УДК 519.872

РАСЧЕТ ХАРАКТЕРИСТИК СМЕШАННОЙ СИСТЕМЫ ОБСЛУЖИВАНИЯ С РЕЗЕРВИРОВАНИЕМ КАНАЛОВ

А.З. МЕЛИКОВ, А.М. ВЕЛИБЕКОВ

Предлагается алгоритм для приближенного расчета характеристик смешанной многоканальной системы обслуживания с двумя типами вызовов. В ней «нетерпеливые» приоритетные вызовы обслуживаются на основе схемы резервирования каналов, а неприоритетные могут ожидать начала обслуживания в конечной и/или бесконечной очереди. Приводятся результаты численных экспериментов.

ВВЕДЕНИЕ

Модели многоканальных систем обслуживания с разнотипными вызовами широко используются в математическом анализе современных беспроводных сетей связи, что объясняется самой сутью этих сетей, так как даже в традиционных беспроводных сотовых сетях передачи речи необходимо различать вызовы двух типов: новые вызовы (o-вызовы) и хэндовер вызовы (h-вызовы). Как известно, в сетях последнего типа при переходе мобильным пользователем границы данной соты он становится хэндовер-вызовом для соседней соты, и если в новой соте имеется хотя бы один свободный канал, то разговор h-вызова возобновляется для него незаметно. В противном случае происходит вынужденное прерывание разговора данного h-вызова.

Поскольку h-вызовы более чувствительны к возможным потерям и задержкам, чем o-вызовы, то зачастую используются различные схемы приоритетного обслуживания h-вызовов. Эти схемы, главным образом, подразумевают использование резервных каналов для h-вызовов и/или рациональную организацию их очереди [1–6].

Вместе с тем, с целью компенсации шансов o-вызовов необходимо организовать их очереди (конечную или бесконечную), сохраняя при этом высокую приоритетность h-вызовов при поступлении в систему, которая обеспечивается за счет резервирования для них определенного числа радиоканалов. Модели последнего типа исследовались в работах [7, 8]. При этом в [7] использован матрично-геометрический подход, а в [8] для расчета характеристик моделей — метод производящих функций.

В настоящей работе для исследования характеристик рассматриваемых моделей используется новый численный метод, с помощью которого возможно эффективное вычисление стационарного распределения двумерных цепей Маркова (ЦМ) [9]. Применение данного подхода позволяет разработать явные формулы для приближенного расчета характеристик исследуемых моделей.

© *А.З. Меликов, А.М. Велибеков*, 2008

МОДЕЛЬ И МЕТОД РАСЧЕТА ЕЕ ХАРАКТЕРИСТИК

На рис. 1 показана схема исследуемой смешанной многоканальной системы, содержащей бесконечный буфер лишь для ожидания в очереди *o*-вызовов [7]. Система является смешанной в том смысле, что вызовы одного типа при определенных ситуациях становятся в очередь, а вызовы другого типа обслуживаются по схеме с явными потерями. Предполагается, что *o*-вызовы (*h*-вызовы) поступают в систему согласно закону Пуассона с интенсивностью $\lambda_o(\lambda_h)$, требуемое время их обслуживания не зависит от типа вызова и распределено экспоненциально со средним μ^{-1}.

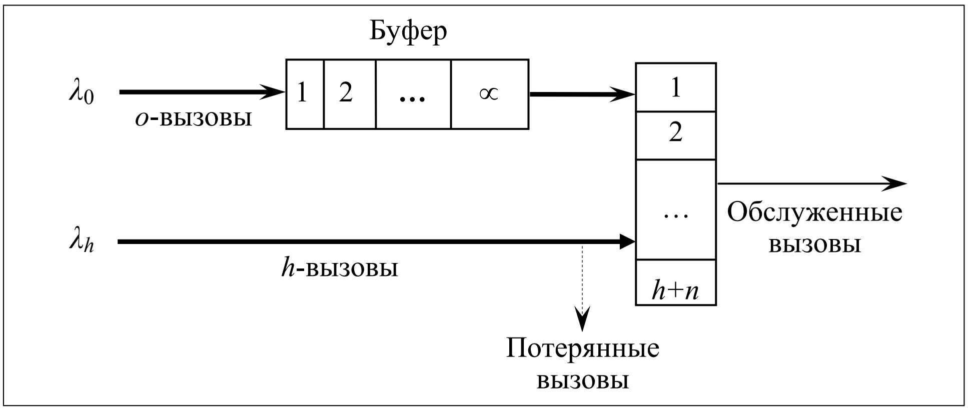

Рис. 1. Структурная схема исследуемой системы

Замечание 1. Идентичность разнотипных вызовов по длительности их обслуживания объясняется отсутствием памяти экспоненциального распределения, так как если в период обслуживания *o*-вызова происходит процедура хэндовер, то оставшееся время обслуживания данного вызова в новой соте (уже в качестве *h*-вызова) также имеет экспоненциальное распределение с тем же средним μ^{-1}.

Обслуживание разнотипных вызовов осуществляется по схеме неизолированного резервирования каналов. Это означает, что поступивший *h*-вызов (приоритетный) принимается при наличии хотя бы одного свободного канала из общего числа $m+n$ каналов. В противном случае *h*-вызов теряется (т.е. *h*-вызовы не буферируются). Поступивший *o*-вызов принимается для обслуживания лишь тогда, когда число свободных каналов больше *n*, где *n* — число резервных каналов. В противном случае *o*-вызов присоединяется к очереди и выбирается из очереди для обслуживания лишь тогда, когда число свободных каналов становится больше *n*. При этом очередь *o*-вызовов обслуживается по схеме FCFS (First Come First Serviced).

Альтернативное описание схемы занятия каналов разнотипными вызовами состоит в следующем [8]. Все $m+n$ каналов разделяются на две группы: первая группа содержит *m*, вторая — *n* каналов. Для обслуживания поступившего *h*-вызова поиск свободного канала сначала осуществляется в первой группе. Если все *m* каналов этой группы заняты, то поиск осуществляется во второй группе. Если все каналы обеих групп заняты, то *h*-вызов теряется. Новые вызовы могут обслуживаться лишь в первой группе. Поэтому если в момент поступления o-вызова все каналы этой группы заняты, то вызов становится в очередь. При освобождении канала первой группы (т.е. после завершения обслуживания одного o- или h-вызова в данной группе) один h-вызов, обслуживаемый в этот момент во второй группе, переключается в первую, независимо от длины очереди o-вызовов (т.е. происходит переназначение h-вызова из второй группы в первую).

Рассмотрим задачу нахождения характеристик описанной системы. При этом под характеристиками понимаются вероятность потери h-вызовов (P_h), средняя длина очереди o-вызовов (L_q^o), а также среднее время ожидания в очереди (W_q^o).

Состояние системы в произвольный момент времени описывается двумерным вектором $\boldsymbol{k}=(k_1,k_2)$, где k_1 указывает число o-вызовов в системе, $k_1=0,1,2,\dots$, и k_2 означает общее число занятых каналов, $k_2=0,1,...,m+n$. Отметим, что фазовое пространство состояний (ФПС) S данной системы (т.е. множество всех возможных состояний) не содержит вектора $\boldsymbol{k}=(k_1,k_2)$, где $k_1>0$, $k_2<m$.

Исходя из механизма функционирования исследуемой системы, заключаем, что элементы производящей матрицы $q(\boldsymbol{k},\boldsymbol{k}')$, $\boldsymbol{k},\boldsymbol{k}'\in S$, соответствующей двумерной ЦМ, определяются из соотношений (рис. 2)

$$q(\boldsymbol{k},\boldsymbol{k}')=\begin{cases}\lambda_o+\lambda_h, & \text{если } k_1=0, \quad k_2\le m-1, \quad \boldsymbol{k}'=\boldsymbol{k}+\boldsymbol{e}_2,\\ \lambda_h, & \text{если } k_2\ge m, \quad \boldsymbol{k}'=\boldsymbol{k}+\boldsymbol{e}_2,\\ \lambda_o, & \text{если } k_2\ge m, \quad \boldsymbol{k}'=\boldsymbol{k}+\boldsymbol{e}_1,\\ k_2\mu, & \text{если } k_2\ne m, \quad \boldsymbol{k}'=\boldsymbol{k}-\boldsymbol{e}_2,\\ m\mu, & \text{если } k_2=m, \quad \boldsymbol{k}'=\boldsymbol{k}-\boldsymbol{e}_1,\\ 0 & \text{в остальных случаях,}\end{cases} \tag{1}$$

где $\mathbf{e}_1=(1,0)$, $\mathbf{e}_2=(0,1)$.

Стационарную вероятность состояния $\boldsymbol{k}\in S$ обозначим $p(\boldsymbol{k})$. Тогда искомые характеристики системы выражаются с помощью стационарного распределения модели

$$P_h=\sum_{k_1=0}^{\infty}p(k_1,m+n), \tag{2}$$

$$L_q^o=\sum_{k_1=1}^{\infty}\sum_{k_2=m}^{m+n}k_1p(k_1,k_2), \tag{3}$$

$$W_q^0=L_q^o/\lambda_o. \tag{4}$$

Следовательно, проблема состоит в нахождении характеристик (2)–(4). Как было отмечено выше, известны достаточно сложные вычислительные процедуры решения этой проблемы. Мы же предлагаем простую численную процедуру ее решения.

Корректное применение разработанной процедуры предполагает $\lambda_h >> \lambda_o$. Отметим, что это допущение всегда выполняется в пикосотах, а также оно вполне приемлемо во многих реальных беспроводных сетях связи [10].

Рассмотрим следующее разбиение ФПС исходной модели:

$$S = \bigcup_{i=0}^{\infty} S_i, \quad S_i \bigcap S_j = \varnothing, \quad i \neq j, \tag{5}$$

где $S_i := \{k \in S : k_1 = i\}$.

Классы микросостояний S_i представляются в виде изолированных укрупненных состояний $< i >$. В исходном ФПС строится функция укрупнения

$$U(\boldsymbol{k}) = < i >, \text{ если } \boldsymbol{k} \in S_i, \; i = 0,1,2,..., \tag{6}$$

определяющая укрупненную модель, которая также является ЦМ с ФПС $\widetilde{S} := \{< i >: \; i = 0,1,2,...\}$.

Элементы производящей матрицы расщепленных моделей с ФПС S_i, обозначаемые $q_i(\boldsymbol{k}, \boldsymbol{k}'), \; \boldsymbol{k}, \boldsymbol{k}' \in S_i$, определяются так (1):

для модели с ФПС S_0

$$q_o(\boldsymbol{k}, \boldsymbol{k}') = \begin{cases} \lambda_o + \lambda_h, & \text{если} \quad k_2 \leq m-1, \quad k_2' = k_2 + 1, \\ \lambda_h, & \text{если} \quad m \leq k_2 \leq m+n-1, \quad k_2' = k_2 + 1, \\ k_2 \mu, & \text{если} \quad k_2' = k_2 - 1, \\ 0 & \text{в остальных случаях;} \end{cases} \tag{7}$$

для модели с ФПС $S_i, \; i \geq 1$

$$q_o(\boldsymbol{k}, \boldsymbol{k}') = \begin{cases} \lambda_h, & \text{если} \quad k_2' = k_2 + 1, \\ k_2 \mu, & \text{если} \quad k_2' = k_2 - 1, \\ 0 & \text{в остальных случаях.} \end{cases} \tag{8}$$

Стационарную вероятность состояния $(i, j) \in S$ внутри расщепленной модели с ФПС S_i обозначим $\rho^i(j)$. Тогда, с помощью (7), (8) находим, что стационарные распределения расщепленных моделей

для $i = 0$

$$\rho^0(j) = \begin{cases} \dfrac{\nu^j}{j!} \rho_0, & \text{если} \quad j = \overline{1, m}, \\ \left(\dfrac{\nu}{\nu_h}\right)^m \dfrac{{\nu_h}^j}{j!} \rho_0, & \text{если} \quad j = \overline{m+1, m+n}; \end{cases} \tag{9}$$

для $i > 0$

$$\rho^i(j) = \frac{m!}{{\nu_h}^m} \frac{{\nu_h}^j}{j!} \rho_1, \quad j = \overline{m+1, m+n}, \tag{10}$$

где

$$\rho_0 = \left(\sum_{i=0}^{m} \frac{\nu'}{i!} + \left(\frac{\nu}{\nu_h} \right)^m \sum_{i=m+1}^{m+n} \frac{\nu_h^i}{i!} \right)^{-1}, \quad \rho_1 = \left(\frac{m!}{\nu_h{}^m} \sum_{i=m}^{m+n} \frac{\nu_h^i}{i!} \right)^{-1}, \quad \nu := \nu_o + \nu_h .$$

Элементы производящей матрицы укрупненной модели, обозначаемые $q(<i>,<i'>)$, $<i>, <i'> \in \tilde{S}$,

$$q(<i>, <i'>) = \begin{cases} \lambda_o \sum_{j=m}^{m+n} \rho^0(j), & \text{если} \quad i = 0, \quad i' = i+1, \\ \lambda_o, & \text{если} \quad i > 0, \quad i' = i+1, \\ m\mu\rho_1, & \text{если} \quad i > 0, \quad i' = i-1, \\ 0 & \text{в остальных случаях.} \end{cases} \tag{11}$$

Из (11) находим условие эргодичности укрупненной модели

$$a := \frac{\nu_o}{m\rho^1(m)} < 1$$

или в явном виде

$$\frac{\nu_o}{m} \frac{m!}{\nu_h^m} \left(\sum_{i=m}^{m+n} \frac{\nu_h^i}{i!} \right) < 1 . \tag{12}$$

При выполнении условия (12) стационарное распределение укрупненной модели $(\pi(<i>) : <i> \in \tilde{S})$

$$\pi(<i>) = a^i b \pi(<0>), \quad i = 1,2,..., \tag{13}$$

где

$$b := \sum_{i=m}^{m+n} \rho^0(i), \quad \pi(<0>) = \frac{1-a}{1-a+ab} . \tag{14}$$

Замечание 2. Очень важно, что условие эргодичности (12) полностью совпадает с аналогичным условием, описанным в работе [7]. Однако в [7] это условие найдено с помощью сложных вероятностных рассуждений, а здесь оно установлено, исходя из эргодичности некоторого одномерного процесса размножения и гибели.

Далее с использованием (9), (10) и (13), (14) приближенно можно записать стационарное распределение исходной модели

$$p(k_1, k_2) \approx \rho^{k_1}(k_2)\pi(<k_1>) . \tag{15}$$

После выполнения необходимых математических преобразований получим следующие приближенные формулы для вычисления характеристик (2) – (4) исследуемой модели:

$$P_h \approx \frac{1}{1-a+ab}((1-a)E_B(\nu_h, m+n) + abE_B{}^T(\nu_h, m)), \tag{16}$$

$$L_q^o \approx \frac{ab}{(1-a+b)(1-a)}, \tag{17}$$

$$W_q^o \approx L_q^o / \lambda_o, \tag{18}$$

где $E_B(\nu_h, m+n)$ — B-формула Эрланга для системы $M/M/m+n$ с нагрузкой ν_h эрл; $E_B{}^T(\nu_h, m)$ — усеченная B-формула Эрланга

$$E_B^T(\nu_h, m) := \frac{\nu_h^{m+n}}{(m+n)!}\left(\sum_{i=m}^{m+n}\frac{\nu_h^i}{i!}\right)^{-1}.$$

Замечание 3. Из формулы (16) видно, что P_h является выпуклой комбинацией двух функций $E_B(\nu_h, m+n)$ и $E_B{}^T(\nu_h, m)$. Следовательно, при любых нагрузках имеет место соотношение

$$E_B(\nu_h, m+n) \le P_h \le E_B{}^T(\nu_h, m). \tag{19}$$

В неравенстве (19) границы достигаются и становятся одинаковыми только в случае $n = 0$, т.е. при отсутствии резервных каналов для h-вызовов. Этот результат вполне ожидаем.

Предложенный метод позволяет определить характеристики данной системы и при наличии лишь ограниченного буфера для ожидания в очереди o-вызовов. Пусть максимально допустимая длина очереди o-вызовов равна N, $N < \infty$. Тогда при любых значениях нагрузочных и структурных параметров в системе существует стационарный режим (т.е. не требуется выполнения условия эргодичности (12)).

Стационарные распределения расщепленных моделей определяются точно так же (9), (10), но в данном случае число таких моделей конечно и равно $N+1$. Применяя описанный выше подход и опуская промежуточные математические преобразования, находим, что для данной модели стационарное распределение укрупненной модели определяется так:

$$\pi_1(<i>) = a^i b \pi_1(<0>), \quad i = \overline{1, N}, \tag{20}$$

где

$$\pi_1(<0>) = \left(1 + ab\frac{1-a^N}{1-a}\right)^{-1}. \tag{21}$$

Следовательно, приближенные значения характеристик (2)–(4) для модели с ограниченной очередью вычисляются следующим образом:

$$P_h(N) \approx \pi_1(<0>)E_B(\nu_h, m+n) + (1-\pi_1(<0>))E_B{}^T(\nu_h, m), \tag{22}$$

$$L_q^o(N) \approx ab\frac{1-a^N(N+1+Na)}{(1-a)^2}\pi(<0>), \tag{23}$$

$$W_q^0(N) \approx \frac{L_q^0(N)}{\lambda_o(1-P_o(N))}, \tag{24}$$

где $P_o(N)$ — вероятность потери o-вызовов

$$P_o(N) \approx \pi_1(N) \text{ или } P_o(N) \approx a^N b \pi_1(<0>). \tag{25}$$

ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

Простота предложенных формул (16)–(18) и (20)–(25) позволяет изучить характеристики исследуемых систем практически во всех допустимых диапазонах изменения их структурных и нагрузочных параметров.

В численных экспериментах для модели с бесконечной очередью o-вызовов (рис. 2) исходные данные выбирались так: $N = 40$, $\lambda_h = 15$, $\mu = 1$.

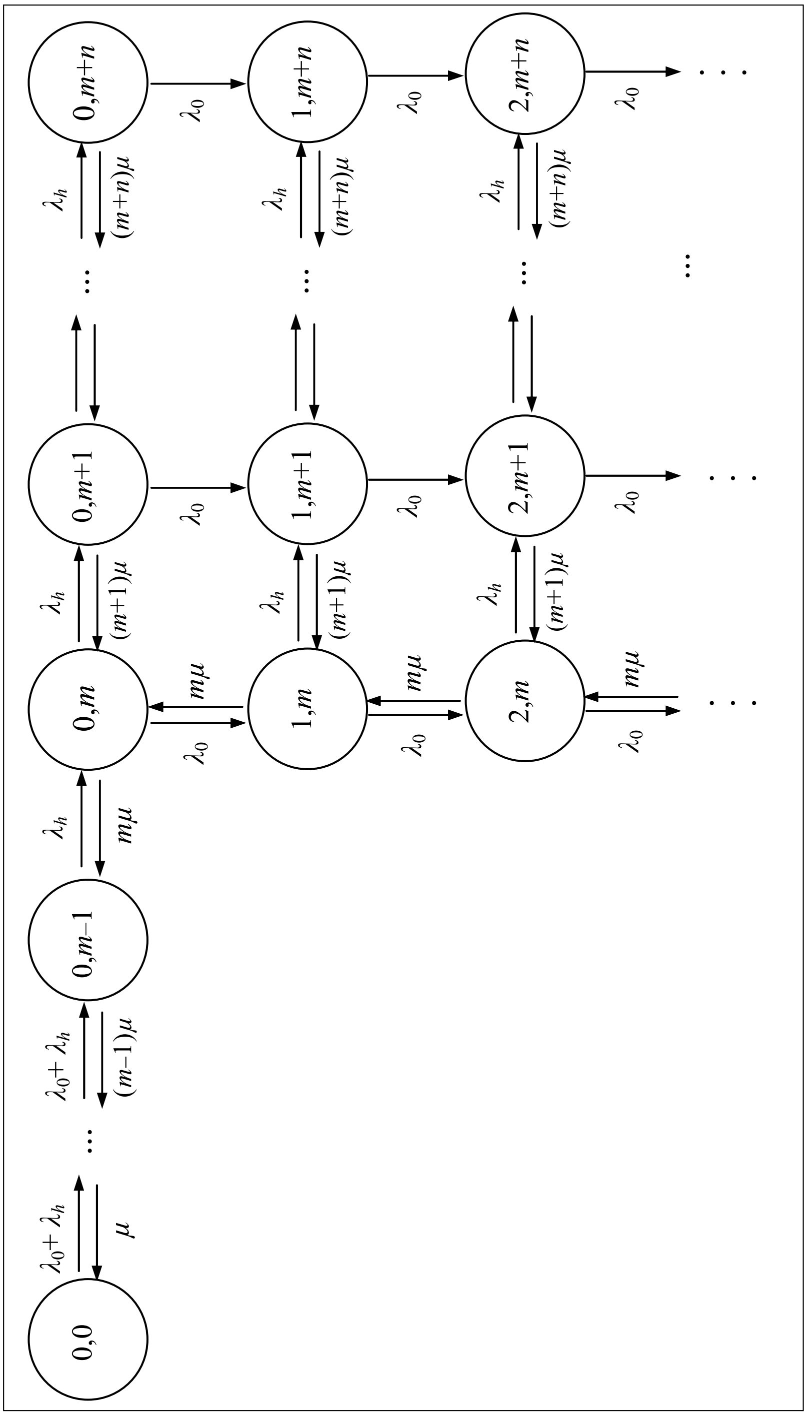

Рис. 2. Граф переходов между состояниями модели с бесконечной очередью o-вызовов

Результаты численных экспериментов полностью подтвердили все теоретические ожидания. Так, функция вероятности потери h-вызовов является убывающей (рис. 3), а функции среднего числа o-вызовов в очереди (рис. 4) и среднего времени их ожидания (рис. 5) — возрастающими относительно числа резервных каналов. При этом все эти функции являются возрастающими относительно интенсивности трафика o-вызовов. Отметим, что для указанных исходных данных свойство эргодичности модели теряется при $n \geq 24$, потому на этих графиках значения n указаны в интервале [1, 23].

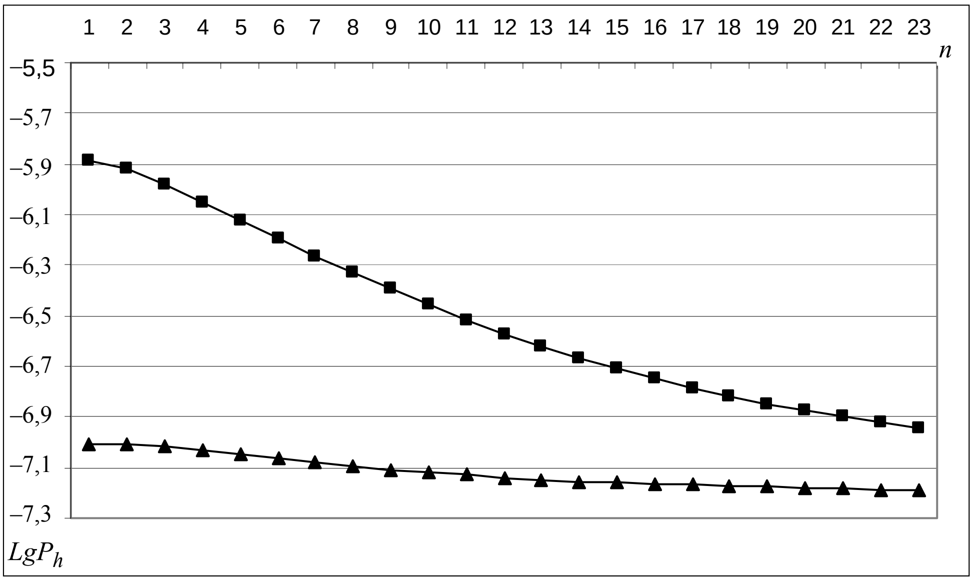

Рис. 3. Зависимость вероятности потери h-вызовов от числа резервных каналов в модели с бесконечной очередью (▲ — λ_0=2; ■ — λ_0=4)

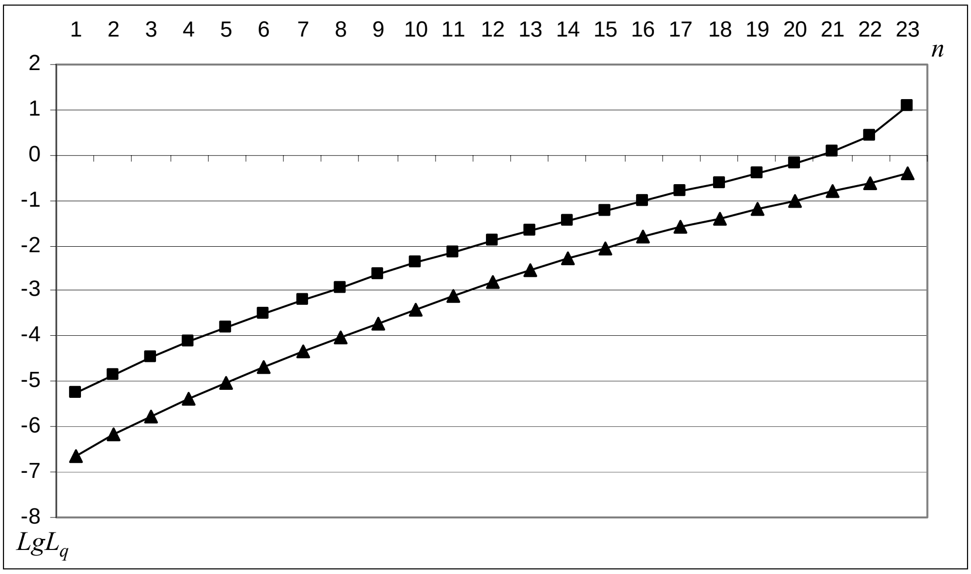

Рис. 4. Зависимость средней длины очереди o-вызовов от числа резервных каналов в модели с бесконечной очередью (▲ — λ_0=2; ■ — λ_0=4)

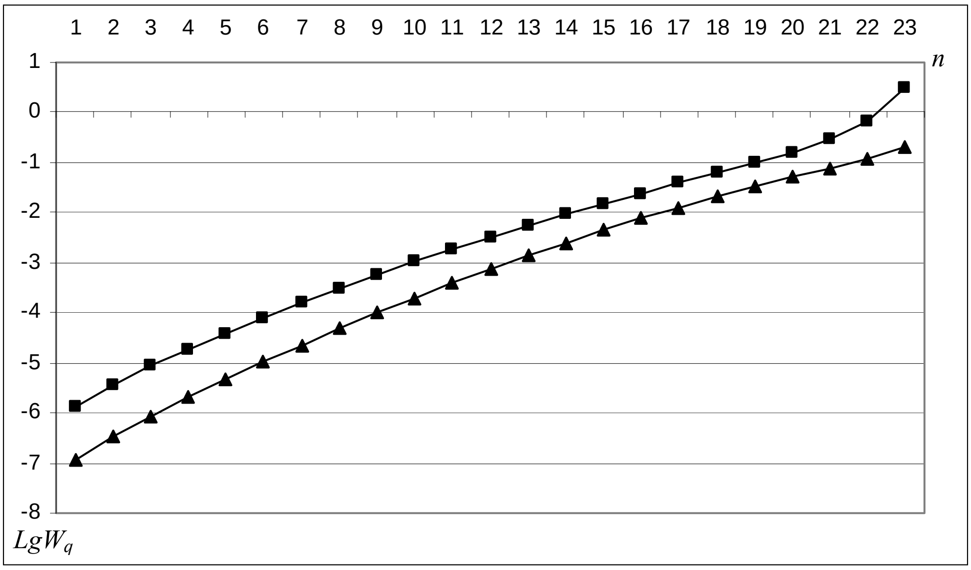

Рис. 5. Зависимость среднего времени ожидания o-вызовов от числа резервных каналов в модели с бесконечной очередью (▲ — λ_0=2; ■ — λ_0=4)

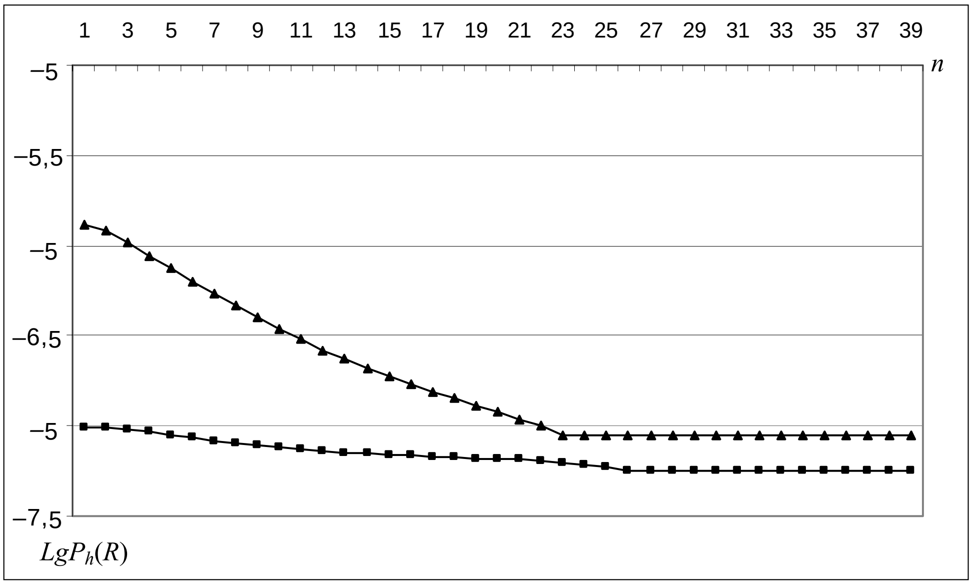

Рис. 6. Зависимость вероятности потери h-вызовов от числа резервных каналов в модели с конечной очередью (▲ — λ_0=4; ■ — λ_0=2)

Некоторые результаты численных экспериментов для модели с конечной очередью o-вызовов показаны на рис. 6−9. В них исходные данные выбирались так: $N=40$, $R=3$, $\lambda_h=15$, $\mu=1$. Для этой модели функция вероятности потери h-вызовов также является убывающей относительно числа резервных каналов (см.рис. 6). Сравнение результатов соответствующих экспериментов показали, что введение ограниченного буфера для o-вызовов почти не влияет на значения вероятности потери h-вызовов (см. рис. 3 и 6). Это объясняется тем, что в выбранных исходных данных интенсивность трафика o-вызовов намного меньше, чем интенсивность трафика h-вызовов. Функция вероятности потери o-вызовов возрастает относительно числа резервных каналов. С ростом же размера буфера, как и следовало ожидать, она уменьшается (рис. 7). С увеличением числа резервных каналов увеличиваются средняя длина очереди o-вызовов (рис. 8) и среднее время их ожидания (рис. 9). При этом функции являются возрастающими относительно интенсивности трафика o-вызовов. Сравнение результатов численных экспериментов показали, что введение ограниченного буфера для o-вызовов почти не влияет на значения средней длины их очереди, особенно при малых интенсивностях этого трафика (см. рис.4 и 8). Те же комментарии относятся к поведению функции «среднее время ожидания o-вызовов» в различных моделях (см. рис. 5 и 9). Эти факты объясняются достаточно низкой интенсивностью o-вызовов по сравнению с интенсивностью h-вызовов.

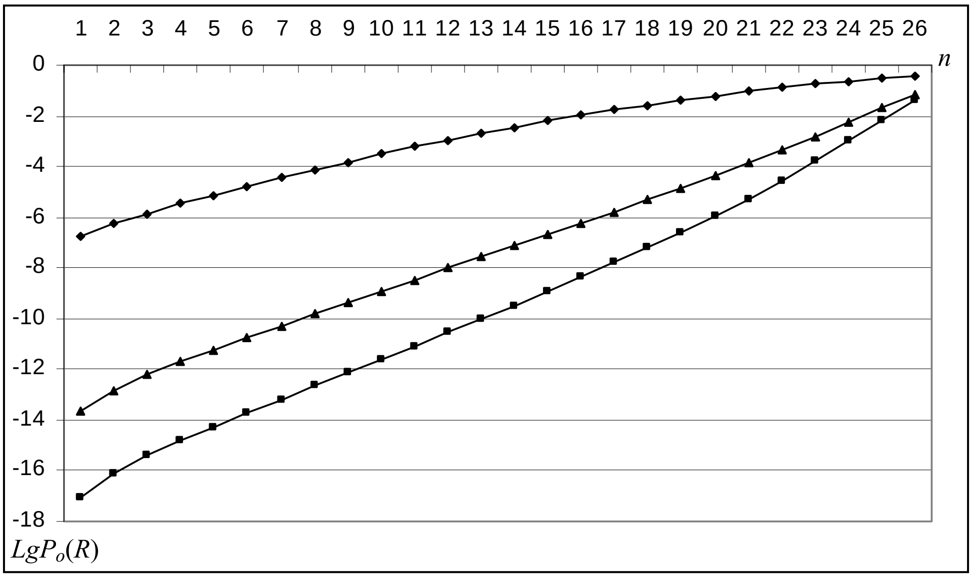

Рис. 7. Зависимость вероятности потери o-вызовов от числа резервных каналов в модели с конечной очередью (▲ — R=7; ■ — R=15; ♦ — R=1)

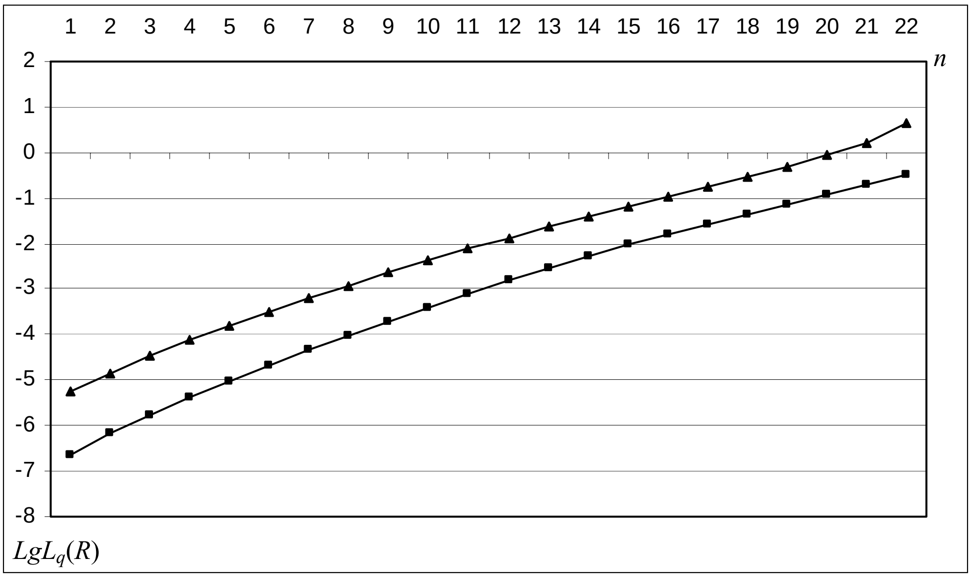

Рис. 8. Зависимость средней длины очереди o-вызовов от числа резервных каналов в модели с конечной очередью (▲ — λ_0=4; ■ — λ_0=2)

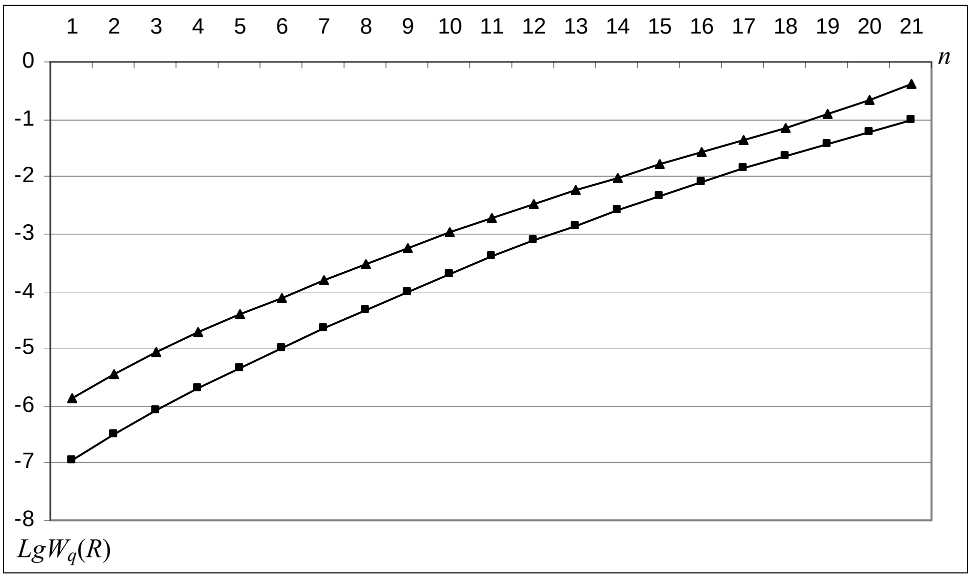

Рис. 9. Зависимость среднего времени ожидания *o*-вызовов от числа резервных каналов в модели с конечной очередью (▲ — λ_0=4; ■ — λ_0=2)

Другая цель выполнения численных экспериментов — оценка точности полученных формул (16)–(18). Так, при выполнении указанного выше допущения относительно соотношений интенсивности трафиков *o*- и *h*-вызовов (т.е. $\lambda_h >> \lambda_o$) наши результаты почти полностью совпадают с результатами работы [7] (ее результаты являются точными). Некоторые сравнения даны в таблице, где в третьем столбце указаны значения P_h, вычисленные с помощью алгоритма [7], а в четвертом — соответствующие значения этой величины, рассчитанные с применением предложенного здесь алгоритма. Аналогичные результаты справедливы и для других характеристик исследуемых моделей.

Сравнительный анализ результатов при $\mu = 2{,}0;\ \lambda_o = 1;\ \lambda_h = 10$

$m+n$	n	P_h [7]	P_h (16)
20	3	8.43E-07	2.98E-07
20	5	7.02E-07	2.97E-07
20	7	5.88E-07	2.96E-07
30	3	1.92E-13	2.79E-14
30	5	1.59E-13	2.75E-14
30	7	1.32E-13	2.72E-14
30	20	4.07E-14	2.71E-14
40	7	1.07E-21	9.42E-23
40	20	3.17E-22	8.54E-23

Важно отметить, что даже при невыполнении указанного допущения относительно соотношений интенсивностей разнотипных вызовов разница между нашими результатами и соответствующими результатами работы [7] в наихудших случаях не превосходит 10^{-5}. Так, для модели с 20-ю каналами $(m+n=20)$ и нагрузочными параметрами $\mu=2$, $\lambda_o=\lambda_h=7$ максимальная разница получена при $n=1$, т.е. для этих параметров значения вероятности потери *h*-вызовов, вычисленные с помощью алгоритма, описанного в работе [7], и нашего алгоритма, равны 1.91E-05 и 4.13E-06, соответственно. Аналогичные результаты получены и при других соотношениях интенсивностей разнотипных трафиков.

ЗАКЛЮЧЕНИЕ

Предложены простые вычислительные процедуры для расчета характеристики смешанных многоканальных систем с классической схемой резервирования каналов для приоритетных *h*-вызов и бесконечной или конечной очередью «терпеливых» *o*-вызовов. Разработанный подход может быть использован и для исследования характеристик аналогичных систем с более сложными механизмами резервирования каналов, а также для исследования моделей с «нетерпеливыми» *o*-вызовами. Эти задачи — предмет дальнейших исследований.

ЛИТЕРАТУРА

1. *Hong D., Rapoport S.S.* Traffic model and performance analysis of cellular mobile radio telephones systems with prioritized and nonprioritized handoff procedures // IEEE Trans. on Vehicular Technology. — 1986. — **35,** № 3. — P. 77–92.
2. *Tekinay S., Jabbari B.* A measurement-based prioritization scheme for handovers in mobile cellular networks // IEEE J. Selected Areas in Commun. — 1992. — **10**, № 8. — P. 1343–1350.
3. *Yoon C.H., Un C.K.* Perfomance of personal portable radio telephone systems with and without guard channels //IEEE J. Selected Areas in Commun. — 1993. — **11**, № 6. — P. 911–917.
4. *Lin Y.B., Mohan S., Noerpel A.* Queueuing priority channel assignment strategies for PCS handoff and initial access // IEEE Trans. on Vehicular Technology. — 1994. — **43,** № 3. — P. 704–712.
5. *Пономаренко Л.А., Меликов А.З., Бабаев А.Т.* Численный метод исследования моделей сотовых сетей связи с ограниченной очередью *h*-вызовов // Проблемы управления и информатики. — 2005. — № 3. — С. 76–88.
6. *Пономаренко Л.А., Меликов А.З., Бабаев А.Т.* Исследование характеристик сетей сотовой связи с ограниченной очередью нетерпеливых *h*-вызовов // Проблемы управления и информатики. —2006. — № 4. — С. 97–107.
7. *Guerin R.* Queueuing-blocking system with two arrival streams and guard channel // IEEE Trans. on Commun. — 1988. — **36**, № 2. — P. 153–163.
8. *Pla V., Casares-Giner V.* A spectral-based analysis of priority channel assignment schemes in mobile cellular communication systems // Int. J. of Wireless Information networks. — 2005. — **12,** № 2. — P. 87–99.
9. *Melikov A.Z., Babayev A.T.* Refined approximations for performance analysis and optimization of queuing model with guard channels for handovers in cellular networks // Computer Communications. — 2006. — **29**, № 9. — P. 1386–1392.
10. *Casares-Giner V.* Integration of dispatch and interconnect traffic in a land mobile trunking systems. Waiting time distribution // Telecommunication Systems. — 2001. — **16,** № 3, 4. — P. 539–554.

Поступила 08.11.2006