



ИЕРАРХИЧЕСКИЙ АЛГОРИТМ ФАЗОВОГО УКРУПНЕНИЯ ДЛЯ АНАЛИЗА ТАНДЕМНЫХ ОТКРЫТЫХ СЕТЕЙ ОБСЛУЖИВАНИЯ

Аннотация. Предложена модель открытой тандемной сети массового обслуживания с ограниченными буферами и обратной связью, в каждый узел которой поступают независимые пуассоновские потоки заявок. Часть заявок после обслуживания в первом узле мгновенно поступает во второй (если в нем имеется свободное место), остальные заявки окончательно покидают сеть. После завершения обслуживания во втором узле заявка покидает сеть либо при наличии свободного места мгновенно поступает в первый узел, либо в орбит и после случайной задержки требует повторного обслуживания. Показано, что математической моделью данной сети является трехмерная цепь Маркова, предложен иерархический алгоритм фазового укрупнения для расчета вероятностей ее состояний. Приведены результаты численных экспериментов.

Ключевые слова: открытая тандемная сеть, обратная связь, трехмерная цепь Маркова, иерархический алгоритм фазового укрупнения.

ВВЕДЕНИЕ

В сетях массового обслуживания (СеМО) каждая заявка после завершения обслуживания в конкретном узле либо мгновенно отправляется в другой узел, либо возвращается обратно согласно маршрутной матрице, т.е. если в сети допускается возврат в прежний узел, то ее можно рассматривать как систему с мгновенной обратной связью. Отметим, что термин «обратная связь» первоначально вводился для однофазных систем массового обслуживания (СМО) [1], которые в последние годы интенсивно изучаются [2, 3].

В работах [4–8] исследуются модели тандемных открытых СеМО с мгновенной обратной связью для анализа интегрированных сотовых сетей связи и беспроводных локальных сетей (Wireless Local Area Network, WLAN). Модель тандемной открытой СеМО с отсроченной обратной связью представлена в работе [9], где предполагается, что перед первой фазой (станцией) сети не имеется места для ожидания, а размер промежуточного буфера между фазами ограничен. В данной работе изучаемая сеть используется для моделирования центра обслуживания вызовов (Call Center).

Основной задачей исследования тандемных открытых СеМО является поиск стационарных вероятностей состояний соответствующих многомерных цепей Маркова, так как показатели качества обслуживания (Quality of Service, QoS) сетей определяются с помощью этих вероятностей. Так, в работах [4–8] изучаемые модели представляют собой двумерные цепи Маркова (2-D Markov Chain, 2-D MC), математическими моделями этих сетей являются квазипроцессы размножения и гибели (Quasi Birth-Death (QBD) Processes). Более детальную информацию о ме-

тодах расчета вероятностей состояний QBD можно найти в [10–14], и поэтому в настоящей статье не будем останавливаться на изложении известных методов.

В работе [9] изучается пятимерная цепь Маркова, для расчета ее стационарного распределения применяется метод, предложенный в [15]. Отметим, что для преодоления проблемы «тирании размерности» при расчете вероятностей состояний QBD, а также цепей более высокой размерности в основном используются матрично-геометрический метод Ньютона [16] и спектральный метод [14], а также их различные модификации [17–21]. Преимущества и недостатки каждого из них хорошо известны. Так, при применении матрично-геометрического метода возникают вычислительные проблемы, связанные с плохой обусловленностью матриц большого размера, используемых в различных этапах алгоритма, а при применении спектрального метода потребуются вычисления собственных чисел и собственных векторов (левых и правых) матриц большого размера, что представляет нетривиальную вычислительную проблему. Кроме того, описанные методы можно применять лишь для расчета вероятностей состояний QBD, которые являются конечными хотя бы по одной компоненте. Отметим, что в последние годы для расчета QBD разработан эффективный приближенный метод [22], основанный на идеях теории фазового укрупнения состояний стохастических систем [23]. Этот метод широко использовался для анализа различных моделей СМО [24–26].

В доступной литературе мало работ, в которых изучаются СеМО с обратными связями обоих типов: мгновенных и отсроченных. Поэтому в настоящей статье исследуются модели тандемных открытых СеМО с мгновенной и отсроченной обратными связями. Учет отсроченной обратной связи приводит к необходимости изучения трехмерной цепи Маркова (3-D MC), для чего разработан эффективный метод вычисления вероятностей состояний 3-D MC в случаях большой размерности ее пространства состояний. Данный метод использует явные формулы для расчета вероятностей состояний соответствующей 3-D MC, и полученные формулы позволяют легко вычислить характеристики изучаемой СеМО. Отметим, что в отличие от обилия публикаций о методах расчета 2-D MC в доступной литературе не имеется работ, посвященных разработке методов расчета 3-D MC.

ОПИСАНИЕ МОДЕЛИ СеМО

Структурная схема изучаемой тандемной открытой СеМО с обратными связями показана на рис. 1. Для простоты получения конечных формул предполагается, что оба узла содержат лишь по одному каналу обслуживания (серверу), но скорости их работы (интенсивности обслуживания) различны, т.е. времена занятия канала заявками в узле i — случайные величины, не зависящие одна от другой и от времени обслуживания в другом узле, а также имеющие одинаковое показательное распределение со средним $1/\mu_i$, $i=1,2$, $\mu_1 \neq \mu_2$. Общая вместимость узла i (т.е. общее число заявок в нем, включая заявки в канале) равна R_i , и в этот узел извне поступает пуассоновский поток заявок с интенсивностью λ_i , $i=1,2$. Если в момент поступления заявки в узел i , $i=1,2$, последний полностью заполнен, то поступившая заявка теряется с вероятностью 1.

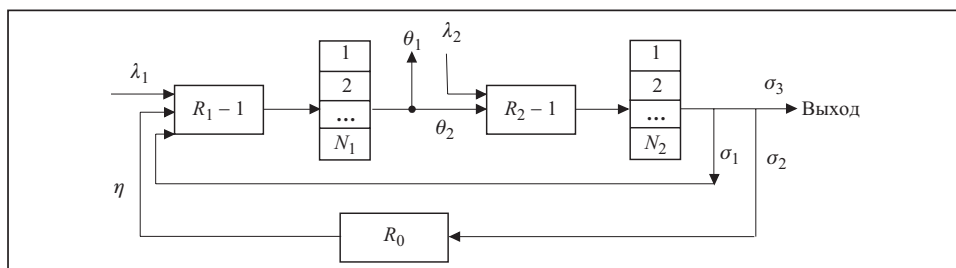


Рис. 1. Структурная схема изучаемой СеМО

После завершения обслуживания в первом узле заявка либо окончательно покидает сеть с вероятностью θ_1 , либо поступает во второй узел с вероятностью $\theta_2 = 1 - \theta_1$. Если в момент завершения обслуживания в первом узле второй узел полностью заполнен, то завершившая обслуживание заявка теряется с вероятностью 1.

После завершения обслуживания конкретной заявки во втором узле возможно принятие следующих решений: данная заявка мгновенно возвращается в первый узел с вероятностью σ_1 ; поступает в орбит с вероятностью σ_2 ; покидает сеть с вероятностью $\sigma_3 = 1 - \sigma_1 - \sigma_2$.

Размер орбита равен R_0 , $0 < R_0 < \infty$. Это означает, что поступившая из второго узла заявка принимается в орбит, если в данный момент число заявок в нем меньше, чем R_0 ; в противном случае она теряется с вероятностью 1. Времена пребывания заявок в орбите независимые случайные величины, показательно распределенные со средним $1/\eta$. Для определенности изложения предположим, что заявки из орбита не являются настойчивыми, т.е. если в момент поступления заявки из орбита все места в первом узле заняты, то поступившая заявка окончательно покидает сеть.

МЕТОД РАСЧЕТА ВЕРОЯТНОСТЕЙ СОСТОЯНИЙ СеМО

Работа изучаемой СеМО описывается 3-D МС, и ее состояния задаются трехмерными векторами $\mathbf{n} = (n_1, n_2, n_3)$, где первая (n_1) и вторая (n_2) компоненты обозначают число заявок в первом и втором узлах соответственно, а третья компонента (n_3) — число заявок в орбите. Множество всех возможных состояний, т.е. фазовое пространство состояний (ФПС) этой 3-D МС, определяется так:

$$S = \{\mathbf{n} : n_1 = 0, 1, \dots, R_1; n_2 = 0, 1, \dots, R_2; n_3 = 0, 1, \dots, R_0\}. \quad (1)$$

Равенство (1) означает, что ФПС модели геометрически представляет собой множество точек с целочисленными координатами внутри параллелепипеда высотой R_0 , основанием которого является прямоугольник со сторонами длиной R_1 и R_2 (рис. 2).

Интенсивность перехода из состояния \mathbf{n} в состояние \mathbf{n}' обозначается $q(\mathbf{n}, \mathbf{n}')$, $\mathbf{n}, \mathbf{n}' \in S$. Эти величины составляют производящую матрицу (Q -матрицу) данной 3-D МС. Из описания работы СеМО заключаем, что они определяются из следующих соотношений:

$$q(\mathbf{n}, \mathbf{n}') = \begin{cases} \lambda_1, & \text{если } \mathbf{n}' = \mathbf{n} + \mathbf{e}_1, \\ \mu_1 \theta_1, & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_1, \\ \mu_1 \theta_2, & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_1 + \mathbf{e}_2, \\ \lambda_2, & \text{если } \mathbf{n}' = \mathbf{n} + \mathbf{e}_2, \\ \mu_2 \sigma_3, & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_2, \\ \mu_2 \sigma_1, & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_2 + \mathbf{e}_1, \\ \mu_2 \sigma_2 & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_2 + \mathbf{e}_3, \\ \eta, & \text{если } n_1 < R_1, \mathbf{n}' = \mathbf{n} - \mathbf{e}_3 + \mathbf{e}_1 \\ & \text{или } n_1 = R_1, \mathbf{n}' = \mathbf{n} - \mathbf{e}_3, \\ 0 & \text{в других случаях.} \end{cases} \quad (2)$$

Здесь и далее \mathbf{e}_i обозначает i -й орт-вектор трехмерного евклидова пространства, $i = 1, 2, 3$.

Из соотношений (2) заключаем, что данная конечная 3-D МС является неприводимой, следовательно, существует ее единственное стационарное распределение. Пусть $p(\mathbf{n})$ обозначает стационарную вероятность состояния $\mathbf{n} \in S$.

Точные значения вероятностей состояний определяются в результате решения соответствующей системы уравнений равновесия (СУР), которые составляются на основе соотношений (2) (вследствие очевидности явный вид СУР в на-

стоящей статье не приводится). Отметим, что, к сожалению, ввиду сложной структуры описанной СУР не удается найти ее аналитического решения и приходится ограничиться численным решением этой СУР, размерность которой равна $(R_0 + 1)(R_1 + 1)(R_2 + 1)$. Однако, как отмечалось ранее, существующие численные алгоритмы решения имеют серьезные вычислительные трудности даже для СеМО небольшой размерности. Поэтому необходимо разработать более эффективные подходы к решению этой проблемы.

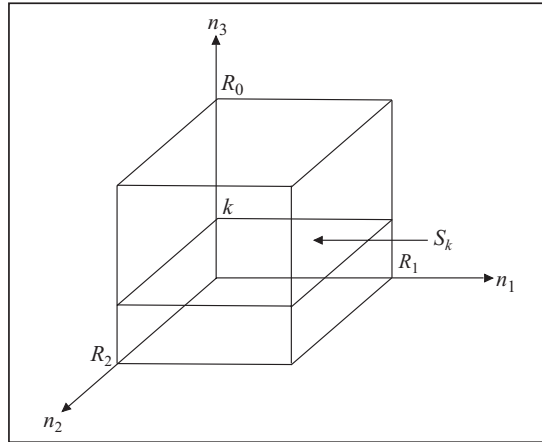


Рис. 2. Вид ФПС модели

Далее предлагается эффективный метод, позволяющий решать поставленную проблему при определенных условиях относительно соотношений нагрузочных параметров сети. Он основан на идеях фазового укрупнения состояний стохастических систем [23]. Следуя терминологии указанной работы, назовем его иерархическим алгоритмом фазового укрупнения (ИАФУ).

Разработанный ИАФУ имеет высокую точность при вычислении вероятностей состояний изучаемой СеМО, если в ней выполняется условие $\sigma_2 \ll \sigma_1 + \sigma_2$. Физический смысл этого условия заключается в том, что после завершения обслуживания во втором узле заявки редко уходят в орбит (по сравнению с уходом из системы и мгновенным возвращением в первый узел). Если указанное условие выполняется, то интенсивности переходов между состояниями, которые находятся на плоскостях, параллельных основанию параллелепипеда, будут намного больше, чем интенсивности переходов между состояниями, находящимися на различных плоскостях (см. рис. 2). Исходя из этого, рассмотрим следующее расщепление пространства состояний (1):

$$S = \bigcup_{k=0}^{R_0} S_k, \quad S_k \cap S_{k'} = \emptyset, \quad \text{если } k \neq k', \quad (3)$$

где $S_k = \{n \in S : n_3 = k\}$, $k = 0, 1, \dots, R_0$. Расщепление (3) означает, что исходное пространство состояний (1) разбито на плоскости (их количество равно $R_0 + 1$), параллельные основанию параллелепипеда (см. рис. 2).

На основе расщепления (3) строится функция укрупнения

$$U_1(n) = \langle k \rangle, \quad \text{если } n \in S_k, \quad (4)$$

где $\langle k \rangle$ обозначает укрупненное состояние, которое включает все состояния из подмножества S_k .

Множество всех укрупненных состояний обозначается $\Omega_1 = \{\langle k \rangle : k = 0, 1, \dots, R_0\}$. Тогда согласно ИАФУ [23] имеем

$$p(n) \approx \rho^k(n_1, n_2) \pi_1(\langle k \rangle), \quad (5)$$

где $\rho^k(n_1, n_2)$ — вероятность состояния (n_1, n_2) внутри расщепленной модели с пространством состояний S_k и $\pi_1(\langle k \rangle)$ — вероятность укрупненного состояния $\langle k \rangle \in \Omega_1$.

Из соотношений (5) можно видеть, что для решения поставленной задачи потребуется нахождение вероятностей состояний двумерных цепей Маркова

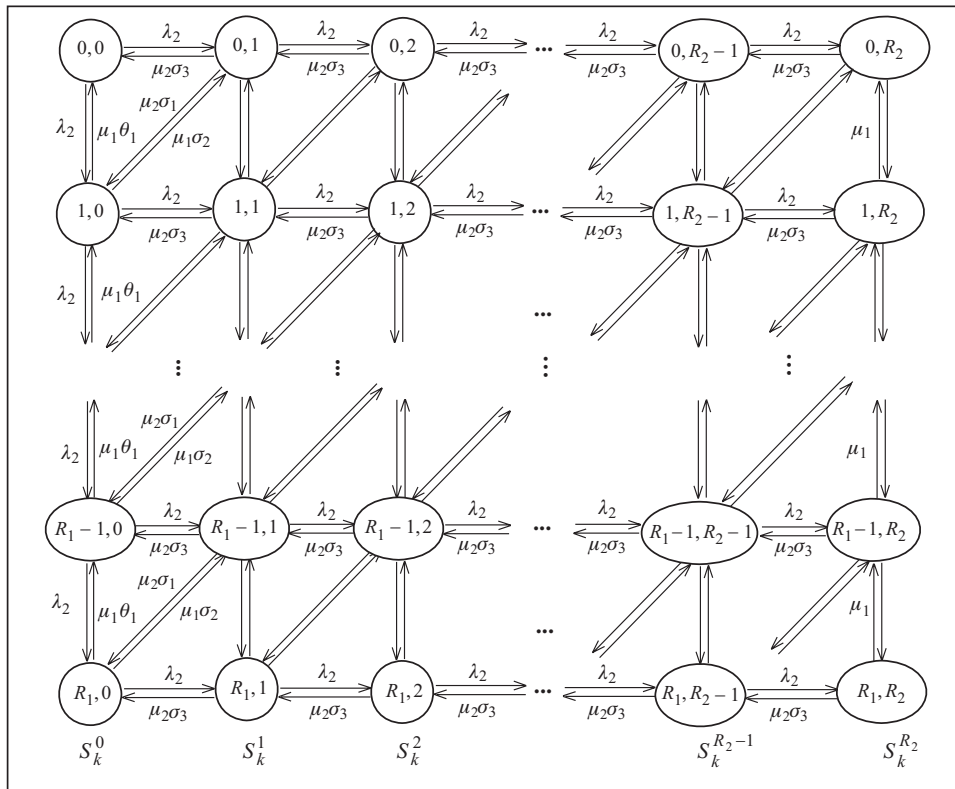


Рис. 3. Граф переходов расщепленной модели с ФПС S_k , $k = 0, 1, \dots, R_0$

(их количество равно $R_0 + 1$) и одной одномерной цепи Маркова. Отметим, что если удастся найти вероятности состояний $\rho^k(n_1, n_2)$ указанных двумерных цепей Маркова, то проблема легко решается; иначе проблема «проклятия размерности» возникает и при нахождении вероятностей состояний этих 2-D МС. Для ее решения предлагается еще раз, но теперь уже к указанным 2-D МС с пространством состояний S_k , $k = 0, 1, \dots, R_0$, применить ИАФУ, т.е. рассматривается иерархия укрупненных моделей.

Важно, что все расщепленные модели с пространством состояний S_k , $k = 0, 1, \dots, R_0$, идентичны 2-D МС (рис. 3), поэтому далее фиксируется индекс k и рассматривается расщепленная модель с пространством состояний S_k .

Предложенный метод позволяет найти приближенные значения вероятностей состояний $\rho^k(n_1, n_2)$ при асимметричных нагрузках заявок в различных узлах, т.е. будем различать два условия: 1) $\lambda_1 \gg \lambda_2$, $\mu_1 \gg \mu_2$; 2) $\lambda_1 \ll \lambda_2$, $\mu_1 \ll \mu_2$.

Замечание 1. Отметим, что указанное ранее допущение асимметричности нагрузок в различных узлах можно заменить и другим условием. Так, например, приведенные два условия можно заменить соответственно условиями $\lambda_1 \gg \max\{\lambda_2, \mu_2\}$ и $\lambda_2 \gg \max\{\lambda_1, \mu_1\}$. При этом важно обеспечить, чтобы интенсивности переходов между состояниями, которые находятся на одинаковых столбцах (строках), были намного большими, чем интенсивности переходов между строками (столбцами) (см. рис. 3).

Вначале рассмотрим условие 1, когда можно провести расщепление пространства состояний S_k по столбцам, т.е. во множестве S_k рассматривается расщепление

$$S_k = \bigcup_{i=0}^{R_2} S_k^i, S_k^i \cap S_k^j = \emptyset, \text{ если } i \neq j, \quad (6)$$

где $S_k^i = \{n \in S_k : n_2 = i\}$, $i = 0, 1, \dots, R_2$.

Аналогично расщеплению (6) подобно (4) строится следующая функция укрупнения:

$$U_2(\mathbf{n}) = \langle i \rangle, \text{ если } \mathbf{n} \in S_k^i, \quad (7)$$

где $\langle i \rangle$ обозначает укрупненное состояние, которое включает все состояния из класса S_k^i . Множество всех укрупненных состояний, полученных в (7), обозначим $\Omega_2 = \{\langle i \rangle : i = 0, 1, \dots, R_2\}$.

Согласно теории ИАФУ имеем

$$\rho^k(n_1, n_2) \approx \rho_{n_2}^k(n_1) \pi_2^k(\langle n_2 \rangle), \quad (8)$$

где $\rho_{n_2}^k(n_1)$ — вероятность состояния (n_1, n_2) внутри расщепленной модели с пространством состояний $S_k^{n_2}$ и $\pi_2^k(\langle n_2 \rangle)$ — вероятность укрупненного состояния $\langle n_2 \rangle \in \Omega_2$.

Рассмотрим задачу нахождения вероятностей состояний внутри расщепленной модели с пространством состояний S_k^i . Внутри этого класса вторая компонента для всех состояний $(n_1, i) \in S_k^i$ постоянная, и поэтому каждое состояние можно задать лишь первой компонентой, т.е. в дальнейшем состояние $(n_1, i) \in S_k^i$ обозначается $n_1, n_1 = 0, 1, \dots, R_1$. Из соотношений (2) получаем, что интенсивности переходов между состояниями n_1 и n_1' расщепленной модели с пространством состояний S_k^i не зависят от индекса k , потому далее в обозначениях вероятностей состояний этой модели верхний индекс k опускается.

Также из (2) и (8) находим, что вероятности $\rho_{n_2}(n_1)$ для значений $n_2 = 0, 1, \dots, R_2 - 1$ совпадают с вероятностями состояний СМО типа $M(\lambda_1) / M(\mu_1 \theta_1) / 1 / R_1$, а при $n_2 = R_2$ эти вероятности совпадают с вероятностями состояний СМО типа $M(\lambda_1) / M(\mu_1) / 1 / R_1$ (здесь и далее используются модифицированные обозначения Кендалла, где в скобках указаны параметры соответствующих распределений). Следовательно, искомые вероятности состояний $\rho_i(j), j = 0, 1, \dots, R_1$, вычисляются так:

$$\rho_i(j) = \begin{cases} \frac{1-v_1}{1-v_1^{R_1+1}} v_1^j, & \text{если } i = 0, 1, \dots, R_2 - 1, \\ \frac{1-v_2}{1-v_2^{R_1+1}} v_2^j, & \text{если } i = R_2, \end{cases} \quad (9)$$

где $v_1 = \lambda_1 / \mu_1 \theta_1, v_2 = \lambda_1 / \mu_1$.

Замечание 2. При вычислении вероятностей состояний расщепленной модели с пространством состояний $S_{R_0}^i$ следует учитывать, что $\sigma_2 = 0$.

После определенных алгебраических преобразований находим, что интенсивности переходов между укрупненными состояниями $\langle i \rangle$ и $\langle j \rangle, \langle i \rangle, \langle j \rangle \in \Omega_2$, определяются так:

$$q(\langle i \rangle, \langle j \rangle) = \begin{cases} \lambda_2 + \mu_1 \theta_2 (1 - \rho_i(0)), & \text{если } i = 0, 1, \dots, R_2 - 1, j = i + 1, \\ \mu_2 (\sigma_3 + \sigma_1 (1 - \rho_i(R_1))), & \text{если } i = 1, \dots, R_2, j = i - 1, \\ 0 & \text{в других случаях.} \end{cases} \quad (10)$$

Следовательно, из соотношений (10) получим формулу для вычисления вероятностей укрупненных состояний $\pi_2(\langle n_2 \rangle), \langle n_2 \rangle \in \Omega_2$:

$$\pi_2(\langle n_r \rangle) = \prod_{i=0}^{n_2-1} \frac{q(\langle i \rangle, \langle i+1 \rangle)}{q(\langle i+1 \rangle, \langle i \rangle)} \pi_2(\langle 0 \rangle), \quad n_2 = 1, \dots, R_2, \quad (11)$$

где $\pi_2(\langle 0 \rangle)$ находится из условия нормировки, т.е. $\sum_{n_2=0}^{R_2} \pi_2(\langle n_2 \rangle) = 1$.

С учетом соотношений (2), (9) и (11) после определенных алгебраических преобразований находим, что интенсивности переходов между классами S_k и $S_{k'}$ определяются так:

$$q(S_k, S_{k'}) = \begin{cases} \Lambda_1, & \text{если } k' = k + 1, \\ k\eta, & \text{если } k' = k - 1, \\ 0 & \text{в других случаях,} \end{cases} \quad (12)$$

где $\Lambda_1 = \mu_2 \sigma_2 (1 - \pi_2(\langle 0 \rangle))$.

Иными словами, из соотношений (12) получается, что вероятности укрупненных состояний $\pi_1(\langle k \rangle)$, $\langle k \rangle \in \Omega_1$, находятся из классических формул Эрланга для СМО типа $M(\Lambda_1)/M(\eta)/R_0/0$, т.е.

$$\pi_1(\langle k \rangle) = \frac{\varphi^k}{k!} \left(\sum_{i=0}^{R_0} \frac{\varphi^i}{i!} \right)^{-1}, \quad k = 1, \dots, R_0, \quad (13)$$

где $\varphi = \Lambda_1 / \eta$.

Окончательно с учетом формул (9), (11) и (13) искомые вероятности состояний изучаемой 3-D MC определяются так:

$$p(n_1, n_2, n_3) \approx \rho_{n_2}(n_1) \pi_2(\langle n_2 \rangle) \pi_1(\langle n_3 \rangle). \quad (14)$$

Аналогично изучается условие 2, т.е. когда выполняются соотношения $\lambda_1 \ll \lambda_2$, $\mu_1 \ll \mu_2$. При этом следует провести расщепление пространства состояний S_k по строкам, т.е. рассматривать следующее расщепление в пространстве состояний S_k (см. рис. 3):

$$S_k = \bigcup_{i=0}^{R_1} S_k^i, \quad S_k^i \cap S_k^j = \emptyset, \quad \text{если } i \neq j, \quad (15)$$

где $S_k^i = \{n \in S_k : n_1 = i\}$, $i = 0, 1, \dots, R_1$.

Замечание 3. Для удобства изложения в (15) приняты обозначения, аналогичные используемым в условии 1.

Далее выполняются все этапы описанного алгоритма. Не повторяя их, кратко изложим ключевые моменты вычисления вероятностей состояний. Так, вероятности состояний внутри всех классов S_k^i совпадают с вероятностями состояний СМО типа $M(\lambda_2)/M(\mu_2 \sigma_3)/1/R_2$. При этом интенсивности переходов $q(\langle i \rangle, \langle j \rangle)$, $i, j \in \{0, 1, \dots, R_1\}$, определяются следующим образом:

$$q(\langle i \rangle, \langle j \rangle) = \begin{cases} \lambda_1 + \mu_2 \sigma_1 (1 - \rho(0)), & \text{если } j = i + 1, \\ \mu_1 (\rho(R_2) + \theta_1 (1 - \rho(R_2))), & \text{если } j = i - 1, \\ 0 & \text{в других случаях,} \end{cases} \quad (16)$$

где $\rho(0)$ и $\rho(R_2)$ — вероятности того, что СМО типа $M(\lambda_2)/M(\mu_2 \sigma_3)/1/R_2$ является пустой и полностью заполненной соответственно.

Следовательно, из соотношений (16) находим, что вероятности укрупненных состояний на второй ступени иерархии $\pi_2(\langle n_1 \rangle)$, $n_1 \in \{0, 1, \dots, R_1\}$, определяются как вероятности состояний СМО типа $M(\lambda_1 + \mu_2 \sigma_1 (1 - \rho(0)))/M(\mu_1 (\rho(R_2) + \theta_1 (1 - \rho(R_2))))/1/R_1$.

Интенсивности переходов между классами S_k и $S_{k'}$ в этом случае определяются аналогично (12), однако величина Λ_1 заменяется величиной $\Lambda_2 = \mu_2 \sigma_2 (1 - \rho(0))$. Иными словами, вероятности укрупненных состояний на первой ступени иерархии $\pi_1(\langle k \rangle)$, $k \in \{0, 1, \dots, R_0\}$, определяются с помощью классических формул Эрланга (13).

ПОКАЗАТЕЛИ КАЧЕСТВА ОБСЛУЖИВАНИЯ

После определения вероятностей состояний 3-D МС можно вычислить точные значения показателей качества обслуживания изучаемой СМО. Так, поскольку поток заявок в каждый узел пуассоновский, вероятности потери заявок в первом (P_1) и во втором (P_2) узле определяются из формул

$$P_i = \sum_{n \in S} p(n) \delta(n_i, R_i), \quad i=1, 2, \quad (17)$$

где $\delta(i, j)$ — символы Кронекера.

Поступившая из орбита заявка покидает сеть повторно необслуженной (т.е. теряется), если в этот момент все места в первом узле заняты. Следовательно, вероятность потери заявок из орбита (P_0) определяется так:

$$P_0 = \sum_{n \in S} p(n) \delta(n_1, R_1) (1 - \delta(n_3, 0)). \quad (18)$$

Точные значения средней длины очереди заявок в первом (L_1) и во втором (L_2) узлах определяются как математические ожидания соответствующих дискретных случайных величин:

$$L_i = \sum_{j=1}^{R_i} j \Phi_i(j), \quad (19)$$

где $\Phi_i(j) = \sum_{n \in S} p(n) \delta(n_i, j)$, $i=1, 2$.

С учетом (14) для вычисления приближенных значений указанных выше показателей QoS после определенных алгебраических преобразований получим следующие формулы:

— для условия $\lambda_1 \gg \lambda_2$, $\mu_1 \gg \mu_2$ имеем

$$P_1 \approx \rho_0(R_1)(1 - \pi_2(\langle R_2 \rangle)) + \rho_{R_2}(R_1)\pi_2(\langle R_2 \rangle), \quad (20)$$

$$P_2 \approx \pi_2(\langle R_2 \rangle), \quad (21)$$

$$P_0 \approx (\rho_0(R_1)(1 - \pi_2(\langle R_2 \rangle)) + \rho_{R_2}(R_1)\pi_2(R_2))(1 - \pi_1(\langle 0 \rangle)), \quad (22)$$

$$L_1 \approx \sum_{k=1}^{R_1} k(\rho_0(k)(1 - \pi_2(\langle R_2 \rangle)) + \rho_{R_2}(k)\pi_2(\langle R_2 \rangle)), \quad (23)$$

$$L_2 \approx \sum_{k=1}^{R_2} k\pi_2(\langle k \rangle); \quad (24)$$

— для условия $\lambda_1 \ll \lambda_2$, $\mu_1 \ll \mu_2$ имеем

$$P_1 \approx \pi_2(\langle R_1 \rangle), \quad (25)$$

$$P_2 \approx \rho(R_2), \quad (26)$$

$$P_0 \approx \pi_2(\langle R_1 \rangle)(1 - \pi_1(\langle 0 \rangle)), \quad (27)$$

$$L_1 \approx \sum_{k=1}^{R_1} k\pi_2(\langle k \rangle), \quad (28)$$

тогда L_2 определяется как средняя длина очереди в СМО $M(\lambda_2)/M(\mu_2\sigma_3)/1/R_2$, т.е.

$$L_2 = \begin{cases} \frac{\omega}{1-\omega} - \frac{R_2+1}{1-\omega^{R_2+1}} \omega^{R_2+1}, & \text{если } \omega \neq 1, \\ \frac{R_2}{2}, & \text{если } \omega = 1, \end{cases} \quad (29)$$

где $\omega = \lambda_2 / \mu_2\sigma_3$.

ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

Рассмотрим результаты некоторых численных экспериментов, проведенных с помощью предложенных алгоритмов. Последние позволяют изучать поведение показателей QoS сети при изменении ее структурных параметров R_0, R_1, R_2 , нагрузочных параметров $\lambda_1, \lambda_2, \mu_1, \mu_2, \eta$, а также вероятностных параметров $\theta_1, \theta_2, \sigma_1, \sigma_2, \sigma_3$, которые определяют поведение заявок после завершения их обслуживания в различных фазах сети. Поскольку число этих параметров достаточно большое, далее для краткости изложения рассмотрим результаты, иллюстрирующие поведение описанных показателей QoS лишь относительно изменения размера первого узла сети.

Исходные параметры сети следующие:

$$R_0 = 10, R_2 = 50, \lambda_1 = 0.2, \lambda_2 = 0.01, \mu_2 = 0.02, \\ \eta = 0.05, \theta_1 = 0.4, \sigma_1 = 0.3, \sigma_2 = 0.05.$$

Рассматриваются две серии экспериментов: в первой $\mu_1 = 0.1$, а во второй $\mu_1 = 0.3$. Выбранные значения исходных данных соответствуют условию $\lambda_1 \gg \lambda_2$, $\mu_1 \gg \mu_2$, т.е. вычисления показателей QoS осуществляются с помощью формул (20)–(24).

Соответствующие графики показаны на рис. 4–8, где \times и \circ обозначены кривые, соответствующие изменениям показателя QoS в первой и второй серии экспериментов соответственно.

Из рис. 4 видно, что во второй серии экспериментов значения вероятностей потери заявок в первом узле существенно меньше, чем в первой серии экспериментов. При этом лишь при малых значениях параметра R_1 показатель QoS уменьшается с очень малой скоростью, и уже при $R_1 \geq 20$ в обеих сериях он становится почти постоянным, а именно, в первой серии при $R_1 \geq 20$ имеем $P_1 = 0.5668$, а во второй серии экспериментов — $P_1 = 0.0039$. Обратный результат наблюдается для вероятности потери заявок во втором узле, т.е. значения этого показателя QoS во второй серии экспериментов существенно больше, чем в первой серии (см. рис. 5). Этого следовало ожидать, так как увеличение интенсивности обслуживания заявок в первом узле приводит к уменьшению вероятности потери заявки в нем, но при этом увеличивается интенсивность поступления заявок во второй узел, и тем самым увеличивается вероятность потери заявок в нем. Увеличение размера буфера в первом узле почти не влияет на значения вероятности потери заявок во втором узле, т.е. в первой серии при $R_1 \geq 20$ имеем $P_2 = 0.7771$, а во второй серии экспериментов — $P_2 = 0.9012$.

Зависимость вероятности потери заявок из орбита от размера буфера первого узла показана на рис. 6. Из данного рисунка, а также из рис. 4 видно, что по-

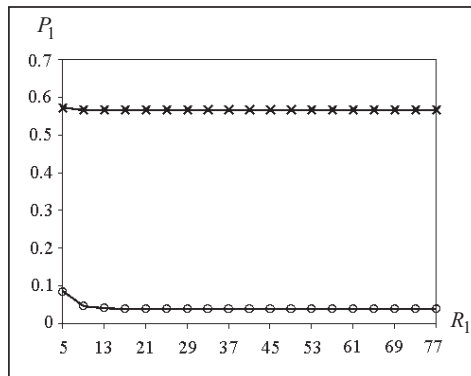


Рис. 4. Зависимость вероятности потери заявок в первом узле от размера его буфера

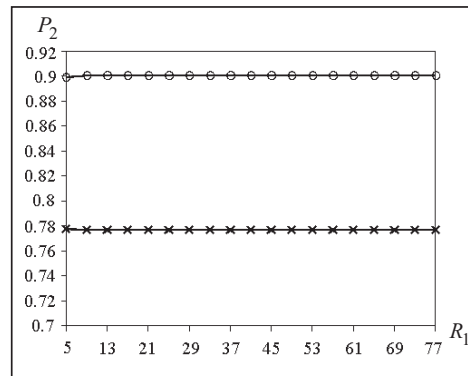


Рис. 5. Зависимость вероятности потери заявок во втором узле от размера буфера первого узла

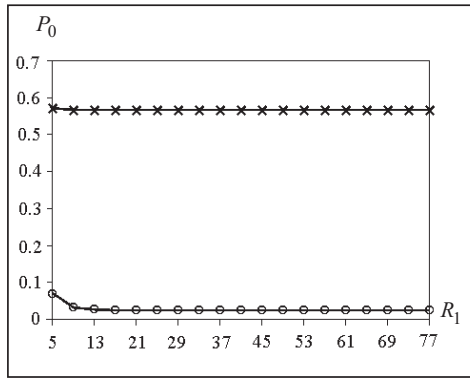


Рис. 6. Зависимость вероятности потери заявок из орбита от размера буфера первого узла

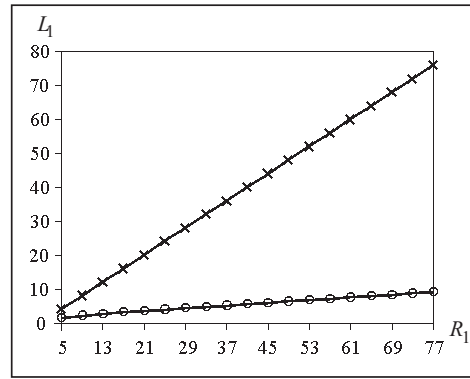


Рис. 7. Зависимость среднего числа заявок в первом узле от размера его буфера

ведение рассматриваемого показателя идентично поведению вероятности потери заявок в первом узле. Этого следовало ожидать (см. формулы (20) и (22)), причем для всех значений параметра R_1 имеет место соотношение $P_0 < P_1$. Отметим, что в первой серии экспериментов разность $P_1 - P_0$ меньше, чем во второй серии, т.е. в первой серии экспериментов указанная разность приблизительно равна 10^{-4} , а во второй — 10^{-2} . Это объясняется тем, что вероятность того, что орбит пустой, в первой серии экспериментов больше, чем во второй (см. также формулы (20) и (22)).

Действительно, с увеличением интенсивности обслуживания заявок в первом узле увеличивается и интенсивность поступления заявок во второй узел, и следовательно, увеличивается интенсивность загрузки орбита.

Из рис. 7 видно, что во второй серии экспериментов среднее число заявок в первом узле существенно меньше, чем в первой серии. Это ожидаемый результат, так как увеличение интенсивности заявок в первом узле приводит к уменьшению среднего числа заявок в нем, при этом скорость роста указанного показателя в первой серии экспериментов достаточно высока относительно изменения размера буфера первого узла. Вместе с тем в обеих сериях экспериментов среднее число заявок во втором узле почти одинаковое и постоянно, т.е. второй узел почти полностью загружен (см. рис. 8).

Целью проведения численных экспериментов также была оценка точности предложенных (приближенных) формул для вычисления вероятностей состояния сети, так как показатели QoS представляют собой некоторые линейные формы этих вероятностей (см. формулы (17)–(19)), и поэтому невозможно гарантировать высокой точности их вычисления. Отметим, что теоретически высокая точность предложенных формул для моделей СеМО большой размерности гарантируется предельными теоремами теории фазового укрупнения состояний марковских цепей [23]. Вместе с тем эта проблема исследована с помощью численных экспериментов, при этом их точность оценивается нормой $\Delta = \max_{n \in S} |p(n) - \tilde{p}(n)|$, где $\tilde{p}(n)$ — приближенное значение вероятности состояния $n \in S$. Точные значения

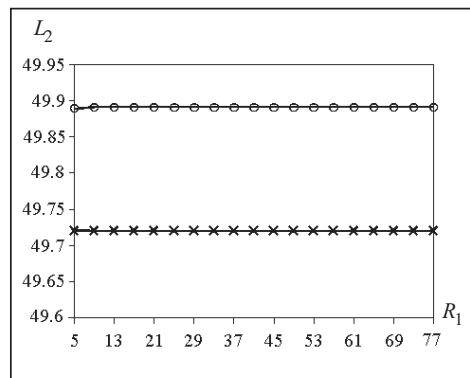


Рис. 8. Зависимость среднего числа заявок во втором узле от размера буфера первого узла

вероятностей состояний для моделей умеренной размерности определяются из соответствующей СУР. Проведенные эксперименты показали, что в широком интервале изменения параметров сети эта норма была в допустимых пределах, принятых в инженерных расчетах. Более того, установлено, что с ростом размерности сети эта норма приближается к нулю. Последнее очень важно, так как предложенный приближенный подход предназначен именно для сетей большой размерности.

Наконец, отметим важность условий $\lambda_1 \gg \lambda_2$, $\mu_1 \gg \mu_2$ и $\lambda_1 \ll \lambda_2$, $\mu_1 \ll \mu_2$ при выборе надлежащих приближенных вычислительных алгоритмов. Значения изучаемых показателей QoS найдены при выбранных исходных данных с помощью формул (25)–(29) (как описано ранее, для этих исходных данных выполняется условие 1). Сравнительный анализ результатов численных экспериментов показал, что абсолютные значения разности значений показателей P_0 и P_1 , вычисленных с помощью различных алгоритмов, находятся в пределах [0.01; 0.06] и [0.03; 0.08] соответственно. Вместе с тем разность значений остальных показателей, вычисленных с помощью различных алгоритмов, слишком большая, т.е. абсолютные значения разности значений показателей P_2 и L_2 — постоянные величины и примерно равны 0.8 и 46 соответственно, а аналогичная величина для показателя L_1 изменяется в пределах [2; 66]. Иными словами, выбор алгоритма вычисления показателей QoS существенно зависит от того, какое из приведенных условий между нагрузочными параметрами сети выполняется.

ЗАКЛЮЧЕНИЕ

В данной работе предложен иерархический алгоритм фазового укрупнения для анализа тандемных открытых сетей обслуживания с единственными каналами в каждой фазе и ограниченными буферами перед ними. В отличие от классических моделей подобного типа в настоящей статье предполагается, что после завершения обслуживания во втором узле заявка может поступать в первый узел не только мгновенно, но и после определенной задержки в орбите ограниченного размера. Отметим, что предложенный подход можно использовать для изучения моделей с многоканальными узлами и бесконечными размерами буферов в каждой фазе и бесконечным размером орбита.

СПИСОК ЛИТЕРАТУРЫ

1. Takacs L. A single-server queue with feedback // Bell System Technical Journal. — 1963. — **42**. — P. 505–519.
2. Melikov A.Z., Ponomarenko L.A., Rustamov A.M. Methods for analysis of queueing models with instantaneous and delayed feedbacks // Communications in Computer and Information Sciences. — 2015. — **564**. — P. 185–199.
3. Koroliuk V.S., Melikov A.Z., Ponomarenko L.A., Rustamov A.M. Methods for analysis of multi-channel queueing models with instantaneous and delayed feedbacks // Cybernetics and System Analysis. — 2016. — **52**, N 1. — P. 58–70.
4. Gemikonakli O., Ever E., Kocyigit A. Approximate solution for two stage open networks with Markov-modulated queues minimizing the state space explosion problem // Journal of Computational and Applied Mathematics. — 2009. — **223**, N 1. — P. 519–533.
5. Kirsal Y., Gemikonakli E., Ever E., Mapp G., Gemikonakli O. An analytical approach for performance analysis of handoffs in the next generation integral cellular networks and WLANs // Proc. of 19th IEEE International Conference on Computer Communications and Networks. — Zurich, Switzerland, Aug. 2–5, 2010. — P. 1–6. — DOI: 10.1109/ICCCN.2010.5560093.

6. Ever E., Gemikonakli O., Kocyigit A., Gemikonakli E. A hybrid approach to minimize state explosion problem for the solution of two stage tandem queues // *Journal of Network and Computer Applications*. — 2013. — **36**. — P. 908–926.
7. Kirsal Y., Ever E., Kocyigit A., Gemikonakli O., Mapp G. A generic analytical modeling approach for performance evaluation of the handover schemes in heterogeneous environments // *Wireless Personal Communications*. — 2014. — **79**. — P. 1247–1276.
8. Kirsal Y., Ever E., Kocyigit A., Gemikonakli O., Mapp G. Modeling and analysis of vertical handover in highly mobile environments // *Journal of Supercomputing*. — 2015. — **71**. — P. 4352–4380.
9. Kim C.S., Klimenok V.I., Dudin A.N. Priority tandem queueing system with retrials and reservation of channels as a model of call center // *Computers and Industrial Engineering*. — 2016. — **96**. — P. 61–71.
10. Ciardo G., Smirni E. ETAQA: an efficient technique for the analysis of QBD-processes by aggregation // *Performance Evaluation*. — 1999. — **36–37**. — P. 71–93.
11. Ciardo G., Mao W., Riska A., Smirni E. ETAQA-MG1: an efficient technique for the analysis of a class M/G/1-type processes by aggregation // *Performance Evaluation*. — 2004. — **57**. — P. 235–260.
12. Mitrani I., Chakka R. Spectral expansion solution for a class of Markov models: application and comparison with the matrix-geometric method // *Performance Evaluation*. — 1995. — **23**. — P. 241–260.
13. Mitrani I. Approximate solutions for heavily loaded Markov-modulated queues // *Performance Evaluation*. — 2005. — **62**. — P. 117–131.
14. Chakka R. Spectral expansion solution for some finite capacity queues // *Annals of Operations Research*. — 1998. — **79**. — P. 27–44.
15. Klimenok V.I., Dudin A.N. Multi-dimensional asymptotically quasi-Teoplitz Markov chains and their application in queueing theory // *Queueing Systems*. — 2006. — **54**. — P. 245–259.
16. Neuts M.F. *Matrix-geometric solutions in stochastic models: An algorithmic approach*. — Baltimore: John Hopkins Univ. Press, 1981. — 332 p.
17. Hung T.T., Do T.V. Computational aspects for steady state analysis of QBD processes // *Periodica Polytechnica. Ser. Electrical Engineering*. — 2001. — **44**, Iss. 2. — P. 179–200.
18. Baumann H., Sandmann W. Numerical solution of level dependent QBD processes // *Procedia Computer Science*. — 2010. — **1**. — P. 1555–1563.
19. Servi L.D. Algorithmic solution of two-dimensional BDP with application to capacity planning // *Telecommunication Systems*. — 2002. — **21**, Iss. 2–4. — P. 205–212.
20. Ye J., Li S.Q. Folding algorithm: A computational method for finite QBD processes with level dependent transitions // *IEEE Transactions on Communications*. — 1994. — **42**. — P. 625–639.
21. De Nitto Persone V., Grassi V. Solution of finite QBD processes // *Applied Probability*. — 1996. — **33**. — P. 1003–1010.
22. Ponomarenko L., Kim C.S., Melikov A. *Performance analysis and optimization of multi-traffic on communication networks*. — Heidelberg; Dortrecht; London; New York: Springer, 2010. — 208 p.
23. Korolyuk V.S., Korolyuk V.V. *Stochastic models of systems*. — Boston: Kluwer, 1999. — 185 p.
24. Liang C., Luh H. Cost estimation queueing model for large-scale file delivery service // *International Journal of Electronic Commerce Studies*. — 2011. — **2**, Iss. 1. — P. 19–34.
25. Liang C., Luh H. Optimal services for content delivery based on business priority // *Journal of the Chinese Institute of Engineers*. — 2013. — **36**, Iss. 4. — P. 422–440.
26. Liang C., Luh H. Efficient method for solving a two-dimensional Markov chain model for call centers // *Industrial Management & Data Systems*. — 2015. — **115**, Iss. 5. — P. 901–922.

Надійшла до редакції 12.04.2016

А.З. Меліков, Л.А. Пономаренко, А.М. Рустамов
ІЄРАРХІЧНИЙ АЛГОРИТМ ФАЗОВОГО УКРУПНЕННЯ ДЛЯ АНАЛІЗУ
ТАНДЕМНИХ ВІДКРИТИХ МЕРЕЖ ОБСЛУГОВУВАННЯ

Анотація. Запропоновано модель відкритої тандемної мережі масового обслуговування з обмеженими буферами та зворотним зв'язком, на кожний вузол якої надходять незалежні пуасонівські потоки заявок. Частина заявок після обслуговування у першому вузлі миттєво надходить у другий (якщо там є вільні місця), інші заявки остаточно покидають мережу. Після завершення обслуговування у другому вузлі заявка покидає мережу або за наявності вільного місця миттєво надходить у перший вузол, або в орбіт і після випадкової затримки вимагає повторного обслуговування. Показано, що математичною моделлю цієї мережі є тривимірний ланцюг Маркова, запропоновано ієрархічний алгоритм фазового укрупнення для розрахунку ймовірностей її станів. Наведено результати чисельних експериментів.

Ключові слова: відкрита тандемна мережа, зворотний зв'язок, тривимірний ланцюг Маркова, ієрархічний алгоритм фазового укрупнення.

A.Z. Melikov, L.A. Ponomarenko, A.M. Rustamov
HIERARCHICAL SPACE MERGING ALGORITHM FOR THE ANALYSIS
OF OPEN TANDEM QUEUING NETWORKS

Abstract. The Markov model of two-stage queuing network with feedback is proposed. Poisson flows arrive to both stages from outside. A part of already serviced calls at the first node instantaneously enter the second node (if there is free space here) while the other calls leave the network. After the service is completed at the second node, there are three possibilities: (1) it leaves the network; (2) it instantaneously feeds back to the first node (if there is free space here); (3) it feeds back to the first node after some delay in orbit. All feedbacks are determined by known probabilities. Both nodes have finite capacities. The mathematical model of the investigated network is a three-dimensional Markov chain, and a hierarchical space merging algorithm is developed to calculate its steady-state probabilities. The results of numerical experiments are demonstrated.

Keywords: open two stage queuing network, feedback, three dimensional Markov chain, hierarchical space merging algorithm.

Меліков Агаси Зарбали оглы,
чл.-кор. НАН Азербайджана, доктор техн. наук, профессор, заведующий отделом Института проблем управления НАН Азербайджана, Баку, e-mail: agassi.melikov@gambler.ru.

Пonomаренко Леонид Анатольевич,
доктор техн. наук, профессор, главный научный сотрудник Международного научно-учебного центра информационных технологий и систем НАН и МОН Украины, Киев, e-mail: laponomarenko@ukr.net.

Рустамов Анар Муршид оглы,
кандидат техн. наук, доцент Университета «Кавказ», Баку, Азербайджан,
e-mail: anar.rustamov@gmail.com.