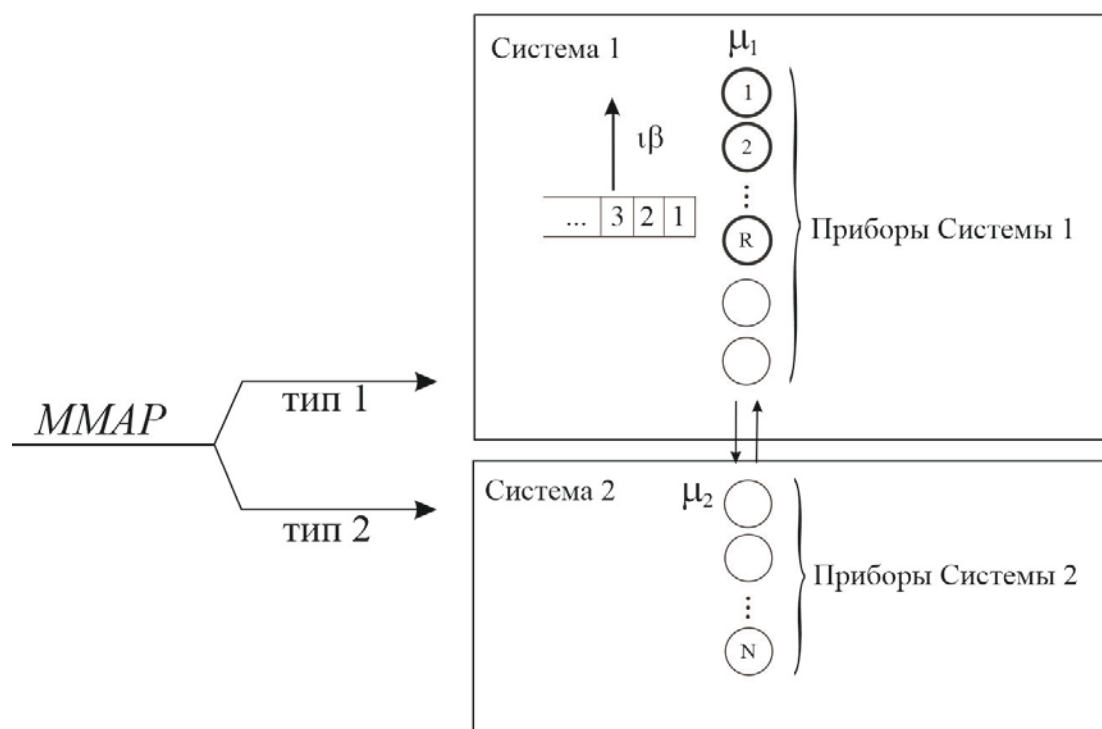


Рисунок – Структура системы

Процесс поступления запросов двух типов описывается ММАР-поток. Данный поток задается управляющим процессом $\nu_t, t \geq 0$, который является неприводимой цепью Маркова с непрерывным временем и конечным пространством состояний $\{1, \dots, W\}$, и матрицами D_0 , $D_1^{(1)}$ и $D_1^{(2)}$. Обозначим среднюю интенсивность поступления запросов k -го типа как λ_k . Считаем, что запросы k -го типа поступают в k -ю систему, $k=1,2$.

Система 1 имеет бесконечный буфер, а Система 2 не имеет буфера. Общее количество приборов в обеих системах равно N . Системы совместно используют существующие приборы следующим образом. Количество приборов, зарезервированных исключительно для использования Системой 1 (для обслуживания запросов первого типа), равно R . Количество приборов, зарезервированных исключительно для использования Системой 2, равно $M, 1 \leq M \leq N - R - 1$. Оставшийся пул из $N - R - M$ приборов может использоваться обеими системами, когда все их зарезервированные приборы заняты. Определенный приоритет в доступе к общему пулу имеет Система 1. А именно, если количество запросов, требующих обслуживания в Системе 1, невелико, то все приборы $N - R$, которые не зарезервированы исключительно для использования Системой 1, доступны для использования Системой 2. Однако при

увеличении количества запросов в Системе 1 эта система может последовательно (по одному) забирать приборы из Системы 2. Это может привести к прекращению обслуживания запросов второго типа, получающих обслуживание на изымаемом приборе.

Правило вывода приборов Системой 1 из общего пула определяется набором порогов $(J_1, J_2, \dots, J_{N-R-M})$, где $R < J_1 < J_2 < \dots < J_{N-R-M}$. Если число запросов в Системе 1 меньше порога J_1 , то только R приборов находятся (обслуживают или простаивают) в Системе 1. Если число заявок в Системе 1 принадлежит интервалу $[J_k, J_{k+1})$, $k = 1, N - R - M - 1$, то $R + k$ приборов предоставляют обслуживание в Системе 1. Если число запросов в Системе 1 превышает порог J_{N-R-M} , то количество приборов, работающих в Системе 1, равно $N - M$. Если в момент, когда необходимо забрать прибор из Системы 2, в ней есть свободные приборы, то один из свободных приборов начнет работу в Системе 1. Если все приборы Системы 2 заняты и общее количество приборов позволяет изъятие прибора из нее, то один из этих приборов прекращает обслуживание и начинает обслуживание запроса типа 1 из буфера. Запрос, обслуживание которого было прекращено, теряется. Когда все приборы из общего пула обслуживают запросов первого типа, дальнейшее изъятие приборов становится невозможным. При уменьшении очереди заявок первого типа соответствующие приборы из общего пула снова становятся доступными для заявок второго типа в зависимости от соотношения текущего количества заявок первого типа в системе и пороговых значений J_k , $k = 1, N - R - M$.

Полагаем, что время обслуживания в Системе r , $r = 1, 2$, распределено экспоненциально с параметром μ_r , $\mu_r > 0$. Запросы первого типа, ожидающие обслуживания, могут проявлять нетерпеливость. Каждый такой запрос может покинуть буфер после экспоненциально распределенного с параметром β , $\beta \geq 0$, время. В ситуации, когда из-за ухода запроса из буфера возникает необходимость передать из Системы 1 в Систему 2 прибор, которые в данный момент осуществляет обслуживание, один из занятых приборов прекращает обслуживание запроса первого типа, переходит во вторую систему, а запрос, обслуживание которого было прервано, становится на первую позицию в буфере, отодвигая стоящие там запросы.

2. Процесс изменения состояний системы

Процесс изменения состояний системы исследуемой системы можно описать регулярной неприводимой цепью Маркова с непрерывным временем

$$\xi_t = \{i_t, r_t, v_t\}, t \geq 0,$$

где, в момент времени t ,

i_t – число запросов в Системе 1, $i_t \geq 0$;

r_t – число занятых приборов в Системе 2, $r_t = \overline{0, N-R}$, если $i_t < J_1$;
 $r_t = \overline{0, N-R-k}$, если $J_k \leq i_t < J_{k+1}$ и $r_t = \overline{0, M}$, если $i_t \geq J_{N-R-M}$;

v_t – состояние управляющего процесса ММАР, $v_t = \overline{1, W}$.

Пронумеруем состояния цепи Маркова $\xi_t, t \geq 0$, в лексикографическом порядке. Интенсивности переходов процесса ξ_t определяются элементами его инфинитезимального генератора G .

Теорема 1. Генератор G цепи Маркова $\xi_t, t \geq 0$, имеет следующую блочно-трехдиагональную структуру

$$G = \begin{pmatrix} G_{0,0} & G_{0,1} & O & O & \dots \\ G_{1,0} & G_{1,1} & G_{1,2} & O & \dots \\ O & G_{2,1} & G_{2,2} & G_{2,3} & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Ненулевые блоки генератора определяются следующим образом:

$$G_{0,0} = I_{N-R+1} \otimes D_0 + \mu_2 C_{N-R} E_{N-R}^- \otimes I_W - \mu_2 C_{N-R} \otimes I_W + E_{N-R}^+ \otimes D_2,$$

$$G_{i,i} = G_{0,0} - i\mu_1 I_{(N-R+1)W}, i = \overline{1, R},$$

$$G_{i,i} = G_{0,0} - (R\mu_1 + (i-R)\beta) I_{(N-R+1)W}, i = \overline{R+1, J_1-1},$$

$$G_{i,i} = I_{N-R-k+1} \otimes D_0 + \mu_2 C_{N-R-k} E_{N-R-k}^- \otimes I_W - \mu_2 C_{N-R-k} \otimes I_W +$$

$$+ E_{N-R-k}^+ \otimes D_2 - ((R+k)\mu_1 + (i-(R+k))\beta) I_{(N-R-k+1)W},$$

$$i = \overline{J_k, J_{k+1}-1}, k = \overline{1, N-R-M-1},$$

$$\begin{aligned}
G_{i,i} &= I_{M+1} \otimes D_0 + \mu_2 C_M E_M^- \otimes I_W - \mu_2 C_M \otimes I_W + \\
&E_M^+ \otimes D_2 - ((N-M)\mu_1 + (i - (N-M))\beta) I_{(M+1)W}, i \geq J_{N-R-M}, \\
G_{i,i-1} &= (\min\{i, R\}\mu_1 + (i - \min\{i, R\})\beta) I_{(N-R+1)W}, 0 < i < J_1, \\
G_{i,i-1} &= ((R+k)\mu_1 + (i - (R+k))\beta) \tilde{E}_{N-R-k}^+ \otimes I_W, i = J_k, k = \overline{1, N-R-M}, \\
G_{i,i-1} &= ((R+k)\mu_1 + (i - (R+k))\beta) I_{(N-R-k+1)W}, \\
&i = \overline{J_k + 1, J_{k+1} - 1}, k = \overline{1, N-R-M-1}, \\
G_{i,i-1} &= ((N-M)\mu_1 + (i - (N-M))\beta) I_{(M+1)W}, i > J_{N-R-M}, \\
G_{i,i+1} &= I_{N-R+1} \otimes D_1, i = \overline{0, J_1 - 2}, \\
G_{i,i+1} &= \tilde{E}_{N-R-k+1}^- \otimes D_1, i = J_k - 1, k = \overline{1, N-R-M}, \\
G_{i,i+1} &= I_{N-R-k+1} \otimes D_1, i = \overline{J_k, J_{k+1} - 2}, k = \overline{1, N-R-M-1}, \\
G_{i,i+1} &= I_{M+1} \otimes D_1, i \geq J_{N-R-M},
\end{aligned}$$

где

$\otimes$ – символ Кронекера произведения матриц, см. [2];
 $C_i = \text{diag}\{0, 1, \dots, i-1, i\}$, где $\text{diag}\{\dots\}$ обозначает диагональную матрицу, диагональные элементы которой указаны в скобках;
 E_k^- – квадратная матрица размера $k+1$ со всеми нулевыми элементами, кроме элементов $(E_k^-)_{l,l-1} = 1, l = \overline{1, k}$;
 E_k^+ – квадратная матрица размера $k+1$ со всеми нулевыми элементами, кроме элементов $(E_k^+)_{l,l+1} = 1, l = \overline{0, k-1}$ и $(E_k^+)_{k,k} = 1$;
 $\tilde{E}_k^-$ – матрица размера $(k+1) \times k$ со всеми нулевыми элементами, кроме элементов $(\tilde{E}_k^-)_{l,l} = 1, l = \overline{0, k-1}$ и $(\tilde{E}_k^-)_{k,k-1} = 1$;
 $\tilde{E}_k^+$ – матрица размера $(k+1) \times (k+2)$ со всеми нулевыми

элементами, кроме элементов $(\tilde{E}_k^+)_{l,l} = 1, l = \overline{0, k}$.

Доказательство теоремы проводится путем анализа всех вариантов переходов цепи Маркова $\xi_t, t \geq 0$, на интервале бесконечно малой длины.

Библиографические ссылки

1. Lee, S., Dudin, A., Dudina, O., Kim, C. Analysis of a priority queueing system with the enhanced fairness of servers scheduling // Journal of Ambient Intelligence and Humanized Computing, 2022. P. 1–13.
2. Graham, A. Kronecker Products and Matrix Calculus with Applications. Cichester: Ellis Horwood. 1981. 130 p.