

МЕТОДЫ АНАЛИЗА МНОГОКАНАЛЬНОЙ СИСТЕМЫ ОБСЛУЖИВАНИЯ С МГНОВЕННОЙ И ОТСРОЧЕННОЙ ОБРАТНЫМИ СВЯЗЯМИ

Аннотация. Предложена математическая модель многоканальной системы массового обслуживания с обратной связью, в которой одна часть вызовов для получения повторного обслуживания мгновенно поступает в систему, а другая — либо обращается для повторного обслуживания через некоторое случайное время, либо окончательно покидает систему. Поведение обслуженных вызовов имеет рандомизированный характер. Разработаны методы точного и асимптотического анализа характеристик предложенной модели. Приведены результаты численных экспериментов.

Ключевые слова: многоканальная система обслуживания, обратная связь, качество обслуживания.

ВВЕДЕНИЕ

В классических моделях систем массового обслуживания (СМО) не учитывается эффекта обратной связи с системой. Это обстоятельство существенно уменьшает адекватность классических моделей СМО, которые используются для математического анализа реальных процессов обработки вызовов, когда уже прошедшая стадию обслуживания часть вызовов требует повторного обслуживания по различным причинам, например, в зависимости от качества обслуживания, от времени пребывания в системе и т.д. Так, в мультиагентных системах запросы, даже получившие удовлетворительное обслуживание, могут снова поступить на повторное обслуживание этими агентами [1–3].

Среди моделей СМО с обратной связью выделяют два типа: модели без орбита и модели с орбитом. В моделях без орбита некоторые вызовы после получения первичного обслуживания мгновенно возвращаются на повторное, в то время как в моделях с орбитами, находящиеся в них вызовы прежде чем поступить на повторное обслуживание «обдумывают» свое решение. В соответствии с этими свойствами модели без орбита называются моделями с мгновенной обратной связью, а модели с орбитом — моделями с отсроченной обратной связью.

Модели СМО с обратной связью мало изучены. После классических работ Такача [4, 5], в которых методом производящих функций изучены двумерные марковские модели одноканальных СМО с неограниченной очередью и бесконечным объемом орбита, модели СМО с обратной связью долгое время не привлекали внимания исследователей.

Модели без орбита изучены в работах [6–8]. Так, модели с бесконечным числом каналов исследованы в [6, 7], где для расчета распределения (стационарного и нестационарного) первичного, повторного, а также суммарного потоков используется метод производящих функций. Модель одноканальной СМО с ограниченной очередью изучена в [8], где для расчета стационарного распределения соответствующей четырехмерной цепи Маркова (ЦМ) применяется матрично-геометрический метод Ньютона [9]. (Здесь и далее под стационарным распределением модели, в частности ЦМ, понимается совокупность стационарных вероятностей состояний описанной модели.) Отметим, что в работах [4–8] предполагается, что время обработки первичных и повторных вызовов идентично, а вероятности выхода из системы и возвращения вызовов в нее для получения повторного обслуживания — постоянные величины (т.е. они не зависят от состояния системы). Эти до-

пушения ограничивают области приложения предложенных моделей. Во всех приведенных работах модели СМО без орбита и с орбитом изучались отдельно, что также сужает области приложения полученных в них результатов.

В связи с этим в настоящей статье рассматривается модель СМО с обратной связью, в которой, кроме первичных, имеются мгновенные повторные вызовы и повторные вызовы из орбита (т.е. отсроченные повторные вызовы). Отметим, что учет наличия этих вызовов существенно усложняет исследования модели, так как приходится рассматривать трехмерную ЦМ с достаточно сложной производящей матрицей. Использование системы уравнений равновесия (СУР) для расчета стационарного распределения соответствующей трехмерной ЦМ малопродуктивно для моделей большой размерности. Потому далее предложены эффективные вычислительные процедуры, позволяющие осуществить асимптотический анализ изучаемых моделей, когда число каналов системы и размер орбита достаточно большие, в том числе и бесконечные, величины.

Отметим, что модели СМО с обратной связью близки к моделям СМО с повторными вызовами [10]. Вместе с тем в отличие от последних в моделях СМО с обратной связью орбит образуется не вновь поступившими вызовами, а уже получившими обслуживание. Это свойство рассматриваемых в [10] систем требует разработки специального подхода к их изучению.

В работах [11, 12] исследуются модели СМО с обратной связью других типов, однако в них термин «обратная связь» обозначает иные свойства этих моделей. Так, в [11] он используется для обозначения СМО, в которой функции распределения (ф.р.) времени обслуживания вызовов или интервалов между их поступлениями зависят от длины очереди, а в работе [12] — для обозначения взаимосвязи между процессами обслуживания вызовов в различных фазах двухфазной СМО.

ОПИСАНИЕ МОДЕЛИ СМО С ОБРАТНОЙ СВЯЗЬЮ

Структурная схема описанной СМО с обратной связью показана на рис. 1. На вход системы, содержащей $N > 1$ автономных и идентичных каналов, поступает пуассоновский поток первичных вызовов (p -вызовы) с интенсивностью λ_p . Времена их обработки — независимые и одинаково распределенные случайные величины, и их ф.р. являются экспоненциальными с общим средним $1/\mu_p$.

После окончания процесса обработки p -вызова он либо с вероятностью $\sigma_1(x)$ покидает систему, либо с вероятностью $\sigma_2(x)$ требует мгновенной повторной обработки, либо с вероятностью $\sigma_3(x) = 1 - \sigma_1(x) - \sigma_2(x)$ уходит в орбит. Эти вероятности зависят от параметра x , который определяет состояние некоторой внешней среды, стохастически связанной с рассматриваемой системой.

Предполагается, что максимальный размер орбита для повторных вызовов равен R , $0 < R \leq \infty$. Ограниченность объема орбита (при $R < \infty$) означает, что вызовы в него принимаются, если в момент поступления последнего общее число повторных вызовов (r -вызовов) меньше R , иначе он покидает систему. В случае неограниченного объема орбита (при $R = \infty$) уходящий в него вызов с вероятностью 1 принимается.

Время пребывания r -вызовов в орбите ограничено некоторой случайной величи-

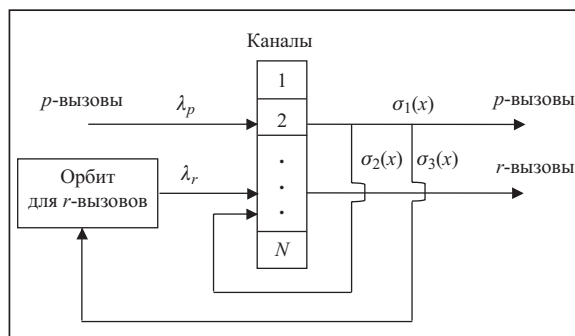


Рис. 1. Структурная схема СМО с обратной связью

ной, а именно эти вызовы поступают для получения повторного обслуживания через случайные моменты времени, которые подчиняются показательному закону распределения со средним $1/\lambda_r$. Предполагается, что повторные вызовы из орбита не являются настойчивыми, т.е. если в момент поступления повторного вызова из орбита все каналы заняты, то он окончательно покидает систему. Заметим, что можно исследовать модель, в которой эти вызовы в такие моменты не теряются, а возвращаются обратно в орбит.

Времена обработки повторных вызовов (мгновенных или из орбита) также независимые и одинаково распределенные случайные величины, при этом их ф.р. для всех r -вызовов являются экспоненциальными с общим средним $1/\mu_r$. Для общности модели предполагается, что, вообще говоря, $\mu_p \neq \mu_r$, а для простоты модели — что мгновенные повторные вызовы и повторные вызовы из орбита после завершения их обслуживания больше не повторяются (хотя, как будет ясно из дальнейшего изложения, предложенный подход позволяет учесть и многократные повторения). Проблема состоит в определении основных характеристик рассматриваемой модели, а именно вероятностей потери первичных P_p и повторных вызовов P_r , среднего числа первичных L_p и повторных вызовов в каналах L_r , среднего числа повторных вызовов в орбите L_o , а также коэффициента использования C_u каналов.

МЕТОД ТОЧНОГО АНАЛИЗА

Исходя из вида ф.р. случайных величин, используемых в формировании модели, определяем, что изучаемая система описывается трехмерной ЦМ. Состояние этой ЦМ в произвольный момент времени описывается трехмерным вектором $\mathbf{n} = (n_p, n_r, n_o)$, где компоненты n_p и n_r соответствуют числу первичных и повторных вызовов в каналах, а n_o — числу повторных вызовов в орбите. Фазовое пространство состояний (ФПС) указанной ЦМ задается так:

$$S = \{ \mathbf{n} : n_p = 0, 1, \dots, N; n_r = 0, 1, \dots, N; n_p + n_r \leq N; n_o = 0, 1, \dots, R \}. \quad (1)$$

Следовательно, геометрически ФПС модели (рис. 2) задается точками с целочисленными координатами внутри призмы высотой R , основанием которой является равнобедренный прямоугольный треугольник с катетами длиной N .

Предположим, что состояние внешней среды x , которое определяет поведение уже получившего обслуживание первичного вызова, задается числом по-

вторных вызовов в орбите, т.е. $x = 0, 1, 2, \dots, R$. Вначале рассмотрим модель с конечным объемом орбита при $R < \infty$. Поскольку вызов принимается в орбит лишь тогда, когда в момент его поступления общее число повторных вызовов меньше R , вероятность $\sigma_3(R) = 0$.

Интенсивность перехода из состояния \mathbf{n} в состояние \mathbf{n}' обозначим $q(\mathbf{n}, \mathbf{n}')$, $\mathbf{n}, \mathbf{n}' \in S$. Совокупность этих величин составляет производящую матрицу (Q -матрицу) данной ЦМ. Они определяются из следующих соотношений:

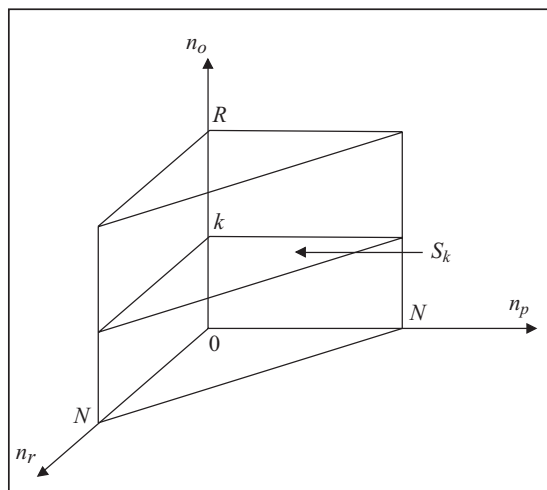


Рис 2. Вид ФПС модели

$$q(\mathbf{n}, \mathbf{n}') = \begin{cases} \lambda_p, & \text{если } \mathbf{n}' = \mathbf{n} + \mathbf{e}_1, \\ n_p \mu_p \sigma_1(n_o), & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_1, \\ n_p \mu_p \sigma_2(n_o), & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_1 + \mathbf{e}_2, \\ n_p \mu_p \sigma_3(n_o), & \text{если } n_o < R, \mathbf{n}' = \mathbf{n} - \mathbf{e}_1 + \mathbf{e}_3, \\ n_r \mu_r, & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_2, \\ n_o \lambda_r, & \text{если } n_p + n_r < N, \mathbf{n}' = \mathbf{n} + \mathbf{e}_2 - \mathbf{e}_3 \text{ или } n_p + n_r = N, \mathbf{n}' = \mathbf{n} - \mathbf{e}_3, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (2)$$

Здесь \mathbf{e}_i обозначает i -й ортвектор трехмерного евклидова пространства, $i=1,2,3$. Данная трехмерная ЦМ с конечным числом состояний является неприводимой, что определяет существование в ней стационарного режима. Пусть $p(\mathbf{n})$ — стационарная вероятность состояния $\mathbf{n} \in S$. Эти величины удовлетворяют соответствующей системе уравнений равновесия (СУР), которая составляется на основе соотношений (2) и имеет следующий вид:

$$\begin{aligned} (\lambda_p I(n_p + n_r < N) + n_p \mu_p + n_r \mu_r + n_o \lambda_r) p(\mathbf{n}) = & \lambda_p p(\mathbf{n} - \mathbf{e}_1) I(n_p > 0) + \\ & + (n_p + 1) \mu_p (p(\mathbf{n} + \mathbf{e}_1) \sigma_1(n_o) + p(\mathbf{n} + \mathbf{e}_1 - \mathbf{e}_2) \sigma(n_o) + \\ & + p(\mathbf{n} + \mathbf{e}_1 - \mathbf{e}_3) \sigma_3(n_o - 1) I(n_o > 0)) + (n_r + 1) \mu_r p(\mathbf{n} + \mathbf{e}_2) + \\ & + (n_o + 1) \lambda_r (p(\mathbf{n} + \mathbf{e}_3) \delta(n_p + n_r, N) + p(\mathbf{n} - \mathbf{e}_2 + \mathbf{e}_3) I(n_o < R, n_r > 0)). \end{aligned} \quad (3)$$

Здесь и далее $I(A)$ — индикаторная функция события A и $\delta(i, j)$ — символы Кронекера. К этой СУР добавляется условие нормировки

$$\sum_{\mathbf{n} \in S} p(\mathbf{n}) = 1. \quad (4)$$

После нахождения решения СУР (3), (4) характеристики изучаемой системы задаются через маргинальные распределения данной трехмерной цепи. Так, поскольку поток первичных вызовов пуассоновский, вероятность потери первичных вызовов можно определить с помощью теоремы PASTA [13]:

$$P_p = \sum_{\mathbf{n} \in S} p(\mathbf{n}) \delta(n_p + n_r, N). \quad (5)$$

На основании результатов работы [14] определяем, что поток повторных вызовов можно считать пуассоновским, и поэтому для вычисления вероятности потери повторных вызовов также можно использовать теорему PASTA [13]. Поскольку повторные вызовы генерируются лишь тогда, когда орбит не является пустым, искомая характеристика определяется так:

$$P_r = \sum_{\mathbf{n} \in S} p(\mathbf{n}) \delta(n_p + n_r, N) (1 - \delta(n_o, 0)). \quad (6)$$

Среднее число первичных и повторных вызовов в каналах и повторных вызовов в орбите определяется как математическое ожидание соответствующих дискретных случайных величин

$$L_x = \sum_{j=1}^N j \Phi_x(j), \quad (7)$$

где $\Phi_x(j) = \sum_{\mathbf{n} \in S} p(\mathbf{n}) \delta(n_x, j)$, $x \in \{p, r\}$;

$$L_o = \sum_{j=1}^R j \psi(j), \quad (8)$$

где $\psi(j) = \sum_{\mathbf{n} \in S} p(\mathbf{n}) \delta(n_o, j)$.

Коэффициент использования каналов C_u определяется с учетом формул (7):

$$C_u = (L_p + L_r) / N. \quad (9)$$

Относительно нахождения решения СУР (3), (4) отметим, что, к сожалению, вследствие сложной структуры Q -матрицы не удастся найти ее аналитического решения, и поэтому для этой цели следует использовать численные методы линейной алгебры (зачастую для решения подобных СУР эффективно применение известного метода Гаусса–Зейделя).

В случае бесконечного размера орбита необходимо использовать подход, основанный на применении метода двумерных производящих функций, хотя методологические и вычислительные недостатки данного подхода хорошо известны.

МЕТОД АСИМПТОТИЧЕСКОГО АНАЛИЗА

Размерность СУР (3), (4) определяется размерностью фазового пространства состояний (1), которое состоит из $(N+1)(N+2)(R+1)/2$ состояний, т.е. она оценивается как $O(N^2R)$. Поэтому описанный ранее точный метод позволяет вычислить характеристики (5)–(9) лишь для систем умеренной размерности, а с ростом размерности ФПС (1) возникают большие вычислительные трудности. Для их устранения можно использовать метод теории фазового укрупнения состояний стохастических систем из [15], где предложен иерархический алгоритм фазового укрупнения (АФУ) для расчета стационарного распределения изучаемой трехмерной ЦМ при выполнении определенных асимптотических условий.

Для корректного применения данного метода принимается следующее асимптотическое условие: $\lambda_p \gg \max\{\mu_p, \mu_r\}$, т.е. предполагается, что система работает в условиях большой нагрузки.

Отметим, что в исследуемых моделях, как правило, интенсивность поступления p -вызовов существенно превышает интенсивность поступления r -вызовов из орбита, т.е. условие $\lambda_p \gg \lambda_r$ естественно, так как орбит образуется частью первичных вызовов. Тогда в силу указанного допущения заключаем, что интенсивности переходов между состояниями внутри плоскостей, параллельных основанию призмы, намного превышают интенсивности переходов между состояниями из различных плоскостей (см. рис. 2). Таким образом, можно рассматривать расщепление ФПС (1)

$$S = \bigcup_{k=0}^R S_k, \quad S_k \cap S_{k'} = \emptyset, \quad \text{если } k \neq k', \quad (10)$$

где $S_k = \{\mathbf{n} \in S : n_o = k\}$, $k = 0, 1, \dots, R$, иными словами, — расслоение ФПС модели по плоскостям, параллельным основанию призмы (см. рис. 2).

Вводимая в ФПС (1) функция укрупнения определяется на основе расщепления (10):

$$U_1(\mathbf{n}) = \langle k \rangle, \quad \text{если } \mathbf{n} \in S_k, \quad (11)$$

где $\langle k \rangle$ — укрупненное состояние, включающее все состояния из класса (слоя) S_k . Обозначим $\Omega_1 = \{\langle k \rangle : k = 0, 1, \dots, R\}$.

Согласно АФУ из работы [16 и приложения в ней] вероятности состояний исходной модели определяются следующим образом:

$$p(\mathbf{n}) \approx \rho_k(n_p, n_r) \pi_1(\langle k \rangle), \quad (12)$$

где $\rho_k(n_p, n_r)$ — вероятность состояния (n_p, n_r) внутри расщепленной модели с пространством состояний S_k , а $\pi_1(\langle k \rangle)$ — вероятность укрупненного состояния $\langle k \rangle \in \Omega_1$.

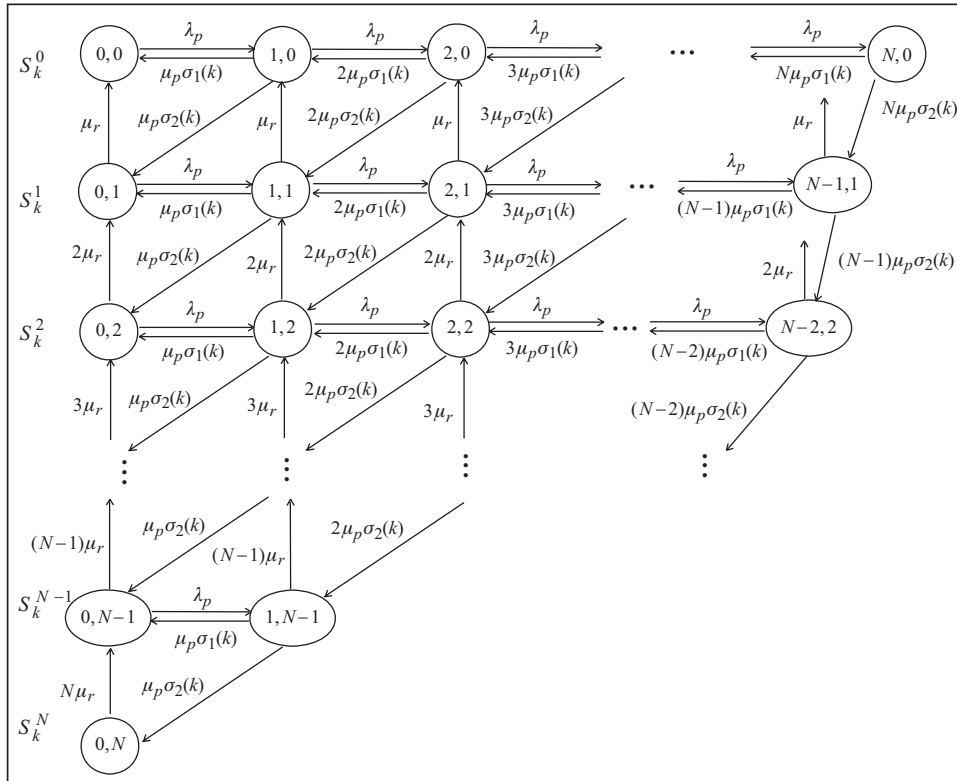


Рис 3. Граф переходов расщепленной модели с ФПС $S_k, k = 0, 1, \dots, R$

Следовательно, для расчета стационарного распределения исходной трехмерной ЦМ потребуется найти соответствующие распределения вероятностей состояний $R + 1$ двумерных и одной одномерной цепей. При больших значениях числа каналов аналогичные вычислительные трудности возникают и при расчете стационарных распределений описанных двумерных цепей с пространствами состояний $S_k, k = 0, 1, \dots, R$. Поэтому для расчета стационарных распределений внутри классов $S_k, k = 0, 1, \dots, R$, к каждому из них применяется АФУ, иными словами, рассматривается иерархия укрупненных моделей.

Из схемы разбиения (10) видно, что все расщепленные модели с пространствами состояний $S_k, k = 0, 1, \dots, R$, представляют собой идентичные двумерные ЦМ (рис. 3). Поэтому в дальнейшем фиксируется значение k и рассматривается расщепленная модель с пространством состояний S_k .

В пространстве состояний S_k рассматривается разбиение

$$S_k = \bigcup_{i=0}^N S_k^i, S_k^i \cap S_k^j = \emptyset, \text{ если } i \neq j, \quad (13)$$

где $S_k^i = \{n \in S_k : n_r = i\}, i = 0, 1, \dots, N$, иными словами — разбиение пространства состояний расщепленной модели по строкам (см. рис. 3).

Далее на основе разбиения (13) в пространстве состояний S_k определяется следующая функция укрупнения:

$$U_2(n) = \langle i \rangle, \text{ если } n \in S_k^i, \quad (14)$$

где $\langle i \rangle$ — укрупненное состояние, которое включает все состояния из класса S_k^i . Обозначим $\Omega_2 = \{\langle i \rangle : i = 0, 1, \dots, N\}$.

Согласно АФУ имеем

$$\rho_k(n_p, n_r) \approx \rho_{n_r}^k(n_p) \pi_2^k(\langle n_r \rangle), \quad (15)$$

где $\rho_{n_r}^k(n_p)$ — вероятность состояния (n_p, n_r) внутри расщепленной модели с пространством состояний $S_k^{n_r}$, а $\pi_2^k(\langle n_r \rangle)$ — вероятность укрупненного состояния $\langle n_r \rangle \in \Omega_2$.

Рассмотрим задачу вычисления вероятностей состояний внутри классов S_k^i . Отметим, что, поскольку класс состояний S_k^N содержит лишь одно состояние $(0, N)$, далее принимается, что $\rho_N^k(0) = 1$.

В классе состояний S_k^i , $i \neq N$, вторая компонента является постоянной. Поэтому при изучении расщепленных моделей с ФПС S_k^i микросостояние $(n_p, i) \in S_k^i$ можно задать лишь первой компонентой. Далее для удобства изложения при изучении расщепленной модели с ФПС S_k^i ее состояние (n_p, i) обозначается как $n_p, n_p = 0, 1, \dots, N - i$.

Интенсивность перехода между состояниями n_p и n'_p расщепленной модели с ФПС S_k^i обозначается $q_k(n_p, n'_p)$. Из соотношений (2) получаем, что эти параметры определяются так (см. рис. 3):

$$q_k(n_p, n'_p) = \begin{cases} \lambda_p, & \text{если } n'_p = n_p + 1, \\ n_p \mu_p \sigma_1(k), & \text{если } n'_p = n_p - 1, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (16)$$

Следовательно, из соотношений (16) получаем, что вероятности состояний внутри расщепленной модели с ФПС S_k^i , $i = 0, 1, \dots, N - 1$, вычисляются как вероятности состояний одномерного процесса размножения и гибели:

$$\rho_i^k(j) = \frac{(v_p / \sigma_1(k))^j}{j!} \rho_i^k(0), \quad j = 1, \dots, N - i, \quad (17)$$

где $v_p = \lambda_p / \mu_p$, а $\rho_i^k(0)$ находится из условия нормировки, т.е. $\sum_{j=0}^{N-i} \rho_i^k(j) = 1$.

Интенсивность перехода из укрупненного состояния $\langle i \rangle$ в другое укрупненное состояние $\langle j \rangle$ обозначим $q_k(\langle i \rangle, \langle j \rangle)$, $\langle i \rangle, \langle j \rangle \in \Omega_2$. Тогда после определенных математических преобразований с применением АФУ получаем

$$q_k(\langle i \rangle, \langle j \rangle) = \begin{cases} \mu_p \alpha_k(i), & \text{если } j = i + 1, \\ i \mu_r, & \text{если } j = i - 1, \\ 0 & \text{в остальных случаях,} \end{cases} \quad (18)$$

где $\alpha_k(i) = \sigma_2(k) \sum_{j=1}^{N-i} j \rho_i^k(j)$, $i = 0, 1, \dots, N - 1$.

Таким образом, из соотношений (18) для вычисления вероятностей укрупненных состояний $\pi_2^k(\langle n_r \rangle)$, $\langle n_r \rangle \in \Omega_2$, получаем следующее выражение:

$$\pi_2^k(\langle n_r \rangle) = \frac{(\mu_p / \mu_r)^{n_r}}{n_r!} \prod_{i=0}^{n_r-1} \alpha_k(i) \pi_2^k(\langle 0 \rangle), \quad n_r = 1, \dots, N, \quad (19)$$

где $\pi_2^k(\langle 0 \rangle)$ находится из условия нормировки, т.е. $\sum_{j=0}^N \pi_2^k(\langle j \rangle) = 1$.

Замечание 1. При вычислении стационарного распределения расщепленной модели с ФПС S_R^i следует учитывать, что $\sigma_3(R) = 0$, т.е. $\sigma_1(R) + \sigma_2(R) = 1$.

Теперь для вычисления стационарного распределения исходной трехмерной ЦМ потребуется нахождение вероятностей укрупненных состояний $\pi_1(\langle k \rangle)$, $\langle k \rangle \in \Omega_1$ (12). Для этого необходимо определить интенсивности переходов между классами (слоями) S_k , $k = 0, 1, \dots, R$ (10).

Обозначим $q(S_k, S_{k'})$ интенсивность перехода между классами S_k и $S_{k'}$. Эти интенсивности определяются с учетом соотношений (2), (17) и (19). Тогда с помощью АФУ после математических преобразований имеем

$$q(S_k, S_{k'}) = \begin{cases} \Lambda(k), & \text{если } k' = k + 1, \\ k\lambda_r, & \text{если } k' = k - 1, \\ 0 & \text{в остальных случаях,} \end{cases} \quad (20)$$

где $\Lambda(k) = \mu_p \sigma_3(k) \sum_{i=0}^{N-1} \pi_2^k(\langle i \rangle) \sum_{j=1}^{N-i} j \rho_i^k(j)$, $k = 0, 1, \dots, R - 1$.

Следовательно, из соотношений (20) получаем, что искомые вероятности укрупненных состояний $\pi_1(\langle k \rangle)$, $\langle k \rangle \in \Omega_1$, определяются как стационарное распределение одномерного процесса размножения и гибели с переменными параметрами:

$$\pi_1(\langle k \rangle) = \frac{1}{k! \lambda_r^k} \prod_{i=0}^{k-1} \Lambda(i) \pi_1(\langle 0 \rangle), \quad k = 1, \dots, R, \quad (21)$$

где $\pi_1(\langle 0 \rangle)$ находится из условия нормировки, т.е. $\sum_{k=0}^R \pi_1(\langle k \rangle) = 1$.

Окончательно стационарное распределение исходной трехмерной ЦМ определяется из (12) с помощью соотношений (17), (19) и (21):

$$p(n_p, n_r, n_o) \approx \rho_{n_r}^{n_o} (n_p) \pi_1(\langle n_o \rangle) \pi_2^{n_o}(\langle n_r \rangle). \quad (22)$$

Резюмируя, получаем следующие выражения для приближенного вычисления искомых характеристик изучаемой системы:

$$P_p \approx \sum_{k=0}^R \pi_1(\langle k \rangle) \sum_{i=0}^N \rho_i^k (N-i) \pi_2^k(\langle i \rangle); \quad (23)$$

$$P_r \approx \sum_{k=1}^R \pi_1(\langle k \rangle) \sum_{i=0}^N \rho_i^k (N-i) \pi_2^k(\langle i \rangle); \quad (24)$$

$$L_p \approx \sum_{k=0}^R \pi_1(\langle k \rangle) \sum_{j=1}^N j \sum_{i=0}^{N-j} \rho_i^k(j) \pi_2^k(\langle i \rangle); \quad (25)$$

$$L_r \approx \sum_{k=0}^R \pi_1(\langle k \rangle) \sum_{i=1}^N i \pi_2^k(\langle i \rangle); \quad (26)$$

$$L_o \approx \sum_{k=1}^R k \pi_1(\langle k \rangle). \quad (27)$$

Коэффициент использования каналов C_u определяется приближенно с учетом формул (9) и (25), (26).

АСИМПТОТИЧЕСКИЙ АНАЛИЗ МОДЕЛИ С БЕСКОНЕЧНЫМИ СТРУКТУРНЫМИ ПАРАМЕТРАМИ

Рассмотрим модель с бесконечными параметрами, т.е. с бесконечным размером орбита и/или с бесконечным числом каналов обслуживания. Отметим, что

если вероятности $\sigma_i(k)$, $i=1,2,3$, не зависят от состояния орбита k , $k=0,1,\dots,R$, то предложенные формулы (23)–(27) существенно упрощаются. Действительно, в этом случае стационарные распределения внутри расщепленных моделей с ФПС S_k^i не зависят от параметра k , $k=0,1,\dots,R$. Иными словами, из соотношений (17) получаем, что вероятности состояний внутри всех расщепленных моделей с ФПС S_k^i , $i=0,1,\dots,N-1$, вычисляются так:

$$\rho_i(j) = \frac{(v_p / \sigma_1)^j}{j!} \rho_i(0), \quad j=1, \dots, N-i, \quad (28)$$

где $\rho_i(0)$ находится из условия нормировки.

Далее из соотношений (18) и (19) получаем следующие выражения для вычисления вероятностей укрупненных состояний на второй ступени иерархии:

$$\pi_2(\langle n_r \rangle) = \frac{(\mu_p / \mu_r)^{n_r}}{n_r!} \prod_{i=0}^{n_r-1} \alpha(i) \pi_2(\langle 0 \rangle), \quad n_r = 1, \dots, N, \quad (29)$$

где $\alpha(i) = \sigma_2 \sum_{j=1}^{N-i} j \rho_i(j)$, $i=0,1,\dots,N-1$, и $\pi_2(\langle 0 \rangle)$ находится из условия нормировки.

Замечание 2. Согласно замечанию 1 здесь также следует учитывать, что для расщепленной модели с ФПС S_R^i имеет место соотношение $\sigma_1 + \sigma_2 = 1$.

Для вычисления вероятностей укрупненных состояний в первой ступени иерархии из соотношений (20) и (21) получим

$$\pi_1(\langle k \rangle) = \frac{(\Lambda / \lambda_r)^k}{k!} \pi_1(\langle 0 \rangle), \quad k=1, \dots, R, \quad (30)$$

где $\Lambda = \mu_p \sigma_3 \sum_{i=0}^{N-1} \sum_{j=1}^{N-i} j \rho_i(j) \pi_2(\langle i \rangle)$ и $\pi_1(\langle 0 \rangle)$ находится из условия нормировки.

С учетом (28)–(30) из (23)–(27) получаем следующие выражения для приближенного вычисления характеристик модели с постоянными вероятностями ухода и возвращения вызовов:

$$P_p \approx \sum_{i=0}^N \rho_i(N-i) \pi_2(\langle i \rangle); \quad (31)$$

$$P_r \approx (1 - \pi_1(\langle 0 \rangle)) \sum_{i=0}^N \rho_i(N-i) \pi_2(\langle i \rangle); \quad (32)$$

$$L_p \approx \sum_{j=1}^N j \sum_{i=0}^{N-j} \rho_i(j) \pi_2(\langle i \rangle); \quad (33)$$

$$L_r \approx \sum_{i=1}^N i \pi_2(\langle i \rangle); \quad (34)$$

$$L_o \approx \sum_{k=1}^R k \pi_1(\langle k \rangle). \quad (35)$$

Из формул (31), (33) и (34) видно, что в данной модели вероятности потери первичных вызовов, их среднее число в каналах, а также среднее число повторных вызовов в каналах напрямую не зависят от интенсивности λ_r поступления повторных вызовов из орбита и его размера R . Эти факты объясняются тем, что данные формулы основаны на асимптотическом условии того, что интенсив-

ность p -вызовов существенно превышает интенсивности r -вызовов из орбита. Отметим, что описанные характеристики зависят от параметров σ_1 и σ_2 , следовательно, и от σ_3 — вероятности поступления первичных вызовов в орбит после завершения обслуживания. Иными словами, данные характеристики косвенно зависят от интенсивности заполнения орбита R , следовательно, и от интенсивности λ_r генерации повторных вызовов. Вместе с тем вероятности потери повторных вызовов, их среднее число в каналах и в орбите напрямую зависят от интенсивности поступления и обслуживания первичных и повторных вызовов, а также от размера орбита (32), (34) и (35).

Если вероятности $\sigma_i(k)$, $i=1, 2, 3$, не зависят от состояния орбита k , $k=0, 1, \dots, R$, то предложенный подход позволяет определить характеристики системы и при бесконечном размере орбита и/или при бесконечном числе каналов обслуживания.

Пусть число каналов является конечным, $N < \infty$, а размер орбита — бесконечным, $R = \infty$. Тогда ФПС этой модели представляет собой призму с бесконечной высотой (см. рис. 2). В этом случае вероятности состояний расщепленных моделей также вычисляются из (28) и вероятности укрупненных состояний на второй ступени иерархии вычисляются из (29). Из (20), (21) и (30) получаем, что вероятности укрупненных состояний на первой ступени иерархии вычисляются так:

$$\pi_1(\langle k \rangle) = \frac{\theta^k}{k!} e^{-\theta}, \quad k=0, 1, \dots, \quad (36)$$

где $\theta = \Lambda / \lambda_r$. Тогда вероятность потери первичных вызовов снова определяется как свертка двух распределений (31), а вероятность потери повторных вызовов из орбита вычисляется так:

$$P_r \approx (1 - e^{-\theta}) \sum_{i=0}^N \rho_i(N-i) \pi_2(\langle i \rangle). \quad (37)$$

Среднее число первичных и повторных вызовов в каналах определяются из формул (33) и (34) соответственно. Из (27), (35) и (36) получаем, что в данной модели среднее число повторных вызовов в орбите $L_o = \theta$.

Предположим, что, кроме размера орбита, число каналов также является бесконечным, т.е. $R = \infty$ и $N = \infty$. Тогда ФПС этой модели представляет собой весь первый (положительный) октант трехмерного евклидова пространства, а стационарные вероятности состояний внутри всех расщепленных моделей с ФПС S_k^i , $k=0, 1, 2, \dots$, являющиеся одинаковыми и не зависящими от параметров k и i , определяются так:

$$\rho(i) = \frac{\beta^i}{i!} e^{-\beta}, \quad i=0, 1, 2, \dots, \quad (38)$$

где $\beta = \nu_p / \sigma_1$.

Вероятности укрупненных состояний на различных уровнях иерархии вычисляются следующим образом:

$$\pi_2(\langle i \rangle) = \frac{\gamma^i}{i!} e^{-\gamma}, \quad i=0, 1, 2, \dots, \quad (39)$$

где $\gamma = \frac{\lambda_p}{\mu_r} \cdot \frac{\sigma_2}{\sigma_1}$;

$$\pi_1(\langle i \rangle) = \frac{\tau^i}{i!} e^{-\tau}, \quad i=0, 1, 2, \dots, \quad (40)$$

где $\tau = \frac{\lambda_p}{\lambda_r} \cdot \frac{\sigma_3}{\sigma_1}$.

Очевидно, что в данной модели вероятности потерь первичных вызовов и повторных вызовов из орбита совпадают и равны нулю. С учетом формул (38)–(40) после определенных преобразований получим, что другие характеристики данной мо-

дели определяются так:

$$L_p \approx \beta; L_r \approx \gamma; L_o \approx \tau. \quad (41)$$

Из формулы (41) видно, что среднее число первичных вызовов L_p в каналах зависит лишь от нагрузочных параметров этого трафика и от вероятности ухода этих вызовов из системы после получения обслуживания и не зависит от нагрузочных параметров повторных вызовов. Этого следовало ожидать, так как система имеет бесконечное число каналов, и согласно принятому ранее асимптотическому условию имеем $\lambda_p \gg \max\{\mu_p, \mu_r\}$. Аналогично интерпретируются остальные формулы из (41).

ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

Разработанные методы (точный и приближенный) позволяют изучить характеристики СМО с мгновенной и отсроченной обратными связями относительно изменения ее структурных (число каналов и размер орбита) и нагрузочных параметров (интенсивности поступления и обслуживания первичных и повторных вызовов), а также параметров, которые определяют поведение вызовов после получения первичного обслуживания, т.е. вероятностей $\sigma_i(k)$, $i=1, 2, 3$. При этом, как описано ранее, точный анализ возможен лишь для моделей умеренной размерности.

Проведение экспериментов имело две цели: во-первых, изучить поведение характеристик системы относительно ее структурных параметров, и во-вторых, установить точность разработанных приближенных формул, а также сравнить временные показатели выполнения точных и приближенных алгоритмов.

Для краткости изложения рассматриваются результаты для модели СМО с конечными структурными параметрами, при этом изучаются зависимости характеристик системы от числа каналов обслуживания при фиксированных значениях других параметров.

Результаты численных экспериментов для гипотетической модели показаны на рис. 4–7 (для удобства графического представления в некоторых случаях на оси ординат используется логарифмическая шкала), где приведены графики исследуемых характеристик для следующих значений нагрузочных параметров модели: $\lambda_p = 4$, $\mu_p = 0.5$, $\lambda_r = 2$, $\mu_r = 0.2$. Для простоты в них предполагается, что вероятности $\sigma_i(k)$, $i=1, 2, 3$, являются постоянными величинами, иными словами, принимается, что $\sigma_1(k) = 0.4$, $\sigma_2(k) = \sigma_3(k) = 0.3$, для любого $k = 0, 1, \dots, R-1$ и $\sigma_1(R) = \sigma_2(R) = 0.5$.

Увеличение числа каналов приводит к систематическому уменьшению вероятности потери первичных вызовов P_p ; вместе с тем вероятность потери повторных вызовов P_r растет при малых значениях числа каналов N , и по достижении определенного максимального значения данная характеристика системы также систематически уменьшается (см. рис. 4). Этот феномен объясняется наличием эффекта обратной связи при малых значениях числа каналов. С ростом числа каналов влияние эффекта обратной связи почти полностью исчезает и уже при очень больших значениях N значение функции P_r асимптотически приближается к значениям функции P_p , при этом всегда имеет место соотношение $P_r < P_p$.

Отметим, что увеличение объема орбита для повторных вызовов приводит к уменьшению значений функций P_p и P_r . Иными словами, значения указанных функций при $R = 3$ оказываются большими, чем их соответствующие значения при $R = 10$ (см. рис. 4). Если это свойство для функции P_r имеет вполне логическое объяснение, то для функции P_p оно на первый взгляд кажется нелогичным, так как увеличение объема орбита должно привести к увеличению интенсивности повторных вызовов, и следовательно, увеличится вероятность потери первичных вызовов. Этот факт объясняется тем, что для выбранных исходных данных гипотетической модели при любых значениях числа каналов N среднее число повторных вызовов в орбите L_o при $R = 3$ оказывается существенно большим, чем при $R = 10$.

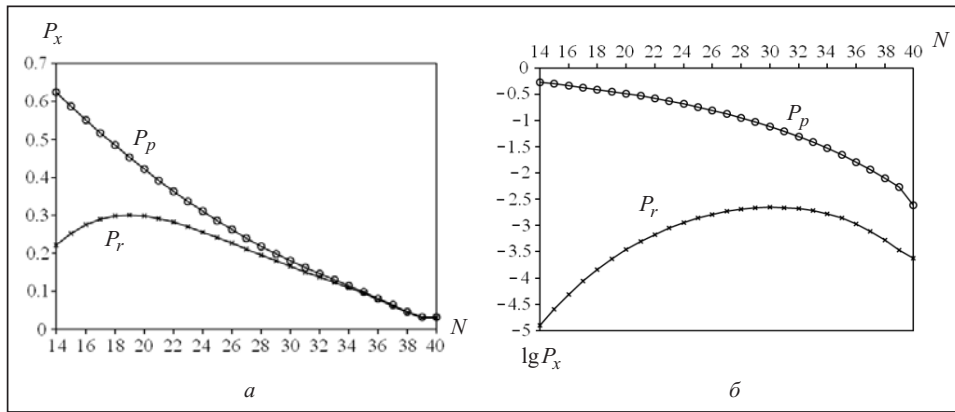


Рис. 4. Зависимость вероятности потери первичных и повторных вызовов от числа каналов при $R = 3$ (а) и $R = 10$ (б)

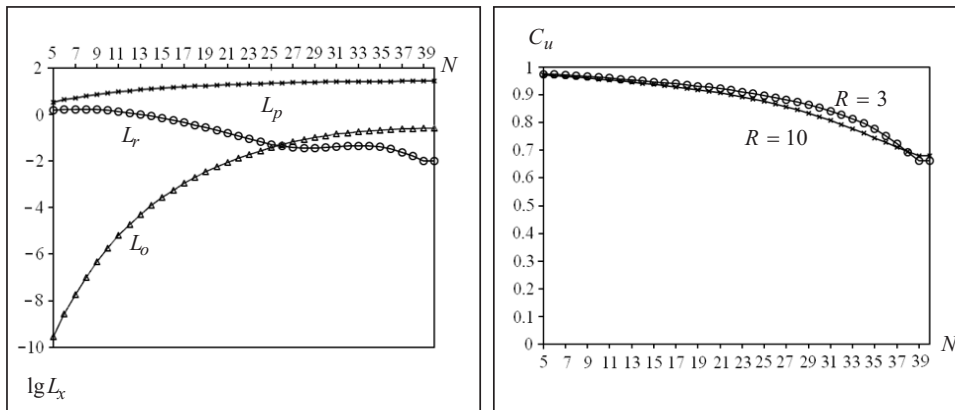


Рис. 5. Зависимости среднего числа первичных и повторных вызовов в каналах, а также числа повторных вызовов в орбите от числа каналов системы при $R = 10$

Рис. 6. Зависимость коэффициента использования каналов от числа каналов системы

Зависимости среднего числа первичных L_p и повторных вызовов в каналах L_r , а также среднего числа повторных вызовов в орбите L_o от числа каналов N системы показаны на рис. 5. Из него видно, что для выбранных исходных данных функции L_p и L_o систематически растут, в то время как функция L_r невозрастающая; при этом число повторных вызовов в каналах оказывается существенно меньшим, чем число первичных вызовов в каналах.

Зависимость коэффициента использования каналов от их числа показана на рис. 6. Как и следовало ожидать, функция C_u убывающая, и при этом даже более чем трехкратное увеличение объема орбита почти не влияет на ее значения.

Относительно точности разработанных приближенных формул отметим, что они существенно зависят от размерности модели, от абсолютных значений ее нагрузочных параметров, а также от их соотношения. Для приведенных ранее исходных данных изучаемой модели значения стационарных вероятностей состояний при точном и приближенном подходах очень близки. Поскольку искомые характеристики определяются как маргинальные распределения исходной модели, следует ожидать, что значения этих характеристик в некоторых случаях достаточно сильно отличаются. Однако поведение этих характеристик относительно структурных параметров при точном и приближенном подходах идентично. Более того, численные эксперименты показали, что с ростом размерности модели абсолютная

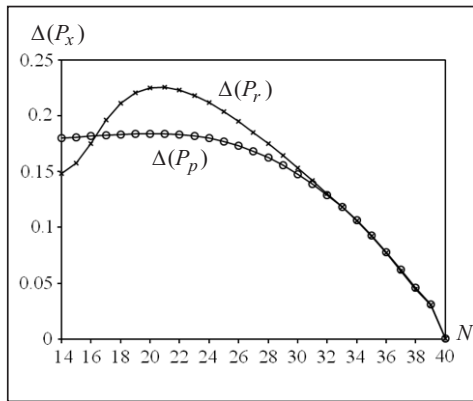


Рис. 7. Зависимость абсолютной погрешности вычисления вероятностей потери первичных и повторных вызовов от числа каналов системы при $R = 3$

Таблица 1. Показатели времени выполнения различных алгоритмов при $R = 3$

N	Время выполнения алгоритмов, мин	
	Приближенный подход	Точный подход
3	0.12	0.14
5	0.34	0.36
10	3.11	4.23
20	5.01	6.35
30	6.32	9.12
40	8.23	13.10

- memory available for all arrays: 2163 MB (2.268e+009 bytes, limited by virtual address space available);
- memory used by MATLAB: 1743 MB (1.828e + 009 bytes);
- physical memory (RAM): 8092 MB (8.485e + 009 bytes).

ЗАКЛЮЧЕНИЕ

В работе изучена многоканальная СМО с мгновенной и отсроченной обратными связями. В ней одна часть вызовов может требовать повторного обслуживания мгновенно после получения первичного обслуживания, а другая — либо возвращаться через случайное время из орбита, либо окончательно покинуть систему. Рассматривается достаточно общая модель, где предполагается, что времена обслуживания первичных и повторных вызовов отличаются. Вероятности ухода вызовов после получения первичного обслуживания, их мгновенного возвращения в каналы или поступления в орбит зависят от числа повторных вызовов в орбите. Изучены системы с конечным и бесконечным размером орбита. Показано, что математической моделью исследуемой СМО является некоторая трехмерная цепь Маркова. Разработаны методы точного и асимптотического анализов характеристик указанной системы. Точный анализ основан на использовании системы уравнений равновесия для вероятностей состояний, а асимптотический — использует принципы фазового укрупнения состояний стохастических систем. Разработан ие-

погрешность вычисления искомых характеристик асимптотически приближается к нулю, что существенно, так как приближенный подход предназначен именно для изучения моделей очень большой размерности.

На рис. 7 показана зависимость поведения абсолютной погрешности вычисления вероятностей потери первичных ΔP_p и повторных ΔP_r вызовов от числа каналов системы.

Как отмечалось ранее, точный подход позволяет исследовать модели с умеренными размерностями их пространства состояний. По данным, приведенным в табл. 1, можно сделать сравнительный анализ времени выполнения алгоритмов точного и приближенного подходов, из которого следует, что алгоритмы приближенного подхода выполняются гораздо быстрее, а алгоритмы точного подхода становятся непригодными для $R = 10$ уже при $N > 38$.

Для выполнения численных экспериментов использовался компьютер со следующими характеристиками:

- memory command in MatLab;
- maximum possible array: 562 MB

(5.890e + 008 bytes, limited by contiguous virtual address space available);

рархический алгоритм фазового укрупнения для расчета стационарного распределения изучаемой трехмерной цепи Маркова.

Предложенный подход можно использовать для исследования аналогичных моделей, где применяются более сложные схемы доступа с приоритетами для разнотипных вызовов, а также моделей, в которых вероятности ухода из системы, возвращения в каналы или поступления в орбит зависят от числа разнотипных вызовов в каналах или их суммы. Не менее актуальны задачи оптимизации изучаемой модели СМО с обратной связью. Эти проблемы являются объектами дальнейших исследований.

СПИСОК ЛИТЕРАТУРЫ

1. Lee M.H., Birukou A., Dudin A., Klimenok V., Kostyukova O., Choe C.H. Queueing model of a single-level single-mediator with cooperation of the agents // *Agent and Multi-Agent Systems: Technology and Applications* / N.T. Nguyen (Ed). — Heidelberg; Dordrecht: Springer, 2007. — P. 447–455.
2. Gnanasambandam N., Lee S., Gautam N., Kumara S.R.T., Peng W., Manikonda V., Brinn M., Greaves M. Reliable MAS performance prediction using queueing models // *Proceeding of the IEEE 1st Symposium on Multi-Agent Security and Survivability*. Philadelphia, Pennsylvania at Drexel University, August 30–31, 2004. — P. 55–64.
3. Gnanasambandam N., Lee S., Kumara S.R.T. An autonomous performance control framework for distributed multi-agent systems: A queueing theory based approach // *Proceeding of AAMAS'05*, July 25–29, 2005, Utrecht, Netherlands. — P. 1313–1314.
4. Takacs L. A single-server queue with feedback // *Bell System Technical Journal*. — 1963. — **42**. — P. 505–519.
5. Takacs L. A queueing model with feedback // *Operations Research*. — 1977. — **11**, N 4. — P. 345–354.
6. Назаров А.А., Моисеева С.П., Морозова А.С. Исследования СМО с повторным обслуживанием и неограниченным числом обслуживающих приборов методом предельной декомпозиции // *Вычислительные технологии*. — 2008. — **13**. — Вып. 5. — С. 88–92.
7. Моисеева С.П., Захорольная И.А. Математическая модель параллельного обслуживания кратных заявок с повторными обращениями // *Автоматика*. — 2011. — **47**, Вып. 6. — С. 51–58.
8. Dudin A.N., Kazimirsky A.V., Klimenok V.I., Breuer L., Krieger U. The queueing model MAP/PH/1/N with feedback operating in a Markovian random environment // *Austrian Journal of Statistics*. — 2005. — **34**, N 2. — P. 101–110.
9. Neuts M.F. *Matrix-geometric solutions in stochastic models: An algorithmic approach*. — Baltimore: John Hopkins University Press, 1981. — 332 p.
10. Artalejo J.R., Gomez-Corral A. *Retrial queueing. A computational approach*. — Heidelberg; Dordrecht: Springer, 2008. — 318 p.
11. Абольников Л.М., Джалалов Е.А. Система массового обслуживания с обратной связью. Принцип двойственности и оптимизации // *Автоматика и телемеханика*. — 1978. — № 1. — С. 17–28.
12. Громов А.И. О двухфазной системе массового обслуживания с обратной связью // *Автоматика и телемеханика*. — 1988. — № 12. — С. 93–97.
13. Wolff R.W. Poisson arrivals see time averages // *Operations Research*. — 1992. — **30**, N 2. — P. 223–231.
14. Pekoz E.A., Joglekar N. Poisson traffic flow in a general feedback // *Journal of Applied Probability*. — 2002. — **39**, N 3. — P. 630–636.
15. Korolyuk V.S., Korolyuk V.V. *Stochastic models of systems*. — Boston: Kluwer, 1999. — 185 p.
16. Ponomarenko L., Kim C.S., Melikov A. *Performance analysis and optimization of multi-traffic on communication networks*. — Heidelberg; Dordrecht; London; New York: Springer, 2010. — 208 p.

Поступила 18.05.2015