Информационные процессы, Том 11, № 2, 2011, стр. 179–195
© *2011 Жерновый.*

Стационарные характеристики системы $\mathrm{M}^{\theta}/\mathrm{G}/1/\mathrm{m}$ с пороговой стратегией функционирования

К. Ю. Жерновый

Львовский национальный университет им. Ивана Франко, Львов, Украина

Поступила в редколлегию 15.02.2011

Аннотация—Рассмотрена система обслуживания типа $\mathrm{M}^{\theta}/\mathrm{G}/1/\mathrm{m}$ с групповым поступлением заявок, переключениями режимов обслуживания и пороговой блокировкой потока заявок. Блокировка входного потока осуществляется, если в момент начала обслуживания очередной заявки число заявок в системе превышает заданный пороговый уровень h. Если в момент t начала обслуживания заявки число заявок в системе $\xi(t)$ удовлетворяет условию $h_i < \xi(t) \le h_{i+1}$ $(i = \overline{1, r})$, то времени обслуживания этой заявки соответствует функция распределения $F_i(t)$. При $1 \le \xi(t) \le h = h_1$ время обслуживания заявки распределено по закону $F(t)$ (основной режим обслуживания). Для случая одного переключения $(r = 1)$ определены средняя продолжительность частей периода занятости с отсутствием и наличием блокировки входного потока, вероятность обслуживания заявок и стационарные характеристики очереди. Исследован характер зависимостей средней продолжительности периода занятости и вероятности обслуживания от параметров m и h. Для случая $m = \infty$ решены некоторые задачи оптимального синтеза систем с заданными характеристиками и задача минимизации затрат на обслуживание.

1. ВВЕДЕНИЕ

Цель настоящей работы — продолжение изучения системы обслуживания типа $\mathrm{M}^{\theta}/\mathrm{G}/1/\mathrm{m}$ с переключениями режимов обслуживания и пороговой блокировкой потока заявок, начатое в статье [1].

Приведём формальное описание рассматриваемой системы. Пусть заданы последовательности случайных величин $\{\alpha_n\}$, $\{\theta_n\}$, $\{\beta_n\}$ $(n \ge 1)$, где α_n представляет время между поступлением $(n-1)$–ой и n–ой группы, $\{\theta_n\}$ – число заявок в n–ой группе, а $\{\beta_n\}$ — время обслуживания n–ой заявки. Все перечисленные выше величины независимы, причём $\mathbf{P}\{\alpha_n < x\} = 1 - e^{-\lambda x}$ $(\lambda > 0)$, $\mathbf{P}\{\theta_n = i\} = a_i$ $(i \ge 1)$, и $\mathbf{P}\{\beta_n < x\} = F(x)$ $(x \ge 0)$, $F(0) = 0$. Если $\mathbf{P}\{\theta_n = 1\} = a_1 = 1$, то заявки в систему поступают по одной.

Заявки обслуживаются по одной, обслуженная заявка покидает систему, а обслуживающее устройство немедленно начинает обслуживать заявку из очереди при её наличии или же ждёт поступления очередной группы заявок. Применяется дисциплина обслуживания FIFO.

Пусть m — максимальное число заявок, которые одновременно могут находиться в очереди. Итак, если в систему, в которой уже находится $k \in [0, m+1]$ заявок, поступает группа из θ_n заявок, то только $\min\{\theta_n, m+1-k\}$ из них присоединяются к очереди, а остальные теряются.

Обозначим через $\xi(t)$ число заявок в системе в момент времени t и введём для рассматриваемой системы так называемый пороговый уровень h. Если t — момент начала обслуживания n–ой заявки и $\xi(t) > h$ $(h = \overline{1, m-1})$, то во время обслуживания этой заявки осуществляется блокировка потока заявок (поступившие заявки не допускаются на вход системы). Предположим также, что $F(x)$ является функцией распределения времени обслуживания заявки только в том случае, когда в момент t начала обслуживания этой заявки выполнено условие

$1 \le \xi(t) \le h$. Если же $\xi(t) > h$, то обслуживание заявки может осуществляться в одном из r режимов ($r \ge 1$), каждому из которых соответствует своя функция распределения времени обслуживания. Обозначим через $F_i(t)$ ($i = \overline{1, r}$) функцию распределения времени обслуживания заявки в i–ом режиме. Условие переключения на этот режим зададим в виде $h_i < \xi(t) \le h_{i+1}$ ($i = \overline{1, r}$), где t — время начала обслуживания заявки, h_i ($i = \overline{1, r}$) — фиксированные натуральные числа, удовлетворяющие неравенствам $1 < h = h_1 < h_2 < \ldots < h_r < h_{r+1} = m$. Описанную систему обозначим через $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_\mathrm{r}/1/\mathrm{m}$.

Для системы $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_\mathrm{r}/1/\mathrm{m}$ в [1] определена, в частности, средняя продолжительность периода занятости и получены формулы для эргодического распределения числа заявок в системе.

В настоящей работе мы, в основном, будем рассматривать систему с одним переключением $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_1/1/\mathrm{m}$, для которой будем решать задачу определения тех частей периода занятости, когда не осуществляется блокировка входного потока, и когда поток заявок заблокирован, а также стационарных вероятностей соответствующих состояний системы. Это позволит найти вероятность обслуживания заявок в системе. Определим также стационарные характеристики очереди (среднюю длину очереди, среднее время ожидания) и изучим поведение некоторых стационарных характеристик как функций параметров m и h. Как показано в работах [2–4], определение характера монотонной зависимости характеристик систем обслуживания от входных параметров открывает путь к решению задач оптимального синтеза систем с заданными характеристиками. Решения некоторых из таких задач будут найдены в явном виде.

2. ПЕРИОД ЗАНЯТОСТИ: ЧАСТИ С ОТСУТСТВИЕМ И НАЛИЧИЕМ БЛОКИРОВКИ

Сохраняя все обозначения, введённые в статье [1], рассмотрим сначала случай одного переключения режимов обслуживания ($r = 1$, $h_1 = h$, $h_{r+1} = h_2 = m$). Тогда формулы для средней продолжительности периода занятости и эргодического распределения числа заявок в системе принимают вид

$$\begin{aligned}\mathbf{M}\,\tau(m) = {} & \sum_{i=1}^{h} R_i\Bigl(m_1 + m_{11}\Bigl((m-h)\overline{p}_{m-i} + \sum_{j=h+1}^{m-1}(j-h)p_{j-i}\Bigr)\Bigr) \\ & - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i\Bigl(m_1 + m_{11}\Bigl((m-h)\overline{p}_{m-n-i} + \sum_{j=h+1}^{m-1}(j-h)p_{j-n-i}\Bigr)\Bigr) \\ & + m_{11}\Bigl(\sum_{n=h+1}^{m}(n-h)a_n + (m+1-h)\overline{a}_{m+1}\Bigr);\end{aligned} \tag{1}$$

$$\begin{aligned}\rho_0(m) &= \frac{1}{1+\lambda \mathbf{M}\,\tau(m)}\,; \\ \rho_k(m) &= \frac{\lambda}{1+\lambda \mathbf{M}\,\tau(m)}\left(\sum_{i=1}^{k} R_i q_{k-i} - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{k-n} R_i q_{k-n-i}\right) \qquad (k = \overline{1, h}\,); \\ \rho_k(m) &= \frac{\lambda}{1+\lambda \mathbf{M}\,\tau(m)}\left(\sum_{i=1}^{h} R_i\bigl(q_{k-i} + m_{11}\overline{p}_{k-i}\bigr)\right. \\ & \left. - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i\bigl(q_{k-n-i} + m_{11}\overline{p}_{k-n-i}\bigr) + m_{11}\overline{a}_k\right) \qquad (k = \overline{h+1, m}\,); \\ \rho_{m+1}(m) &= \frac{\lambda}{1+\lambda \mathbf{M}\,\tau(m)}\left(\sum_{i=1}^{h} R_i\overline{q}_{m+1-i} - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i\overline{q}_{m+1-n-i} + m_{11}\overline{a}_{m+1}\right).\end{aligned} \tag{2}$$

Напомним, что здесь

$$m_1 = \int_0^\infty x dF(x) < \infty, \qquad m_{11} = \int_0^\infty x dF_1(x) < \infty$$

обозначают соответственно среднюю продолжительность обслуживания для основного и послепорогового режимов.

Очевидно, что те слагаемые в правых частях соотношений (1) и (2), которые содержат множитель m_{11}, соответствуют периоду блокировки входного потока, но только той её части, которая начиналась с момента t начала обслуживания очередной заявки, для которой выполнялось условие $\xi(t) > h$. Поэтому для получения из (1) формулы для средней продолжительности той части периода занятости, когда входной поток блокируется (обозначим её через $\tau_{\mathrm{b}}(m)$), необходимо к указанной сумме добавить выражение

$$\mathbf{M}\,\widetilde{\tau}_{m+1}(m) = \sum_{i=1}^{h} R_i \overline{q}_{m+1-i} - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i \overline{q}_{m+1-n-i}, \tag{3}$$

которое находится в числителе формулы для вероятности $\rho_{m+1}(m)$ и отображает среднюю продолжительность той части периода занятости, когда входной поток заблокирован из-за превышения количества заявок в системе числа m, и осуществляется дообслуживание той заявки, в момент t начала обслуживания которой выполнялось условие $\xi(t) \le h$. В то же время, выражение (3) необходимо вычесть от суммы тех слагаемых правой части формулы (1), которые не содержат сомножителя m_{11}, чтобы получить среднюю продолжительность той части периода занятости, когда отсутствует блокировка входного потока (обозначим её через $\tau_{\mathrm{nb}}(m)$).

Обозначим через $\rho_{k(\mathrm{b})}(m)$ и $\rho_{k(\mathrm{nb})}(m)$ части эргодического распределения числа заявок в системе, соответствующие состояниям системы, когда входной поток заявок заблокирован и, соответственно, незаблокирован. Из вышеизложенного и очевидного равенства

$$\sum_{i=1}^{h} R_i - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i = \sum_{i=1}^{h} R_i \overline{a}_{h+1-i}$$

получим следующее утверждение.

Теорема 1. *Средние значения случайных величин $\tau_{nb}(m)$, $\tau_b(m)$, соответствующие части эргодического распределения числа заявок в системе $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_1/1/\mathrm{m}$ и вероятность обслуживания заявок $\mathbf{P}_{\mathrm{sv}}(m)$ определяются по формулам:*

$$\mathbf{M}\,\tau_{nb}(m) = m_1 \sum_{i=1}^{h} R_i \overline{a}_{h+1-i} - \mathbf{M}\,\widetilde{\tau}_{m+1}(m); \tag{4}$$

$$\begin{aligned} \mathbf{M}\,\tau_b(m) = m_{11}\Bigg(& \sum_{i=1}^{h} R_i \bigg((m-h)\overline{p}_{m-i} + \sum_{j=h+1}^{m-1} (j-h) p_{j-i} \bigg) \\ & - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i \bigg((m-h)\overline{p}_{m-n-i} + \sum_{j=h+1}^{m-1} (j-h) p_{j-n-i} \bigg) \\ & + \sum_{n=h+1}^{m} (n-h) a_n + (m+1-h)\overline{a}_{m+1} \Bigg) + \mathbf{M}\,\widetilde{\tau}_{m+1}(m); \end{aligned} \tag{5}$$

$$\rho_{0(nb)}(m) = \rho_0(m); \qquad \rho_{k(nb)}(m) = \rho_k(m) \quad (\,k = \overline{1,\,h}\,); \qquad \rho_{m+1(nb)}(m) = 0;$$
$$\rho_{k(nb)}(m) = \frac{\lambda}{1+\lambda\mathbf{M}\,\tau(m)}\left(\sum_{i=1}^{h} R_i q_{k-i} - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i q_{k-n-i}\right) (\,k = \overline{h+1,\,m}\,); \tag{6}$$

$$\rho_{k(b)}(m) = 0 \quad (\,k = \overline{0,\,h}\,); \qquad \rho_{m+1(b)}(m) = \rho_{m+1}(m);$$
$$\rho_{k(b)}(m) = \frac{\lambda m_{11}}{1+\lambda\mathbf{M}\,\tau(m)}\left(\sum_{i=1}^{h} R_i \overline{p}_{k-i} - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i \overline{p}_{k-n-i} + \overline{a}_k\right) \quad (\,k = \overline{h+1,\,m}\,); \tag{7}$$

$$\mathbf{P}_{\mathrm{sv}}(m) = \frac{1+\lambda\mathbf{M}\,\tau_{nb}(m)}{1+\lambda\mathbf{M}\,\tau(m)} \qquad (a_1 = 1;\; a_n = 0,\; n \ge 2). \tag{8}$$

Рассмотрим систему $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_1/1$ с неограниченным объёмом накопителя ($m = \infty$). Положив $m \to \infty$ в соотношениях (4)–(8), получим следующее утверждение.

Теорема 2. *Средние значения случайных величин $\tau_{nb}(\infty)$, $\tau_b(\infty)$, соответствующие части эргодического распределения числа заявок в системе $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_1/1$ и вероятность обслуживания заявок $\mathbf{P}_{\mathrm{sv}}(\infty)$ определяются по формулам:*

$$\mathbf{M}\,\tau_{nb}(\infty) = m_1 \sum_{i=1}^{h} R_i \overline{a}_{h+1-i};$$
$$\mathbf{M}\,\tau_b(\infty) = m_{11}\left(\sum_{i=1}^{h} R_i \sum_{j=h+1}^{\infty}(j-h)p_{j-i} - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i \sum_{j=h+1}^{\infty}(j-h)p_{j-n-i} + \sum_{n=h+1}^{\infty}(n-h)a_n\right);$$
$$\mathbf{M}\,\tau(\infty) = \mathbf{M}\,\tau_{nb}(\infty) + \mathbf{M}\,\tau_b(\infty); \qquad \rho_{0(nb)}(\infty) = \frac{1}{1+\lambda\mathbf{M}\,\tau(\infty)};$$
$$\rho_{k(nb)}(\infty) = \frac{\lambda}{1+\lambda\mathbf{M}\,\tau(\infty)}\left(\sum_{i=1}^{k} R_i q_{k-i} - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{k-n} R_i q_{k-n-i}\right) \quad (\,k = \overline{1,\,h}\,);$$
$$\rho_{k(nb)}(\infty) = \frac{\lambda}{1+\lambda\mathbf{M}\,\tau(\infty)}\left(\sum_{i=1}^{h} R_i q_{k-i} - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i q_{k-n-i}\right) \quad (\,k \ge h+1\,);$$
$$\rho_{k(b)}(\infty) = 0 \qquad (\,k = \overline{0,\,h}\,);$$
$$\rho_{k(b)}(\infty) = \frac{\lambda m_{11}}{1+\lambda\mathbf{M}\,\tau(\infty)}\left(\sum_{i=1}^{h} R_i \overline{p}_{k-i} - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i \overline{p}_{k-n-i} + \overline{a}_k\right) \quad (\,k \ge h+1\,);$$

$$\mathbf{P}_{\mathrm{sv}}(\infty) = \frac{1+\lambda\mathbf{M}\,\tau_{nb}(\infty)}{1+\lambda\mathbf{M}\,\tau(\infty)}\,. \tag{9}$$

3. ИССЛЕДОВАНИЕ ЗАВИСИМОСТЕЙ $\mathbf{M}\,\tau(m)$ ОТ ПАРАМЕТРОВ m И h

Напомним некоторые обозначения из [1]:

$$\rho = \lambda m_1 b_1, \qquad b_1 = \sum_{n=1}^{\infty} n a_n < \infty.$$

Лемма 1. *Для последовательностей* $\{p_i\}$, $\{R_i\}$, *введёных в* [1], *выполняются следующие соотношения:*

$$\sum_{j=0}^{\infty}\overline{p}_j=\rho;\qquad \sum_{i=1}^{k}R_i\sum_{j=0}^{k-i}\overline{p}_j=\sum_{i=1}^{k}R_i-k. \tag{10}$$

Доказательство. Учитывая, что согласно определению последовательности вероятностей $\{p_i\}$ $(i\geq -1)$

$$\sum_{i=-1}^{\infty}(i+1)p_i=\rho,$$

получим

$$\sum_{i=-1}^{\infty}(i+1)p_i=\sum_{j=0}^{\infty}(j+1)p_j=\sum_{j=0}^{\infty}\overline{p}_j=\rho.$$

Используя равенство из [1]

$$\sum_{i=1}^{k}R_i\overline{p}_{k-i}=R_k-1, \tag{11}$$

находим

$$\sum_{i=1}^{k}R_i\sum_{j=0}^{k-i}\overline{p}_j=\sum_{j=0}^{k-1}\sum_{i=1}^{k-j}R_i\overline{p}_{k-j-i}=\sum_{j=0}^{k-1}(R_{k-j}-1)=\sum_{i=1}^{k}R_i-k.$$

Лемма доказана. □

Лемма 2. *Средняя продолжительность периодов занятости систем* $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_1/1/\mathrm{m}$ *и* $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_1/1$ *определяется по формулам*

$$\begin{aligned}\mathbf{M}\,\tau(m)=b_1m_{11}+\big(m_1-m_{11}(1-\rho)\big)\sum_{i=1}^{h}R_i\overline{a}_{h+1-i}-m_{11}\Bigg(\sum_{i=1}^{h}R_i\sum_{j=m+1-i}^{\infty}\overline{p}_j\\ -\sum_{n=1}^{h-1}a_n\sum_{i=1}^{h-n}R_i\sum_{j=m+1-n-i}^{\infty}\overline{p}_j+\sum_{n=m+2}^{\infty}(n-m-1)a_n\Bigg);\end{aligned} \tag{12}$$

$$\mathbf{M}\,\tau(\infty)=b_1m_{11}+\big(m_1-m_{11}(1-\rho)\big)\sum_{i=1}^{h}R_i\overline{a}_{h+1-i}. \tag{13}$$

Доказательство. Для отыскания $\mathbf{M}\,\tau(m)$ используем соотношение (1). Учитывая равенства (10), после несложных преобразований получим

$$(m-h)\overline{p}_{m-i}+\sum_{j=h+1}^{m-1}(j-h)p_{j-i}=\sum_{j=h+1-i}^{\infty}\overline{p}_j-\sum_{j=m+1-i}^{\infty}\overline{p}_j=\rho-\sum_{j=0}^{h-i}\overline{p}_j-\sum_{j=m+1-i}^{\infty}\overline{p}_j;$$

$$\begin{aligned}\sum_{i=1}^{h}R_i\Bigg((m-h)\overline{p}_{m-i}+\sum_{j=h+1}^{m-1}(j-h)p_{j-i}\Bigg)=\sum_{i=1}^{h}R_i\Bigg(\rho-\sum_{j=0}^{h-i}\overline{p}_j-\sum_{j=m+1-i}^{\infty}\overline{p}_j\Bigg)\\ =h+\sum_{i=1}^{h}R_i\Bigg(\rho-1-\sum_{j=m+1-i}^{\infty}\overline{p}_j\Bigg);\end{aligned} \tag{14}$$

$$\sum_{n=h+1}^{m}(n-h)a_n+(m+1-h)\overline{a}_{m+1}=b_1-\sum_{n=1}^{\infty}na_n+\sum_{n=h+1}^{m}na_n-h\overline{a}_{h+1}+(m+1)\overline{a}_{m+1}$$
$$=b_1-\sum_{n=m+2}^{\infty}(n-m-1)a_n-\sum_{n=1}^{h}na_n-h\overline{a}_{h+1}.$$

После аналогичных к выполненным в (14) преобразований выражения

$$\sum_{i=1}^{h-n}R_i\Big((m-h)\overline{p}_{m-n-i}+\sum_{j=h+1}^{m-1}(j-h)p_{j-n-i}\Big)$$

найдём

$$\mathbf{M}\,\tau(m)=\big(m_1-m_{11}(1-\rho)\big)\sum_{i=1}^{h}R_i\overline{a}_{h+1-i}+m_{11}\Big(h-\sum_{n=1}^{h-1}(h-n)a_n$$
$$-\sum_{i=1}^{h}R_i\sum_{j=m+1-i}^{\infty}\overline{p}_j+\sum_{n=1}^{h-1}a_n\sum_{i=1}^{h-n}R_i\sum_{j=m+1-n-i}^{\infty}\overline{p}_j+b_1-\sum_{n=m+2}^{\infty}(n-m-1)a_n-\sum_{n=1}^{h}na_n-h\overline{a}_{h+1}\Big).$$

Отсюда, учитывая, что

$$h-\sum_{n=1}^{h-1}(h-n)a_n=h-\sum_{n=1}^{h}(h-n)a_n=h\overline{a}_{h+1}+\sum_{n=1}^{h}na_n,$$

получим формулу (12), а после перехода в ней к пределу при $m\to\infty$ — соотношение (13). Лемма доказана. □

По такой же схеме доказывается следующее аналогичное утверждение для системы с r переключениями ($r\geq 1$).

Лемма 3. *Средняя продолжительность периодов занятости систем* $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_\mathrm{r}/1$ *и* $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_\mathrm{r}/1/\mathrm{m}$ *определяется по формулам*

$$\mathbf{M}\,\tau(\infty)=b_1m_{r1}+\big(m_1-m_{r1}(1-\rho)\big)\sum_{i=1}^{h}R_i\overline{a}_{h+1-i}$$
$$+\sum_{i=1}^{r-1}(m_{i1}-m_{r1})\Big(\sum_{n=h_i+1}^{h_{i+1}}na_n+h_{i+1}\overline{a}_{h_{i+1}+1}-h_i\overline{a}_{h_i+1}$$
$$+\sum_{j=1}^{h}R_j\sum_{s=h_i+1-j}^{h_{i+1}-j}\overline{p}_s-\sum_{n=1}^{h-1}a_n\sum_{j=1}^{h-n}R_j\sum_{s=h_i+1-n-j}^{h_{i+1}-n-j}\overline{p}_s\Big);$$

$$\mathbf{M}\,\tau(m)=\mathbf{M}\,\tau(\infty)-m_{r1}\Big(\sum_{i=1}^{h}R_i\sum_{j=m+1-i}^{\infty}\overline{p}_j$$
$$-\sum_{n=1}^{h-1}a_n\sum_{i=1}^{h-n}R_i\sum_{j=m+1-n-i}^{\infty}\overline{p}_j+\sum_{n=m+2}^{\infty}(n-m-1)a_n\Big).\qquad(15)$$

Теорема 3. *Для системы обслуживания* $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_{\mathrm{r}}/1/\mathrm{m}$ *средняя продолжительность периода занятости* $\mathbf{M}\,\tau(m)$ *возрастает как функция параметра* m.

Доказательство. Из формулы (15) видим, что утверждение теоремы вытекает из неравенства

$$\sum_{i=1}^{h} R_i \sum_{j=m+1-i}^{\infty} \overline{p}_j - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i \sum_{j=m+1-n-i}^{\infty} \overline{p}_j + \sum_{n=m+2}^{\infty} (n-m-1)a_n > 0.$$

Докажем его. Имеем:

$$\begin{aligned}
&\sum_{i=1}^{h} R_i \sum_{j=m+1-i}^{\infty} \overline{p}_j - \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i \sum_{j=m+1-n-i}^{\infty} \overline{p}_j \\
&= \overline{a}_h \sum_{i=1}^{h} R_i \sum_{j=m+1-i}^{\infty} \overline{p}_j + \sum_{n=1}^{h-1} a_n \left(\sum_{i=1}^{h} R_i \sum_{j=m+1-i}^{\infty} \overline{p}_j - \sum_{i=1}^{h-n} R_i \sum_{j=m+1-n-i}^{\infty} \overline{p}_j \right) \\
&= \overline{a}_h \sum_{i=1}^{h} R_i \sum_{j=m+1-i}^{\infty} \overline{p}_j + \sum_{n=1}^{h-1} a_n \left(\sum_{i=1}^{n} R_i \sum_{j=m+1-i}^{\infty} \overline{p}_j + \sum_{i=1}^{h-n} (R_{n+i} - R_i) \sum_{j=m+1-n-i}^{\infty} \overline{p}_j \right) > 0,
\end{aligned}$$

поскольку $R_{n+1} - R_n > 0$ для всех натуральных значений n. Теорема доказана. □

Теорема 4. *А) Если для системы обслуживания* $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_1/1/\mathrm{m}$ *выполняется условие*

$$m_1 - m_{11}\left(1 - \rho + \sum_{j=m-1}^{\infty} \overline{p}_j\right) \geq 0, \tag{16}$$

то $\mathbf{M}\,\tau(m)$ *возрастает как функция параметра* h *для* $h = \overline{2,\, m-1}$. *Если же выполняется неравенство*

$$m_1 - m_{11}(1-\rho) \leq 0, \tag{17}$$

то $\mathbf{M}\,\tau(m)$ *убывает как функция параметра* h.

Б) Если для системы обслуживания $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_1/1$ *выполняется условие*

$$m_1 - m_{11}(1-\rho) > 0, \tag{18}$$

то $\mathbf{M}\,\tau(\infty)$ *возрастает как функция параметра* h. *Если же выполняется противоположное неравенство (17), то* $\mathbf{M}\,\tau(\infty)$ *не возрастает как функция параметра* h.

Доказательство. А) Обозначим через $\mathbf{M}\,\tau_{h+1}(m)$ значение правой части (12) после замены h на $h+1$. Тогда из (12) получим

$$\mathbf{M}\,\tau_{h+1}(m) - \mathbf{M}\,\tau(m) = \left(m_1 - m_{11}\left(1 - \rho + \sum_{j=m-h}^{\infty} \overline{p}_j\right)\right)\left(R_{h+1} - \sum_{i=1}^{h} R_i a_{h+1-i}\right). \tag{19}$$

Поскольку

$$R_{h+1} - \sum_{i=1}^{h} R_i a_{h+1-i} > 0,$$

то знак разности $\mathbf{M}\,\tau_{h+1}(m) - \mathbf{M}\,\tau(m)$ совпадает со знаком выражения

$$A(m,h) = m_1 - m_{11}\biggl(1-\rho+\sum_{j=m-h}^{\infty}\overline{p}_j\biggr).$$

Из очевидных неравенств

$$m_1 - m_{11}\biggl(1-\rho+\sum_{j=m-1}^{\infty}\overline{p}_j\biggr) < A(m,h) < m_1 - m_{11}(1-\rho)$$

следуют условия (16) и (17). При выполнении (16) имеем $A(m,h) > 0$, а при выполнении (17) $A(m,h) < 0$, поэтому первая часть теоремы доказана.

Б) Из (19) получаем

$$\mathbf{M}\,\tau_{h+1}(\infty) - \mathbf{M}\,\tau(\infty) = \bigl(m_1 - m_{11}(1-\rho)\bigr)\biggl(R_{h+1} - \sum_{i=1}^{h} R_i a_{h+1-i}\biggr).$$

В этом случае знак разности $\mathbf{M}\,\tau_{h+1}(\infty) - \mathbf{M}\,\tau(\infty)$ совпадает со знаком выражения $m_1 - m_{11}(1-\rho)$. Теорема доказана. □

Отметим, что условие (18) выполняется, если $\rho \geq 1$, а также при одновременном выполнении неравенств

$$\rho < 1, \qquad m_{11} < \frac{m_1}{1-\rho}\,.$$

Условие (17) выполняется при одновременном выполнении неравенств

$$\rho < 1, \qquad m_{11} \geq \frac{m_1}{1-\rho}\,.$$

4. ИССЛЕДОВАНИЕ ЗАВИСИМОСТЕЙ ВЕРОЯТНОСТИ ОБСЛУЖИВАНИЯ ОТ ПАРАМЕТРОВ m И h

Теорема 5. *А) Для системы обслуживания* $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_1/1$ *вероятность обслуживания заявок* $\mathbf{P}_{\mathrm{sv}}(\infty)$ *возрастает как функция параметра* h.

Б) Если $a_1 = 1$ *(рассматривается система обслуживания* $\mathrm{M}_{\mathrm{h}}/\mathrm{G}_1/1/\mathrm{m}$*), то при выполнении условия*

$$1 + \lambda R_h(m_1 - m_{11}) > 0 \tag{20}$$

вероятность обслуживания заявок $\mathbf{P}_{\mathrm{sv}}(m)$ *возрастает как функция параметра* m*, и не возрастает при выполнении противоположного к (20) неравенства*

$$1 + \lambda R_h(m_1 - m_{11}) \leq 0. \tag{21}$$

Доказательство. А) Соотношение (9) можно записать в виде

$$\mathbf{P}_{\mathrm{sv}}(\infty) = \frac{1 + \lambda m_1 \sum\limits_{i=1}^{h} R_i \overline{a}_{h+1-i}}{1 + \lambda \mathbf{M}\,\tau(\infty)}\,. \tag{22}$$

Обозначим через $\mathbf{P}_{\mathrm{sv}(h+1)}(\infty)$ значение правой части (22) после замены h на $h+1$. Тогда с помощью (22) получим

$$\big(\mathbf{P}_{\mathrm{sv}(h+1)}(\infty)-\mathbf{P}_{\mathrm{sv}}(\infty)\big)\big(1+\lambda\mathbf{M}\,\tau(\infty)\big)\big(1+\lambda\mathbf{M}\,\tau_{h+1}(\infty)\big)$$
$$=\left(1+\lambda m_1\sum_{i=1}^{h+1}R_i\overline{a}_{h+2-i}\right)\big(1+\lambda\mathbf{M}\,\tau(\infty)\big)-\left(1+\lambda m_1\sum_{i=1}^{h}R_i\overline{a}_{h+1-i}\right)$$
$$\times\big(1+\lambda\mathbf{M}\,\tau_{h+1}(\infty)\big)=(1-\rho)\lambda m_{11}\left(R_{h+1}-\sum_{i=1}^{h}R_ia_{h+1-i}\right)$$
$$+\lambda^2m_1\sum_{i=1}^{h+1}R_i\overline{a}_{h+2-i}\left(b_1m_{11}+\big(m_1-m_{11}(1-\rho)\big)\sum_{i=1}^{h}R_i\overline{a}_{h+1-i}\right)$$
$$-\lambda m_1^2\sum_{i=1}^{h}R_i\overline{a}_{h+1-i}\left(b_1m_{11}+\big(m_1-m_{11}(1-\rho)\big)\sum_{i=1}^{h+1}R_i\overline{a}_{h+2-i}\right)$$
$$=\lambda m_{11}\left(R_{h+1}-\sum_{i=1}^{h}R_ia_{h+1-i}\right)>0,$$

что и доказывает первую часть теоремы.

Б) В случае, когда заявки поступают по одной ($a_1=1$), из соотношений (8) и (12) получим

$$\mathbf{P}_{\mathrm{sv}}(m)=\frac{1+\lambda\left(m_1R_h-R_h\overline{q}_{m+1-h}+\sum\limits_{i=1}^{h-1}R_iq_{m-i}\right)}{1+\lambda\mathbf{M}\,\tau(m)}\,; \tag{23}$$

$$\mathbf{M}\,\tau(m)=\big(m_1-m_{11}(1-\rho)\big)R_h+m_{11}\left(1-R_h\sum_{j=m+1-h}^{\infty}\overline{p}_j+\sum_{i=1}^{h-1}R_i\overline{p}_{m-i}\right). \tag{24}$$

Обозначим числитель правой части формулы (23) через $\mathbf{P}_{\mathrm{num}}(m)$. Если $a_1=1$, то из определений последовательностей q_i, p_i [1] можно вывести следующие соотношения

$$\lambda q_i=\overline{p}_i\qquad(\,i\ge 0\,). \tag{25}$$

Тогда с помощью (23)–(25) получим

$$\mathbf{P}_{\mathrm{num}}(m+1)-\mathbf{P}_{\mathrm{num}}(m)=R_h\overline{p}_{m+1-h}-\sum_{i=1}^{h-1}R_ip_{m-i};$$
$$\mathbf{M}\,\tau(m+1)-\mathbf{M}\,\tau(m)=m_{11}\left(R_h\overline{p}_{m+1-h}-\sum_{i=1}^{h-1}R_ip_{m-i}\right);$$
$$\big(\mathbf{P}_{\mathrm{sv}}(m+1)-\mathbf{P}_{\mathrm{sv}}(m)\big)\big(1+\lambda\mathbf{M}\,\tau(m)\big)\big(1+\lambda\mathbf{M}\,\tau(m+1)\big)$$
$$=\big(\mathbf{P}_{\mathrm{num}}(m+1)-\mathbf{P}_{\mathrm{num}}(m)\big)\big(1+\lambda\mathbf{M}\,\tau(m)\big)-\lambda\mathbf{P}_{\mathrm{num}}(m)\big(\mathbf{M}\,\tau(m+1)$$
$$-\mathbf{M}\,\tau(m)\big)=\left(R_h\overline{p}_{m+1-h}-\sum_{i=1}^{h-1}R_ip_{m-i}\right)\big(1+\lambda\mathbf{M}\,\tau(m)-\lambda m_{11}\mathbf{P}_{\mathrm{num}}(m)\big)$$
$$=\left(R_h\overline{p}_{m+1-h}-\sum_{i=1}^{h-1}R_ip_{m-i}\right)\big(1+\lambda R_h(m_1-m_{11})\big).$$

Поскольку $R_h\overline{p}_{m+1-h} - \sum\limits_{i=1}^{h-1} R_i p_{m-i} > 0$, то знак разности $\mathbf{P}_{\rm sv}(m+1) - \mathbf{P}_{\rm sv}(m)$ совпадает со знаком выражения $1 + \lambda R_h(m_1 - m_{11})$, поэтому характер монотонности $\mathbf{P}_{\rm sv}(m)$ определяется условиями (20) и (21). Теорема доказана. □

5. ОПРЕДЕЛЕНИЕ СТАЦИОНАРНЫХ ХАРАКТЕРИСТИК ОЧЕРЕДИ

Для системы с одним переключением режимов обслуживания ($r = 1$) рассмотрим стационарные характеристики очереди: среднюю длину очереди $\mathbf{M}\,Q(m)$ и среднее время ожидания $\mathbf{M}\,w(m)$. Для системы с ограниченным объёмом накопителя эти характеристики находим по формулам

$$\mathbf{M}\,Q(m) = \sum_{k=1}^{m} k\rho_{k+1}(m); \qquad \mathbf{M}\,w(m) = \frac{\mathbf{M}\,Q(m)}{\lambda b_1 \mathbf{P}_{\rm sv}(m)}\,. \tag{26}$$

Соотношение для $\mathbf{M}\,w(m)$ вытекает из формулы Литтла для систем обслуживания с потерями заявок. Перед использованием формул (26) необходимо предварительно вычислить вероятности $\rho_i(m)$ и $\mathbf{P}_{\rm sv}(m)$ с помощью соотношений (2) и (8) (при $a_1 = 1$) соответственно.

Формула (8) не может быть использована в случае группового поступления заявок ($a_1 < 1$), поскольку для системы с ограниченной очередью некоторые заявки, прибывшие на вход системы в момент, когда входной поток не блокируется, могут быть потеряны. Соответствующую формулу для $\mathbf{P}_{\rm sv}(m)$ можно получить как предел при $T \to \infty$ отношения числа обслуженных заявок к числу прибывших за время T. Среднее число прибывших на вход системы за время T заявок равно $\lambda b_1 T$, а среднее число обслуженных за это же время составляет $\big(1 - \rho_0(m)\big)T/\widetilde{m}_1$, где $\widetilde{m}_1$ — среднее время обслуживания одной заявки, которое можно определить по формуле

$$\frac{1}{\widetilde{m}_1} = \frac{\mathbf{M}\,\tau_1(m)}{m_1 \mathbf{M}\,\tau(m)} + \frac{\mathbf{M}\,\tau_{11}(m)}{m_{11}\mathbf{M}\,\tau(m)}\,.$$

Здесь $\mathbf{M}\,\tau_1(m)$ и $\mathbf{M}\,\tau_{11}(m)$ — средние продолжительности частей периода занятости, соответствующих основному и послепороговому режимам обслуживания соответственно. В результате с помощью соотношения (12) для $\mathbf{M}\,\tau(m)$ получим следующую формулу для вероятности обслуживания

$$\mathbf{P}_{\rm sv}(m) = \frac{1}{b_1\big(1 + \lambda\mathbf{M}\,\tau(m)\big)}\Bigg(b_1 + \rho\sum_{i=1}^{h} R_i \overline{a}_{h+1-i} - \sum_{i=1}^{h} R_i \sum_{j=m+1-i}^{\infty} \overline{p}_j$$
$$+ \sum_{n=1}^{h-1} a_n \sum_{i=1}^{h-n} R_i \sum_{j=m+1-n-i}^{\infty} \overline{p}_j - \sum_{n=m+2}^{\infty} (n - m - 1)a_n\Bigg).$$

В случае системы с неограниченным объёмом накопителя из (26) получим равенства

$$\mathbf{M}\,Q(\infty) = \sum_{k=1}^{\infty} k\rho_{k+1}(\infty); \qquad \mathbf{M}\,w(\infty) = \frac{\mathbf{M}\,Q(\infty)}{\lambda b_1 \mathbf{P}_{\rm sv}(\infty)}\,. \tag{27}$$

Итак, для отыскания $\mathbf{M}\,Q(\infty)$ необходимо найти сумму бесконечного ряда.

Определим $\mathbf{M}\,Q(\infty)$ для случая, когда заявки поступают по одной ($a_1 = 1$), и формулы для эргодического распределения $\rho_i(\infty)$, полученные из (2) при $m \to \infty$, упрощаются:

$$\rho_0(\infty) = \frac{1}{1 + \lambda\mathbf{M}\,\tau(\infty)}\,; \qquad \rho_1(\infty) = \frac{R_1}{1 + \lambda\mathbf{M}\,\tau(\infty)}\,;$$

$$\rho_k(\infty) = \frac{R_k - R_{k-1}}{1 + \lambda \mathbf{M}\,\tau(\infty)} \qquad (\,k = \overline{1,\,h}\,); \qquad \rho_k(\infty) = \frac{\lambda}{1 + \lambda \mathbf{M}\,\tau(\infty)} \Bigg(R_h(q_{k-h} + m_{11}\overline{p}_{k-h}) + \sum_{i=1}^{h-1} R_i (q_{k-i} - q_{k-1-i} - m_{11} p_{k-1-i}) \Bigg) \qquad (\,k > h\,). \tag{28}$$

Теорема 6. *Если $a_1 = 1$ (заявки прибывают по одной, $\rho = \lambda m_1$), то средняя длина очереди для системы* $\mathrm{M_h/G_1/1}$ *определяется по формуле*

$$\mathbf{M}\,Q(\infty) = \frac{1}{1 + \lambda \mathbf{M}\,\tau(\infty)} \Bigg(\sum_{i=1}^{h-1} (R_h - R_i) + (1 + \lambda m_{11}) \Bigg((h-1) \times \big(1 - R_h(1-\rho)\big) + \sum_{k=1}^{\infty} \Bigg(R_h \sum_{i=k}^{\infty} \overline{p}_i - \sum_{i=1}^{h-1} R_i \overline{p}_{h+k-1-i} \Bigg) \Bigg) \Bigg). \tag{29}$$

Доказательство. Используя обозначение $\overline{\rho}_k(\infty) = \sum\limits_{i=k}^{\infty} \rho_i(\infty)$, первую из формул (27) запишем в виде

$$\mathbf{M}\,Q(\infty) = \sum_{k=2}^{\infty} \overline{\rho}_k(\infty). \tag{30}$$

Учитывая первое из равенств (10), а также соотношения (11) и (25), с помощью (28) последовательно находим

$$\overline{\rho}_{h+k}(\infty) = \frac{1 + \lambda m_{11}}{1 + \lambda \mathbf{M}\,\tau(\infty)} \Bigg(R_h \sum_{i=k}^{\infty} \overline{p}_i - \sum_{i=1}^{h-1} R_i \overline{p}_{h+k-1-i} \Bigg) \qquad (\,k \geq 1\,);$$

$$\overline{\rho}_{h+1}(\infty) = \frac{1 + \lambda m_{11}}{1 + \lambda \mathbf{M}\,\tau(\infty)} \big(1 - R_h(1-\rho)\big);$$

$$\overline{\rho}_k(\infty) = \frac{1}{1 + \lambda \mathbf{M}\,\tau(\infty)} \big(R_h - R_{k-1} + (1 + \lambda m_{11})\big(1 - R_h(1-\rho)\big)\big) \quad (\,k = \overline{2,\,h}\,).$$

Подставляя найденные выражения для $\overline{\rho}_k(\infty)$ в формулу (30), получим соотношение (29). Теорема доказана. □

Для некоторых распределений $F(x)$ времени обслуживания основного режима суммы бесконечных рядов в (29) удаётся найти в явном виде. Рассмотрим два таких случая.

Теорема 7. *Пусть $a_1 = 1$ (заявки прибывают по одной). А) В случае показательного распределения времени обслуживания основного режима ($F(x) = 1 - e^{-\mu x}$, $x \geq 0$; $\rho = \lambda m_1 = \lambda/\mu$)*

$$\mathbf{M}\,Q(\infty) = \frac{1}{1 + \lambda \mathbf{M}\,\tau(\infty)} \Bigg(\sum_{i=1}^{h-1} (R_h - R_i) + (1 + \lambda m_{11}) \Bigg((h-1) \times \big(1 - R_h(1-\rho)\big) + R_h \rho^2 - \rho \sum_{i=1}^{h-1} R_i \left(\frac{\rho}{1+\rho} \right)^{h-i} \Bigg) \Bigg). \tag{31}$$

Б) Если время обслуживания основного режима распределено по закону Эрланга второго порядка ($F(x) = 1 - (1+\mu x)e^{-\mu x}$, $x \geq 0$; $\rho = \lambda m_1 = 2\lambda/\mu$), то

$$\mathbf{M}\,Q(\infty) = \frac{1}{1+\lambda\mathbf{M}\,\tau(\infty)}\Bigg(\sum_{i=1}^{h-1}(R_h - R_i) + (1+\lambda m_{11})\bigg((h-1) \times \big(1 - R_h(1-\rho)\big) + \frac{3}{4}R_h\rho^2 - \sum_{i=1}^{h-1}R_i(\rho+h+2-i)\left(\frac{\rho}{2+\rho}\right)^{h+1-i}\bigg)\Bigg). \tag{32}$$

Доказательство. А) Для времени обслуживания, распределённого по показательному закону, из определения последовательности $\{p_i\}$ (см. [1]) получаем

$$p_i = \frac{\mu\Lambda^{i+1}}{\lambda+\mu}, \quad \overline{p}_i = \Lambda^{i+1} \quad (\,i \geq -1\,); \quad \Lambda = \frac{\lambda}{\lambda+\mu}; \quad \sum_{i=k}^{\infty}\overline{p}_i = \rho\Lambda^k; \quad \sum_{k=1}^{\infty}\sum_{i=k}^{\infty}\overline{p}_i = \rho^2,$$

и формула (29) приобретает вид (31).

Б) Если $F(x) = 1 - (1+\mu x)e^{-\mu x}$, $x \geq 0$, то

$$p_i = \frac{(i+2)\mu^2\Lambda^{i+1}}{(\lambda+\mu)^2}, \quad \overline{p}_i = \frac{\big(\lambda + (i+2)\mu\big)\Lambda^{i+1}}{\lambda+\mu} \quad (\,i \geq -1\,);$$

$$\sum_{i=k}^{\infty}\overline{p}_i = \left(\frac{\lambda}{\mu}\right)^2\Lambda^{k-1}\big(k+2-2(k+1)\Lambda + k\Lambda^2\big); \quad \sum_{k=1}^{\infty}\sum_{i=k}^{\infty}\overline{p}_i = \Lambda\sum_{k=1}^{\infty}k\Lambda^k + \rho\sum_{k=1}^{\infty}\Lambda^k = \frac{3}{4}\rho^2,$$

и из (29) получаем соотношение (32). Теорема доказана. □

6. ПРИМЕРЫ ВЫЧИСЛЕНИЯ СТАЦИОНАРНЫХ ХАРАКТЕРИСТИК

Предположим, что время обслуживания основного режима распределено по закону Эрланга второго порядка с параметром μ ($m_1 = 2/\mu$). Для системы $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_1/1$ рассмотрим случаи: $a_1 = 0,75$, $a_2 = 0,25$ (заявки прибывают по одной и по две, **пример 1**); $a_1 = 1$ (заявки поступают по одной, **пример 2**).

Обозначим через $\mathbf{P}_{\mathrm{LS}}(\infty)$ вероятность потери заявки для системы $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_1/1$, тогда

$$\mathbf{P}_{\mathrm{LS}}(\infty) = 1 - \mathbf{P}_{\mathrm{sv}}(\infty). \tag{33}$$

Пусть $\lambda = 2$, $\mu = 3$, а $m_{11} = 0,2$ — среднее время обслуживания, соответствующее функции распределения $F_1(x)$ послепорогового режима, явный вид которой не влияет на результаты расчётов. Результаты вычисления стационарных характеристик системы, полученные для данных примеров 1 и 2 с использованием формул (13), (22), (27), (32) и (33), приведены в таблицах 1 и 2 соответственно. Для сравнения в этих таблицах записаны также значения некоторых характеристик, полученные с помощью системы имитационного моделирования GPSS World [5, 6] для значений времени работы системы обслуживания $t = 10^5$ (табл. 1) и $t = 3\cdot 10^5$ (табл. 2) при равномерном распределении на отрезке $[0;\,0,4]$ времени обслуживания по закону $F_1(x)$. Программы GPSS World, которые использовались для расчётов, приведены в [1]. В них внесены незначительные изменения, обусловленные особенностями поступления заявок (распределение $\{a_i\}$) для примеров 1 и 2.

Полученные результаты подтверждают выводы теорем 4 и 5 о характере монотонной зависимости характеристик $\mathbf{M}\,\tau(\infty)$ и $\mathbf{P}_{\mathrm{sv}}(\infty)$ от параметра h. Возрастание $\mathbf{M}\,\tau(\infty)$ как функции h является следствием выполнения условия (18) теоремы 4. Сравнение данных таблиц 1 и 2 показывает, что средняя продолжительность периода занятости $\mathbf{M}\,\tau(\infty)$ и вероятность потери заявки $\mathbf{P}_{\mathrm{LS}}(\infty)$ уменьшаются при снижении нагрузки на систему (нагрузка больше для примера 1, когда заявки могут прибывать и по две).

Таблица 1. Стационарные характеристики системы $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_1/1$ (пример 1)

h	$\mathbf{M}\,\tau(\infty)$	$\mathbf{P}_{\mathrm{LS}}(\infty)$	$\mathbf{P}_{\mathrm{LS}}(\infty)$ (GPSS)
1	2,472	0,209	0,207
2	5,645	0,187	0,189
3	10,955	0,176	0,177
4	20,387	0,173	0,173
5	37,122	0,170	0,170
6	66,807	0,169	0,169
7	119,457	0,168	0,168

Таблица 2. Стационарные характеристики системы $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_1/1$ (пример 2)

h	$\mathbf{M}\,\tau(\infty)$	$\mathbf{P}_{\mathrm{LS}}(\infty)$	$\mathbf{P}_{\mathrm{LS}}(\infty)$ (GPSS)	$\mathbf{M}\,w(\infty)$	$\mathbf{M}\,w(\infty)$ (GPSS)
1	2,237	0,141	0,140	0,551	0,550
2	4,229	0,120	0,119	0,855	0,853
3	7,190	0,109	0,108	1,226	1,221
4	11,570	0,102	0,102	1,639	1,638
5	18,040	0,098	0,098	2,083	2,087
6	27,595	0,096	0,095	2,550	2,550
7	41,707	0,094	0,094	3,037	3,029

7. О ВОЗМОЖНОСТИ РЕШЕНИЯ ЗАДАЧ ОПТИМАЛЬНОГО СИНТЕЗА

Информацию о характере монотонной зависимости характеристики системы обслуживания от одного из параметров системы можно использовать для решения задачи оптимального синтеза системы с заданной характеристикой. Покажем это на примере вероятности потери заявки $\mathbf{P}_{\mathrm{LS}}(\infty)$ для системы $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_1/1$.

Сформулируем задачу оптимального синтеза, которую будем называть задачей $(h,\ \mathbf{P}_{\mathrm{LS}}(\infty))$: *для фиксированных значений параметров λ, m_1, m_{11} и a_i $(i \geq 1)$ найти такое наименьшее значение порога h, при котором $\mathbf{P}_{\mathrm{LS}}(\infty)$ не превышает заданного значения P_0*. Из теоремы 5 и формулы (33) следует, что $\mathbf{P}_{\mathrm{LS}}(\infty)$ убывает как функция параметра h. Поэтому, очевидно, что решение сформулированной задачи определяется с помощью алгоритма

$$h^* = \min\{\, h \in \mathbb{N} :\ \mathbf{P}_{\mathrm{LS}(h)}(\infty) \leq P_0 \,\}, \tag{34}$$

где $\mathbf{P}_{\mathrm{LS}(h)}(\infty) = \mathbf{P}_{\mathrm{LS}}(\infty)$ для конкретного значения h. Для реализации алгоритма (34) достаточно иметь результаты вычислений значений $\mathbf{P}_{\mathrm{LS}}(\infty)$ для различных h, такие как, например, приведённые в таблицах 1 и 2.

Решения задач $(h,\ \mathbf{P}_{\mathrm{LS}}(\infty))$ для данных примеров 1 и 2, полученные при помощи таблиц 1 и 2, приведены в таблицах 3 и 4 соответственно.

Таблица 3. Решения задачи оптимального синтеза $(h,\ \mathbf{P}_{\mathrm{LS}}(\infty))$ для данных примера 1

P_0	0,175	0,180	0,185	0,190	0,195	0,200	0,205	0,210
h^*	4	3	3	2	2	2	2	1

Рассмотрим задачи оптимального синтеза для системы $\mathrm{M}_\mathrm{h}^\theta/\mathrm{G}_1/1$, решения которых можно найти в явном виде. Первую из них будем называть задачей $(m_{11},\ \mathbf{P}_{\mathrm{LS}}(\infty))$ и сформулируем так: *для фиксированных значений параметров λ, m_1, h и a_i $(i \geq 1)$ найти такое наибольшее значение среднего времени обслуживания m_{11} послепорогового режима, при котором вероятность потери заявки $\mathbf{P}_{\mathrm{LS}}(\infty)$ не превышает заданного значения P_0.*

Таблица 4. Решения задачи оптимального синтеза $(h,\ \mathbf{P}_{\text{LS}}(\infty))$ для данных примера 2

P_0	0,095	0,100	0,105	0,110	0,115	0,120	0,130	0,140
h^*	7	5	4	3	3	2	2	2

Теорема 8. *Если* $0 < P_0 < 1$, *то решение задачи* $(m_{11},\ \mathbf{P}_{LS}(\infty))$ *для системы обслуживания* $\mathrm{M}_{\mathrm{h}}^{\theta}/\mathrm{G}_1/1$ *определяется в виде*

$$m_{11}^* = \frac{P_0\left(1 + \lambda m_1 \sum\limits_{i=1}^{h} R_i \overline{a}_{h+1-i}\right)}{\lambda(1-P_0)\left(b_1 - (1-\rho)\sum\limits_{i=1}^{h} R_i \overline{a}_{h+1-i}\right)}. \tag{35}$$

Доказательство. Пользуясь соотношениями (22) и (13), записываем неравенство $\mathbf{P}_{\text{LS}}(\infty) \le P_0$ и решаем его относительно m_{11}. Получаем

$$m_{11} \le \frac{P_0\left(1 + \lambda m_1 \sum\limits_{i=1}^{h} R_i \overline{a}_{h+1-i}\right)}{\lambda(1-P_0)\left(b_1 - (1-\rho)\sum\limits_{i=1}^{h} R_i \overline{a}_{h+1-i}\right)}.$$

Поскольку

$$b_1 - (1-\rho)\sum_{i=1}^{h} R_i \overline{a}_{h+1-i} = \frac{\mathbf{M}\,\tau_{\mathrm{b}}(\infty)}{m_{11}} > 0,$$

то теорема доказана. □

Рассмотрим задачу оптимального синтеза $(m_{11},\ \mathbf{M}\,w(\infty))$ для случая, когда заявки прибывают по одной ($a_1 = 1$): *для фиксированных значений параметров* λ, m_1 *и* h *найти такое наибольшее значение среднего времени обслуживания* m_{11} *послепорогового режима, при котором среднее время ожидания в очереди* $\mathbf{M}\,w(\infty)$ *не превышает заданного значения* w_0.

Теорема 9. *Если*

$$w_0 > \frac{\sum\limits_{i=1}^{h-1}(R_h - R_i) + B(\lambda,\ m_1,\ h)}{\lambda(1+\rho R_h)}, \tag{36}$$

где $B(\lambda,\ m_1,\ h) = (h-1)\big(1 - R_h(1-\rho)\big) + \sum\limits_{k=1}^{\infty}\left(R_h \sum\limits_{i=k}^{\infty} \overline{p}_i - \sum\limits_{i=1}^{h-1} R_i \overline{p}_{h+k-1-i}\right)$, *то решение задачи* $(m_{11},\ \mathbf{M}\,w(\infty))$ *определяется в виде*

$$m_{11}^* = \frac{\lambda w_0(1+\rho R_h) - \sum\limits_{i=1}^{h-1}(R_h - R_i) - B(\lambda,\ m_1,\ h)}{\lambda B(\lambda,\ m_1,\ h)}. \tag{37}$$

Доказательство. Используя соотношения (27), (29) и (22), записываем неравенство $\mathbf{M}\,w(\infty) \le w_0$ и решаем его относительно m_{11}. Получаем

$$m_{11} \le \frac{\lambda w_0(1+\rho R_h) - \sum\limits_{i=1}^{h-1}(R_h - R_i) - B(\lambda,\ m_1,\ h)}{\lambda B(\lambda,\ m_1,\ h)}. \tag{38}$$

Поскольку $B(\lambda,\, m_1,\, h) > 0$ (см. доказательство теоремы 6), и при выполнении условия (36) числитель правой части неравенства (38) положителен, то теорема доказана. □

Решения задач оптимального синтеза $(m_{11},\, \mathbf{P}_{\mathrm{LS}}(\infty))$ и $(m_{11},\, \mathbf{M}\, w(\infty))$, полученные по формулам (35) и (37) для данных примеров 1 и 2 соответственно, приведены в таблицах 5 и 6.

Таблица 5. Решения задачи оптимального синтеза $(m_{11},\, \mathbf{P}_{\mathrm{LS}}(\infty))$ для данных примера 1 ($h = 5$)

P_0	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45
m_{11}^*	0,051	0,109	0,172	0,244	0,326	0,419	0,526	0,651	0,799

Таблица 6. Решения задачи оптимального синтеза $(m_{11},\, \mathbf{M}\, w(\infty))$ для данных примера 2 ($h = 5$)

w_0	1,8	1,9	2,0	2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9
m_{11}^*	0,024	0,086	0,149	0,211	0,273	0,335	0,397	0,459	0,522	0,584	0,646	0,708

Используя условие (36), находим, что для данных примера 2 решение задачи $(m_{11},\, \mathbf{M}\, w(\infty))$ существует, если $w_0 > 1,761$.

8. ЗАДАЧА МИНИМИЗАЦИИ ЗАТРАТ НА ОБСЛУЖИВАНИЕ

Рассмотрим экономический функционал качества функционирования системы обслуживания $\mathrm{M_h/G_1/1}$ в виде

$$I(m_{11}) = c_w \mathbf{M}\, w(\infty) + c_{\mathrm{sv}} T(\infty) + c_{\mathrm{sv},1} T_1(\infty), \tag{39}$$

где c_w — стоимость единицы времени ожидания в очереди; $T(\infty)$ и $T_1(\infty)$ — среднее относительное время использования основного и послепорогового режима соответственно; c_{sv} и $c_{\mathrm{sv},1}$ — стоимость единицы времени использования каждого из этих режимов. Зафиксируем среднее время обслуживания m_1 основного режима и предположим, что $c_{\mathrm{sv},1} = m_1 c_{\mathrm{sv}}/m_{11}$. Считая заданными параметры ρ и h, будем искать такое значение среднего времени обслуживания послепорогового режима m_{11}, чтобы минимизировать средние затраты на обслуживание $I(m_{11})$.

После подстановки выражений для $\mathbf{M}\, w(\infty)$, $T(\infty)$ и $T_1(\infty)$ в правую часть (39) получим явную зависимость функционала $I(m_{11})$ от m_{11}. Таким образом, задача сведена к минимизации функции

$$I(m_{11}) = \frac{c_w}{\lambda(1+\rho R_h)} \left(\sum_{i=1}^{h-1} (R_h - R_i) + (1 + \lambda m_{11}) B(\lambda,\, m_1,\, h) \right) + \frac{c_{\mathrm{sv}} \rho (1 + \rho R_h)}{1 + \lambda\big(m_1 R_h + m_{11}\big(1 - (1-\rho) R_h\big)\big)}$$

для всех положительных m_{11}.

Теорема 10. *Если выполнено условие*

$$\frac{c_{\mathrm{sv}}}{c_w} > \frac{B(\lambda,\, m_1,\, h)}{\lambda^2 m_1 \big(1 - (1-\rho) R_h\big)}\,, \tag{40}$$

то минимум функционала $I(m_{11})$ для системы обслуживания $\mathrm{M_h/G_1/1}$ *существует и достигается при*

$$m_{11}^* = \frac{1+\rho R_h}{\lambda\bigl(1-(1-\rho)R_h\bigr)}\left(\lambda\sqrt{\frac{c_{\mathrm{sv}} m_1\bigl(1-(1-\rho)R_h\bigr)}{c_w B(\lambda,\, m_1,\, h)}}-1\right). \tag{41}$$

Доказательство. Производная функции $I(m_{11})$ записывается в виде

$$I'(m_{11}) = A_1 - \frac{A_2}{(A_3+A_4 m_{11})^2}\,; \qquad A_1 = \frac{c_w B(\lambda,\, m_1,\, h)}{1+\rho R_h}\,;$$
$$A_2 = c_{\mathrm{sv}} m_1 \lambda^2 (1+\rho R_h)\bigl(1-(1-\rho)R_h\bigr); \qquad A_3 = 1+\rho R_h; \qquad A_4 = \lambda\bigl(1-(1-\rho)R_h\bigr),$$

где постоянные A_i $(i=\overline{1,\,4})$ положительны. При условии (40) единственное положительное решение уравнения $I'(m_{11}) = 0$ имеет вид (41). Функция $y = y(m_{11}) = I'(m_{11})$ монотонно возрастает, поэтому, поскольку $y(m_{11}^*) = 0$, то $y(m_{11}^* - \varepsilon) < 0$, $y(m_{11}^* + \varepsilon) > 0$ для как угодно малого $\varepsilon > 0$, а это значит, что в точке $m_{11} = m_{11}^*$ функция $I(m_{11})$ достигает минимума. Теорема доказана. □

Проиллюстрируем применение теоремы 10 для системы обслуживания $\mathrm{M_h/G_1/1}$, используя данные примера 2, рассмотренного выше. Предположим, что заявки прибывают по одной, время обслуживания основного режима распределено по закону Эрланга второго порядка с параметром μ $(m_1 = 2/\mu)$, $\lambda = 2$, $\mu = 3$, $h = 5$, , а функция распределения $F_1(x)$ времени обслуживания послепорогового режима произвольна. Условие (40) накладывает такое ограничение на значения стоимостей обслуживания: $c_{\mathrm{sv}}/c_w > 2{,}2136$. Оно выполняется, например, если $c_w = 1$, $c_{\mathrm{sv}} = 3$. Для этих значений по формуле (41) находим значение среднего времени обслуживания послепорогового режима, при котором затраты на обслуживание минимальны, $m_{11}^* = 0{,}3013$. С помощью таблицы 7 убеждаемся, что минимум функционала $I(m_{11})$ достигается при $m_{11} = m_{11}^*$.

Таблица 7. Затраты на обслуживание при различных значениях m_{11} для данных примера 2 ($h = 5$, $c_w = 1$, $c_{\mathrm{sv}} = 3$)

m_{11}	0,1	0,2	0,25	**0,3013**	0,35	0,4	0,5	0,6
$I(m_{11})$	5,7152	5,6897	5,6836	**5,6817**	5,6833	5,6886	5,7088	5,7405

9. ЗАКЛЮЧЕНИЕ

В отличие от системы обслуживания $\mathrm{M^\theta/G/1}$, для которой в условиях большой загрузки ($\rho \geq 1$) очередь с течением времени неограниченно растёт, для системы $\mathrm{M_h^\theta/G_r/1}$ стационарный режим существует при любой загрузке. С помощью математического аппарата, базирующегося, в основном, на методе потенциала В. С. Королюка, в настоящей работе определены стационарные характеристики систем обслуживания $\mathrm{M_h^\theta/G_1/1/m}$ и $\mathrm{M_h^\theta/G_1/1}$ и изучен характер зависимостей некоторых из них от параметров m и h. Решены задачи оптимального синтеза систем с заданными характеристиками, в которых в качестве критерия наилучшего функционирования системы выбиралось заданное значение одной из стационарных характеристик, а также задача минимизации затрат на обслуживание за счёт выбора оптимальной интенсивности обслуживания послепорогового режима.

СПИСОК ЛИТЕРАТУРЫ

1. Жерновый К. Ю. Исследование системы $M^{\theta}/G/1/m$ с переключениями режимов обслуживания и пороговой блокировкой потока заявок. *Информационные процессы*, 2010, т. 10, № 2, стр. 159-180.
2. Жерновий К. Оптимізація режимів обслуговування для систем M/M/1/m та M/M/1 з блокуванням вхідного потоку. *Вісник Львівського університету. Сер. мех-мат.*, 2009, вип. 71, стр. 92-101.
3. Жерновый Ю. В. Решение задач оптимального синтеза для некоторых марковских моделей обслуживания. *Информационные процессы*, 2010, т. 10, № 3, стр. 257-274.
4. Братійчук А. М. *Дослідження систем обслуговування з обмеженою чергою.* Кандидатська дисертація, Київ: Київський національний університет імені Т. Шевченка, 2008.
5. Боев В. Д. *Моделирование систем. Инструментальные средства GPSS World.* Санкт-Петербург: БХВ-Петербург, 2004.
6. Жерновий Ю. В. *Імітаційне моделювання систем масового обслуговування.* Львів: Видавничий центр ЛНУ імені Івана Франка, 2007.