

ОПТИМАЛЬНОЕ МНОГОПороГОВОЕ УПРАВЛЕНИЕ ПОТОКОМ В СИСТЕМЕ МАССОВОГО ОБСЛУЖИВАНИЯ G/MSP/1 С МАР-ПОТОКОМ СБОЕВ

О.В. СЕМЕНОВА

Рассмотрена система массового обслуживания G/MSP/1 с марковским потоком катастрофических сбоев и управляемым потоком запросов. Найдено стационарное распределение вероятностей состояний вложенной цепи Маркова. Разработан алгоритм нахождения оптимальной многопороговой стратегии управления потоком запросов. Приведен численный пример.

ВВЕДЕНИЕ

Необходимость адекватного описания случайных процессов в современных информационных сетях определяет интерес исследователей к моделям систем массового обслуживания, учитывающих особенности потоков информации в этих сетях: разноприоритетность и требование обеспечения высокого качества обработки приоритетных потоков; ненадежность работы сети и возможность потери части информации при сбоях в системе.

Одной из таких моделей является модель управляемой системы массового обслуживания с потоком катастрофических сбоев, приводящих к потере всех запросов в системе, включая обслуживаемый запрос. Частный случай таких систем — системы с несколькими режимами входящего потока запросов. Характерным для них является то, что режимы, в которых интенсивности поступления запросов ниже, более дорогие, чем режимы, запросы в которых поступают интенсивнее. Выбор режима входящего потока происходит в соответствии с некоторой стратегией (например, одно-, многопороговой или гистерезисной) с целью минимизации экономического критерия качества, оценивающего эффективность работы системы [1, 5].

Сбои, происходящие в реальных системах массового обслуживания, в том числе и сетях связи, нарушают работу систем и, в частности, приводят к потере нескольких или всех запросов. Сбои, вызывающие потерю всех запросов в системе (disasters), являются важным частным случаем так называемого отрицательного запроса. Это понятие в 1991 г. ввел Э. Геленбе [6]. Список работ по исследованию систем с отрицательными запросами и сбоями можно найти в [6–10]. Для этих систем экономический критерий качества содержит штраф за потерю запросов в единицу времени, а динамическое управление системой (в частности, переключение на режим с менее интенсивным потоком запросов) позволяет уменьшить длину очереди в системе, что снижает расходы от потерь запросов из-за поступления сбоя.

Марковский процесс обслуживания (MSP) является обобщением обслуживания фазового типа. Система G/MSP/1/r с конечным и бесконечным буферами исследована в [11].

В данной работе рассмотрена система с многорежимным потоком запросов, марковским обслуживанием и потоком катастрофических сбоев, приводящих к мгновенному уходу всех запросов из системы. Для управления потоком запросов используется многопороговая стратегия.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Рассмотрим однолинейную систему массового обслуживания с неограниченным буфером и потоком катастрофических сбоев.

Входящий поток запросов имеет n режимов, $n \geq 2$. Интервалы времени между моментами поступления запросов при работе системы в k -м режиме имеют функцию распределения $A^{(k)}(t)$, $k = \overline{1, n}$.

Интенсивность потока запросов в k -м режиме λ_k определяется следующим образом:

$$\lambda_k^{-1} = \int_0^{\infty} t dA^{(k)}(t), \quad k = \overline{1, n}.$$

Процесс поступления сбоев — МАР-поток (Markov Arrival Process), задаваемый цепью Маркова η_t , $t \geq 0$, с непрерывным временем, пространством состояний $\{0, \dots, N\}$ и матричной производящей функцией

$$F(z) = F_0 + zF_1, \quad |z| \leq 1.$$

Обозначим $\vec{\varphi}$ вектор стационарного распределения цепи Маркова η_t , $t \geq 0$. Вектор-строка $\vec{\varphi}$ — единственное решение системы уравнений

$$\vec{\varphi}F(1) = \vec{0}, \quad \vec{\varphi}\vec{e} = 1,$$

где $\vec{0}$ — вектор-строка соответствующей размерности, состоящая из нулей; \vec{e} — вектор-столбец, состоящий из единиц.

Средняя интенсивность МАР-потока сбоев δ подсчитывается по формуле

$$\delta = \vec{\varphi} \left. \frac{dF(z)}{dz} \right|_{z=1} 1.$$

Подробное описание МАР-потока как частного случая ВМАР-потока дано в работе Лукантони [12]. Переходы цепи η_t , $t \geq 0$, не вызывающие поступления сбоя, управляются матрицей F_0 , а переходы, сопровождающиеся поступлением сбоя в систему, — матрицей F_1 . Так же, как Джейн и Сигман [8], полагаем, что приход сбоя в систему вызывает немедленный уход из системы всех запросов, включая обслуживаемый запрос. Считаем также, что сбои, поступающие в пустую систему, не накапливаются в ней и не оказывают никакого влияния на ее последующую работу.

Процесс обслуживания запросов — MSP (Markovian Service Process) — управляется цепью Маркова m_t , $t \geq 0$, с непрерывным временем, пространством состояний $\{0, \dots, M\}$ и матричной производящей функцией

$$B(z) = B_0 + B_1z, \quad |z| \leq 1.$$

Описание MSP-процесса обслуживания запросов приведено в работе [11]. Переходы цепи $m_t, t \geq 0$, не приводящие к завершению обслуживания, управляются матрицей B_0 , а переходы, приводящие к завершению обслуживания, — матрицей B_1 . Полагаем, что цепь Маркова $m_t, t \geq 0$, «замораживает» свое состояние в момент окончания периода занятости системы при успешном завершении обслуживания запроса и меняет свое состояние в соответствии со стохастической матрицей P в момент поступления сбоя.

Работа системы оценивается экономическим критерием качества

$$C = C(j_1, \dots, j_{n-1}) = aU + \sum_{k=1}^n c_k Y^{(k)} + gV, \quad (1)$$

где U — среднее время ожидания в системе; $Y^{(k)}$ — средняя доля использования k -го режима потока запросов, $k = \overline{1, n}$; V — среднее число запросов, потерянных в единицу времени; $a, c_k, k = \overline{1, n}, g$ — стоимостные коэффициенты, причем $a > 0, c_1 \leq c_2 \leq \dots \leq c_n, g \geq 0$.

Поток запросов может изменять свой режим только в моменты их поступления. Для управления входящим потоком используется многопороговая стратегия, определяемая таким образом.

Фиксируется набор порогов $(j_1, j_2, \dots, j_{n-1})$, причем $-1 = j_0 < j_1 \leq j_2 \leq \dots \leq j_{n-1} < j_n = \infty$. Если число запросов в системе в данный момент прихода запроса удовлетворяет неравенству $j_{k-1} + 1 \leq i \leq j_k$, то запросы будут поступать в соответствии с k -м режимом, $k = \overline{1, n}$.

В данной работе предложен алгоритм нахождения оптимальной многопороговой стратегии управления, состоящий в следующем: фиксируем набор порогов и находим стационарное распределение вероятностей вложенной цепи Маркова, описывающей поведение системы. Далее вычисляем значение критерия качества при заданном наборе порогов и находим оптимальный их набор, минимизируя значение критерия качества.

СТАЦИОНАРНОЕ РАСПРЕДЕЛЕНИЕ ВЕРОЯТНОСТЕЙ ВЛОЖЕННОЙ ЦЕПИ МАРКОВА

Пусть $(j_1, j_2, \dots, j_{n-1})$ — фиксированный набор порогов и t_k есть k -й момент поступления запроса в систему, $k \geq 1$. Рассмотрим случайный процесс $\xi_k = \{i_k, m_k, \eta_k\}$, $k \geq 1$, где i_k — число запросов в системе в момент $t_k - 0$, $i_k \geq 0$; m_k — состояние процесса обслуживания m_t в момент $t_k - 0$, $m_k = \overline{0, N}$; η_k — состояние процесса поступления сбоев η_t в момент t_k , $\eta_k = \overline{0, N}$. Случайный процесс $\xi_k, k \geq 1$ является цепью Маркова. Занулируем состояния этого процесса в лексикографическом порядке и введем стационарные вероятности

$$\pi(i, m, \eta, \nu) = \lim_{k \rightarrow \infty} P\{i_k = i, m_k = m, \eta_k = \eta, \nu_k = \nu\},$$

$$i \geq 0, m = \overline{0, M}, \eta = \overline{0, N}, \nu = \overline{1, W}$$

и векторы

$$\begin{aligned}\bar{\pi}(i, m) &= (\bar{\pi}(i, m, 0), \dots, \bar{\pi}(i, m, N)), \\ \bar{\pi}_i &= (\bar{\pi}(i, 0), \dots, \bar{\pi}(i, W)).\end{aligned}$$

Обозначим $\Phi_i^{(k)}$ матрицу, элементы которой имеют следующий вероятностный смысл: при работе системы в k -м режиме за время между моментами поступления запросов будут обслужены ровно i запросов, в систему не поступит сбой, и случайный процесс $\{m_k, \eta_k, \nu_k\}$ перейдет из состояния $\{m, \eta\}$ в состояние $\{m', \eta'\}$ при условии, что в начальный момент времени в системе было не менее i запросов, $i \geq 0, k = \overline{1, n}$. Матрицы $\Phi_i^{(k)}, i \geq 0$ задаются формулой

$$\Phi_i^{(k)} = \int_0^\infty Q(i, t) \otimes e^{F_0 t} dA^{(k)}(t), \quad i \geq 0, k = \overline{1, n},$$

где \otimes — символ кронекерова произведения (определение и свойства см., например, в работе [13, с. 318]); $Q(i, t)$ — матрица вероятностей переходов процесса $m_t, t \geq 0$ за время x , при котором будут обслужены ровно i запросов, $i \geq 0$. Матрицы $\Phi_i^{(k)}, i \geq 0$ определяются как коэффициенты матричного разложения

$$\sum_{i=0}^\infty \Phi_i^{(k)} z^i = \int_0^\infty e^{B(z)x} \otimes e^{F_0 x} dA^{(k)}(x), \quad k = \overline{1, n}. \quad (2)$$

Обозначим $\Omega_i^{(k)}$ матрицу, элементы которой имеют такой вероятностный смысл: при работе системы в k -м режиме за время между моментами поступления запросов систему покинет ровно $i+1$ запрос, и случайный процесс $\{m_k, \eta_k\}$ перейдет из состояния $\{m, \eta\}$ в состояние $\{m', \eta'\}$ при условии, что в начальный момент времени в системе было ровно i запросов, $i \geq 0, k = \overline{1, n}$. Матрицы $\Omega_i^{(k)}, i \geq 0$, определяются как коэффициенты матричного разложения

$$\begin{aligned}\sum_{i=0}^\infty \Omega_i^{(k)} z^i (1-z) &= \\ &= \int_0^\infty \int_0^t e^{B(z)x} \otimes e^{F_0 x} e^{F(1)(t-x)} dx dA^{(k)}(t) (B_1 + B_0 P + z B_1 (P - I_{M+1})) \otimes I_{N+1} - \\ &\quad - \int_0^\infty e^{B(z)t} P \otimes e^{F_0 t} dA^{(k)}(t) + \int_0^\infty P \otimes e^{F(1)t} dA^{(k)}(t), \quad k = \overline{1, n}.\end{aligned} \quad (3)$$

Из равенства (3) получаем формулы для вычисления матриц $\Omega_i^{(k)}, i \geq 0, k = \overline{1, n}$

$$\begin{aligned}\Omega_i^{(k)} &= \sum_{l=0}^i \Gamma_l^{(k)} (B(1)P \otimes I_{N+1}) + \Gamma_i^{(k)} (B_1(I_{M+1} - P) \otimes I_{N+1}) + \\ &\quad + \left[H^{(k)} - \sum_{l=0}^i \Phi_l^{(k)} \right] (P \otimes I_{N+1}), \quad i \geq 0, k = \overline{1, n},\end{aligned}$$

где

$$H^{(k)} = \int_0^\infty e^{(I_{M+1} \otimes F(1))t} dA^{(k)}(t), \quad k = \overline{1, n},$$

а матрицы $\Gamma_i^{(k)}$, $i \geq 0$, $k = \overline{1, n}$ — коэффициенты матричного разложения

$$\sum_{i=0}^{\infty} \Gamma_i^{(k)} z^i = \int_0^{\infty} \int_0^t e^{B(z)x} \otimes e^{F_0 x} e^{F(1)(t-x)} dx dA^{(k)}(t) \quad (4)$$

и подсчитываются с помощью рекуррентной процедуры

$$\Gamma_0^{(k)} = [B_0 \oplus (-F_1)]^{-1} (\Phi_0^{(k)} - H^{(k)}),$$

$$\Gamma_i^{(k)} = [B_0 \oplus (-F_1)]^{-1} (\Phi_i^{(k)} - B_1 \otimes I_{N+1} \Gamma_{i-1}^{(k)}), \quad i \geq 1, k = \overline{1, n}.$$

Формула полной вероятности и определения многопороговой стратегии дает следующий результат.

Лемма 1. Векторы стационарных вероятностей $\bar{\pi}_i$, $i \geq 0$ удовлетворяют уравнениям равновесия

$$\bar{\pi}_0 = \sum_{v=0}^{n-1} \sum_{l=j_v+1}^{j_{v+1}} \bar{\pi}_l \Omega_l^{(v+1)}, \quad (5)$$

$$\bar{\pi}_i = \sum_{l=i-1}^{j_t} \bar{\pi}_l \Phi_{l-i+1}^{(t)} + \sum_{v=t}^{n-1} \sum_{l=j_v+1}^{j_{v+1}} \bar{\pi}_l \Phi_{l-i+1}^{(v+1)}, \quad j_{t-1} + 1 \leq i \leq j_t, \quad t = \overline{1, n}. \quad (6)$$

Теорема 1. Векторы стационарных вероятностей $\bar{\pi}_i$, $i \geq 0$ вычисляются таким образом:

$$\bar{\pi}_i = \bar{c} \Pi_i^{(t)}, \quad j_{t-1} + 1 \leq i \leq j_t, \quad t = \overline{1, n-1}, \quad j_0 = -1, \quad (7)$$

$$\bar{\pi}_i = \bar{c} R^i, \quad i > j_{n-1}, \quad (8)$$

где

$$\Pi_i^{(t)} = \sum_{r_i=i}^{j_t} \sum_{r_{t+1}=r_t}^{j_{t+1}} \dots \sum_{r_{n-1}=r_{n-2}}^{j_{n-1}} R^{r_{n-1}} \Lambda_{r_{n-1}-r_{n-2}}^{(n-1)} \Lambda_{r_{n-2}-r_{n-3}}^{(n-2)} \dots \Lambda_{r_t-i}^{(t)} ;$$

матрица R — единственное решение уравнения

$$R = \sum_{l=0}^{\infty} R^l \Phi_l^{(n)} \quad (9)$$

во множестве неотрицательных матриц спектрального радиуса не больше 1; матрицы $\Lambda_l^{(v)}$, $l \geq 0$, вычисляются рекуррентно

$$\Lambda_0^{(v)} = \Phi_0^{(v+1)} (\Phi_0^{(v)})^{-1},$$

$$\Lambda_1^{(v)} = (\Lambda_0^{(v)} - \Lambda_0^{(v)} \Phi_1^{(v)} - I + \Phi_1^{(v+1)}) (\Phi_0^{(v)})^{-1}, \quad (10)$$

$$\Lambda_l^{(v)} = \left(\Lambda_{l-1}^{(v)} - \sum_{i=0}^{l-1} \Lambda_i^{(v)} \Phi_{l-i}^{(v)} + \Phi_l^{(v+1)} \right) (\Phi_0^{(v)})^{-1}, \quad l \geq 2, \quad v = \overline{1, n-1},$$

а вектор \bar{c} — решение системы

$$\bar{c} \left[-\Pi_0^{(1)} + \sum_{v=0}^{n-2} \sum_{l=j_v+1}^{j_{v+1}} \Pi_l^{(v+1)} \Omega_l^{(v+1)} + R - \sum_{l=0}^{j_{n-1}} R^l \Phi_l^{(n)} \right] = \bar{0}, \quad (11)$$

$$\bar{c} \left[\sum_{v=1}^{n-1} \sum_{i=j_{v-1}+1}^{j_v} \Pi_i^{(v)} + R^{j_{n-1}+1} (I - R)^{-1} \right] \bar{e} = 1. \quad (12)$$

Доказательство. Формула (8) проверяется непосредственной подстановкой в уравнение (6) для $i > j_{n-1} + 1$. Исследование матричного уравнения (9) приведено в работе [14]. Справедливость формул (7) можно проверить, подставив их в соотношения (6) с учетом (10).

Систему (11) для вектора \bar{c} получаем, подставляя в равенство (5) выражения (7), (8) для векторов $\bar{\pi}_i, i \geq 0$. В силу однородности этой системы для нахождения вектора \bar{c} также используем условие нормировки (12). Теорема доказана.

Замечание. Формулы (5) – (12) представляют собой алгоритм для вычисления стационарного распределения вложенной цепи Маркова для любого фиксированного набора порогов $(j_1, j_2, \dots, j_{n-1})$.

ЗАВИСИМОСТЬ ЗНАЧЕНИЯ КРИТЕРИЯ КАЧЕСТВА ОТ ВЕЛИЧИНЫ ПОРОГОВ

С помощью эргодических теорем для цепей Маркова [15, с. 11] получаем следующий результат.

Лемма 2. Среднее время τ между моментами поступления запросов

$$\tau = \bar{c} \left[\sum_{k=1}^{n-1} \lambda_k^{-1} \sum_{i=j_{k-1}+1}^{j_k} \Pi_i^{(k)} + \lambda_n^{-1} R^{j_{n-1}+1} (I - R)^{-1} \right] \bar{e}.$$

Теорема 2. Преобразование Лапласа–Стилтьеса $w(s)$ распределения времени ожидания в системе определяется равенством

$$w(s) = \bar{c} \left(\Pi_0^{(1)} + \sum_{l=1}^{j_1} \Pi_l^{(1)} T_l(s) + \sum_{v=1}^{n-2} \sum_{l=j_v+1}^{j_{v+1}} \Pi_l^{(v+1)} T_l(s) + \sum_{l=j_{n-1}+1}^{\infty} R^l T_l(s) \right) \bar{e}, \quad \text{Re } s > 0,$$

где

$$T_l(s) = G_{l-1}(s)(B_1 \otimes I_{N+1}) + \sum_{k=0}^l G_k(s)(P \otimes F_1), \quad l \geq 1,$$

а матрицы $G_i(s), i \geq 0$ — коэффициенты матричного разложения

$$\sum_{i=0}^{\infty} G_i(s) z^i = -[B(z) \oplus (F_0 - sI_{N+1})]^{-1}.$$

Следствие. Среднее время ожидания в системе

$$U = \bar{c} \left(\sum_{l=1}^{j_1} \Pi_l^{(1)} S_l + \sum_{v=1}^{n-2} \sum_{l=j_v+1}^{j_{v+1}} \Pi_l^{(v+1)} S_l + \sum_{l=j_{n-1}+1}^{\infty} R^l S_l \right) \bar{e}, \quad (13)$$

где

$$S_l = \Delta_{l-1}(B_1 \otimes I_{N+1}) + \sum_{k=0}^l \Delta_k(P \otimes F_1), \quad l \geq 1,$$

$$\Delta_0 = L^2, \quad \Delta_i = \Delta_{i-1}(B_1 \otimes I_{N+1})L + L[(B_1 \otimes I_{N+1})L]^i L, \quad i \geq 1,$$

$$L = -[B_0 \oplus F_0]^{-1}.$$

Теорема 3. Средняя доля $Y^{(k)}$ использования k -го режима ($k = \overline{1, n}$) определяется как

$$Y^{(k)} = \lambda_k^{-1} \bar{c} \sum_{i=j_{k-1}+1}^{j_k} \Pi_i^{(k)} \bar{e} \tau^{-1}, \quad k = \overline{1, n-1},$$

$$Y^{(n)} = \lambda_n^{-1} \bar{c} R^{j_{n-1}+1} (I - R)^{-1} \bar{e} \tau^{-1}.$$
(14)

Теорема 4. Среднее число V запросов, потерянных в единицу времени, вычисляется с помощью формулы

$$V = \tau^{-1} \bar{c} \sum_{i=1}^{\infty} i \left(\sum_{l=i-1}^{j_i} \Pi_l^{(i)} D_{l-i+1}^{(i)} + \sum_{v=i}^{n-2} \sum_{l=j_v+1}^{j_{v+1}} \Pi_l^{(v+1)} D_{l-i+1}^{(v+1)} + \sum_{l=j_{n-1}+1}^{\infty} R^l D_{l-i+1}^{(n)} \right) \bar{e}, \quad (15)$$

где

$$D_0^{(k)} = \Gamma_0^{(k)} (B_0 \otimes I_{N+1}) - \Phi_0^{(k)} + H^{(k)},$$

$$D_i^{(k)} = \Gamma_i^{(k)} (B_0 \otimes I_{N+1}) + \Gamma_{i-1}^{(k)} (B_1 \otimes I_{N+1}) - \Phi_i^{(k)}, \quad i \geq 1, \quad k = \overline{1, n}.$$
(16)

Доказательство. Пусть $\vec{\beta}_i$ — вектор, задающий вероятности того, что между моментами поступления запросов систему покинет ровно i запросов из-за поступления сбоя, $i \geq 1$. Для векторов $\vec{\beta}_i$, $i \geq 1$ запишем

$$\vec{\beta}_i = \sum_{l=i-1}^{\infty} \bar{\pi}_l D_{l-i+1}^{(k)}, \quad j_{k-1} + 1 \leq l \leq j_k, \quad k = \overline{1, n},$$
(17)

где

$$D_l^{(k)} = \int_0^{\infty} \int_0^t Q(l, v) P \otimes e^{F_0 v} F_1 e^{F^{(1)}(t-v)} dv dA^{(k)}(t), \quad l \geq 0, \quad k = \overline{1, n}.$$
(18)

Преобразуем подынтегральное выражение в (18) с помощью формулы интегрирования по частям и с учетом соотношений (2) и (4) получим равенства (16).

Среднее число запросов, теряемых за время между моментами их поступления, определяется как $\sum_{i=0}^{\infty} i \vec{\beta}_i \bar{e}$, а среднее число запросов, потерянных в единицу времени, задается формулой

$$V = \tau^{-1} \sum_{i=0}^{\infty} i \vec{\beta}_i \bar{e}.$$
(19)

Подставляя в (19) формулу (17) и выражения (5), (6) для $\bar{\pi}_i$, $i \geq 1$, получаем равенство (15). Теорема доказана.

Подставляя выражения (13)–(15) в (1), получаем значение критерия качества при заданных значениях порогов $(j_1, j_2, \dots, j_{n-1})$. Имея алгоритм для вычисления значения критерия качества при любом фиксированном значении порога, находим оптимальный набор порогов $(j_1^*, j_2^*, \dots, j_{n-1}^*)$, минимизируя значение критерия качества.

Оптимизацию значения критерия качества производим прямым перебором в области $[0, \hat{J}]$ значений каждого порога j_k , $k = \overline{1, n}$, где величина \hat{J} определяется как значение порогов $j_1 = j_2 = \dots = j_{n-1} = \hat{J}$, при котором

$\sum_{i=0}^j \bar{\pi}_i \bar{e} = 1 - \varepsilon$, а ε — точность представления чисел в ЭВМ. Получаемое таким образом значение $C(j_1^*, \dots, j_{n-1}^*)$ является оптимальным (по крайней мере, при заданной точности вычислений).

ЧИСЛЕННЫЙ ПРИМЕР

Положим $n = 4$.

Пусть функция распределения времени между моментами поступления запросов в k -м режиме $A^{(k)}(t)$ имеет вид

$$A^{(k)}(t) = \int_0^t \frac{\gamma^{(k)} (\gamma^{(k)} \tau)^{m^{(k)}-1}}{(m^{(k)} - 1)!} e^{-\gamma^{(k)} \tau} d\tau, \quad k = \overline{1, n},$$

где $m^{(1)} = 4$, $m^{(2)} = m^{(4)} = 3$, $m^{(3)} = 4$, $\gamma^{(1)} = 45$, $\gamma^{(2)} = 20$, $\gamma^{(3)} = \gamma^{(4)} = 10$. Интенсивность входного потока запросов равна 11,25; 6,66; 5; 3,33 соответственно в каждом режиме.

MSP-процесс обслуживания запросов определяется матрицами

$$B_0 = \begin{pmatrix} -12,25 & 2,25 \\ 2,64 & -17,64 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 10 & 0 \\ 0 & 15 \end{pmatrix}.$$

Среднее время обслуживания одного запроса равно 0,0811.

Поток сбоев задается матрицами

$$F_0 = \begin{pmatrix} -0,53 & 0,037 \\ 0,019 & -0,319 \end{pmatrix}, \quad F_1 = \begin{pmatrix} 0,5 & 0 \\ 0 & 0,3 \end{pmatrix}.$$

Средняя интенсивность поступления сбоев равна 0,37. Стоимостные коэффициенты имеют значения $a = 2$, $c_1 = 5$, $c_2 = 8$, $c_3 = 10$, $c_4 = 13$, $g = 5$.

Оптимальные значения критерия качества (1) при различных комбинациях использования режимов входного потока и соответствующие оптимальные наборы порогов

Используемые режимы	Оптимальный набор порогов	Оптимальное значение критерия качества
1,2	2	7,697
1,3	3	7,816
1,4	3	8,016
2,3	4	9,583
2,4	6	9,590
3,4	8	11,051
1,2,3	2,4	7,683
1,2,4	2,6	7,691
1,3,4	2,8	9,314
2,3,4	3,7	7,815
1,2,3,4	2,4,7	7,683

Приведем также значения критерия качества C_k при использовании только k -го режима работы, $k = \overline{1,4}$: $C_1 = 9,192$; $C_2 = 9,591$; $C_3 = 11,051$; $C_4 = 13,600$.

Таким образом, использование многопороговой стратегии управления входным потоком запросов позволяет снизить стоимость работы системы. В частности, сравнивая значения $C_{1,2,3,4} = 7,683$ и $\min\{C_1, C_2, C_3, C_4\}$, получаем, что выигрыш от использования многопороговой стратегии управления входным потоком по сравнению с использованием только одного режима входного потока составляет более 16%.

ЗАКЛЮЧЕНИЕ

Таким образом, для рассматриваемой системы найдено стационарное распределение вероятностей состояний вложенной цепи Маркова и предложен алгоритм нахождения оптимальной пороговой стратегии управления входящим потоком.

ЛИТЕРАТУРА

1. Рыков В.В. Управляемые системы массового обслуживания // Теория вероятностей. Математическая статистика. Теоретическая кибернетика (Итоги науки и техники). — М.: ВИНТИ, 1975. — **10**. — С. 43–153.
2. Dudin A.N. Optimal control for a $M^x/G/1$ queue with two operation modes // Probab. Engin. Inform. Scie. — 1997. — **11**. — P. 225–265.
3. Dudin A.N., Nishimura S. Optimal control for a BMAP/G/1 queue with two service modes // Math. Probl. Engin. — 1999. — **5**. — P. 255–273.
4. Dudin A.N., Klimenok V.I. Optimal admission control in a queueing system with heterogeneous traffic // Operations Research Letters. — 2003. — **31**. — P. 108–118.
5. Нобель Р. Регенеративный подход для анализа очереди $M^x/G/1$ с двумя видами обслуживания // Автоматика и вычислительная техника. — 1998. — № 1. — С. 3–14.
6. Gelenbe E. Product form networks with negative and positive customers // J. Appl. Prob. — 1991. — **28**. — P. 655–663.
7. Artalejo J. G-networks: A versatile approach for work removal in queueing networks // Eur. J. Oper. Res. — 2000. — **126**. — P. 233–249.
8. Jain G., Sigman K. A Pollaczek–Khinchine formula for $M/G/1$ queues with disasters // J. Appl. Prob. — 1996. — **33**. — P. 1191–1200.
9. Dudin A.N., Nishimura S. A BMAP/SM/1 queueing system with Markovian arrival of disasters // J. Appl. Prob. — 1999. — **36**, № 3. — P. 868–881.
10. Dudin A.N., Karolik A.V. BMAP/SM/1 queue with Markovian input of disasters and non-instantaneous recovery // Performance Evaluation. — 2001. — **45**. — P. 19–32.
11. Стационарные характеристики системы массового обслуживания $G/MSP/1/r$ / П.П. Бочаров, Ч. Д’Апиче, А.В. Печинкин, С. Салерно // Автоматика и телемеханика. — 2003. — № 2. — С. 127–142.
12. Lucantoni D.M. New results on the single server queue with a batch Markovian arrival process // Commun. Stat. Stochastic Models. — 1991. — **7**. — P. 1–46.
13. Graham A. Kronecker Products and Matrix Calculus with Applications. — Chichester: Ellis Horwood Ltd., 1981. — 489 с.
14. Neuts M.F. Markov chain with applications in queueing theory, which have a matrix-geometric invariant probability vector // Adv. Appl. Probab. — 1978. — **10**. — P. 185–212.
15. Скороход А.В. Теория вероятностей и случайных процессов. — Киев: Вышш. шк., 1980. — 226 с.