Информационные процессы, Том 15, № 1, 2015, стр. 66–77

Метод потенциалов для системы M/G/1/m с активным управлением очередью

Ю. В. Жерновый, К. Ю. Жерновый

Львовский национальный университет им. Ивана Франко, Львов, Украина
e-mail: yu.zhernovyi@gmail.com
Поступила в редколлегию 10.03.2015

Аннотация—Предложен метод определения характеристик системы обслуживания M/G/1/m с функцией случайного отбрасывания заявок и распределением времени обслуживания, зависящим от длины очереди. Найдены преобразования Лапласа для распределения числа заявок в системе в течение периода занятости и для функции распределения периода занятости. Полученные формулы для стационарных характеристик системы проверены на примерах с помощью имитационных моделей, построенных с привлечением инструментальных средств GPSS World. Приведены примеры сравнения результатов применения различных средств управления параметрами системы.
КЛЮЧЕВЫЕ СЛОВА: система обслуживания со случайным отбрасыванием заявок, время обслуживания, зависящее от длины очереди, метод потенциалов, стационарные характеристики.

1. ВВЕДЕНИЕ

С целью предотвращения перегрузок в узлах сетей с коммутацией пакетов (ATM, TCP/IP и др.) используются алгоритмы активного управления очередью (англ. *Active Queue Management — AQM*). В соответствующей системе обслуживания, моделирующей работу узла сети, каждый поступающий пакет может быть отброшен с определённой вероятностью, даже если буфер ещё полностью не заполнен. Вероятность отбрасывания зависит от длины очереди в момент поступления пакета. Зависимость вероятности отбрасывания пакетов от длины очереди называют функцией отбрасывания [1].

В алгоритмах AQM используют различные функции отбрасывания, например в известном алгоритме RED (Random Early Detection — случайное раннее обнаружение) [2] эта функция является линейной. Благодаря профилактическому случайному отбрасыванию пакетов алгоритм AQM косвенно информирует отправителя, использующего протокол TCP, о приближающейся перегрузке. Применение такого алгоритма в маршрутизаторе может принести много полезных эффектов, в том числе сокращение очереди и времени сетевой задержки [3].

Использование теории массового обслуживания для анализа алгоритмов активного управления очередью началось в последние годы [1, 4–7]. Результаты исследований показывают, что механизм функции отбрасывания является мощным средством для регулирования параметров системы обслуживания. С его помощью можно регулировать не только длину очереди, вероятность потери заявок, время ожидания, дисперсию длины очереди, но и несколько из этих параметров одновременно.

Для управления параметрами системы обслуживания, наряду с отбрасыванием заявок, применяются пороговые стратегии изменения интенсивности (времени) обслуживания. В общем случае сущность такой стратегии состоит в том, что распределение времени обслуживания зависит от числа заявок в системе и определяется в момент начала обслуживания каждой заявки [8]. Представляет интерес сравнение возможностей для повышения показателей производительности системы, получаемых в результате совместного применения механизмов отбрасывания заявок и порогового изменения интенсивности обслуживания, и в результате использования этих средств управления параметрами системы по отдельности.

Эффективные алгоритмы вычисления стационарных характеристик одноканальных систем с пороговыми стратегиями функционирования удалось разработать с помощью метода потенциалов [8–12]. Метод потенциалов является обобщением подхода, предложенного В. С. Королюком [13] для изучения непрерывного снизу случайного блуждания, на случай нескольких случайных блужданий. Метод потенциала с одним базовым случайным блужданием применялся для исследования системы $\mathrm{M}^{\alpha}/\mathrm{G}/1/\mathrm{N}$ с одним фиксированным распределением времени обслуживания [14]. Для изучения систем с несколькими режимами функционирования возникает необходимость использования стольких базовых случайных блужданий и их потенциалов, сколько применяется различных режимов функционирования.

В настоящей работе разработан вариант метода потенциалов, предназначенный для исследования системы M/G/1/m с функцией случайного отбрасывания заявок и распределением времени обслуживания, вид которого зависит от числа заявок в системе.

Мы рассматриваем систему M/G/1/m, где m — максимальное число заявок, которые одновременно могут находиться в очереди. Если в момент начала обслуживания заявки число заявок в системе равно n, то время обслуживания этой заявки распределено по закону $F_n(x)$, $n \in \{1, 2, \dots, m+1\}$, причём $F_n(x) = 0$ для $x \le 0$.

В отличие от работы [7], мы используем функцию отбрасывания, которая изменяет свои значения в моменты изменения числа заявок в системе. Каждая поступающая заявка может быть принята на обслуживание или отброшена согласно правилу: если в момент прибытия заявки в системе находится n заявок, то эта заявка принимается на обслуживание с вероятностью β_n и покидает систему (получает отказ, отбрасывается) с вероятностью $1-\beta_n$. В [7] значение функции отбрасывания не изменяется на промежутке времени от начала до завершения обслуживания каждой заявки.

Параметр показательного распределения промежутков времени между моментами прибытия заявок обозначим через λ.

2. БАЗОВЫЕ СЛУЧАЙНЫЕ БЛУЖДАНИЯ

Обозначим через $\mathbf{P}_n$ условную вероятность при условии, что в начальный момент времени в системе находится $n \in \{0, 1, 2, \dots, m+1\}$ заявок и через $\mathbf{E}$ ($\mathbf{P}$) условное математическое ожидание (условную вероятность) при условии, что система начинает работать в момент прибытия первой заявки. Пусть $\eta(x)$ — число заявок, поступивших в систему на промежутке времени $[0; x)$.

Для $n \in \{1, 2, \dots, m+1\}$ положим:

$$f_n(s) = \int\limits_0^\infty e^{-sx} dF_n(x), \qquad M_n = \int\limits_0^\infty x dF_n(x) < \infty, \qquad \overline{F}_n(x) = 1 - F_n(x).$$

Для $\operatorname{Re} s \ge 0$ и $n \in \{1, 2, \dots, m\}$ рассмотрим последовательности $\pi_{ni}(s)$ и $q_{ni}(s)$, определяемые с помощью соотношений

$$\pi_{ni}(s) = \frac{1}{f_n(s)} \int\limits_0^\infty e^{-sx}\, \mathbf{P}_n\{\eta(x) = i+1\}\, dF_n(x), \quad i \in \{-1, 0, 1, \dots, m-n\}; \tag{1}$$

$$q_{ni}(s) = \int\limits_0^\infty e^{-sx}\, \mathbf{P}_n\{\eta(x) = i\,\}\, \overline{F}_n(x)\, dx, \quad i \in \{0, 1, 2, \ldots, m - n + 1\}. \tag{2}$$

Функции $R_n(s)$ $(n \in \{1, 2, \ldots, m\})$ зададим в виде

$$R_n(s) = \frac{1}{f_n(s)\pi_{n,-1}(s)}\,. \tag{3}$$

Последовательность $\pi_{ni}(s)$ при $s > 0$ и фиксированном n можно трактовать как распределение скачков некоторого полунепрерывного снизу случайного блуждания, соответствующего функции распределения $F_n(x)$ n-го режима обслуживания и вероятностям $\mathbf{P}_n\{\eta(x) = i + 1\}$.

Пусть T_n и T обозначают показательно распределённые случайные величины с параметрами $\lambda_n = \lambda\beta_n$ и λ соответственно, а Z_n – случайная величина, распределённая по закону Паскаля, то есть $\mathbf{P}\{Z_n = k\} = \beta_n^k(1 - \beta_n)^{n-k}$, $k = 1, 2, \ldots$. Известно [15, с. 95], что $Z_nT = T_n$, то есть в результате случайного прореживания простейшего потока получаем простейший поток.

При вычислении вероятностей $\mathbf{P}_n\{\eta(x) = i\}$ будем учитывать, что в момент начала обслуживания заявки A_1, совпадающий с началом отсчёта времени до значения x, в системе находится n $(1 \le n \le m)$ заявок. Поэтому промежуток времени от момента начала обслуживания заявки A_1 до момента принятия в систему следующей заявки (обозначим её через A_{n+1}) при условии, что заявка A_{n+1} принята в систему не позже, чем завершилось обслуживание заявки A_1, — это случайная величина $Z_nT = T_n$, распределённая по показательному закону с параметром $\lambda_n = \lambda\beta_n$.

Если $n + 1 \le m$, то промежуток времени от момента принятия в систему заявки A_{n+1} до момента принятия следующей (обозначим её через A_{n+2}) распределён показательно с параметром $\lambda_{n+1} = \lambda\beta_{n+1}$ при условии, что заявка A_{n+2} принята в систему не позже, чем завершилось обслуживание заявки A_1. Если $n + 2 \le m$, то распределение промежутка времени между моментами принятия в систему заявок A_{n+2} и A_{n+3} показательное с параметром $\lambda_{n+2} = \lambda\beta_{n+2}$ при условии, что заявка A_{n+3} принята в систему не позже, чем завершилось обслуживание заявки A_1, и так далее.

Учитывая вышеизложенное, получаем

$$\begin{aligned}
&\mathbf{P}_n\{\eta(x) = 0\} = \mathbf{P}\{T_n > x\} = e^{-\lambda_n x}, \quad n \in \{1, 2, \ldots, m\};\\
&\mathbf{P}_n\{\eta(x) = j\} = \mathbf{P}\Big\{\sum_{i=n}^{n+j-1} T_i < x < \sum_{i=n}^{n+j} T_i\Big\} = \mathbf{P}\Big\{\sum_{i=n}^{n+j} T_i > x\Big\} - \mathbf{P}\Big\{\sum_{i=n}^{n+j-1} T_i > x\Big\}\\
&\quad = (-1)^j \prod_{k=n}^{n+j-1} \lambda_k \sum_{i=n}^{n+j} \frac{e^{-\lambda_i x}}{\prod\limits_{\substack{s=n\\ s\ne i}}^{n+j} (\lambda_i - \lambda_s)}\,, \;\; n \in \{1, 2, \ldots, m-1\}, \;\; j \in \{1, 2, \ldots, m-n\};\\
&\mathbf{P}_n\{\eta(x) = m - n + 1\} = 1 - \sum_{j=0}^{m-n} \mathbf{P}_n\{\eta(x) = j\} = \mathbf{P}\Big\{\sum_{i=n}^{m} T_i < x\Big\}\\
&\quad = (-1)^{m-n} \prod_{k=n}^{m} \lambda_k \sum_{i=n}^{m} \frac{1 - e^{-\lambda_i x}}{\lambda_i \prod\limits_{\substack{s=n\\ s\ne i}}^{m} (\lambda_i - \lambda_s)}\,, \qquad n \in \{1, 2, \ldots, m-1\};\\
&\mathbf{P}_m\{\eta(x) = 1\} = \mathbf{P}\Big\{T_m < x\Big\} = 1 - e^{-\lambda_m x}.
\end{aligned} \tag{4}$$

С учётом выражений (4) равенства (1), (2) приводятся к виду

$$
\begin{aligned}
&\pi_{n,-1}(s) = \frac{f_n\big(\lambda_n + s\big)}{f_n(s)}\,, \qquad n \in \{1, 2, \dots, m\};\\
&\pi_{n,j-1}(s) = \frac{(-1)^j}{f_n(s)} \prod_{k=n}^{n+j-1} \lambda_k \sum_{i=n}^{n+j} \frac{f_n\big(\lambda_i + s\big)}{\prod\limits_{\substack{\nu=n\\ \nu\neq i}}^{n+j} (\lambda_i - \lambda_\nu)}\,,\ n \in \{1, 2, \dots, m-1\},\ j \in \{1, 2, \dots, m-n\};\\
&\pi_{n,m-n}(s) = \frac{(-1)^{m-n}}{f_n(s)} \prod_{k=n}^{m} \lambda_k \sum_{i=n}^{m} \frac{f_n(s) - f_n\big(\lambda_i + s\big)}{\lambda_i \prod\limits_{\substack{\nu=n\\ \nu\neq i}}^{m} (\lambda_i - \lambda_\nu)}\,, \qquad n \in \{1, 2, \dots, m-1\};\\
&\pi_{m0}(s) = 1 - \frac{f_m\big(\lambda_m + s\big)}{f_m(s)}\,; \qquad q_{n0}(s) = \frac{1 - f_n\big(\lambda_n + s\big)}{\lambda_n + s}\,, \qquad n \in \{1, 2, \dots, m\};\\
&q_{nj}(s) = (-1)^j \prod_{k=n}^{n+j-1} \lambda_k \sum_{i=n}^{n+j} \frac{1 - f_n\big(\lambda_i + s\big)}{\big(\lambda_i + s\big) \prod\limits_{\substack{\nu=n\\ \nu\neq i}}^{n+j} (\lambda_i - \lambda_\nu)}\,, \qquad n \in \{1, 2, \dots, m-1\},\\
&\qquad\qquad j \in \{1, 2, \dots, m-n\};\\
&q_{n,m-n+1}(s) = (-1)^{m-n} \prod_{k=n}^{m} \lambda_k \sum_{i=1}^{m} \frac{\frac{1-f_n(s)}{s} - \frac{1-f_n\left(\lambda_i+s\right)}{\lambda_i+s}}{\lambda_i \prod\limits_{\substack{\nu=n\\ \nu\neq i}}^{n+j} (\lambda_i - \lambda_\nu)}\,, \qquad n \in \{1, 2, \dots, m-1\};\\
&q_{m1}(s) = \frac{1 - f_m(s)}{s} - \frac{1 - f_m\big(\lambda_m + s\big)}{\lambda_m + s}\,.
\end{aligned}
\tag{5}
$$

Отметим, что

$$
\lim_{s\to+0} f_n(s) = 1, \qquad \lim_{s\to+0} \frac{1}{1 - f_n(s)} = M_n, \qquad n \in \{1, 2, \dots, m+1\}. \tag{6}
$$

Введя обозначения

$$
\pi_{ni} = \lim_{s\to+0} \pi_{ni}(s), \qquad q_{ni} = \lim_{s\to+0} q_{ni}(s), \qquad R_n = \lim_{s\to+0} R_n(s),
$$

с помощью равенств (3), (5) и (6) получим соотношения:

$$
\begin{aligned}
&\pi_{n,-1} = f_n\big(\lambda_n\big), \quad n \in \{1, 2, \dots, m\};\\
&\pi_{n,j-1} = (-1)^j \prod_{k=n}^{n+j-1} \lambda_k \sum_{i=n}^{n+j} \frac{f_n\big(\lambda_i\big)}{\prod\limits_{\substack{\nu=n\\ \nu\neq i}}^{n+j} (\lambda_i - \lambda_\nu)}\,,\ n \in \{1, 2, \dots, m-1\},\ j \in \{1, 2, \dots, m-n\};\\
&\pi_{n,m-n} = (-1)^{m-n} \prod_{k=n}^{m} \lambda_k \sum_{i=n}^{m} \frac{1 - f_n\big(\lambda_i\big)}{\lambda_i \prod\limits_{\substack{\nu=n\\ \nu\neq i}}^{m} (\lambda_i - \lambda_\nu)}\,, \qquad n \in \{1, 2, \dots, m-1\};\\
&\pi_{m0} = 1 - f_m\big(\lambda_m\big); \qquad q_{n0} = \frac{1 - f_n\big(\lambda_n\big)}{\lambda_n}\,, \qquad n \in \{1, 2, \dots, m\};
\end{aligned}
$$

$$q_{nj}=(-1)^j\prod_{k=n}^{n+j-1}\lambda_k\sum_{i=n}^{n+j}\frac{1-f_n\big(\lambda_i\big)}{\lambda_i\prod\limits_{\substack{\nu=n\\ \nu\neq i}}^{n+j}(\lambda_i-\lambda_\nu)}\,,\quad n\in\{1,2,\ldots,m-1\},\;\; j\in\{1,2,\ldots,m-n\};$$

$$q_{n,m-n+1}=(-1)^{m-n}\prod_{k=n}^{m}\lambda_k\sum_{i=1}^{m}\frac{M_n-\frac{1-f_n(\lambda_i)}{\lambda_i}}{\lambda_i\prod\limits_{\substack{\nu=n\\ \nu\neq i}}^{n+j}(\lambda_i-\lambda_\nu)}\,,\qquad n\in\{1,2,\ldots,m-1\};$$

$$q_{m1}=M_m-\frac{1-f_m\big(\lambda_m\big)}{\lambda_m}\,;\qquad R_n=\frac{1}{\pi_{n,-1}}\,,\qquad n\in\{1,2,\ldots,m\}.$$

3. РАСПРЕДЕЛЕНИЕ ЧИСЛА ЗАЯВОК В СИСТЕМЕ В ТЕЧЕНИЕ ПЕРИОДА ЗАНЯТОСТИ

Пусть $\xi(t)$ — число заявок в системе в момент времени t, и $\tau(m)=\inf\{\,t\geq 0:\xi(t)=0\,\}$ обозначает первый период занятости для рассматриваемой системы обслуживания.

Для $n,\,k\in\{1,2,\ldots,m+1\}$ введём обозначения:

$$\varphi_n(t,\,k)=\mathbf{P}_n\{\,\xi(t)=k,\;\tau(m)>t\,\};\qquad \Phi_n(s,\,k)=\int\limits_0^\infty e^{-st}\varphi_n(t,\,k)\,dt,\quad \operatorname{Re}\,s>0.$$

Очевидно, что $\varphi_0(t,\,k)=0$. С помощью формулы полной вероятности получим равенства:

$$\begin{aligned}\varphi_n(t,\,k)=&\sum_{j=0}^{m-n+1}\int\limits_0^t\mathbf{P}_n\{\eta(x)=j\,\}\,\varphi_{n+j-1}(t-x,\,k)\,dF_n(x)\\ &+I\{n\leq k\leq m+1\}\mathbf{P}_n\{\eta(t)=k-n\,\}\overline{F}_n(t),\quad n\in\{1,2,\ldots,m\};\\ \varphi_{m+1}(t,\,k)=&\int\limits_0^t\varphi_m(t-x,\,k)\,dF_{m+1}(x)+I\{k=m+1\}\overline{F}_{m+1}(t).\end{aligned}\tag{7}$$

Здесь $I\{A\}$ равно 1 или 0, в зависимости от того, состоялось событие A или нет.

Введя обозначение $f_{(n)}(s,k,m)=I\{n\leq k\leq m+1\}q_{n,k-n}(s)$, и учитывая соотношения (1) и (2), из (7) получим систему уравнений для определения функций $\Phi_n(s,k)$:

$$\Phi_n(s,k)=f_n(s)\sum_{j=0}^{m-n+1}\pi_{n,j-1}(s)\Phi_{n+j-1}(s,k)+f_{(n)}(s,k,m),\qquad n\in\{1,2,\ldots,m\};\tag{8}$$

$$\Phi_{m+1}(s,k)=f_{m+1}(s)\Phi_m(s,k)+I\{k=m+1\}\frac{1-f_{m+1}(s)}{s}\,,\tag{9}$$

с граничным условием

$$\Phi_0(s,k)=0.\tag{10}$$

Найдём функции $\Phi_n(s,k)$, решив систему уравнений (8)–(10).

Будем использовать функции $\mathcal{R}_{ni}(s)$, определяемые с помощью рекуррентных соотношений:

$$\begin{gathered}\mathcal{R}_{n1}(s)=R_{n+1}(s);\;\;\mathcal{R}_{n,j+1}(s)=R_{n+1}(s)\bigg(\mathcal{R}_{n+1,j}(s)-f_{n+1}(s)\sum_{i=0}^{j-1}\pi_{n+1,i}(s)\mathcal{R}_{n+1+i,j-i}(s)\bigg),\\ n\in\{0,1,2,\ldots,m-1\},\quad j\in\{1,2,\ldots,m-n-1\}.\end{gathered}\tag{11}$$

Теорема 1. *Для всех* $k \in \{1, 2, \ldots, m+1\}$ *и* $\operatorname{Re} s > 0$ *функции* $\Phi_n(s,k)$ *определяются в виде*

$$\begin{aligned}\Phi_n(s,k) = \Bigg(\mathcal{R}_{n,\,m-n}(s) - \sum_{i=1}^{m-n} \mathcal{R}_{ni}(s) f_{n+i}(s) \pi_{n+i,m-n-i}(s)\Bigg)\Phi_m(s,k) \\ - \sum_{i=1}^{m-n} \mathcal{R}_{ni}(s) f_{(n+i)}(s,k,m), \qquad n \in \{1, 2, \ldots, m-1\};\end{aligned} \tag{12}$$

$$\Phi_{m+1}(s,k) = f_{m+1}(s)\Phi_m(s,k) + I\{k = m+1\}\frac{1 - f_{m+1}(s)}{s}\,, \tag{13}$$

где

$$\Phi_m(s,k) = \frac{\sum\limits_{i=1}^{m} \mathcal{R}_{0i}(s) f_{(i)}(s,k,m)}{\mathcal{R}_{0m}(s) - \sum\limits_{i=1}^{m} \mathcal{R}_{0i}(s) f_i(s) \pi_{i,m-i}(s)}\,. \tag{14}$$

Доказательство. Равенства (12) доказываются методом математической индукции с использованием уравнений (8) и соотношений (11). Доказательство аналогично доказательству теоремы 1 работы [10] и теоремы 3.1 монографии [11]. Равенства (13) совпадают с (9), а для получения формулы (14) достаточно положить $n = 0$ в (12) и использовать граничное условие (10). Теорема доказана. □

4. ПЕРИОД ЗАНЯТОСТИ И СТАЦИОНАРНОЕ РАСПРЕДЕЛЕНИЕ

Если система начинает работать в момент прибытия первой заявки, то

$$\begin{aligned}&\int_0^\infty e^{-st}\,\mathbf{P}\{\,\xi(t) = k,\ \tau(m) > t\,\}\,dt = \Phi_1(s,k) \\ &= \Bigg(\mathcal{R}_{1,\,m-1}(s) - \sum_{i=1}^{m-1} \mathcal{R}_{1i}(s) f_{i+1}(s) \pi_{i+1,m-1-i}(s)\Bigg)\Phi_m(s,k) - \sum_{i=1}^{m-1} \mathcal{R}_{1i}(s) f_{(i+1)}(s,k,m).\end{aligned} \tag{15}$$

Для отыскания $\int\limits_0^\infty e^{-st}\,\mathbf{P}\{\,\tau(m) > t\,\}\,dt$ необходимо сложить равенства (15) для всех k от 1 до $m+1$. Учитывая определения $f_{(n)}(s,k,m)$ и $q_{nj}(s)$, нетрудно убедиться, что

$$\sum_{k=1}^{m+1} f_{(n)}(s,k,m) = \sum_{k=n}^{m+1} f_{(n)}(s,k,m) = \sum_{j=0}^{m+1-n} q_{nj}(s) = \frac{1 - f_n(s)}{s}\,, \qquad n \in \{1, 2, \ldots, m\}.$$

Итак, из (15) вытекает следующее утверждение.

Теорема 2. *Преобразование Лапласа от функции распределения периода занятости определяется в виде*

$$\begin{aligned}\int_0^\infty e^{-st}\,\mathbf{P}\{\,\tau(m) > t\,\}\,dt = \Bigg(\mathcal{R}_{1,\,m-1}(s) - \sum_{i=1}^{m-1} \mathcal{R}_{1i}(s) f_{i+1}(s) \pi_{i+1,m-1-i}(s)\Bigg)\Phi_m(s) \\ - \frac{1}{s}\sum_{i=1}^{m-1} \mathcal{R}_{1i}(s)\big(1 - f_{i+1}(s)\big),\end{aligned} \tag{16}$$

где

$$\Phi_m(s) = \frac{\frac{1}{s}\sum\limits_{i=1}^{m} \mathcal{R}_{0i}(s)\bigl(1 - f_i(s)\bigr)}{\mathcal{R}_{0m}(s) - \sum\limits_{i=1}^{m} \mathcal{R}_{0i}(s) f_i(s) \pi_{i,m-i}(s)}\,.$$

Для отыскания $\int\limits_0^\infty \mathbf{P}\{\,\tau(m) > t\,\}\,dt = \mathbf{E}\,\tau(m)$ необходимо перейти к пределу в равенстве (16) при $s \to +0$. Для этого будем использовать последовательности $\{\pi_{ni}\}$, $\{q_{ni}\}$ и $\{R_n\}$, а также последовательности $\{\mathcal{R}_{ni}\}$, полученные вследствие предельного перехода $\mathcal{R}_{ni} = \lim\limits_{s\to+0} \mathcal{R}_{ni}(s)$. Для $\mathcal{R}_{ni}$ из (11) вытекают рекуррентные соотношения

$$\mathcal{R}_{n1} = R_{n+1}; \qquad \mathcal{R}_{n,j+1} = R_{n+1}\Biggl(\mathcal{R}_{n+1,j} - \sum_{i=0}^{j-1} \pi_{n+1,i}\mathcal{R}_{n+1+i,j-i}\Biggr),$$
$$n \in \{0,1,2,\ldots,m-1\}, \quad j \in \{1,2,\ldots,m-n-1\}. \tag{17}$$

Используя соотношения (17), получаем равенства

$$\mathcal{R}_{n,m-n} - \sum_{i=1}^{m-n} \mathcal{R}_{ni}\pi_{n+i,m-n-i} = 1, \qquad n \in \{0,1,2,\ldots,m-1\}. \tag{18}$$

Учитывая (6) и (18), с помощью (16) приходим к следующему утверждению.

Теорема 3. *Средняя продолжительность периода занятости определяется в виде*

$$\mathbf{E}\,\tau(m) = \sum_{i=0}^{m} \mathcal{R}_{0i}M_i - \sum_{i=1}^{m-1} \mathcal{R}_{1i}M_{i+1}. \tag{19}$$

Введём обозначения: $\lim\limits_{t\to\infty} \mathbf{P}\{\xi(t) = k\} = p_k(m)$, $k \in \{0,1,2,\ldots,m+1\}$. Рассуждая так же, как в работе [8], и учитывая, что промежутки времени пребывания системы в сосотоянии без заявок распределены показательно с параметром λ, из (15) получаем формулы для стационарного распределения числа заявок в системе.

Теорема 4. *Стационарное распределение числа заявок в системе определяется по формулам:*

$$p_0(m) = \frac{1}{1 + \lambda\,\mathbf{E}\,\tau(m)}\,;$$
$$p_k(m) = \lambda p_0(m)\left(\mathcal{R}_{0k}q_{k0} + \sum_{i=1}^{k-1}\Bigl(\mathcal{R}_{0i}q_{i,k-i} - \mathcal{R}_{1i}q_{i+1,k-i-1}\Bigr)\right), \quad k \in \{1,2,\ldots,m\}; \tag{20}$$
$$p_{m+1}(m) = \lambda p_0(m)\left(\mathcal{R}_{0m}q_{m1} + \sum_{i=1}^{m-1}\Bigl(\mathcal{R}_{0i}q_{i,m+1-i} - \mathcal{R}_{1i}q_{i+1,m-i}\Bigr)\right).$$

Вычислив с помощью (19) отношение среднего числа обслуженных заявок за единицу времени к среднему числу всех прибывающих заявок за единицу времени в стационарном режиме, получим формулу для стационарной вероятности обслуживания поступившей заявки (относительной пропускной способности системы)

$$\mathbf{P}_{\mathrm{sv}}(m) = \frac{\mathcal{R}_{0m} + \sum\limits_{i=1}^{m-1} \left(\mathcal{R}_{0i} - \mathcal{R}_{1i}\right)}{1 + \lambda\, \mathbf{E}\, \tau(m)}\,. \tag{21}$$

Стационарные характеристики очереди — среднюю длину очереди $\mathbf{E}\, Q(m)$ и среднее время ожидания $\mathbf{E}\, w(m)$ — находим по формулам

$$\mathbf{E}\, Q(m) = \sum_{k=1}^{m} k p_{k+1}(m); \qquad \mathbf{E}\, w(m) = \frac{\mathbf{E}\, Q(m)}{\lambda\, \mathbf{P}_{\mathrm{sv}}(m)}\,. \tag{22}$$

5. ПРИМЕРЫ ВЫЧИСЛЕНИЯ СТАЦИОНАРНЫХ ХАРАКТЕРИСТИК

Пусть $m = 5$, $\lambda = 10$. Рассмотрим 4 системы типа M/G/1/m с различными механизмами управления параметрами этих систем. В системе 1 применяются механизмы случайного отбрасывания заявок и порогового изменения времени обслуживания, в системе 2 используется только случайное отбрасывание заявок, в системе 3 — только пороговая стратегия изменения времени обслуживания, а в системе 4 заявки отбрасываются только в случае превышения очереди значения m.

Система 1. Рассмотрим систему с одним порогом h переключения распределения времени обслуживания. Пусть

$$F_n(x) = F(x), \quad n \in \{1, 2, \ldots, h\}; \qquad F_n(x) = \widetilde{F}(x), \quad n \in \{h+1, h+2, \ldots, m+1\}.$$

Введём обозначения:

$$f(y) = \int\limits_0^\infty e^{-yx} dF(x), \quad \widetilde{f}(y) = \int\limits_0^\infty e^{-yx} d\widetilde{F}(x); \qquad M = \int\limits_0^\infty x dF(x), \quad \widetilde{M} = \int\limits_0^\infty x d\widetilde{F}(x).$$

Тогда

$$f_n(y) = f(y), \quad n \in \{1, 2, \ldots, h\}; \qquad f_n(y) = \widetilde{f}(y), \quad n \in \{h+1, h+2, \ldots, m+1\}.$$

Предположим, что применяется механизм случайного отбрасывания заявок с вероятностями $\beta_n = (6 - n)/5$, $1 \le n \le 6$; $h = 3$; функции распределения времени обслуживания $F(x)$ соответствует равномерное распределение на промежутке $(0; 0,5]$, а функции распределения $\widetilde{F}(x)$ — равномерное распределение на промежутке $(0; 0,25]$. Таким образом,

$$M = 0,25, \quad \widetilde{M} = 0,125, \quad f(y) = \frac{2}{y}\left(1 - e^{-0,5y}\right), \quad \widetilde{f}(y) = \frac{4}{y}\left(1 - e^{-0,25y}\right).$$

Система 2. Рассмотрим систему с распределением времени обслуживания, не зависящим от числа заявок в системе: $F_n(x) = F(x)$, $f_n(y) = f(y)$, $n \in \{1, 2, \ldots, m+1\}$. Пусть вероятности β_n и функция распределения $F(x)$ те же, что и для системы 1, то есть

$$M = 0,25, \qquad f(y) = \frac{2}{y}\left(1 - e^{-0,5y}\right).$$

Система 3. Рассмотрим систему с одним порогом h переключения распределения времени обслуживания с теми же значениями h и функциями $F(x)$, $\widetilde{F}(x)$, что и в системе 1. Случайное отбрасывание заявок не применяется.

Система 4. Рассмотрим обычную систему M/G/1/m, в которой не применяется случайное отбрасывание заявок, $F_n(x) = F(x)$, $n \in \{1, 2, \ldots, m+1\}$, и функция распределения $F(x)$ та же, что и для системы 2.

В строках "$p_k(5)$" табл. 1 и 2 записаны стационарные вероятности $p_k(5)$, вычисленные по формулам (20) для систем 1 и 2 соответственно, а в нижних строках этих таблиц для сравнения приведены значения соответствующих вероятностей, полученные с помощью системы имитационного моделирования GPSS World [16]. Значения стационарных характеристик систем 1 и 2, найденные по формулам (19), (21) и (22), а также при помощи GPSS World, представлены в табл. 3. В этой таблице приведены также значения стационарных характеристик систем 3 и 4, вычисленные с помощью GPSS World. Все результаты, которые представлены в табл. 1–3 и получены с помощью имитационных моделей, соответствуют значению времени моделирования $t = 10^6$. Программы GPSS World для построения имитационных моделей систем 1 и 2 приведены в Приложении.

Таблица 1. Стационарное распределение числа заявок в системе 1

Число заявок (k)	0	1	2	3	4	5	6
$p_k(5)$	0,00800	0,03229	0,11646	0,29836	0,33314	0,17687	0,03486
$p_k(5)$ (GPSS World)	0,00808	0,03243	0,11603	0,29880	0,33305	0,17673	0,03489

Таблица 2. Стационарное распределение числа заявок в системе 2

Число заявок (k)	0	1	2	3	4	5	6
$p_k(5)$	0,00539	0,02174	0,07841	0,20089	0,32796	0,28134	0,08427
$p_k(5)$ (GPSS World)	0,00543	0,02177	0,07863	0,20093	0,32757	0,28137	0,08430

Таблица 3. Стационарные характеристики систем 1–4

№ системы	Характеристика	$\mathbf{E}\,\tau(5)$	$\mathbf{E}\,Q(5)$	$\mathbf{E}\,w(5)$	$\mathbf{P}_{\mathrm{sv}}(5)$
1	Аналитическое значение	12,393	2,594	0,539	0,481
1	Значение согласно GPSS World	12,357	2,593	0,539	0,481
2	Аналитическое значение	18,456	3,011	0,757	0,398
2	Значение согласно GPSS World	18,511	3,010	0,756	0,398
3	Значение согласно GPSS World	4,823	3,824	0,575	0,665
4	Значение согласно GPSS World	282,161	4,457	1,114	0,400

Анализируя полученные результаты, видим, что применение пороговой стратегии изменения времени обслуживания и механизма случайного отбрасывания заявок совместно или по отдельности позволяет значительно уменьшить длину очереди и продолжительность периода занятости и, соответственно, повысить пропускную способность системы. Сравнивая данные для систем 2 и 4, приходим к выводу, что случайное отбрасывание заявок позволяет значительно уменьшить среднюю продолжительность периода занятости и длину очереди при почти неизменном значении относительной пропускной способности системы. Если приоритетной целью является повышение пропускной способности системы, то наилучший результат достигается в случае применения пороговой стратегии изменения времени обслуживания.

6. ЗАКЛЮЧЕНИЕ

В настоящей работе с помощью метода потенциалов получены простые и удобные для числовой реализации формулы для отыскания стационарных характеристик систем типа M/G/1/m, в которых с целью улучшения показателей производительности применяется пороговая стратегия изменения времени обслуживания и механизм случайного отбрасывания заявок. Полученные с помощью аналитических моделей результаты подтверждаются данными имитационного моделирования.

7. ПРИЛОЖЕНИЕ. ПРОГРАММЫ GPSS WORLD ДЛЯ СИСТЕМ 1 И 2

```
Lam EQU 10    ; значение λ
Tm EQU 1000000    ; время моделирования
Em EQU 5    ; значение m
AH EQU 3    ; значение h (только для системы 1)
DIS TABLE (F$Sys+Q1) 0,1,8    ; параметры таблицы распределения числа заявок
Tau TABLE X$Tau 0,1,2000    ; параметры таблицы распределения периода занятости
GENERATE 1
TABULATE DIS    ; вычислить распределение числа заявок
TERMINATE
GENERATE (Exponential(5,0,1/Lam))    ; входящий поток
L0 TEST LE (F$Sys+Q1),Em,OUT    ; ограничение на число заявок
TEST LE (F$Sys+Q1),1,L2
TRANSFER ,L1
L2 TEST E (F$Sys+Q1),2,L3
TRANSFER 800,OUT,L1    ; задание вероятности β2
L3 TEST E (F$Sys+Q1),3,L4
TRANSFER 600,OUT,L1    ; задание вероятности β3
L4 TEST E (F$Sys+Q1),4,L5
TRANSFER 400,OUT,L1    ; задание вероятности β4
L5 TRANSFER 200,OUT,L1    ; задание вероятности β5
L1 QUEUE 1    ; занять очередь
TEST E F$Sys,0,L6
SAVEVALUE Tau,AC1    ; записать текущее значение времени
L6 SEIZE Sys    ; войти в систему
DEPART 1    ; покинуть очередь
TEST LE (F$Sys+Q1),AH,LT ; превышено ли значение h? (только для системы 1)
ADVANCE (Uniform(15,0,0.5))    ; распределение времени обслуживания согл. закону F(x)
TEST E Q1,0,L7    ; равна ли длина очереди нулю?
SAVEVALUE Tau,(AC1-X$Tau)    ; записать значение периода занятости
TABULATE Tau    ; вычислить распределение периода занятости
TRANSFER ,L7
LT ADVANCE (Uniform(15,0,0.25)) ; распределение F̃(x) (только для системы 1)
L7 RELEASE Sys    ; покинуть систему
OUT TERMINATE
```

```
GENERATE Tm ; реализовать время моделирования
SAVEVALUE Psv,(N$L7/N$L0) ; записать значение вероятности обслуживания
TERMINATE 1
START 1
```

СПИСОК ЛИТЕРАТУРЫ

1. Chydziński A. *Nowe Modele Kolejkowe Dla Węzłów Sieci Pakietowych.* Gliwice: Pracownia Komputerowa Jacka Skalmierskiego, 2013.
2. Floyd S., Jacobson V. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1993, vol. 1, № 4, pp. 397–413.
3. Bohacek S., Shah K., Arce G. R., Davis M. Signal processing challenges in active queue management. *IEEE Signal Processing Magazine Transactions on Networking*, 2004, vol. 21, № 5, pp. 69–79.
4. Bonald T., May M., Bolot J.-C. Analytic evaluation of RED performance. *Proceedings of INFOCOM*, 2000, vol. 3, pp. 1415–1424.
5. Tikhonenko O., Kempa W. M. The generalization of AQM algorithms for queueing system with bounded capacity. *Lecture Notes in Computer Sciences*, 2012, vol. 7204, pp. 242–251.
6. Zhernovyi Yu. V. The simplest models of queue control at nodes of packet-switched networks. *Journal of Communications Technology and Electronics*, 2014, т. 59, № 12, pp. 1475–1482.
7. Zhernovyi Yu., Kopytko B., Zhernovyi K. On characteristics of the $\mathrm{M}^{\theta}/\mathrm{G}/1/\mathrm{m}$ and $\mathrm{M}^{\theta}/\mathrm{G}/1$ queues with queue-size based packet dropping. *Journal of Applied Mathematics and Computational Mechanics*, 2014, № 13(4), pp. 163–175.
8. Zhernovyi K. Yu., Zhernovyi Yu. V. $\mathrm{M}^{\theta}/\mathrm{G}/1/\mathrm{m}$ and $\mathrm{M}^{\theta}/\mathrm{G}/1$ systems with the service time dependent on the queue length. *Journal of Communications Technology and Electronics*, 2013, vol. 58, № 12, pp. 1267–1275.
9. Zhernovyi K. Yu. Stationary characteristics of the $\mathrm{M}^{\theta}/\mathrm{G}/1/\mathrm{m}$ system with the threshold functioning strategy. *Journal of Communications Technology and Electronics*, 2011, vol. 56, № 12, pp. 1585—1596.
10. Zhernovyi K. Yu., Zhernovyi Yu. V. $\mathrm{M}^{\theta}/\mathrm{G}/1/\mathrm{m}$ and $\mathrm{M}^{\theta}/\mathrm{G}/1$ queues with operating parameters depending on the queue length. *Journal of Communications Technology and Electronics*, 2014, vol. 59, № 6, pp. 605–613.
11. Жерновый Ю., Жерновый К. *Метод потенциалов для пороговых стратегий обслуживания.* Saarbrücken: LAP Lambert Academic Publishing, 2015.
12. Zhernovyi Yu. *Insensitivity of the queueing systems characteristics.* Saarbrücken: LAP Lambert Academic Publishing, 2015.
13. Королюк В. С. *Граничные задачи для сложных пуассоновских процессов.* Киев: Наукова думка, 1975.
14. Bratiychuk M., Borowska B. Explicit formulae and convergence rate for the system $\mathrm{M}^{\alpha}/\mathrm{G}/1/\mathrm{N}$ as $N \to \infty$. *Stochastic Models*, 2002, vol. 18, № 1, pp. 71–84.
15. Вентцель Е. С., Овчаров Л. А. *Теория случайных процессов и её инженерные приложения.* М.: Высшая школа, 2000.
16. Жерновый Ю. *Создание моделей систем обслуживания в среде GPSS World.* Saarbrücken: Palmarium Academic Publishing, 2014.

Method of potentials for the system M/G/1/m with active queue management

Zhernovyi Yu. V., Zhernovyi K. Yu.

We propose a method for determining the characteristics of an M/G/1/m queueing system with the function of the random dropping of customers and distribution of the service time depending on the queue length. The Laplace transforms for the distribution of the number of customers in the system during the busy period and for the distribution function of the busy period are found. Obtained formulas for the stationary characteristics of the system are tested on examples using simulation models constructed with the assistance of the GPSS World tools. Examples of comparison of the results of the use of different control tools of system parameters are given.

KEYWORDS: queueing system with the random dropping of customers, service time depending on the queue length, potentials method, stationary characteristics.