УДК 519.872: 621.321.1

**В. Ш. Фейзиев, к. т. н.; В. Ш. Фейзиев; Г. И. Бuниятова**

# ИМИТАЦИОННЫЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ НАХОЖДЕНИЯ СУБОПТИМАЛЬНОЙ СТРАТЕГИИ ДОСТУПА В МНОГОСКОРОСТНЫХ СИСТЕМАХ ОБСЛУЖИВАНИЯ

*В работе предлагается приближенный метод решения задачи нахождения субоптимальной стратегии доступа в многоскоростной системе обслуживания. В зависимости от текущей ситуации в системе могут быть включены резервные каналы, при этом их включения связаны с определенными экономическими издержками. Предложенный метод базируется на идеях имитационного моделирования. Даны результаты численных экспериментов.*

***Ключевые слова:*** *многоскоростная система обслуживания, резервные каналы, стратегия доступа, субоптимальная стратегия, имитационное моделирование.*

## Введение

Многоскоростные системы обслуживания (Multi Rate Queue, MRQ) являются математическими моделями процессов обработки разнотипной информации в коммуникационных сетях последнего поколения [1 – 3]. В известных работах изучены модели таких систем при простых стратегиях доступа, т. е. в них предполагается, что все каналы системы равноправно используются разнотипными вызовами. Вместе с тем, из-за ограниченности ресурсов сетей, а также вследствие того, что разнотипные вызовы имеют различные степени важности, равноправное использование каналов системы не всегда является эффективным. Поэтому некоторые каналы резервируются и используются лишь при возникновении определенных конфликтных ситуаций. Для разрешения этих ситуаций наиболее эффективным средством являются Марковские процессы принятия решений (МППР). Такой подход ранее был использован в работе [4] для нахождения оптимальной стратегии доступа. Здесь аналогичный подход используется для исследования моделей MRQ с резервными каналами.

## Постановка задачи

Рассмотрим модель многоскоростной системы обслуживания (Multi Rate Queue, MRQ), в которой все каналы разделены на две группы: активные и резервные. При этом активные каналы используются согласно полнодоступной схеме, а включение резервных каналов поддаётся управлению. Последнее означает, что использование резервных каналов связано с определенными экономическими издержками, и потому в моменты поступления разнотипных вызовов необходимо принимать решения об использовании таких каналов. При этом такие решение принимаются лишь тогда, когда количество свободных каналов оказывается недостаточными для обслуживания поступившего вызова. Цель управления включением резервных каналов состоит в минимизации суммарных экономических издержек в единицу времени стационарного режима, связанных с потерями вызовов и использованием резервных каналов.

Все $N>1$ каналы системы разделены на две группы, т. е. $N=A+R$, где $A>1$ указывает количество активных каналов и $R>1$ означает количество резервных каналов, при этом все каналы являются идентичными. Входящий поток вызовов является пуассоновским с параметром $\Lambda$, при этом каждый поступивший вызов с вероятностью $\alpha_i$ требует $b_i$ каналов одновременно, где $\alpha_1 +\ldots\alpha_K=1$, т. е. исходный поток представляет собой суперпозицию $K$ независимых пуассоновских потоков с интенсивностями $\lambda_i := \Lambda\alpha_i$, $i=1,\ldots,K$, где вызовы из $i$-

го потока требуют одновременно $b_i$ каналов, при этом все каналы начинают и завершают обслуживание одновременно. Время обслуживания вызовов $i$-го типа является экспоненциально распределенной случайной величиной со средним $\mu_i^{-1}, i = \overline{1, K}$.

Если в момент поступления вызова любого типа количество свободных активных каналов является достаточным, то необходимое количество активных каналов назначается для его обслуживания. В противном случае свободные резервные каналы могут быть использованы для этой цели. Вместе с тем, если суммарное количество свободных каналов (активных и резервных) окажется недостаточным для обслуживания поступившего вызова, то он с вероятностью 1 теряется (блокируется).

Механизм включения и отключения резервных каналов состоит в следующем. Если в момент завершения обслуживания вызова любого типа количество занятых каналов окажется не меньше, чем $A$, то все освобожденные каналы переключаются в резервную группу; в противном случае любые $A$ каналы ставятся в активную группу, а остальные остаются в резервной группе.

Предположим, что потеря одного вызова $i$-го типа оценивается штрафом в $c(i)$ условных единиц, $i = 1,\dots,K$, а включение $j$ резервных каналов в единицу времени приводит к штрафам $d(j)$ условных единиц, $j = 1,\dots,R$.. Тогда задача нахождения оптимальной стратегии включения резервных каналов сформулируется следующим образом: требуется найти такую стратегию включения резервных каналов, чтобы минимизировать суммарные штрафы в единицу времени стационарного режима, связанные с потерями вызовов различных типов и включением резервных каналов.

Следовательно, искомая оптимальная стратегия доступа (Call Admission Control, CAC) представляет собой последовательность решений, принимаемых в моменты поступления вызовов. При этом, в каждый момент поступления вызовов с учетом их типа и текущего состояния системы необходимо принимать одно из двух решений: либо поступивший вызов теряется, либо определенное количество резервных каналов используется для его обслуживания.

**Расчет характеристик модели**

Состояние данной системы в произвольный момент времени можно описать $K$-мерным вектором $\boldsymbol{n} = (n_1,\dots,n_K)$, где $n_i$ указывает количество вызовов $i$-го типа в системе. Поскольку каждый вызов $i$-го типа одновременно требует $b_i$ каналов системы, то максимальное количество вызовов в системе ограничено величиной $[N/b_i]$, где $[x]$ означает целую часть $x$, $i = \overline{1, K}$. Суммарное количество занятых каналов системы в состоянии $\boldsymbol{n}$, определяется как скалярное произведение векторов $\boldsymbol{n}$ и $\boldsymbol{b} = (b_1,\dots,b_K)$. Таким образом, множество возможных состояний системы определяется так:

$$S := \left\{ \boldsymbol{n} : n_i = 0, \left[ N / b_i \right], i = \overline{1, k}, (\boldsymbol{n}, \boldsymbol{b}) \le N \right\}. \quad (1)$$

Для описания класса стратегий, в котором находится искомая оптимальная стратегия включения резервных каналов, рассмотрим моменты поступления вызовов.

Пусть в момент поступления вызова $i$-го типа система находится в состоянии $\boldsymbol{n} \in S$. Количество свободных активных каналов в этом состоянии определяется как $f(\boldsymbol{n}) = A - (\boldsymbol{n}, \boldsymbol{b})$, если $f(\boldsymbol{n}) \ge 0$. Поэтому если $f(\boldsymbol{n}) < 0$, то это означает, что в состоянии $\boldsymbol{n}$ количество используемых резервных каналов равно $-f(\boldsymbol{n})$. Отсюда делаем вывод, что величина $f(\boldsymbol{n}) + R$ указывает суммарное количество свободных активных и резервных каналов в состоянии $\boldsymbol{n} \in S$, если $f(\boldsymbol{n}) > 0$, а в случае $f(\boldsymbol{n}) \le 0$ указанная величина означает количество свободных резервных каналов в состоянии $\boldsymbol{n} \in S$.

Поскольку активные каналы системы используются согласно полнодоступной схеме, то если в момент поступления вызова $i$-го типа система находится в состоянии $\boldsymbol{n} \in S$, в котором

$f(\boldsymbol{n})\geq b_i$, то поступивший вызов принимается с вероятностью 1, и для его обслуживания выделяются любые $b_i$ свободные активные каналы. Если в этот момент компоненты вектора состояния $\boldsymbol{n}$ удовлетворяют неравенству $b_i> f(\boldsymbol{n})+R$, то поступивший вызов $i$-го типа теряется с вероятностью 1, так как в этот момент количество свободных каналов (активных и резервных) оказывается недостаточным для обслуживания поступившего вызова. Альтернативные решения возможны в моменты поступления вызовов $i$-го типа, если в эти моменты система находится в одном из подклассов множества возможных состояний:

$$S_i^* := \{\boldsymbol{n} \in S : f(\boldsymbol{n}) \leq 0, b_i \leq f(\boldsymbol{n}) + R\}; \tag{2}$$

$$S_i^{**} := \{\boldsymbol{n} \in S : f(\boldsymbol{n}) > 0, f(\boldsymbol{n}) < b_i \leq f(\boldsymbol{n}) + R\}. \tag{3}$$

В обоих подклассах состояний (2) и (3) возможны следующие решения: d1 – поступивший вызов теряется и d2 – резервные каналы используются для обслуживания поступившего вызова $i$-го типа. Отметим, что если принимается решение d2, то в подклассе (2) используются $b_i$ резервных каналов, а в подклассе (3) количество резервных каналов, выделяемых для обслуживания поступившего вызова $i$-го типа, равно $b_i$-$f(\boldsymbol{n})$. Вероятности принятия решений d1 и d2 обозначаются $\alpha_i^-(\boldsymbol{n})$ и $\alpha_i^+(\boldsymbol{n})$ соответственно. Поскольку эти вероятности составляют полную группу, то имеем:

$$\alpha_i^-(\boldsymbol{n}) + \alpha_i^+(\boldsymbol{n}) = 1 \text{ для всех } \boldsymbol{n} \in S_i, \tag{4}$$

где $S_i = S_i^* \bigcup S_i^{**}, i = \overline{1,K}$.

Теперь рассмотрим моменты отправки вызовов системой. Пусть непосредственно перед отправкой вызова $i$-го типа из системы она была в состоянии $\boldsymbol{n}$, где $n_i>0$. Тогда в момент отправки вызова следующим состоянием системы будет $\boldsymbol{n}$-$\boldsymbol{e}_i$, где $\boldsymbol{e}_i$ – $K$-мерный вектор, все компоненты, кроме $i$-го, равны 0, а $i$-ая компонента равна 1. Если $(\boldsymbol{n}\text{-}\boldsymbol{e}_i,\boldsymbol{b})\geq A$, то все освобожденные $b_i$ каналов становятся резервными; в противном случае любые $A$ каналы становятся активными, а остальные переключаются в резервную группу. После отправки из системы вызова $i$-го типа система переходит в состояние $\boldsymbol{n}$-$\boldsymbol{e}_i$ с интенсивностью $n_i\mu_I$, $i=1,\ldots,K$.

Можно показать, что указанные штрафы за единицу времени вычисляются так:

$$\begin{aligned} G\left(p(\boldsymbol{n}), \alpha_i^{\pm}(\boldsymbol{n})\right) := \sum_{\boldsymbol{n}\in S} \sum_{i=I}^{K} \Big(\lambda_i p(\boldsymbol{n}) \Big(c(i)\Big(I(b_i > f(\boldsymbol{n}) + R) + I(\boldsymbol{n} \in S_i)\alpha_i^-(\boldsymbol{n})\Big) + \\ + \left(d(b_i) I\left(\boldsymbol{n} \in S_i^*\right) + d(b_i - f(\boldsymbol{n})) I\left(\boldsymbol{n} \in S_i^{**}\right)\right)\alpha_i^+(\boldsymbol{n})\Big)\Big), \end{aligned} \tag{5}$$

где $p(\boldsymbol{n})$ означает стационарную вероятность состояния $\boldsymbol{n}\in S$.

Исходя из изложенного заключаем, что целью исследования данной системы является решение следующей задачи:

$$G\left(p(n), \alpha_i^{\pm}(n)\right) \xrightarrow[\alpha_i^{\pm}(n)]{} \min. \tag{6}$$

Ограничениями этой задачи является система уравнений равновесия. Следовательно, задача определения оптимальной стратегии включения резервных каналов сводится к определенной задаче марковского программирования. Она имеет нерандомизированное оптимальное решение. Для нахождения оптимальной стратегии включения резервных каналов при малых значениях $N$ и $K$ может быть использован метод линейного программирования. При больших значениях этих параметров могут быть использованы приближенные методы (4).

Вместе с тем, на практике, особенно при исследовании моделей MRQ с большим количеством типов вызовов, состояние системы не наблюдается полностью, т. е. имеется частичная информация о ее состоянии, а именно, наблюдается лишь общее количество занятых (свободных) каналов. Поэтому искомая САС должна принимать решение на основе такой неполной информации. Оптимальную САС, основанную лишь на информации о количестве занятых (свободных) каналов, назовем субоптимальной.

Пусть активные каналы, как и прежде, используются согласно полнодоступной схеме, а количество резервных каналов, которые могут быть использованы для обслуживания вызовов $i$-го типа, ограничены величиной $r_i$, при этом $r_1+...+r_K \geq R$. Задача оптимизации системы заключается в нахождении таких значений $r_i$, $i=1,...,K$**,** чтобы минимизировать суммарные штрафы (6).

Отметим, что субоптимальная стратегия не будет лучше, чем оптимальная стратегия. Это объясняется тем, что оптимальная САС при принятии решения учитывает детальную информацию о состоянии системы, в то время как субоптимальная стратегия основывается лишь на частичной информации о состоянии системы (т. е. учитывается лишь информация об общем количестве занятых каналов). Иными словами, два разных состояния, в которых количество занятых каналов является одинаковым, рассматриваются как одно состояние с точки зрения субоптимальной стратегии включения резервных каналов.

Программа имитационного моделирования разработана и использована для нахождения субоптимальной стратегии в модели MRQ с параметрами $A = 20$, $R = 10$, $K = 2$, $b_1 = 1$, $b_2 = 6$, $c(1) = c(2) = 1$, $d(i) = I$, $i = 1, \ldots ,6$. В каждом прогоне имитационной программы были использованы 100000 вызовов, полностью завершивших обслуживание. Количество повторений каждого эксперимента – 5, и их среднее значение было выбрано в качестве основных показателей QoS системы. В результате этих экспериментов найдена субоптимальная стратегия включения резервных каналов. Соответствующие результаты показаны в табл. 1. Следует отметить, что с целью уменьшения количества переключений различных субоптимальных стратегий отмеченные звездочкой стратегии могут быть заменены стратегией (10, 1), так как максимальная разница между минимальными значениями соответствующих целевых функций не превышает 0,5%. Так, например, из табл. 1 видно, что при $(\rho_1,\rho_2) = (4,4)$ и $(\rho_1,\rho_2) = (4,6)$ оптимальная субоптимальная стратегия $(r_1,r_2) = (10,1)$, а при $(\rho_1,\rho_2) = (4,5)$ соответствующая стратегия определяется как $(r_1,r_2) = (9,1)$. Последнее означает, что если и при $(\rho_1,\rho_2) = (4,5)$ в качестве субоптимальной стратегии принимать $(r_1,r_2) = (10,1)$, то количество переключений различных субоптимальных стратегий существенным образом уменьшается, и при этом не допускаются большие ошибки (так как разность не превышает 0,5%).

Таблица 1

**Субоптимальная стратегия включения резервных каналов**

○: $(r_1, r_2) = (10,1)$, □ : $(r_1, r_2) = (9,1)$, Δ : $(r_1, r_2) = (8,2)$

| $\rho_2$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 | Δ | ○ | □ | ○ | □ | ○ | ○ | ○ | ○ | ○ |
| 9 | Δ | ○ | □ | □* | □ | □* | □* | ○ | ○ | ○ |
| 8 | Δ | ○ | ○ | ○ | □ | ○ | ○ | ○ | ○ | ○ |
| 7 | ○ | ○ | ○ | ○ | □ | □* | □* | ○ | ○ | ○ |
| 6 | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | □* | ○ |
| 5 | ○ | ○ | ○ | □* | ○ | ○ | □* | □* | □* | ○ |
| 4 | ○ | □* | ○ | ○ | □* | □* | ○ | □* | ○ | ○ |
| 3 | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | □* |
| 2 | ○ | ○ | □* | ○ | ○ | ○ | □* | ○ | ○ | ○ |
| 1 | ○ | ○ | ○ | ○ | ○ | ○ | ○ | □ | □ | □ |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 $\rho_1$ |

**Выводы**

В работе предлагается подход имитационного моделирования для решения задачи нахождения субоптимальной стратегии включения резервных каналов в многоскоростных системах, когда точное решение задачи нахождения оптимальной стратегии не представляется возможным из-за огромной размерности исходной задачи. Исследованная модель имеет широкое применение в современных коммуникационных сетях, в которых обрабатывается разнотипная информация.

СПИСОК ЛИТЕРАТУРЫ

1. Меликов А. З. Математические модели многопотоковых систем обслуживания. / А. З. Меликов, Л. А. Пономаренко, Н. А. Рюмшин. – К.: Техника, 1991 – 117 с.

2. Меликов А. З. Приближенный расчет характеристик совместной передачи речи и данных в беспроводных сетях сотовой связи / А. З. Меликов, В. Ш. Фейзиев // Электронное моделирование. – 2007, – Т. 29, № 6, – с. 47 – 59.

3. Kim C. S. Analysis of multirate systems with shared reservation of channels and queues of wideband calls / C. S. Kim, A. Z. Melikov, V. S. Feyziyev // Proc. of Int. conf. “Problems of Cybernetics and Informatics”, Baku. – 2008. –Vol. 2. – P. 210 – 213.

4. Melikov A. Z. Markov decision process approach to finding state-dependent CAC algorithm in broadband integrated network node / A. Z. Melikov, V. S. Feyziyev // Proc. of Int. conf. “Mathematical Methods for Increasing Efficiency of Information Telecommunication Networks”, Minsk. – 2007. – Vol. 19. – P. 142 – 146.

***Фейзиев Васиф Шейдулла оглы*** – к. т. н., старший научный сотрудник.

***Фейзиев Вагиф Шейдулла оглы*** – научный сотрудник.

***Буниятова Гюлшен Иса кызы*** – младший научный сотрудник.

Институт кибернетики НАН Азербайджана, Баку.