

А. С. Власкина, С. А. Бурцева, И. А. Кочеткова, С. Я. Шоргин, Управляемая система массового обслуживания с эластичным трафиком и сигналами для анализа нарезки ресурсов в сети радиодоступа, Информ. и её примен., 2022, том 16, выпуск 3, 90–96

DOI: 10.14357/19922264220312

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением <a href="http://www.mathnet.ru/rus/agreement">http://www.mathnet.ru/rus/agreement</a>

Параметры загрузки: IP: 18.217.230.237

28 сентября 2024 г., 01:17:24



# УПРАВЛЯЕМАЯ СИСТЕМА МАССОВОГО ОБСЛУЖИВАНИЯ С ЭЛАСТИЧНЫМ ТРАФИКОМ И СИГНАЛАМИ ДЛЯ АНАЛИЗА НАРЕЗКИ РЕСУРСОВ В СЕТИ РАДИОДОСТУПА\*

А. С. Власкина<sup>1</sup>, С. А. Бурцева<sup>2</sup>, И. А. Кочеткова<sup>3</sup>, С. Я. Шоргин<sup>4</sup>

Аннотация: Построена управляемая система массового обслуживания с нетерпеливым эластичным трафиком и минимально гарантированной скоростью обслуживания для анализа динамической нарезки радиоресурсов беспроводной сети. Управление перераспределением ресурса между двумя сегментами сети осуществляется при поступлении сигнала контроллера и в случае простаивания ресурса одного сегмента и ожидающих начала обслуживания пользователей другого сегмента. В зависимости от состояния системы принимается решение о том, на какой объем перераспределить ресурс. Принятие решения основано на показателях эффективности нарезки, учитывающих использование ресурса каждого из сегментов, отклонение от рекомендуемого «начального» разделения, дополнительную сигнальную нагрузку от контроллера. Предложена соответствующая целевая функция управления, применен итерационный алгоритм расчета параметров оптимальной политики распределения ресурса.

**Ключевые слова:** 5G; нарезка ресурсов; управляемая система массового обслуживания; марковский процесс принятия решения; эластичный трафик

**DOI:** 10.14357/19922264220312

#### 1 Введение

Одна из ключевых концепций сетей пятого и последующих поколений — технология нарезки сети (англ. network slicing) — позволяет создавать логические сети (сегменты), каждая из которых имеет соответствующие ресурсы для обслуживания пользователей определенных услуг [1—3]. Общая конструкция системы обеспечивает экономичность и гибкость настройки и развертывания сегментов сети [4—6].

В работе исследуется модель сети, где имеются базовый оператор с ресурсом, а также два виртуальных оператора, которые запрашивают доступ к этому ресурсу (рис. 1). В качестве услуг, которые предоставляют виртуальные операторы своим пользователям, рассматриваются услуги передачи данных с минимально допустимой скоростью обслуживания, а также с ограничениями на время ожидания начала обслуживания. Отношения между базовым и виртуальными операторами регулирует соглашение о качестве обслуживания (англ. Service Level Agreement, SLA). В соответствии с этим соглашением базовый оператор планирует распределение ресурса в условиях нормальной загрузки

(«начальное» разделение ресурса). Управление заключается в перераспределении ресурсов между классами трафика для предотвращения ухудшения качества обслуживания в одном сегменте сети из-за нехватки ресурса в другом.

Стратегией управления радиоресурсами занимается контроллер, который с определенной периодичностью направляет сигналы о проверке необходимости перераспределения. Задачей является поиск оптимальных способов планирования радиоресурсов (по каким критериям и сколько ресурсов зарезервировать под конкретную услугу) в зависимости от соответствия начальному распределению, максимального использования ресурсов, а также числа сигналов контроллера, приведших к перераспределению.

Данная работа является продолжением исследований [7, 8], анализирующих эффективность динамической нарезки радиоресурсов между сегментами сети, с учетом перечисленных выше трех принципов. Исследована модель управления ресурсами при поступлении сигнала контроллера в виде управляемой системы массового обслуживания, применен итерационный алгоритм вычисления оптимальной стратегии.

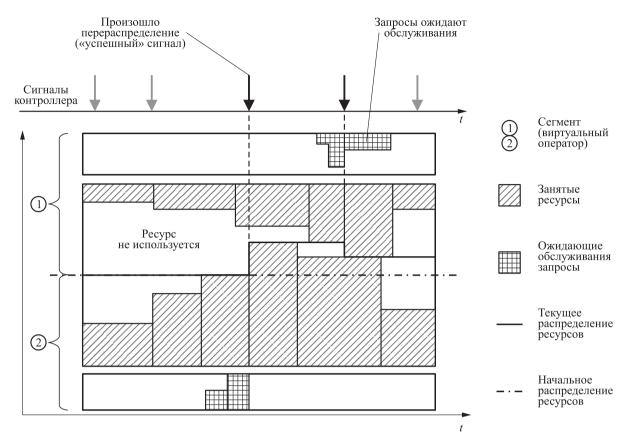
<sup>\*</sup>Публикация выполнена при поддержке Программы стратегического академического лидерства РУДН.

<sup>&</sup>lt;sup>1</sup> Российский университет дружбы народов, vlaskina-as@rudn.ru

<sup>&</sup>lt;sup>2</sup>Российский университет дружбы народов, sofiya\_burceva@inbox.ru

<sup>&</sup>lt;sup>3</sup> Российский университет дружбы народов; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, kochetkova-ia@rudn.ru

<sup>&</sup>lt;sup>4</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru



**Рис. 1** Схема перераспределения ресурса: 1— сегмент (виртуальный оператор); 2— занятые ресурсы; 3— ожидающие обслуживания запросы; 4— текущее распределение ресурсов; 5— начальное распределение ресурсов

# 2 Управляемая система массового обслуживания

Рассмотренную модель с двумя классами услуг можно описать в виде управляемой системы массового обслуживания [9] с нетерпеливым эластичным трафиком, минимально гарантированной скоростью обслуживания и потоком сигналов, которые моделируют сообщения от контроллера. Схематическое изображение системы представлено на рис. 2.

Будем считать, что поток заявок на обслуживание пользователей виртуальных операторов пуассоновский с интенсивностями  $\lambda_1$  и  $\lambda_2$ . Обслуживание — экспоненциальное со средним объемом передаваемых данных  $\mu_1^{-1}$  и  $\mu_2^{-1}$ . С учетом минимально гарантированной скорости обслуживания запросов b рассматривается многоканальная система с N обслуживающими приборами. По внешнему пуассоновскому потоку сигналов с параметром  $\delta$  начальное распределение ресурсов  $\tilde{m}_1$  (для 1-го класса трафика) может изменяться динамически и в произвольный момент времени равняться  $m_1$ . В случае отсутствия свободных приборов, соответ-

ствующих классу поступившего запроса, последний попадает в очередь длины  $R_1$  или  $R_2$  соответственно. С интенсивностью  $\varepsilon_1$  и  $\varepsilon_2$  запросы покидают очередь по причине превышения допустимого порога на экспоненциальное время задержки в очереди.

Опишем функционирование системы при помощи марковского процесса принятия решений в непрерывном времени, который задается четырьмя параметрами (S,  $A_s$ ,  $Q_a(s,s')$ , R(s)). Состояние системы может быть представлено в виде вектора  $s=(m_1,l_1,l_2)$ , где  $m_1$  — текущее число приборов для 1-го класса трафика;  $l_1$  и  $l_2$  — число запросов 1-го и 2-го классов трафика в системе, в пространстве

$$S = \{s = (m_1, l_1, l_2) : m_1 = 0, \dots, N, l_1 = 0, \dots, m_1 + R_1, l_2 = 0, \dots, N - m_1 + R_2\}.$$

Перераспределение ресурсов инициируется при поступлении сигнала контроллера, а также простаивании ресурсов одного класса запросов и наличии ожидающих обслуживания запросов другого класса (будем называть такой сигнал «успешным», т. е.

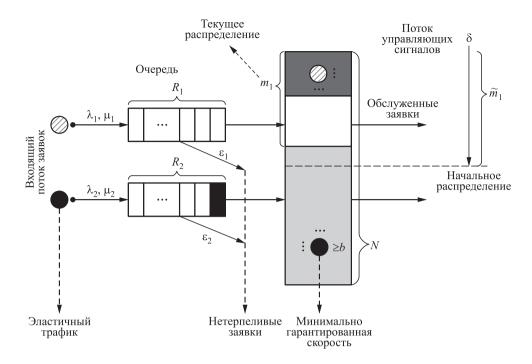


Рис. 2 Схема системы массового обслуживания

приведшим к перераспределению). Если же все очереди пусты или, наоборот, полностью заняты обслуживающие ресурсы, при поступлении сигнала контроллера перераспределения ресурсов не произойдет.

Разбиение множества S на подмножества проводится в зависимости от возможности перераспределения ресурсов, а также соотношения между числом запросов в очередях и на приборах. Это состояния, в которых при поступлении сигнала произойдет  $S_{\delta}$  и не произойдет  $\overline{S_{\delta}}$  перераспределение ресурсов,  $S = S_{\delta} \cup \overline{S_{\delta}}$ . Подмножество состояний с числом приборов 1-го класса  $m_1$ 

$$S(m_1) = \{(i, l_1, l_2) \in S : i = m_1\},\$$
  
 $S = \bigcup_{m_1=0}^{N} S(m_1),$ 

тоже разбивается на те, в которых произойдет перераспределение ресурсов и в которых не произойдет:

$$S(m_1) = S_{\delta}(m_1) \bigcup \overline{S_{\delta}}(m_1).$$

В свою очередь,

$$\mathsf{S}_{\delta}(m_1) = \mathsf{S}^1_{\delta}(m_1) \bigcup \mathsf{S}^2_{\delta}(m_1),$$

где  $\mathsf{S}^1_\delta(m_1) = \{i, l_1, l_2\} \in \mathsf{S}(m_1): l_1 > m_1, l_2 < N - m_1\}$  — состояния, в которых при поступлении сигнала ресурсы 1-го класса трафика будут увеличены;  $\mathsf{S}^2_\delta(m_1) = \{(i, l_1, l_2) \in \mathsf{S}(m_1): l_1 < m_1, l_2 > N - m_1\}$  — где будут увеличены ресурсы 2-го

класса. Следовательно, в случае наличия необходимого числа свободных ресурсов для обслуживания ожидающих запросов увеличение сегментов осуществляется на число ожидающих запросов

$$\begin{split} \mathsf{S}^k_{\delta}(m_1) &= \mathsf{S}^{k1}_{\delta}(m_1) \bigcup \mathsf{S}^{k2}_{\delta}(m_1), \\ \mathsf{S}^{k1}_{\delta}(m_1) &= \{(i,l_1,l_2) \in \mathsf{S}^k_{\delta}(m_1) : l_1 + l_2 \leq N \} \end{split}$$

или на все свободные ресурсы

$$\mathsf{S}^{k2}_{\delta}(m_1) = \{(i,l_1,l_2) \in \mathsf{S}^k_{\delta}(m_1) : l_1 + l_2 > N\}.$$

Задачей управления системой массового обслуживания будет выбор нового объема ресурса для 1-го класса трафика при поступлении сигнала контроллера. Множество допустимых стратегий управления зависит от состояния системы  $s \in S$  и определяется в виде:

$$A_{s} = \begin{cases} \{m_{1}, \dots, l_{1}\}, & s \in \mathsf{S}_{\delta}^{11}; \\ \{m_{1}, \dots, N - l_{2}\}, & s \in \mathsf{S}_{\delta}^{12}; \\ \{N - l_{2}, \dots, m_{1}\}, & s \in \mathsf{S}_{\delta}^{21}; \\ \{l_{1}, \dots, m_{1}\}, & s \in \mathsf{S}_{\delta}^{22}, \end{cases}$$
$$a(s) \in A_{s}, s \in \mathsf{S}.$$

Далее зададим инфинитезимальную матрицу  $Q_a(s,s')=[q_a(s,s')]$  перехода из состояния s в состояние s' при управляющем действии a. Матрица является консервативной, т. е.  $q_a(s,s')\geq 0, s\neq s',$   $q_a(s,s')=-q_a(s)=-\sum_{s\neq s'}q_a(s,s'), \ q_a(s)<\infty,$  где для  $s\neq s'$  элементы матрицы имеют вид:

$$\begin{split} q_a(s,s') &= \\ \begin{cases} \lambda_1\,,s' = (m_1,l_1+1,l_2)\;,\; s:l_1+1 \leq R_1\;;\\ \lambda_2\,,s' = (m_1,l_1,l_2+1)\;,\; s:l_2+1 \leq R_2\;;\\ \frac{m_1}{N}\,V\,\mu_1\,,\;\; s' = (m_1,l_1-1,l_2)\;,\\ s:m_1>0\,,\; l_1>0\;;\\ \end{cases} \\ \frac{N-m_1}{N}\,V\,\mu_2\,,\;\; s' = (m_1,l_1,l_2-1)\;,\\ s:N-m_1>0\,,\; l_2>0\;;\\ \{(l_1-m_1)\varepsilon_1\,,\;\; s' = (m_1,l_1,l_2-1)\;,\\ s:l_1\geq m_1\;;\\ \{(l_2-N+m_1)\varepsilon_2\,,s' = (m_1,l_1,l_2-1)\;,\\ s:l_2\geq N-m_1\;;\\ \delta\,,\; s' = (a,l_1,l_2)\;\text{для}\;m_1\leq a\leq l_1\,,\; s\in \mathsf{S}^{12}_\delta,\\ \mathsf{или}\;N-l_2\leq a\leq m_1\,,\; s\in \mathsf{S}^{22}_\delta,\\ \mathsf{или}\;l_1< a< m_1\,,\; s\in \mathsf{S}^{22}_\delta. \end{split}$$

Наконец, система приносит вознаграждение в размере R(s) денежных единиц за единицу времени в течение всего периода ее пребывания в состоянии s.

## 3 Функция вознаграждения

Напомним, что строится управляемая система массового обслуживания, параметром управления которой служит объем выделяемых ресурсов для 1-го класса трафика. Функция вознаграждения будет состоять из трех слагаемых, которые рассчитываются для каждого конкретного состояния системы. Каждая из трех компонент соответствует принципам оптимального распределения ресурсов, которые будут представлены ниже.

Принцип равного деления ресурсов, позволяющий определять число запросов, которые могли бы обслуживаться, но которым приходится ожидать из-за несправедливого деления ресурсов, выглядит следующим образом:

$$\chi(s) = \begin{cases} \frac{N}{2} - m_1, & s: m_1 < \frac{N}{2}, \ l_1 > m_1, \\ & l_1 - m_1 > \frac{N}{2} - m_1; \\ l_1 - m_1, & s: m_1 < \frac{N}{2}, \ l_1 > m_1, \\ & l_1 - m_1 \le \frac{N}{2} - m_1; \\ \frac{N}{2} - N + m_1, \ s: m_1 > \frac{N}{2}, \ l_2 > N - m_1, \\ & l_2 - N + m_1 > \frac{N}{2} - N + m_1; \\ l_2 - N + m_1, \ s: m_1 > \frac{N}{2}, \ l_2 > N - m_1, \\ & l_2 - N + m_1 \le \frac{N}{2}, \ l_2 > N - m_1, \\ & l_2 - N + m_1 \le \frac{N}{2} - N + m_1. \end{cases}$$

Второй компонентой является принцип «успеха» перераспределения, который учитывает вероятность того, что поступит сигнал о перераспределении ресурсов, но перераспределения не произойдет из-за состояния системы, и определяется как

$$\begin{split} \beta_{\delta}(s) &= \delta \left(\delta + \lambda_{1} \cdot 1 \left\{ l_{1} + 1 \leq R_{1} + m_{1} \right\} + \\ &+ \lambda_{2} \cdot 1 \left\{ l_{2} + 1 \leq R_{2} + N - m_{1} \right\} + \\ &+ \frac{m_{1}}{N} V \mu_{1} \cdot 1 \left\{ m_{1} > 0 \, , \, l_{1} > 0 \right\} + \\ &+ \frac{N - m_{1}}{N} V \mu_{2} \cdot 1 \left\{ N - m_{1} > 0 \, , \, l_{2} > 0 \right\} + \\ &+ \left( l_{1} - m_{1} \right) \varepsilon_{1} \cdot 1 \left\{ l_{1} > m_{1} \right\} + \\ &+ \left( l_{2} - N + m_{1} \right) \varepsilon_{2} \cdot 1 \left\{ l_{2} > N - m_{1} \right\} \right)^{-1}. \end{split}$$

Третья компонента использования ресурсов, отражающая число запросов, которые могли бы обслуживаться, но находятся в ожидании из-за простаивания ресурсов, определяется как

$$\gamma(s) = \begin{cases} N - m_1 - l_2 \,, & s \in \mathsf{S}_{\delta}^{12}(m_1) \,; \\ l_1 - m_1 \,, & s \in \mathsf{S}_{\delta}^{11}(m_1) \,; \\ m_1 - l_1 \,, & s \in \mathsf{S}_{\delta}^{22}(m_1) \,; \\ l_2 - N + m_1 \,, & s \in \mathsf{S}_{\delta}^{21}(m_1) \,. \end{cases}$$

Если в первой компоненте задержка обслуживания связана с «несправедливым» занятием всех ресурсов, то здесь, наоборот, с тем, что система выделила недостаточный объем ресурсов для обслуживания трафика.

В итоге функция вознаграждения вычисляется по формуле:

$$R(s) = -\left(c_1\chi(s) \cdot 1\left\{s \in \overline{\mathsf{S}^1_{\delta}(m_1)}\right\} \left[1\left\{m_1 < \frac{N}{2}, s \in \mathsf{S}^1_{\delta}(m_1)\right\} + 1\left\{m_1 > \frac{N}{2}, s \in \mathsf{S}^2_{\delta}(m_1)\right\}\right] + c_2\beta_{\delta}(s) \cdot 1\left\{s \in \overline{\mathsf{S}_{\delta}}\right\} + c_3\gamma(s) \cdot 1\left\{s \in \mathsf{S}_{\delta}\right\}\right), \quad (1)$$

а функция среднего вознаграждения при управлении a определяется как

$$g_a = \sum_{s \in S} R(s) \pi_a(s) , \qquad (2)$$

где  $\pi_a(s)$  — стационарное распределение вероятностей при управлении a. Используя заданную функцию среднего вознаграждения (2), будем решать оптимизационную задачу:

$$g^* = \max_{a} \left( g_a \right) \tag{3}$$

и получим оптимальную стратегию управления распределения ресурсов.

### 4 Оптимальная стратегия

Для решения оптимизационной задачи вида (3) в теории марковских процессов принятия решений используются итерационные алгоритмы, которые существенно сокращают вычисления. Применим итерационный алгоритм Р. Ховарда [10] для вычисления оптимальной стратегии. Система уравнений относительно среднего вознаграждения (2) и весовых коэффициентов v(s) имеет вид:

$$v(s) = \frac{1}{\sum_{s' \in S \setminus s} q_a(s, s')} \left[ R(s) + \sum_{s' \in S \setminus s} q_a(s, s') v(s') - g^a \right], \quad s \in S, \quad (4)$$

где  $q_a(s,s')$  — элементы инфинитезимальной матрицы  $Q_a(s,s')$ ; R(s) — функция вознаграждения (1). Здесь коэффициенты v(s) имеют смысл вознаграждений в случае, если система находится в состоянии s. Для поиска новой (улучшенной) стратегии задается целевая функция в виде

$$a(s) = \underset{a \in A_s}{\operatorname{arg\,max}} \left\{ v\left(a, l_1, l_2\right) \right\}, \ \ s \in \mathsf{S}_{\delta}.$$
 (5)

С помощью алгоритма 1 зададим начальную стратегию перераспределения ресурсов (шаг 3), при которой ресурсы классов трафика увеличиваются либо на все свободные ресурсы, либо на число ожидающих обслуживания запросов. На первом этапе (шаг 4) вычисляются весовые коэффициенты для текущей стратегии управления. На втором этапе (шаг 5) улучшается стратегия изменением управляющих действий a(s) для каждого состояния  $s \in S_\delta$ . Этапы повторяются до тех пор, пока не будет найдена оптимальная стратегия.

#### 5 Заключение

В статье построена модель обслуживания двух классов трафика с минимально гарантированной скоростью обслуживания при применении технологии нарезки сети, которая инициируется при поступлении сигнала контроллера, в виде управляемой системы массового обслуживания с эластичным трафиком и нетерпеливыми заявками. Система представлена в терминах марковского процесса принятия решений в непрерывном времени. В качестве принятия решений выступает выбор объема выделяемых ресурсов для одного из сегментов при динамическом перераспределении ресурсов. Функция вознаграждения представляет собой совокупность предложенных коэффициентов: соответствия начальному распределению ресурса, успеха перераспределения ресурса и использования ресурса.

## Литература

- Ageev K., Sopin E., Samouylov K. Resource sharing model with minimum allocation for the performance analysis of network slicing // Information technologies and mathematical modelling. Queueing theory and applications / Eds. A. Dudin, A. Nazarov, A. Moiseev. Communications in computer and information science ser. Springer, 2021. Vol. 1391. P. 378—389. doi: 10.1007/978-3-030-72247-0\_28.
- 2. Yarkina N., Correia L. M., Moltchanov D., Gaidama-ka Y., Samouylov K. Multi-tenant resource sharing with equitable-priority-based performance isolation of slices for 5G cellular systems // Comput. Commun., 2022. Vol. 188. P. 39–51. doi: 10.1016/j.comcom.2022.02.019.
- 3. Stepanov M. S., Stepanov S. N., Andrabi U., Petrov D., Ndayikunda J. The increasing of resource sharing efficiency in network slicing implementation // Distributed

#### Алгоритм 1. Итерационный алгоритм вычисления оптимальной стратегии

```
1: n \leftarrow 0
2: v^{(0)}(s) = 0, s \in S

3: a^{(0)}(s) = \begin{cases} l_1, & s \in S_\delta^{11}; \\ N - l_2, & s \in S_\delta^{12}; \\ m_1 - l_2 - N + m_1, & s \in S_\delta^{21}; \\ l_1, & s \in S_\delta^{22}, \end{cases} \Rightarrow начальная политика solve (система (4))

5: a^{(n+1)}(s) = (\text{по формуле } (5))
6: \mathbf{if} \ a^{(n+1)}(s) = a^{(n)}(s), s \in S then

7: \mathbf{return} \ a^{(n+1)}(s), v^{(n)}(s), g^{(n)}

8: \mathbf{else} \ n \leftarrow n+1, \mathbf{go} \ \mathbf{to} \ \mathbf{step} \ 5

9: \mathbf{end} \ \mathbf{if}
```

- computer and communication networks / Eds. V. M. Vishnevskiy, K. E. Samouylov, D. V. Kozyrev. Communications in computer and information science ser. Springer, 2022. Vol. 1552. P. 18—35. doi: 10.1007/978-3-030-97110-6-2
- Muhizi S., Ateya A. A., Muthanna A., Kirichek R., Koucheryavy A. A novel slice-oriented network model // Distributed computer and communication networks / Eds. V. Vishnevskiy, D. Kozyrev. Communications in computer and information science ser. Springer, 2018. Vol. 919. P. 421–431. doi: 10.1007/978-3-319-99447-5\_36.
- Yarkina N., Gaidamaka Y., Correia L. M., Samouylov K. An analytical model for 5G network resource sharing with flexible SLA-oriented slice isolation // Mathematics, 2020. Vol. 8. No. 7. P. 1177. 19 p. doi: 10.3390/ math8071177.
- 6. *Shu Z., Taleb T., Song J.* Resource allocation modeling for fine-granular network slicing in beyond 5G systems // Oper. Res., 2022. Vol. 70. No. 2. P. 349–363. doi: 10.1587/transcom.2021WWI0002.
- 7. Kochetkova I., Vlaskina A., Burtseva S., Savich V., Hosek J. Analyzing the effectiveness of dynamic network slicing procedure in 5G network by queuing and simulation mod-

- els // Internet of things, smart spaces, and next generation networks and systems / Eds. O. Galinina, S. Andreev, S. Balandin, Y. Koucheryavy. Lecture notes in computer science ser. Springer, 2020. Vol. 12525. P. 71—85. doi: 10.1007/978-3-030-65726-0-7.
- Кочеткова И. А., Власкина А. С., Ву Н. Н., Шоргин В. С. Система массового обслуживания с управляемым по сигналам перераспределением приборов для анализа нарезки ресурсов сети 5G // Информатика и её применения, 2021. Т. 15. Вып. 3. С. 91–97. doi: 10.14357/19922264210312.
- Efrosinin D., Kochetkova I., Samouylov K., Stepanova N.
   Algorithmic analysis of a two-class multi-server heterogeneous queueing system with a controllable cross-connectivity // Analytical and stochastic modelling techniques and applications / Eds. M. Gribaudo, E. Sopin, I. Kochetkova. Lecture notes in computer science ser. Springer, 2020. Vol. 12023. P. 1–17. doi: 10.1007/978-3-030-62885-7\_1.
- Ховард Р. А. Динамическое программирование и марковские процессы / Пер. с англ. — М.: Советское радио, 1964. 192 с. (Howard R. A. Dynamic programming and Markov processes. — Cambridge, MA, USA: MIT Press, 1960. 136 р.)

Поступила в редакцию 15.07.22

# CONTROLLABLE QUEUING SYSTEM WITH ELASTIC TRAFFIC AND SIGNALS FOR ANALYZING NETWORK SLICING

A. S. Vlaskina<sup>1</sup>, S. A. Burtseva<sup>1</sup>, I. A. Kochetkova<sup>1,2</sup>, and S. Ya. Shorgin<sup>2</sup>

**Abstract:** A controllable queuing system with impatient elastic traffic and minimum bit rate guarantee transmission is built to analyze the dynamic network slicing in wireless network. Resource management of allocation between two segments is carried out by controller signal and in the case of an idle resource of one segment and waiting users of another. Depending on the state of the system, a decision is made how to allocate the resource. The decision-making is based on the efficiency indicators of slicing taking into account the resource usage of each of the segments, deviation from the recommended "initial" slicing, and additional signal load from the controller. The corresponding objective control function is proposed and an iterative algorithm for calculating the parameters of the optimal resource allocation policy is applied.

**Keywords:** 5G; network slicing; controllable queuing system; Markov decision process; elastic traffic

**DOI:** 10.14357/19922264220312

# Acknowledgments

The paper has been supported by the RUDN University Strategic Academic Leadership Program.

#### References

1. Ageev, K., E. Sopin, and K. Samouylov. 2021. Resource sharing model with minimum allocation for the performance analysis of network slicing. *Information technologies* 

and mathematical modelling. Queueing theory and applications. Eds. A. Dudin, A. Nazarov, and A. Moiseev. Communications in computer and information science ser. Springer. 1391:378—389. doi: 10.1007/978-3-030-72247-0\_28.

<sup>&</sup>lt;sup>1</sup>Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

<sup>&</sup>lt;sup>2</sup>Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

- 2. Yarkina, N., L. M. Correia, D. Moltchanov, Y. Gaidamaka, and K. Samouylov. 2022. Multi-tenant resource sharing with equitable-priority-based performance isolation of slices for 5G cellular systems. *Comput. Commun.* 188:39–51. doi: 10.1016/j.comcom.2022.02.019.
- Stepanov, M. S., S. N. Stepanov, U. Andrabi, D. Petrov, and J. Ndayikunda. 2022. The increasing of resource sharing efficiency in network slicing implementation. *Distributed computer and communication networks*. Eds. V. M. Vishnevskiy, K. E. Samouylov, and D. V. Kozyrev. Communications in computer and information science ser. Springer. 1552:18–35. doi: 10.1007/978-3-030-97110-6\_2.
- Muhizi, S., A. A. Ateya, A. Muthanna, R. Kirichek, and A. Koucheryavy. 2018. A novel slice-oriented network model. *Distributed computer and communication networks*. Eds. V. Vishnevskiy and D. Kozyrev. Communications in computer and information science ser. Springer. 919:421– 431. doi: 10.1007/978-3-319-99447-5\_36.
- Yarkina, N., Y. Gaidamaka, L. M. Correia, and K. Samouylov. 2020. An analytical model for 5G network resource sharing with flexible SLA-oriented slice isolation. *Mathematics* 8(7):1177. 19 p. doi: 10.3390/math8071177.
- Shu, Z., T. Taleb, and J. Song. 2022. Resource allocation modeling for fine-granular network slicing in beyond 5G systems. *Oper. Res.* 70(2):349–363. doi: 10.1587/transcom.2021WWI0002.

- Kochetkova, I., A. Vlaskina, S. Burtseva, V. Savich, and J. Hosek. 2020. Analyzing the effectiveness of dynamic network slicing procedure in 5G network by queuing and simulation models. *Internet of things, smart spaces, and next generation networks and systems*. Eds. O. Galinina, S. Andreev, S. Balandin, and Y. Koucheryavy. Lecture notes in computer science ser. Springer. 12525:71–85. doi: 10.1007/978-3-030-65726-0\_7.
- Kochetkova, I.A., A.S. Vlaskina, N.N. Vu, and V.S. Shorgin. 2021. Sistema massovogo obsluzhivaniya s upravlyaemym po signalam pereraspredeleniem priborov dlya analiza narezki resursov seti 5G [Queuing system with signals for dynamic resource allocation for analyzing network slicing in 5G networks]. *Informatika i ee Primeneniya Inform. Appl.* 15(3):91–97. doi: 10.14357/19922264210312.
- Efrosinin, D., I. Kochetkova, K. Samouylov, and N. Stepanova. 2020. Algorithmic analysis of a twoclass multi-server heterogeneous queueing system with a controllable cross-connectivity. *Analytical and stochastic modelling techniques and applications*. Eds. M. Gribaudo, E. Sopin, and I. Kochetkova. Lecture notes in computer science ser. Springer. 12023:1–17. doi: 10.1007/978-3-030-62885-7\_1.
- 10. Howard, R.A. 1960. *Dynamic programming and Markov processes*. Cambridge, MA: MIT Press. 136 p.

Received July 15, 2022

#### Contributors

**Vlaskina Anastasia S.** (b. 1995) — assistant professor, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; vlaskina-as@rudn.ru

**Burtseva Sofiya A.** (b. 1999) — student, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; sofiya\_burceva@inbox.ru

**Kochetkova Irina A.** (b. 1985) — Candidate of Science (PhD) in physics and mathematics, associate professor, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; kochetkova-ia@rudn.ru

**Shorgin Sergey Ya.** (b. 1952) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; sshorgin@ipiran.ru