

Анализ влияния стратегий доступа к ресурсам на доход поставщика инфраструктуры при применении технологии сетевого слайсинга

Л. О. Лапшенкова^{1*}, Ф. А. Москалева¹, Ю. В. Гайдамака^{1,2}

¹ ФГАОУ ВО «Российский университет дружбы народов», г. Москва, Российская Федерация
117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6

* 1032217633@rudn.ru

² ФГУ «Федеральный исследовательский центр «Информатика и управление» Российской академии наук», г. Москва, Российская Федерация

119333, Российская Федерация, г. Москва, ул. Вавилова, д. 44-2

Аннотация

В работе исследуется концепция технологии нарезки радиоресурса сети или т.н. «сетевого слайсинга» (Network Slicing) и особенности, связанные с внедрением данной технологии в сетях 5G. Под сетевым слайсингом авторы понимают одновременное использование несколькими арендаторами принадлежащего поставщику инфраструктуры ресурса для предоставления пользователям услуг, удовлетворяющих требованиям к качеству. Ключевым отличием сетевого слайсинга от классической концепции разделения ресурсов стало гибкое выделение ресурса в соответствии с текущей потребностью арендатора при строгом соблюдении изоляции слайсов, то есть защиты слайса каждого арендатора от всплесков нагрузки в слайсах других арендаторов. В статье представлена модель сетевого слайсинга, позволяющая анализировать влияние применения стратегий доступа к ресурсам и политик приема заявок в систему на средний доход поставщика инфраструктуры. Исследованы две стратегии управления доступом к ресурсам – «перенарезка» ресурса поставщиком инфраструктуры при каждом поступлении нового запроса пользователя (слайсинг по требованию) и «перенарезка» ресурсов через фиксированные интервалы времени (периодический слайсинг). Для каждой стратегии управления доступом рассмотрены две политики приема запроса в систему – прием запроса вне зависимости от дохода, который он принесёт арендатору (Always Admit policy), и прием только тех запросов, доход от которых превысит заранее установленный порог (Above Threshold policy). Для построенной математической модели процесса нарезки радиоресурса сети проведен численный эксперимент, иллюстрирующий зависимость среднего дохода поставщика инфраструктуры, упущенного дохода поставщика инфраструктуры для слайсинга по требованию и периодического слайсинга от структурных и нагрузочных параметров системы, в том числе, от интенсивности поступления заявок в систему и от емкости системы. Результаты численного эксперимента приведены в заключении в виде рекомендаций поставщику инфраструктуры по определению наиболее выгодной для него стратегии доступа к ресурсам и политики приема заявок в систему.

Ключевые слова: сетевой слайсинг, сети нового поколения

Финансирование: публикация выполнена при поддержке Программы стратегического академического лидерства РУДН. Исследование выполнено при частичной финансовой поддержке Российского фонда фундаментальных исследований в рамках научного проекта № 20-07-01064 «Стохастический анализ и максимизация покрытия зоны действия беспроводной воздушной сети на беспилотных летательных аппаратах».

Авторы заявляют об отсутствии конфликта интересов.



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.



Для цитирования: Лапшенкова, Л. О. Анализ влияния стратегий доступа к ресурсам на доход поставщика инфраструктуры при применении технологии сетевого слайсинга / Л. О. Лапшенкова, Ф. А. Москалева, Ю. В. Гайдамака. – DOI 10.25559/SITITO.17.202103.519-530 // Современные информационные технологии и ИТ-образование. – 2021. – Т. 17, № 3. – С. 519-530.

© Лапшенкова Л. О., Москалева Ф. А., Гайдамака Ю. В., 2021



Analysis of Access Strategies Impact on Infrastructure Provider's Revenue in Case of Network Slicing

L. O. Lapshenkova^{a*}, F. A. Moskaleva^a, Yu. V. Gaidamaka^{a,b}

^a Peoples' Friendship University of Russia, Moscow, Russian Federation

6 Miklukho-Maklaya St., Moscow 117198, Russian Federation

* 1032217633@rudn.ru

^b Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russian Federation

44-2 Vavilov St., Moscow 119333, Russian Federation

Abstract

The article analyses the Network slicing concept and its features associated with the implementation of this technology in 5G networks. By network slicing, the authors mean the simultaneous use of a resource owned by an infrastructure provider by several tenants to provide users with services that meet the quality of services requirements. The key difference between network slicing and the classical resource sharing concept is flexible resource allocation in accordance with the current tenant needs while strictly observing the isolation of slices, that is, protecting a slice of each tenant from load spikes in the slices of other tenants. The article presents a network slicing model that allows analyzing the impact of applying access control strategies and accept's policies on the average revenue of an infrastructure provider. Two access control strategies have been investigated - the re-slice of the resource by the infrastructure provider at each new user request (slicing on demand) and the re-slice at fixed time intervals (periodic slicing). For each access control strategy, two policies for accepting a request into the system are considered - accepting a request regardless of the revenue it will bring to the tenant ("Always Admit" policy), and accepting only those requests whose revenue exceeds a predetermined threshold ("Above Threshold" policy). For the constructed mathematical model of the process of slicing the radio resource of the network, a numerical experiment was carried out that illustrates the dependence of the average revenue of the infrastructure provider, the lost revenue of the infrastructure provider for slicing on demand and periodic slicing, on the structural and load parameters of the system, including the intensity of the incoming flow of requests in the system and on the system's capacity. The results of a numerical experiment are given in the conclusion in the form of recommendations for an infrastructure provider to determine the most profitable access control strategy and accept's policies.

Keywords: network slicing, next generation networks, queuing theory

Funding: The publication was supported by the RUDN University Strategic Academic Leadership Program. The study was carried out with partial financial support from the Russian Foundation for Basic Research within the framework of scientific project No. 20-07-01064 "Stochastic Analysis and Maximization of Wireless Air Network Coverage on Unmanned Aerial Vehicles".

The authors declare no conflict of interest.

For citation: Lapshenkova L.O., Moskaleva F.A., Gaidamaka Yu.V. Analysis of Access Strategies Impact on Infrastructure Provider's Revenue in Case of Network Slicing. *Sovremennye informacionnye tehnologii i IT-obrazovanie = Modern Information Technologies and IT-Education*. 2021; 17(3):519-530. DOI: <https://doi.org/10.25559/SITITO.17.202103.519-530>



Введение

Каждые несколько лет технология мобильной связи претерпевает революционные перемены. Такие быстрые изменения связаны со значительным увеличением объема передаваемых данных за последние годы [1, 2, 3]. Наиболее известные преимущества технологии сетей 5G – это высокая скорость обмена данными, в 10 раз превышающая скорость в сетях 4G LTE, большая пропускная способность сети, простое и быстрое массовое подключение устройств к сети. Фактическая скорость загрузки данных будет зависеть от ряда факторов, включая местоположение пользователя и сетевой трафик¹. Постоянный рост числа пользователей при ограниченности ресурсов сети продолжает являться вызовом даже в сетях пятого поколения. Одним из решений проблемы эффективного использования ресурса сети стала концепция сетевого слайсинга (network slicing), которая предусматривает гибкое создание логических сетей путем разделения одной инфраструктуры. Затем виртуальные сети настраиваются в соответствии с конкретными потребностями приложений, служб, устройств, клиентов или операторов. Таким образом, внедрение поставщиком инфраструктуры технологии нарезки сети позволит ему повысить свой доход за счет эффективного управления выделяемым арендаторам ресурсом так, чтобы арендаторы слайсов не нарушали требования к качеству предоставляемых ими различных типов услуг². Целью данной работы является анализ влияния нескольких стратегий доступа к ресурсам и политик приема заявок в систему на средний доход поставщика инфраструктуры, который является главным показателем эффективности системы. Основная задача работы – определить наиболее выгодную для поставщика инфраструктуры стратегию доступа к ресурсам и политику приема заявок в систему.

Приложения сетевого слайсинга

Крупные компании по всему миру борются за сохранение конкурентоспособной позиции на рынке услуг. Такие характеристики, как возможность создания виртуальной среды и простота в программируемости дают основание считать сеть пятого поколения надёжной, высокоэффективной и недорогой [4]. Сеть 5G позволит получить широкий спектр новых специально адаптированных услуг довольно большому кругу компаний, проектами которых являются: умные города, автономное вождение, телемедицина, интернет вещей и т. д. [5, 6, 7].

Однако, стоит отметить, что предоставление услуг посредством технологии пятого поколения обладает жёсткими требованиями, такими как высокая пропускная способность и пиковая скорость передачи данных, энергоэффективность сети, низкое время ожидания [8, 9]. Для борьбы с трудностями

ми, связанными с использованием и внедрением сетей нового поколения, специалисты разработали технологию сетевого слайсинга, при которой сетевая инфраструктура поделена на логические сегменты (сетевые слайсы), работающие в облачной инфраструктуре [9, 10]. Данная технология позволит операторам сотовой связи (Mobile Network Operators, MNOs), являющимся поставщиками инфраструктуры (Infrastructure Provider, InP), создавать отдельные слайсы под услуги, которые предоставляет поставщик услуг (Service Provider, SP) путём обеспечения функционального и операционного разнообразия на общей сетевой инфраструктуре [11]. Сетевой слайс может охватывать несколько сетей, таких как сеть радиодоступа, сеть транспортировки сообщений и опорная сеть интернета [9]. Однако, поставщик инфраструктуры не гарантирует наличие необходимого ресурса для удовлетворения потребностей отдельного поставщика услуг [12].

По расчётам учёных сетевой слайсинг позволит увеличить доходы операторов до 150% в сравнении с классической концепцией разделения ресурсов в рамках одной большой сети [11]. Примером приложений интернета вещей (Internet of Things, IoT), для которых концепция сетевого слайсинга открывает широкие возможности, является «умный город», в котором независимые частные сети поставщиков коммунальных услуг могут быть реализованы как отдельные логические сети на основе общей физической сети [13-15]. Технологию сетевого слайсинга можно считать объединением нескольких инновационных технологий [16, 17], примером которых являются программно-определяемая сеть (Software defined networking, SDN) и виртуализация сетевых функций (Network function virtualization, NFV) наряду с облачными технологиями³. Технология SDN обеспечивает прямое программируемое сетевое управление, при котором администраторы с лёгкостью могут управлять сетями с помощью гибких техник управления, динамически подстраивая сеть под меняющиеся требования на различных ее участках. Более того, применение SDN позволяет представить управляющие устройства как единый коммутатор для различных подсистем обработки политик (policy engines). Технология SDN основана на стандартах и является универсальной, что упрощает проектирование и эксплуатацию сети [18, 19]. В отличие от программно-определяемой сети NFV отделяет сетевые функции (включая функции переадресации и управления сетью) от аппаратного обеспечения⁴. То есть функции традиционно выделенного сетевого оборудования могут предоставляться в качестве программных функций, работающих на виртуальных машинах. Данная технология позволяет экономить как на капитальных, так и на операционных затратах, поскольку выделенное оборудование может работать на стандартных сетевых серверах. Кроме того, отсутствует необходимость в обслуживании центров обработки данных или сервисных центров, так как емкость сервера

¹ Duffy C. What is 5G? Your questions answered [Электронный ресурс] // CNN. March 6, 2020. URL: <https://edition.cnn.com/interactive/2020/03/business/what-is-5g/index.html> (дата обращения: 15.08.2021).

² Kavanagh S. What is network slicing? [Электронный ресурс] // The 5G Comparison Site. August 28, 2018. URL: <https://5g.co.uk/guides/what-is-network-slicing> (дата обращения: 15.08.2021).

³ Network Slicing for 5G Networks & Services. 5G Americas, 2016. [Электронный ресурс]. URL: <https://www.5gamericas.org/network-slicing-for-5g-networks-services> (дата обращения: 15.08.2021).

⁴ Tittel E. SDN vs. NFV: What's the difference? [Электронный ресурс] // Cisco Systems, Inc., 2021. URL: <https://www.cisco.com/c/en/us/solutions/software-defined-networking/sdn-vs-nfv.html> (дата обращения: 15.08.2021).



может быть изменена по требованию с помощью соответствующих настроек программного обеспечения [19, 20]. Итак, технология SDN обеспечивает гибкую переадресацию и управление трафиком в физической или виртуальной сетевой среде, в то время как NFV – гибкое размещение виртуализированных сетевых функций в сети и облаке⁵. Для создания программируемой сети эти две технологии обычно дополняют друг друга. Как было обозначено ранее, в контексте мобильных сетей технологии SDN и NFV обеспечивает воплощение сетевого слайсинга [21, 22]. Действительно, в основном, сети 5G разработаны с учётом отделения плоскости контроля от плоскости пользователя. Ввиду того, что большая часть трафика плоскости пользователя требует довольно простой обработки, он может быть обслужен при недорогом оборудовании. Отдельные объекты плоскости контроля, в свою очередь, требуют более глубокой обработки в сравнении с плоскостью пользователя. Внедрение технологии сетевого слайсинга обеспечивает выгодное масштабирование в зависимости от требований пользовательской плоскости, так как плоскость управления не зависит от пользовательской плоскости⁶.

Помимо SDN и NFV сети 5G используют ряд технологий, включая облачные вычисления (cloud computing) и периферийные вычисления (edge computing), прямое взаимодействие устройств (Device-to-Device communication, d2d), связь в диапазоне миллиметровых волн (Millimeter wave communication, mmWave). Облачные вычисления были введены для удовлетворения растущих потребностей в управлении ресурсами, хранении данных. Периферийные вычисления рассматриваются в качестве перспективной технологии для расширения возможностей экосистем 5G [23]. Прямое взаимодействие устройств позволяет отдельным частям IoT, находящимся в непосредственной близости друг от друга, контактировать друг с другом, используя прямую связь, а не традиционную передачу сигнала через базовую станцию [24]. Связь в диапазоне миллиметровых волн обладает рядом преимуществ, в том числе большой пропускной способностью и узкой направленностью, служащими средством для преодоления проблем, вызванных ростом объема мобильного трафика [25]. Помимо описанных преимуществ с точки зрения показателей функционирования сети внедрение концепции сетевого слайсинга позволяет поставщику инфраструктуры, владеющему оборудованием и предоставляющему его в аренду поставщикам услуг, повысить доход. Далее в статье для системной модели сетевого слайсинга построена и исследована математическая модель для анализа доход поставщика инфраструктуры.

Системная модель сетевого слайсинга

Системная модель [26], включающая в себя одного поставщика инфраструктуры и несколько поставщиков услуг, изображена на Рис. 1. Полагаем, что пользовательское оборудование (ПО, англ. user equipment, UE) находится в пределах зоны об-

служивания базовой станции (БС, англ. base station, BS), которая принадлежит поставщику инфраструктуры (Infrastructure Provider, InP). БС представляет собой точку доступа к другим сетевым ресурсам, например, транзитному соединению, IP-сетям и облачным хранилищам. Пользовательскому оборудованию может быть предоставлен доступ к нескольким услугам, каждая из которых обеспечивается своим поставщиком услуг (Service Provider, SP). Полагаем, что ПО может получать услугу только у одного SP. Это предположение не нарушает общности, если мы считаем, что пользователь, получающий несколько услуг одновременно, имеет отдельно ПО для каждой услуги. Для предоставления услуги своему пользователю поставщик услуги SP арендует ресурсы поставщика инфраструктуры InP. Объем ресурса, который InP выделяет для обслуживания заявки того или иного SP, составляет соответствующий слайс. В общем случае размеры слайсов разных SP могут различаться, при этом внутри одного слайса для каждой заявки соответствующего SP поставщик инфраструктуры всегда выделяет ресурс одного и того же размера. Объем ресурса, имеющегося у поставщика инфраструктуры, измеряется в [бит/с].

На Рис. 1 изображены два SPs – арендаторы ресурса, направляющие свои заявки поставщику инфраструктуры InP, который имеет ресурс C [бит/с], что соответствует N базовым единицам канального ресурса (ЕКР). Для каждого арендатора определены параметры λ_1 и λ_2 [заявок/с] – интенсивность поступления заявок на предоставление соответствующей услуги, r_1 и r_2 [бит/с] – скорости, требуемые для предоставления услуги соответствующим арендатором, μ_1 и μ_2 [заявок/с] – интенсивность обслуживания заявок соответствующего арендатора, β_1 и β_2 [ден. ед./ед. вр.] – случайная величина, цена, которую арендатор предлагает поставщику инфраструктуры обслуживание своей заявки.

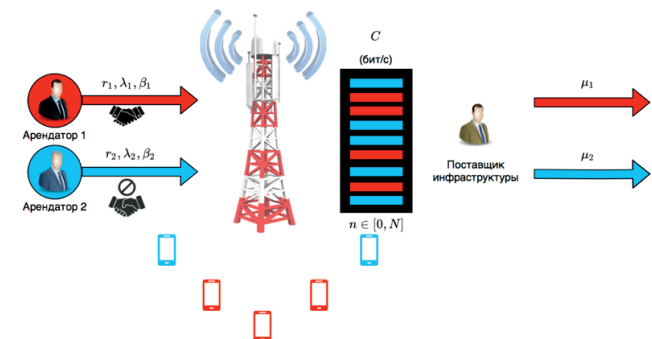


Рис. 1. Системная модель для случая двух поставщиков услуг
Fig. 1. System model for the case of two service providers

⁵ Obraczka K., Rothenberg C., Rostami A. SDN, NFV and their role in 5G. SIGCOM16. ACM SIGCOMM Tutorial. Brazil, 2016. [Электронный ресурс]. URL: <https://www.dca.fee.unicamp.br/~cheesteve/ppt/SIGCOMM16-Tutorial-5G-SDN-NFV-part1.pdf> (дата обращения: 15.08.2021).

⁶ Basilier H., et al. A vision of the 5G core: Flexibility for new business opportunities [Электронный ресурс] // Ericsson Technology Review. 2016. Vol. 93, issue 2. Pp. 1-15. URL: <https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/a-vision-of-the-5g-core-flexibility-for-new-business-opportunities> (дата обращения: 15.08.2021).



Поставщик инфраструктуры решает, принимать ли на обслуживание поступающую заявку, во время проведения торгов, исходя из политики приема заявки в систему. Торги – процедура выбора заявок из поступивших в систему для принятия к обслуживанию. Политика приема заявки в систему учитывает как количество свободных ресурсов на момент принятия решения, так и цену, которую SP предлагает за то, чтобы его заявку допустили к обслуживанию. Во время проведения торгов InP проверяет наличие свободного ресурса, достаточного для обслуживания поступившей заявки, и, при наличии, сравнивает цену β , предлагаемую поступившей заявкой, с имеющимся у него диапазоном цены $[\beta_m, \beta_M]$. Здесь $[\beta_m, \beta_M]$ – диапазон цены, при этом β_m – минимальный тариф, при котором поставщику инфраструктуры выгодно принять заявку на обслуживание, β_M – максимальный тариф, устраивающий арендатора. Примерами политик приема заявки в систему являются политика постоянного допуска (Always Admit policy, AA), когда диапазон цены $[\beta_m, \beta_M]$ поставщика инфраструктуры постоянен, или же политика превышения порога (Above Threshold policy, AT), когда диапазон вариации цены $[\beta, \beta_M]$ имеет плавающий порог допуска заявок в систему β , например, зависящий от числа уже обслуживаемых в системе заявок.

Политика приема заявки в систему \mathcal{P}_n зависит от состояния n системы и представляет собой критерий, по которому InP решает принять или отклонить заявку в состоянии n ($0 \leq n \leq N - 1$, где $N = \lceil C/r \rceil$):

$$\mathcal{P}_n = \begin{cases} \text{принять заявку, } \beta \in \mathcal{D}_n \subset [\beta_m, \beta_M], \\ \text{отклонить заявку,} & \text{иначе.} \end{cases} \quad (1)$$

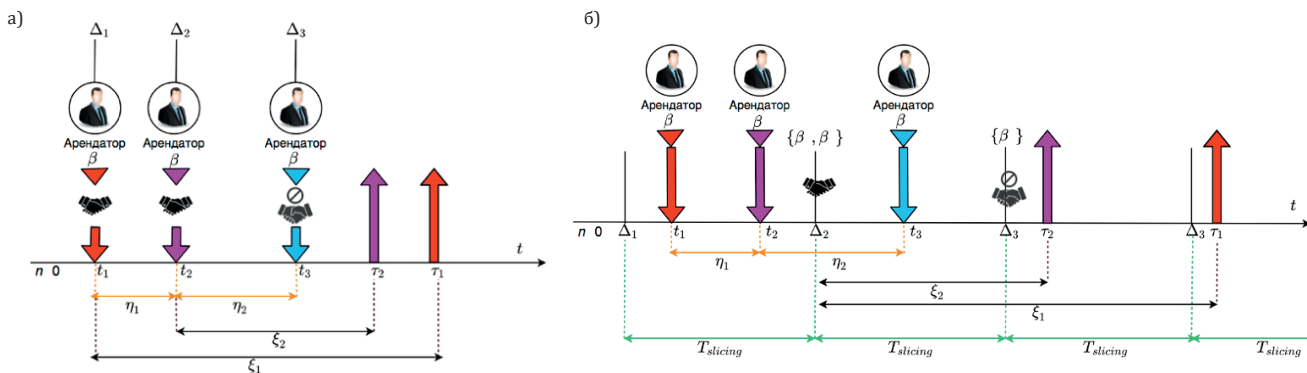
Здесь \mathcal{D}_n – диапазон цены, который устанавливает InP в зависимости от состояния n системы.

Указанные политики представляют собою различные модели ценообразования для SP. В общем случае предлагаемая арендатором цена β – случайная величина с произвольной плотностью f_β на интервале $[\beta_m, \beta_M]$.

Момент проведения торгов, в результате которых принимается управляющее решение, определяется стратегией управления доступом к ресурсу. Далее рассмотрены стратегии управления доступом двух типов – т.н. «слайсинг по требованию» (event-triggered), когда торги

для решения о приеме заявки на обслуживание проводятся в момент её поступления в систему, и т.н. «периодический слайсинг» (time-triggered), когда торги для решения о приеме заявок на обслуживание проводятся периодически одновременно для всех заявок, которые поступили в систему с момента принятия последнего решения.

Для иллюстрации стратегий управления доступом к ресурсу далее будем строить модель для случая одного арендатора, который предоставляет пользователям одну услугу. В этом случае к InP поступает один поток заявок, обладающий следующими характеристиками: r [бит/с] – скорость, требуемая для предоставления услуги арендатором; λ [заявок/с] – интенсивность поступления заявок; μ [заявок/с] – интенсивность обслуживания; β [ден. ед./ед. вр.] – цена, которую предлагает SP за обслуживание своей заявки, случайная величина, подчиняющаяся распределению общего вида с плотностью f_β на фиксированном для SP интервале $[\beta_m, \beta_M]$. Если общий объем ресурса, имеющийся у InP, C [бит/с], тогда максимальное число заявок, которое может быть принято в систему, равно $N = \lceil C/r \rceil$. В рассматриваемом случае одного арендатора для обслуживания каждой заявки поставщик инфраструктуры InP должен выделить одну единицу канального ресурса, соответствующую скорости r [бит/с]. Между SP и InP заключается соглашение SLA (Service Level Agreements), в котором уточняются технические детали, в частности, скорость (или диапазон скоростей), которую поставщик инфраструктуры InP обязуется предоставить арендатору SP для каждой из услуг в соответствии с требованиями к качеству обслуживания (Quality of Service, QoS), цена (или диапазон цен), которую SP обязуется заплатить арендатору за полученный ресурс, стратегия управления доступом к ресурсу. На Рис. 2 приведена иллюстрация двух стратегий управления доступом к ресурсу: а) слайсинг по требованию (event-driven), когда перераспределение (reslicing, перенарезка) ресурсов происходит в моменты наступления некоторых событий, а именно, в моменты поступления заявок в систему; б) периодический слайсинг (time-triggered), при котором решение о приеме на обслуживание заявок, полученных поставщиком инфраструктуры InP в промежуток времени $T_{slicing}$, принимается в начале следующего промежутка времени $T_{slicing}$.



Р и с. 2. Стратегии управления доступом к ресурсу:

а) слайсинг по требованию; б) периодический слайсинг

F i g. 2. Resource access control strategies:

а) slicing on demand; б) periodic slicing



На Рис. 2 для а) слайсинга по требованию и б) периодического слайсинга на оси времени показаны моменты наступления следующих событий: моменты поступления к поставщику инфраструктуры InP от арендатора SP -й заявки на выделение ресурса для предоставления услуги своему пользователю, моменты освобождения ресурса по окончании предоставления услуги -му пользователю, моменты проведения -х торгов. Заметим, что для слайсинга по требованию торги проводятся в момент поступления заявки на выделение слайса, а для периодического слайсинга отмечены постоянные периоды между проведением торгов.

Математическая модель сетевого слайсинга

Математическую модель процесса обслуживания запросов на предоставление услуги от арендатора, поставщика услуг, к поставщику инфраструктуры строим в виде системы массового обслуживания (СМО) M|M|N|0 (Рис. 3). На нее поступает пуассоновский поток заявок интенсивности λ , соответствующий потоку запросов от арендатора, при этом заявка может быть потеряна как из-за отсутствия свободного канального ресурса для её обслуживания, так и вследствие политики приема заявки в систему. Длительность обслуживания соответствует длительности предоставления услуги и является экспоненциально распределенной случайной величиной с параметром μ .

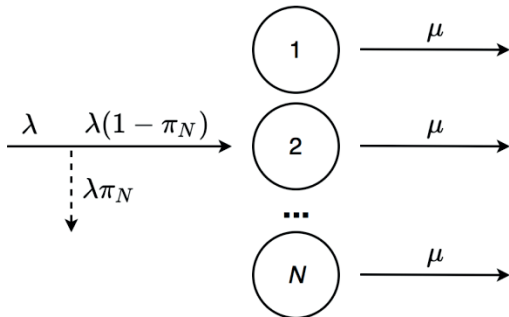


Рис. 3. Система массового обслуживания M|M|N|0
Fig. 3. System of mass servicing M|M|N|0

Процесс занятия ресурса поставщика инфраструктуры может быть описан Марковским процессом (МП) с непрерывным временем $X(t) = (n(t), \mathcal{P}_n(t))$, где $n(t)$ - число обслуживаемых заявок в момент t , $\mathcal{P}_n(t)$ - политика приема заявки в систему в момент t , с пространством состояний $X = \{(n, \mathcal{P}_n), n = 0, 1, \dots, N = \lfloor C/r \rfloor\}$.

Каждому состоянию МП $X(t)$ соответствует критерий для принятия или отклонения входящих заявок на слайсы \mathcal{P}_n , $n = 0, \dots, N - 1$, при этом $\mathcal{P}_{system} = (\mathcal{P}_0, \dots, \mathcal{P}_N)$. Политика \mathcal{P}_n , $n = 0, \dots, N - 1$, может быть определена одним из следующих способов.

Для стратегии слайсинга по требованию:

1) политика AA: заявка будет принята, если цена $\beta \in \mathcal{D}_n^{AA} = [\beta_m, \beta_M]$, где β_m - минимальное значение цены, которую готов получить поставщик инфраструктуры за предоставленный ресурс, β_M - максимальный тариф, по которому готов платить поставщик услуг за используемые ресурс;

2) политика AT: заявка будет принята, если цена $\beta \in \mathcal{D}_n^{AT} = [\beta_n, \beta_M]$, где $\beta_n \geq \beta_m$, β_n - порог приёма; Для стратегии периодического слайсинга:

1) политика BB: из участвующих в торгах заявок будет принята заявка, которая предложит более высокую цену β , при этом политика применяется, пока не исчерпается свободный ресурс;

2) политика FCFS: из участвующих в торгах заявок будет заявка, дольше остальных ожидающая начала торгов, при этом политика применяется, пока не исчерпается свободный ресурс;

Вероятность того, что заявка будет принята в систему в состоянии n , определяется следующим образом:

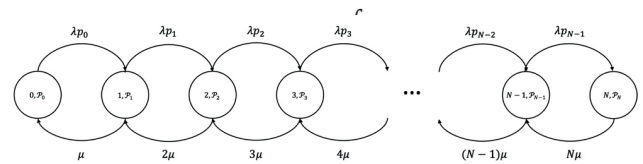


Рис. 4. Граф интенсивностей переходов для МП $X(t)$, реализующего слайсинг по требованию

Fig. 4. Graph of transition intensities for an MP $X(t)$, implementing on-demand slicing

Решение системы уравнений баланса дают выражения для стационарных вероятностей состояний системы:

$$\pi_0(\rho, f_\beta, N, \mathcal{P}_{system}) = \frac{1}{1 + \sum_{i=1}^N \frac{\rho^i}{i!} \prod_{l=0}^{i-1} p_l},$$

$$\pi_n(\rho, f_\beta, N, \mathcal{P}_{system}) = \frac{\frac{\rho^n}{n!} \prod_{l=0}^{n-1} p_l}{1 + \sum_{i=1}^N \frac{\rho^i}{i!} \prod_{l=0}^{i-1} p_l}, n = 1, \dots, N,$$

где $\rho = \frac{\lambda}{\mu}$ - нагрузка на систему.

Заметим, что вектор при этом $\mathcal{P}_{system} = (\mathcal{P}_0, \dots, \mathcal{P}_N)$, где \mathcal{P}_n задается формулой (1), определяет политику, зависящую от состояния системы (state dependent policy, SD).

Зная распределение стационарных вероятностей, можно выписать следующие характеристики системы для случая SD:

- вероятность того, что заявка будет принята в систему:

$$P_{admit}(\rho, f_\beta, N, \mathcal{P}_{system}) = \sum_{n=0}^{N-1} \pi_n \cdot p_n, n = 0, \dots, N - 1;$$

- вероятность блокировки заявки из-за отсутствия свободного ресурса:

$$B(\rho, f_\beta, N, \mathcal{P}_{system}) = \pi_N;$$

- вероятность блокировки заявки из-за несоответствия политике приема заявки в систему:

$$P_{reject} = 1 - P_{admit};$$

- средняя ожидаемая цена при политике приема заявки в систему \mathcal{P}_n , которую арендатор будет платить за выделенный ему ресурс:

$$E[\beta | \beta \in \mathcal{D}_n] = \int_{-\infty}^{+\infty} \beta \cdot p_{\{\beta | \beta \in \mathcal{D}_n\}} d\beta = \frac{1}{p_n} \int_{\mathcal{D}_n} \beta \cdot f(\beta) d\beta;$$



- средний доход, полученный поставщиком инфраструктуры:

$$R_{\beta}(\rho, f_{\beta}, N, \mathcal{P}_{system}) = \rho \cdot \sum_{n=0}^{N-1} \pi_n \cdot p_n \cdot E[\beta | \beta \in \mathcal{D}_n]. \quad (6)$$

Заметим, что случай политики приема заявки в систему, не зависящей от состояния системы (state independent policy, SI), является частным случаем множества политик SD (state dependent policy):

$$\begin{aligned} \mathcal{P}_n &= \mathcal{P}; \\ \mathcal{D}_N &= \mathcal{D}; \\ p_n &= p. \end{aligned} \quad (7)$$

По аналогии со случаем SD, для SI запишем расчетные формулы для тех же величин. Из (7) распределение стационарных вероятностей (2) будет выглядеть следующим образом:

$$\pi_{n(\rho, f_{\beta}, N, \mathcal{P})} = \frac{\frac{(\rho \cdot p)^n}{n!}}{\sum_{i=0}^N \frac{(\rho \cdot p)^i}{i!}}. \quad (8)$$

Зная стационарные вероятности, по аналогии с (3)-(6) можно выписать следующие характеристики системы для случая SI:

- вероятность того, что заявка будет принята в систему:

$$P_{admit}(\rho, f_{\beta}, N, \mathcal{P}) = (1 - \pi_N) \cdot p; \quad (9)$$

- средний доход, полученный поставщиком инфраструктуры:

$$R_{\beta}(\rho, f_{\beta}, N, \mathcal{P}) = \rho \cdot P_{admit} \cdot E[\beta | \beta \in \mathcal{D}]. \quad (10)$$

Возвращаясь к политикам приема заявок в систему, для политики превышения порога AT отказ в обслуживании получает заявка SP с тарифом ниже β , следовательно,

вероятность того, что заявка будет принята в систему в состоянии n имеет вид

$$p_n^{AT}(f_{\beta}, \mathcal{P}_n) = \int_{\beta_n}^{\beta_M} f_{\beta}(\beta) d\beta. \quad (11)$$

Что касается политики постоянного доступа AA, вероятность того, что заявка будет принята в систему в состоянии n имеет вид

$$p_n^{AA}(f_{\beta}, \mathcal{P}_n) = \int_{\beta_m}^{\beta_M} f_{\beta}(\beta) d\beta. \quad (12)$$

Рекомендации о применении той или иной политики поставщиком инфраструктуры могут быть сделаны на основании оценки удельного дохода, то есть среднего дохода на единицу ресурса, и упущенного дохода с точки зрения системы в целом. Расчет удельного дохода R_{unit} произведен по следующей формуле:

$$R_{unit} = \frac{R}{N}, \text{ при } n = 1, \dots, N, \quad (13)$$

упущенный доход R_{lost} вычисляется по следующей формуле:

$$R_{lost} = R_N - R_{\beta}, \text{ при } n = 1, \dots, N \quad (14)$$

Для иллюстрации применения построенной модели к анализу дохода поставщика инфраструктуры в следующем разделе проведен численный эксперимент и сравнение показателей эффективности сетевого слайсинга для нескольких политик приема заявок в систему.

Численный эксперимент

В ходе численного эксперимента проведен анализ зависимости среднего дохода $\ln P$ от структурных и нагрузочных параметров системы. Параметры системы для слайсинга по требованию приведены в Табл. 1.

Таблица 1. Параметры системы для слайсинга по требованию
Table 1. System parameters for on-demand slicing

Параметр	Обозначение	Значение
интенсивность поступления запросов от арендатора	λ	5,100 запрос/с
интенсивность обслуживания запросов	μ	1 запрос/с
общий объем ресурса поставщика инфраструктуры	C	20 Мбит/с
единица канального ресурса	r	2 Мбит/с
емкость системы	N	1, ..., 10 запросов
случайная цена, предлагаемая арендатором поставщику инфраструктуры за обслуживание запроса	β	$\sim U[\beta_m, \beta_M]$ евро/с
диапазон цены	$[\beta_m, \beta_M]$	[3 евро/с, 20 евро/с]
минимальный тариф, принятый поставщиком инфраструктуры	β_m	3 евро/с
максимальный тариф, устраивающий арендатора	β_M	20 евро/с

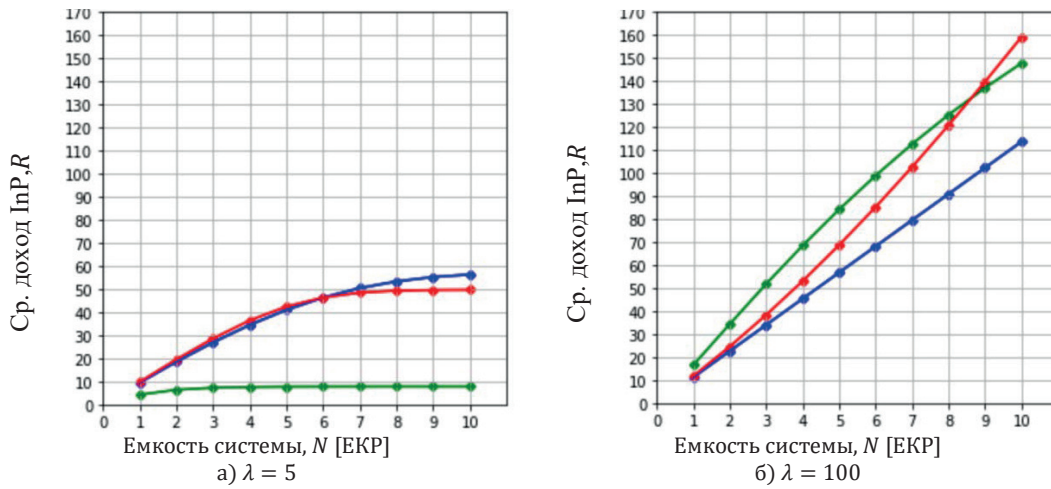
Отметим, для политики приема заявки в систему AA все компоненты вектора пороговых значений равны минимальному тарифу, поэтому $\beta = \beta_m$. Это соответствует множеству политик SI, при котором для анализа используются формулы (8)-(14). Для политики приема заявки в систему AT, соответствующей множеству политик SD, расчет проведен по формулам (2)-(6), при этом рассмотрено 2 случая:

- повышение пороговых значений с ростом N вектор $\beta = \beta_{inc}$;
- понижение пороговых значений с ростом N вектор $\beta = \beta_{dec}$.

Графики на Рис. 5, Рис. 6 и Рис. 7 соответствуют следующим случаям:

- InP применена политика приема заявки в систему AA из множества политик SI, где $\beta = \beta_m$, на графике обозначение политики: AA (синий цвет);
- InP применена политика AT из множества политик SD, где $\beta = \beta_{inc}$, на графике обозначение политики: AT_{inc} (зелёный цвет);
- InP применена политика AT из множества политик SD, где $\beta = \beta_{dec}$, на графике обозначение политики: AT_{dec} (красный цвет).

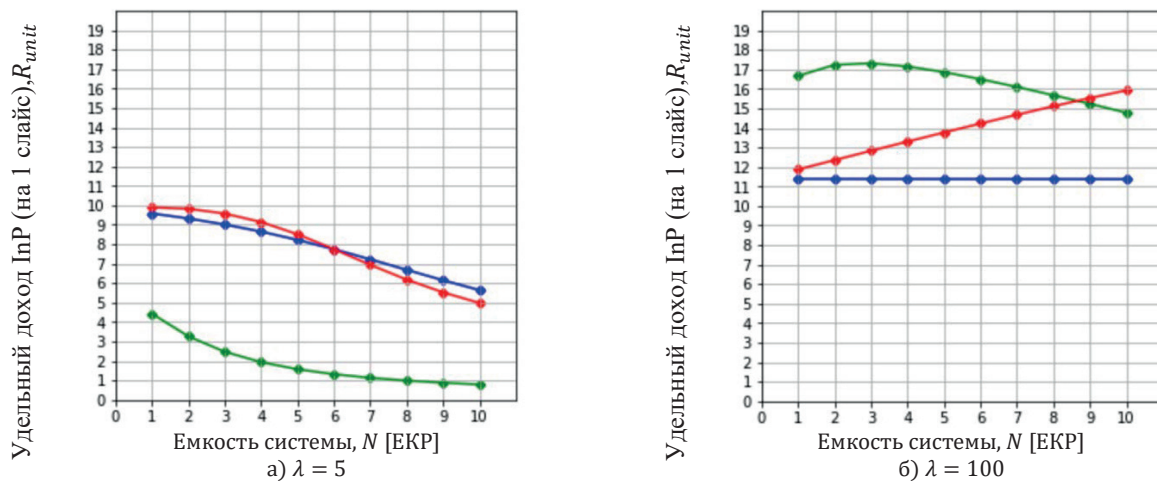




Р и с. 5. Зависимость величины среднего дохода InP от объема ресурса в системе
F i g. 5. Dependence of the average income InP value on the amount of resource in the system

При небольшой нагрузке $\lambda = 5$ (Рис. 5 (а)) зависящая от состояния системы политика приёма заявок в систему AT_{inc} наименее выгодна поставщику инфраструктуры InP . При большой нагрузке $\lambda = 100$ (Рис. 5 (б)) выгодно уменьшать порог приема, поскольку в этом случае при росте интенсивности поступления заявок доход не будет упускаться. Здесь при применении политики приёма заявок в систему AT_{inc} доход поставщика инфраструктуры несколько выше, чем при AT_{dec} , что объясняется тем, что при AT_{inc} InP получает доход за счет приема большего числа поступающих заявок. Применение не зависящей от состояния системы политики приёма заявок в систему AA

будет выгодно для поставщика инфраструктуры в случае низкой нагрузки на систему (Рис. 5 (а)), так как позволяет избежать потери дохода в случае несоответствия предлагаемой арендатором цены и установленного поставщиком порога (для AA пороговое значение $\beta_n = \beta_m$). Таким образом, при рассмотренном наборе значений параметров системы для получения наибольшего дохода поставщику инфраструктуры для случая большой нагрузки на систему рекомендовано использование зависящей от состояния системы политики AT_{inc} , в то время как для малой нагрузки – политики AT_{dec} .

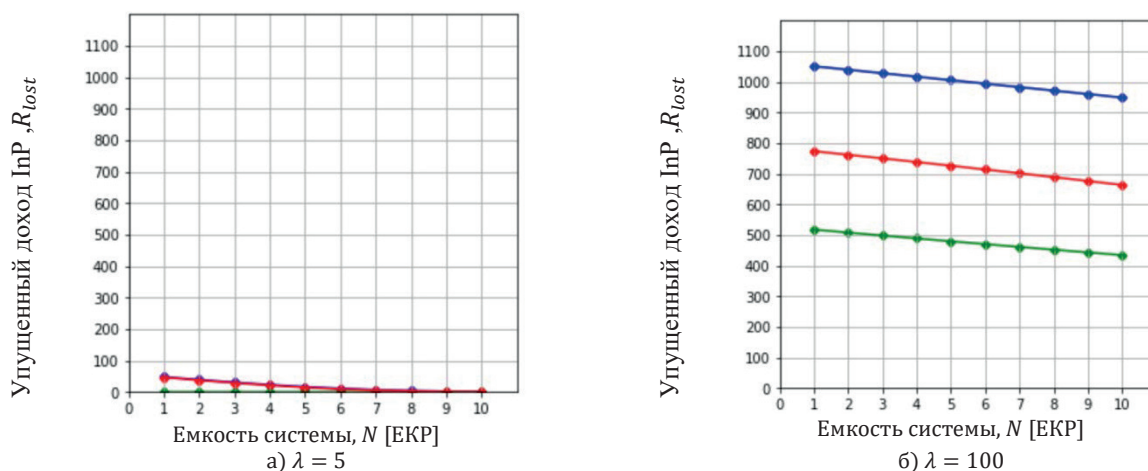


Р и с. 6. Зависимость величины удельного дохода InP от объема ресурса в системе
F i g. 6. Dependence of the value of the specific income InP on the amount of resource in the system



Аналогично ведут себя графики для удельного дохода R_{unit} поставщика инфраструктуры на Рис. 6. Под удельным доходом здесь понимается доход на единицу канального ресурса. Убывание удельного дохода для всех трех политик на Рис. 6 (а) объясняется тем, что невысокой нагрузкой недостаточно для того, чтобы занять весь доступный объем ресурса, и часть ресурса простаивает. Этот же эффект наблюдается для политики AT_{inc} на Рис. (Рис. 6 (б)), когда установленный арендатором высокий порог приема заявок

в систему лишает его дохода при свободном ресурсе. Две остальные политики, AT_{dec} и AA , способствуют занятию всего доступного ресурса, так что даже при большой емкости системы ресурс не простаивает. Как и для случая небольшой нагрузки, при высокой нагрузке выгодно уменьшать пороговое значение для приема заявок в систему, потому что при росте интенсивности поступления заявок доход не опускается, как следствие, удельный доход lnP в таком случае выше.



Р и с. 7. Зависимость величины упущенного дохода lnP от объема ресурса в системе
Fig. 7. Dependence of the amount of lost income lnP on the amount of resource in the system

На Рис. 7 представлены графики зависимостей величины упущенного дохода lnP от объема ресурса в системе. Для случаев высокой и низкой нагрузки на систему при увеличении числа слайсов величина упущенного дохода R_{lost} падает. На Рис. 7 (а) ввиду малой нагрузки на систему с ростом числа слайсов упущенный доход стремится к нулю (для AT_{inc} доход не опускается вовсе). При высокой нагрузке на систему (Рис. 7 (б)) для политики AA поставщик инфраструктуры теряет больше дохода в сравнении с двумя другими политиками. Таким образом, можно сделать аналогичный исследованным выше случаям вывод: как при высокой, так и при низкой нагрузке на систему выгодно уменьшать пороговое значение для приема заявок, что приведет к наименьшим потерям дохода при росте интенсивности поступления заявок. Заметим, что для случая системы с несколькими арендаторами, провайдером услуг, рекомендации могут быть другими. В частности, возможно применение резервирования ресурса для заявок арендатора, приносящих наибольший доход, что можно регулировать выбором соответствующего значения порога в критерии приема запросов в систему.

Заключение

В работе на основе системной модели сетевого слайсинга построена математическая модель в виде моносервисной системы массового обслуживания с потерями, которая позволяет проанализировать влияние стратегии доступа к ресурсам и

политики приема заявок в систему на средний доход поставщика инфраструктуры. Показано отличие стратегий слайсинга по требованию и периодического слайсинга, затем на примере периодического слайсинга дано подробное описание зависящих и не зависящих от состояния системы политик доступа заявок в систему. Получено стационарное распределение вероятностей состояний системы и формулы для анализа основных показателей эффективности системы: вероятность того, что заявка будет принята в систему, вероятность блокировки заявки из-за отсутствия свободного ресурса, вероятность блокировки заявки из-за несоответствия политике приема заявки в систему, средняя ожидаемая цена при политике приема заявки в систему, средний доход, полученный поставщиком инфраструктуры. Для построенной модели проведен численный эксперимент, в котором рассчитаны величины среднего, удельного и упущенного дохода в зависимости от применяемой политики приема заявок в систему для двух стратегий доступа к ресурсам. Задачей дальнейших исследований является расширение модели на случай нескольких арендаторов, провайдеров услуг, с помощью мультисервисной модели с потерями, а также исследование механизма формирования диапазона цены. Верхнее и нижнее значения диапазона цены должно дать решение задачи максимизации дохода поставщика инфраструктуры в ограничениях на емкость системы, требования к ресурсу для различных услуг и прочие структурные и нагрузочные параметры системы.



References

- [1] Shafi M., et al. 5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice. *IEEE Journal on Selected Areas in Communications*. 2017; 35(6):1201-1221. (In Eng.) DOI: <https://doi.org/10.1109/JSAC.2017.2692307>
- [2] Attaran M. The impact of 5G on the evolution of intelligent automation and industry digitization. *Journal of Ambient Intelligence and Humanized Computing*. 2021. vol. 12. 17 p. (In Eng.) DOI: <https://doi.org/10.1007/s12652-020-02521-x>
- [3] Kleine D., Unwin T. Technological Revolution, Evolution and New Dependencies: What's New about ICT4D? *Third World Quarterly*. 2009; 30(5):1045-1067. (In Eng.) DOI: <https://doi.org/10.1080/01436590902959339>
- [4] Lu Y., Chen X., Xi R., Chen Y. An access selection mechanism in 5G network slicing. *2020 IEEE International Conference on Smart Internet of Things (SmartIoT)*. IEEE Press, Beijing, China; 2020. p. 72-78. (In Eng.) DOI: <https://doi.org/10.1109/SmartIoT49966.2020.00020>
- [5] Matencio-Escolar A., Wang Q., Alcaraz Calero J.M. *SliceNetVSwitch: Definition, Design and Implementation of 5G Multi-Tenant Network Slicing in Software Data Paths*. *IEEE Transactions on Network and Service Management*. 2020; 17(4):2212-2225. (In Eng.) DOI: <https://doi.org/10.1109/TNSM.2020.3029653>
- [6] Santos J.F, Kist M., Rochol J., DaSilva L.A. Virtual Radios, Real Services: Enabling RANaaS Through Radio Virtualisation. *IEEE Transactions on Network and Service Management*. 2020; 17(4):2610-2619. (In Eng.) DOI: <https://doi.org/10.1109/TNSM.2020.3009863>
- [7] Liu Y., Ding J., Liu X. A Constrained Reinforcement Learning Based Approach for Network Slicing. *2020 IEEE 28th International Conference on Network Protocols (ICNP)*. IEEE Press, Madrid, Spain; 2020. p. 1-6. (In Eng.) DOI: <https://doi.org/10.1109/ICNP49622.2020.9259378>
- [8] Montero R., Agraz F., Pagès A., Spadaro S. Real-time maintenance of latency-sensitive 5G services through network slicing. *Photonic Network Communications*. 2020; 40(3):221-232. (In Eng.) DOI: <https://doi.org/10.1007/s11107-020-00897-6>
- [9] De Domenico A., Liu Y.-F., Yu W. Optimal Virtual Network Function Deployment for 5G Network Slicing in a Hybrid Cloud Infrastructure. *IEEE Transactions on Wireless Communications*. 2020; 19(12):7942-7956. (In Eng.) DOI: <https://doi.org/10.1109/TWC.2020.3017628>
- [10] Wei F., Feng G., Sun Y., Wang Y., Qin S., Liang Y.-C. Network Slice Reconfiguration by Exploiting Deep Reinforcement Learning With Large Action Space. *IEEE Transactions on Network and Service Management*. 2020; 17(4):2197-2211. (In Eng.) DOI: <https://doi.org/10.1109/TNSM.2020.3019248>
- [11] Gligoroski D., Králevska K. Expanded Combinatorial Designs as Tool to Model Network Slicing in 5G. *IEEE Access*. 2019; 7:54879-54887. (In Eng.) DOI: <https://doi.org/10.1109/ACCESS.2019.2913185>
- [12] Luu Q.-T., Kerboeuf S., Mouradian A., Kieffer M. A Coverage-Aware Resource Provisioning Method for Network Slicing. *IEEE/ACM Transactions on Networking*. 2020; 28(6):2393-2406. (In Eng.) DOI: <https://doi.org/10.1109/TNET.2020.3019098>
- [13] Qin Y., et al. Enabling Multicast Slices in Edge Networks. *IEEE Internet of Things Journal*. 2020; 7(9):8485-8501. (In Eng.) DOI: <https://doi.org/10.1109/JIOT.2020.2991107>
- [14] Zhou F., et al. Automatic Network Slicing for IoT in Smart City. *IEEE Wireless Communications*. 2020; 27(6):108-115. (In Eng.) DOI: <https://doi.org/10.1109/MWC.001.2000069>
- [15] Onur A., Shin'ichi A., Masayuki M. SDN-Based Control of IoT Network by Brain-Inspired Bayesian Attractor Model and Network Slicing. *Applied Sciences*. 2020; 10(17):5773. (In Eng.) DOI: <https://doi.org/10.3390/app10175773>
- [16] Kim D., Kim S. Network slicing as enablers for 5G services: state of the art and challenges for mobile industry. *Telecommunication Systems*. 2019; 71(3):517-527. (In Eng.) DOI: <https://doi.org/10.1007/s11235-018-0525-2>
- [17] Sattar D., Matrawy A. Optimal Slice Allocation in 5G Core Networks. *IEEE Networking Letters*. 2019; 1(2):48-51. (In Eng.) DOI: <https://doi.org/10.1109/LNET.2019.2908351>
- [18] Kreutz D., et al. Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*. 2015; 103:14-76. (In Eng.) DOI: <https://doi.org/10.1109/JPROC.2014.2371999>
- [19] Alam I., et al. A Survey of Network Virtualization Techniques for Internet of Things Using SDN and NFV. *ACM Computing Surveys*. 2020; 53(2):35. (In Eng.) DOI: <https://doi.org/10.1145/3379444>
- [20] Bonfim M.S., Dias K.L., Fernandes S.F.L. Integrated NFV/SDN Architectures: A Systematic Literature Review. *ACM Computing Surveys*. 2019; 51(6):114. (In Eng.) DOI: <https://doi.org/10.1145/3172866>
- [21] Stamou A., Dimitriou N., Kontovasilis K., Papavassiliou S. Autonomic Handover Management for Heterogeneous Networks in a Future Internet Context: A Survey. *IEEE Communications Surveys & Tutorials*. 2019; 21(4):3274-3297. (In Eng.) DOI: <https://doi.org/10.1109/COMST.2019.2916188>
- [22] van der Meer S., Keeney J., Fallon L. 5G networks must be autonomic! *NOMS 2018 – 2018 IEEE/IFIP Network Operations and Management Symposium*. IEEE Press, Taipei, Taiwan; 2018. p. 1-5. (In Eng.) DOI: <https://doi.org/10.1109/NOMS.2018.8406185>
- [23] Khan A.N., Kiah M.L.M., Khan S.U., Madani S.A. Towards secure mobile cloud computing: A survey. *Future Generation Computer Systems*. 2013; 29(5):1278-1299. (In Eng.) DOI: <https://doi.org/10.1016/j.future.2012.08.003>
- [24] Kar U.N., Sanyal D.K. An overview of device-to-device communication in cellular networks. *ICT Express*. 2018; 4(4):203-208. (In Eng.) DOI: <https://doi.org/10.1016/j.icte.2017.08.002>
- [25] Wang X., Kong L., Kong F., Qiu F., Xia M., Arnon S., Chen G. Millimeter Wave Communication: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*. 2018; 20(3):1616-1653. (In Eng.) DOI: <https://doi.org/10.1109/COMST.2018.2844322>
- [26] Vincenzi M., Lopez-Aguilera E., Garcia-Villegas E. Maximizing Infrastructure Providers' Revenue Through Network Slicing in 5G. *IEEE Access*. 2019; 7:128283-128297. (In Eng.) DOI: <https://doi.org/10.1109/ACCESS.2019.2939935>



*Поступила 15.08.2021; одобрена после рецензирования
17.09.2021; принята к публикации 20.09.2021.
Submitted 15.08.2021; approved after reviewing 17.09.2021;
accepted for publication 20.09.2021.*

Об авторах:

Лапшенкова Любовь Олеговна, магистрант кафедры прикладной информатики и теории вероятностей, факультет физико-математических и естественных наук, ФГАОУ ВО «Российский университет дружбы народов» (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6), **ORCID:** <https://orcid.org/0000-0002-1053-4925>, 1032217633@rudn.ru

Москалева Фаина Александровна, аспирант кафедры прикладной информатики и теории вероятностей, факультет физико-математических и естественных наук, ФГАОУ ВО «Российский университет дружбы народов» (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6), **ORCID:** <https://orcid.org/0000-0002-4082-5096>, moskaleva-fa@rudn.ru

Гайдамака Юлия Васильевна, профессор кафедры прикладной информатики и теории вероятностей, факультет физико-математических и естественных наук, ФГАОУ ВО «Российский университет дружбы народов» (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); старший научный сотрудник Института проблем информатики Российской академии наук, ФГУ «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» (119333, Российская Федерация, г. Москва, ул. Вавилова, д. 44-2), доктор физико-математических наук, доцент, **ORCID:** <https://orcid.org/0000-0003-2655-4805>, gaidamaka-yuv@rudn.ru

Все авторы прочитали и одобрили окончательный вариант рукописи.

About the authors:

Lyubov O. Lapshenkova, Master student of the Department of Applied Probability and Informatics, Faculty of Science, Peoples' Friendship University of Russia (6 Miklukho-Maklaya St., Moscow 117198, Russian Federation), **ORCID:** <https://orcid.org/0000-0002-1053-4925>, 1032217633@rudn.ru

Faina A. Moskaleva, Postgraduate Student of the Department of Applied Probability and Informatics, Faculty of Science, Peoples' Friendship University of Russia (6 Miklukho-Maklaya St., Moscow 117198, Russian Federation), **ORCID:** <https://orcid.org/0000-0002-4082-5096>, moskaleva-fa@rudn.ru

Yuliya V. Gaidamaka, Professor of the Department of Applied Informatics and Probability Theory, Faculty of Science, Peoples' Friendship University of Russia (6 Miklukho-Maklaya St., Moscow 117198, Russian Federation); Senior Scientist of the Institute of Informatics Problems of Russian Academy of Sciences, Federal Research Center "Computer Science and Control" of Russian Academy of Sciences (44-2 Vavilov St., Moscow 119333, Russian Federation), Dr.Sci. (Phys.-Math.), Associate Professor, **ORCID:** <https://orcid.org/0000-0003-2655-4805>, gaidamaka-yuv@rudn.ru

All authors have read and approved the final manuscript.

