

Оптимизация разделения радиоресурсов и интенсивностей перехода состояний в двухсервисной модели сети 5G с эластичным типом трафика

С.Н. Гончаренко, И.Р. Радимов

Национальный исследовательский технологический университет «МИСИС»

Аннотация: Беспроводные сети 5G вызывают большой интерес для изучения. Network Slicing — одна из ключевых технологий, которая позволяет эффективно использовать ресурсы в сетях пятого поколения. В этой работе рассматривается метод распределения ресурсов в беспроводных сетях 5G с использованием технологии Network Slicing. В работе была рассмотрена модель доступа к радиоресурсам сети, которая включает ряд решений для повышения эффективности обслуживания за счёт настройки логической части сети. В этой модели применяются технология нарезки сети и эластичный трафик. В практической части работы было построено матрицы интенсивностей перехода для двух различных конфигураций.

Ключевые слова: 5G, двухсервисная СМО, минимальная гарантированная скорость, нарезка сети, распределение ресурса, система массового обслуживания, эластичный трафик.

Введение. При разработке архитектуры сетей пятого поколения учитывались определённые требования, связанные с функциональными особенностями новых сервисов. В отличие от предыдущего поколения мобильных сетей, архитектура сетей пятого поколения характеризуется не только более высокой пиковой скоростью передачи данных, но и рядом концептуальных отличий. Основное внимание уделяется гибкости архитектуры и замене части физического оборудования сетевыми программными функциями.

В данной работе будет рассмотрена модель схема доступа к радиоресурсам сети, содержащая ряд решений по увеличению эффективности обслуживания средствами конфигурации логической части сети. В модели будет использоваться эластичный трафик, а также технология нарезки сети.

Поиск различных конфигураций и анализ моделей сетей, в особенности сетей с эластичным типом трафика, является актуальной задачей. Актуальность исследования таких моделей обусловлена в первую очередь

гибкостью в управлении трафиком. Конфигурация из нескольких слайсов сети позволяет выделять гарантированные ресурсы для конкретного типа заявок. Кроме того, в зависимости от потребностей в данный период времени можно увеличивать или уменьшать количество ресурсов в том или ином слайсе сети. Как следствие – операторы имеют широкие возможности по оптимизации ресурсов связи, сохраняя высокий уровень качества обслуживания.

Также важно подчеркнуть актуальность самой технологии нарезки сети Network Slicing. 3GPP в своей документации (3GPP TS 23.501 V17.4.0) рассматривает разделение сети как одну из ключевых функций всего стандарта 5G. В связи с этим у технологии очевидно большое будущее, а следовательно, и решения связанных с ней задач будут востребованы.

Архитектура сетей пятого поколения. Физическая часть сети представлена следующим образом. Элемент (R)AN (Radio Access Network) означает сеть радиодоступа, которая содержит систему физических базовых станций (рис. 1), для связки UE с 5GC (5G Core). Для построения сетей пятого поколения используется соответствующая сеть радиодоступа NR с базовыми станциями gNB (Next Generation NodeB) и ng-eNB.

gNB и ng-eNB связаны между собой через интерфейс Xn. Согласно этому интерфейсу, сигнальные пакеты имеют гарантированную доставку, в отличие от пакетов пользовательского уровня. Кроме того, Xn выполняет пересылку данных и управление потоком, облегчает управление мобильностью UE и двойное подключение. Узлы 5G NR подключаются к ядру 5G через интерфейс NG.

Помимо потенциала физической части сети, существуют обширные возможности конфигурации 5G Core. Стандарт 5G позволяет повышать эффективность обслуживания абонентов, не применяя сложных и дорогих технологий для NR.

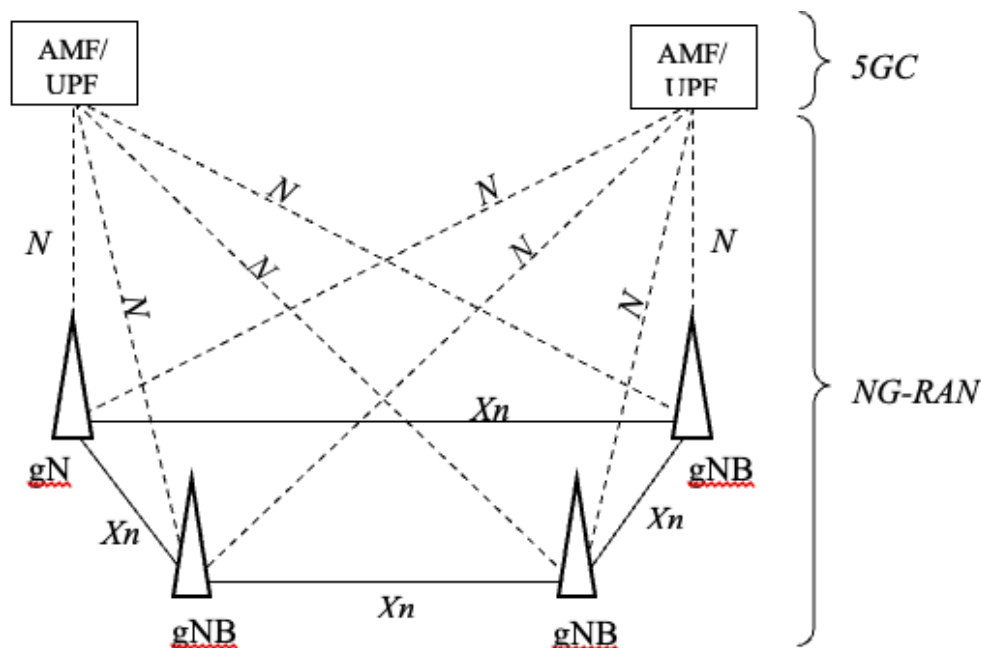


Рисунок 1. Схема сети радиодоступа (R)AN.

Один из базовых способов повышения эффективности обслуживания абонентов – использование эластичного трафика вместо потокового.

Потоковый трафик генерируется услугами с гарантированной скоростью передачи данных. Потоковый трафик характеризуется в первую очередь фиксированной скоростью и фиксированным временем обслуживания. Этот тип трафика малочувствителен к потерям и задержкам. Примерами приложений с потоковым трафиком являются приложения видео и аудио телефония. Эластичный трафик характеризуется переменной скоростью передачи данных, переменным временем обслуживания и заданной фиксированной длиной передаваемого фрагмента файла. Такой тип трафик отличается низкой чувствительностью к задержкам [1, 2].

При эластичном трафике скорость передачи данных зависит от доступных ресурсов соты и во время обслуживания может меняться, а следовательно, при чрезмерной критической загрузке соты возможна потеря пакетов. Эластичный трафик генерируется такими приложениями, как передача файлов по протоколам FTP и P2P, электронная почта, сигнальный

трафик, разного рода службы HTTP, различные мультимедийные услуги [3, 4].

С точки зрения теории массового обслуживания потоковый и эластичный трафик можно описать следующим образом. Пусть пиковая пропускная способность соты равна $C = 6$. Интенсивность поступления запросов на предоставление услуг составляет одну заявку в единицу времени. В момент времени в системе нет заявок, свободны все ресурсы сети [5].

В случае потокового трафика каждая поступающая заявка занимает одну единицу емкости сети. В результате имеем свободные ресурсы в сети, при этом соблюдается гарантированная скорость передачи данных, так как на одну заявку приходится одна единица емкости сети (рис. 2А).

В случае эластичного трафика первая заявка займет сразу все ресурсы сети, соответственно скорость передачи данных будет высокой. Далее ресурсы сети распределяются равномерно между запросами на обслуживание. В итоге можем наблюдать меньшее время обслуживания, но и меньшую скорость передачи данных при загрузке сети (рис. 2Б).

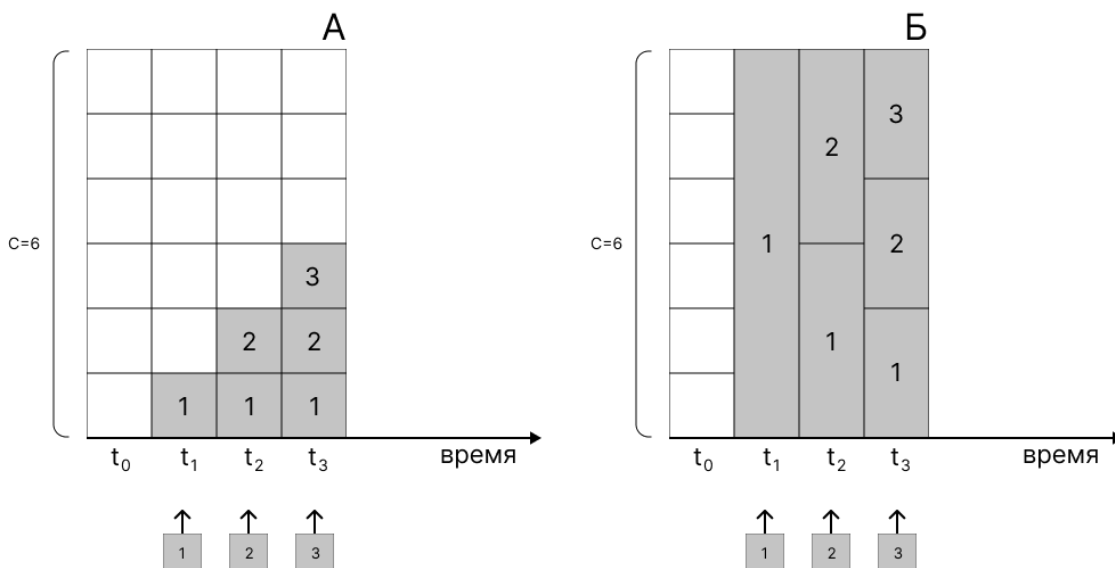


Рисунок 2. Распределение ресурсов сети при потоковом (А) и эластичном трафике (Б).

Наиболее важной технологией для эффективного использования, заложенного в архитектуру опорной сети 5G потенциала, является Network Slicing или нарезка сети. 3GPP рассматривает разделение сети как одну из ключевых функций 5G. Разделение сети на части является важной возможностью для повышения эффективности использования сетевых ресурсов, гибкости развертывания и поддержки быстрорастущих приложений и сервисов. Эта технология позволяет разделять физическую архитектуру на требуемое количество виртуальных слоев. Каждый из слайсов имеет разный набор сетевых функций и ресурсов, необходимых для предоставления требуемых телекоммуникационных услуг [6, 7].

Часть емкости сети может резервироваться как гарантированная для определенного типа запросов. При поступлении запросов на обслуживание заявка занимает выделенные её типу ресурсы в приоритетном порядке, при необходимости вытесняя с гарантированных ресурсов заявки другого типа.

Постановка задачи. Необходимо построить диаграмму интенсивности перехода двухсервисной модели с эластичным типом трафика в общем виде. Рассмотреть примеры функционирования двухсервисной модели с эластичным типом трафика. Построить диаграммы интенсивностей перехода для рассмотренных конфигураций.

Модель должна обеспечивать внутреннюю изоляцию логических ресурсов сети для различных типов заявок, минимальную гарантированную скорость передачи данных для поступающих в сеть заявок.

В качестве примеров функционирования модели рассмотреть следующие конфигурации.

Первая конфигурация. Имеется базовая станция сети с емкостью $C = 8$, которая используется совместно двумя слайсами емкостью $C_1 = 4$, $C_2 = 6$. В соответствии с приоритетными правилами BG трафика емкость первого слайса C_1 имеет гарантированную часть $Q_1 = 3$, в которой приоритет отдаётся

заявкам первого типа; емкость второго слайса C_2 имеет гарантированную часть $Q_2 = 5$, в которой приоритет отдаётся заявкам второго типа.

Также стоит рассмотреть пример работы сети, где часть ресурсов не гарантируется ни одному из слайсов, но принадлежит обоим.

Вторая конфигурация. Рассматривается базовая станция сети с емкостью $C = 8$, которая используется совместно двумя слайсами емкостью $C_1 = 5, C_2 = 6$. В соответствии с приоритетными правилами BG трафика емкость первого слайса C_1 имеет гарантированную часть $Q_1 = 4$, в которой приоритет отдаётся заявкам первого типа; емкость второго слайса C_2 имеет гарантированную часть $Q_2 = 3$, в которой приоритет отдаётся заявкам второго типа.

В обеих конфигурациях в систему обслуживания поступают два потока заявок с интенсивностями λ_1 и λ_2 , каждая заявка требует для обслуживания $b_1 = 1$ и $b_2 = 1$ единиц ресурса соответственно.

Реализация. Ввиду особенностей эластичного трафика схема передачи трафика требует пояснений. Рассматривается система емкостью C (рис. 3). В начальный момент времени t_0 в системе нет ни одной заявки. Далее на обслуживание поступает первая заявка. В момент времени t_1 на её обслуживание выделяются все доступные ресурсы в объеме C .

После этого в момент времени t_2 поступает на обслуживание вторая заявка, ресурсы сети распределяются равномерно между заявками, соответственно $\frac{C}{2}$ на каждую заявку. В момент времени t_3 поступает ещё одна, третья заявка. Тогда на обслуживание каждой выделено $\frac{C}{3}$. Такой формат распределения ресурсов сети между заявками и есть эластичность. Заявки при поступлении на обслуживание не блокируются при таком типе трафика, ресурсы сети делятся равномерно между всеми обслуживаемыми в момент времени $t_k, k=1, 2, \dots$ заявками. По истечению времени обслуживания первой

заявки в системе остаются вторая и третья заявки, а емкость сети перераспределяется между ними в размере $\frac{C}{2}$ для каждой и так далее.

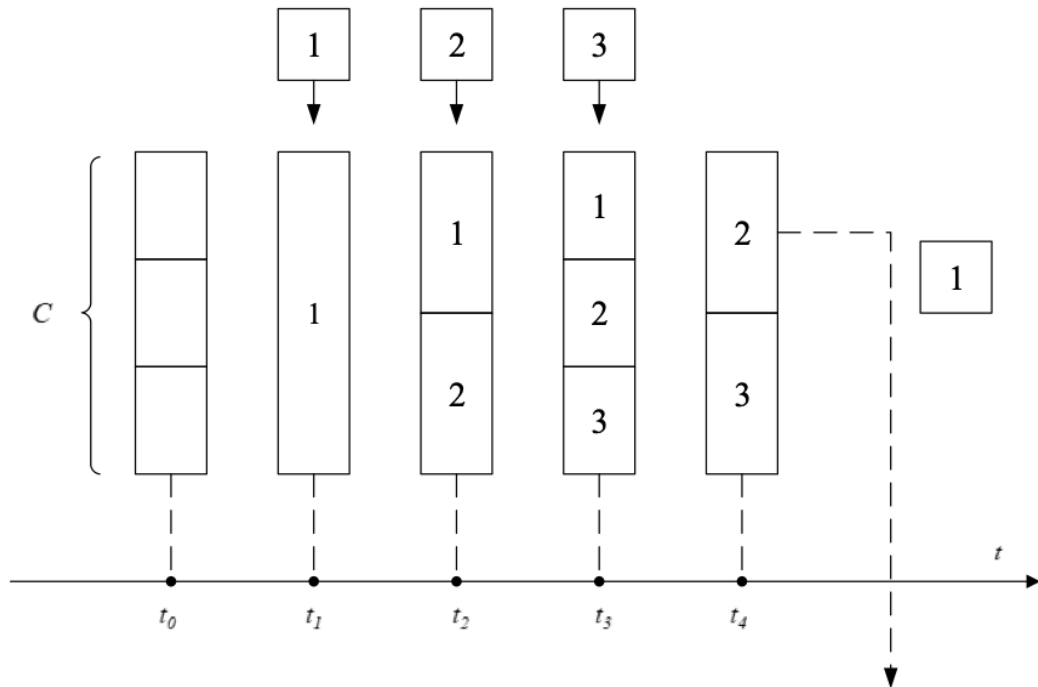


Рисунок 3. Схема передачи эластичного трафика.

Для начала рассмотрим двухсервисную модель с эластичным типом трафика без указания слайсов для разного типа заявок. Пространство состояний системы в силу эластичности трафика состоит из следующих пар чисел $X = \{(x,y): x, y \geq 0\}$. Таким образом диаграмма интенсивностей перехода двухсервисной модели с эластичными трафиком выглядит следующим образом (рис. 4) [8].

Данная модель является простейшей моделью с эластичным трафиком. В ней не регламентируется минимальной скорости передачи пакетов данных, из-за чего при поступлении большого количества заявок на обслуживание они будут приняты, но скорость передачи данных будет крайне низкой.

Следующим этапом к рассмотренной модели необходимо применить технологию нарезки сети. Тем самым установится приоритет обслуживания, обеспечится минимальная гарантированная скорость передачи данных.

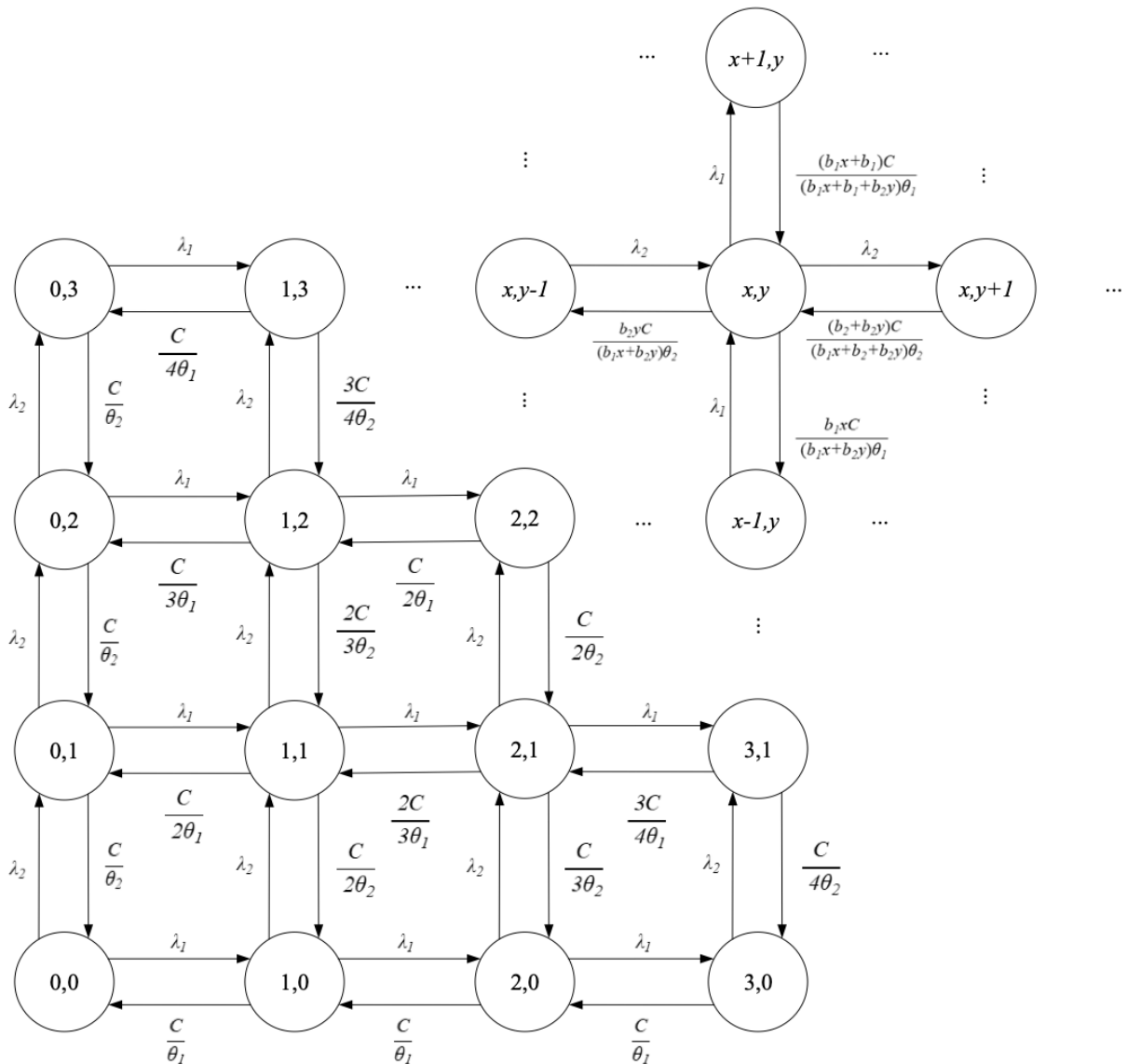


Рисунок 4. Диаграмма интенсивностей перехода двухсервисной модели с эластичными трафиком.

В качестве основы модели будет рассматриваться двухсервисная модель с двумя слайсами со следующими параметрами: общая емкость участка сети составляет C [ед. емкости], емкость первого слайса составляет C_1 [ед. емкости], емкость второго слайса составляет C_2 [ед. емкости], $C_1 + C_2 \geq C$. Слайс C_k состоит из гарантированной (для обслуживания k -го типа заявок) части емкостью Q_k и общедоступной части емкостью $C_k - Q_k$. Подробно данная модель описана в статье [9]. В общем виде схема модели представлена на рисунке 5.

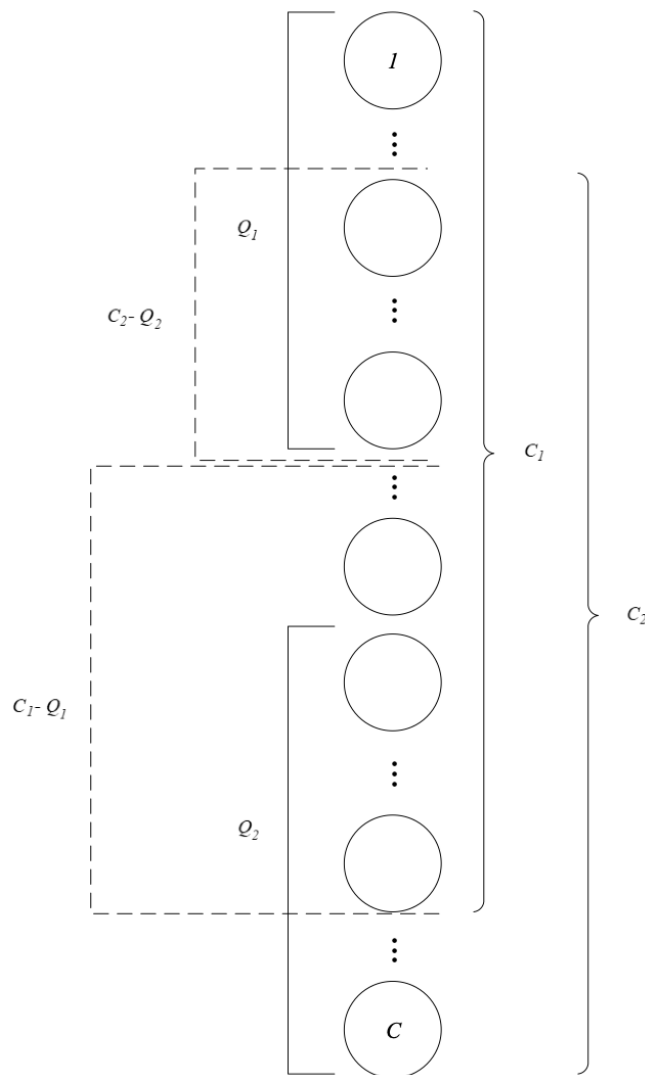


Рисунок 5. Схема доступа к радиоресурсам беспроводной сети с двумя слайсами и трафиком BG

Рассмотрим первую конфигурацию данной модели. Особенности этой конфигурации следующие:

- Гарантированные части слайсов в сумме содержат ресурсов не более, чем вся емкость участка сети;
- Слайсы в сумме содержат ресурсов не менее, чем вся емкость участка сети;
- Гарантированные части меньше размера самого слайса;
- Общая емкость сети составляет 8 условных единиц ресурса;

- Емкость первого слайса составляет 4 условных единиц ресурса, при этом гарантированная часть составляет 3 условные единицы;
- Емкость второго слайса составляет 6 условных единиц ресурса, при этом гарантированная часть составляет 5 условных единиц;
- Для обслуживания заявок обоих типов необходимо по 1 единице ресурса.

Таким образом, пользуясь выведенными статье [10] формулами расчёта элементов матрицы интенсивности, изобразим матрицу интенсивности переходов для заданной конфигурации (рис. 6). Пользуясь построенной матрицей, можем определить пространства состояний. Заявки первого типа блокируются в шести состояниях: (3,5), (4,0), (4,1), (4,2), (4,3), (4,4). Заявки второго типа блокируются в четырех состояниях: (0,6), (1,6), (2,6), (3,5). Заявки первого типа вытесняются заявками второго типа в состоянии (4, 4). Заявки второго типа вытесняются заявками первого типа в состоянии (2, 6).

Рассмотрим вторую конфигурацию модели. Особенности этой конфигурации следующие:

- Гарантированные части слайсов в сумме содержат ресурсов не более, чем вся емкость участка сети;
 - Слайсы в сумме содержат ресурсов не менее, чем вся емкость участка сети;
 - Гарантированные части меньше размера самого слайса;
 - Общая емкость сети составляет 8 условных единиц ресурса;
 - Емкость первого слайса составляет 5 условных единиц ресурса, при этом гарантированная часть составляет 4 условные единицы;
 - Емкость второго слайса составляет 6 условных единиц ресурса, при этом гарантированная часть составляет 3 условных единиц;
-

- Для обслуживания заявок обоих типов необходимо по 1 единице ресурса.

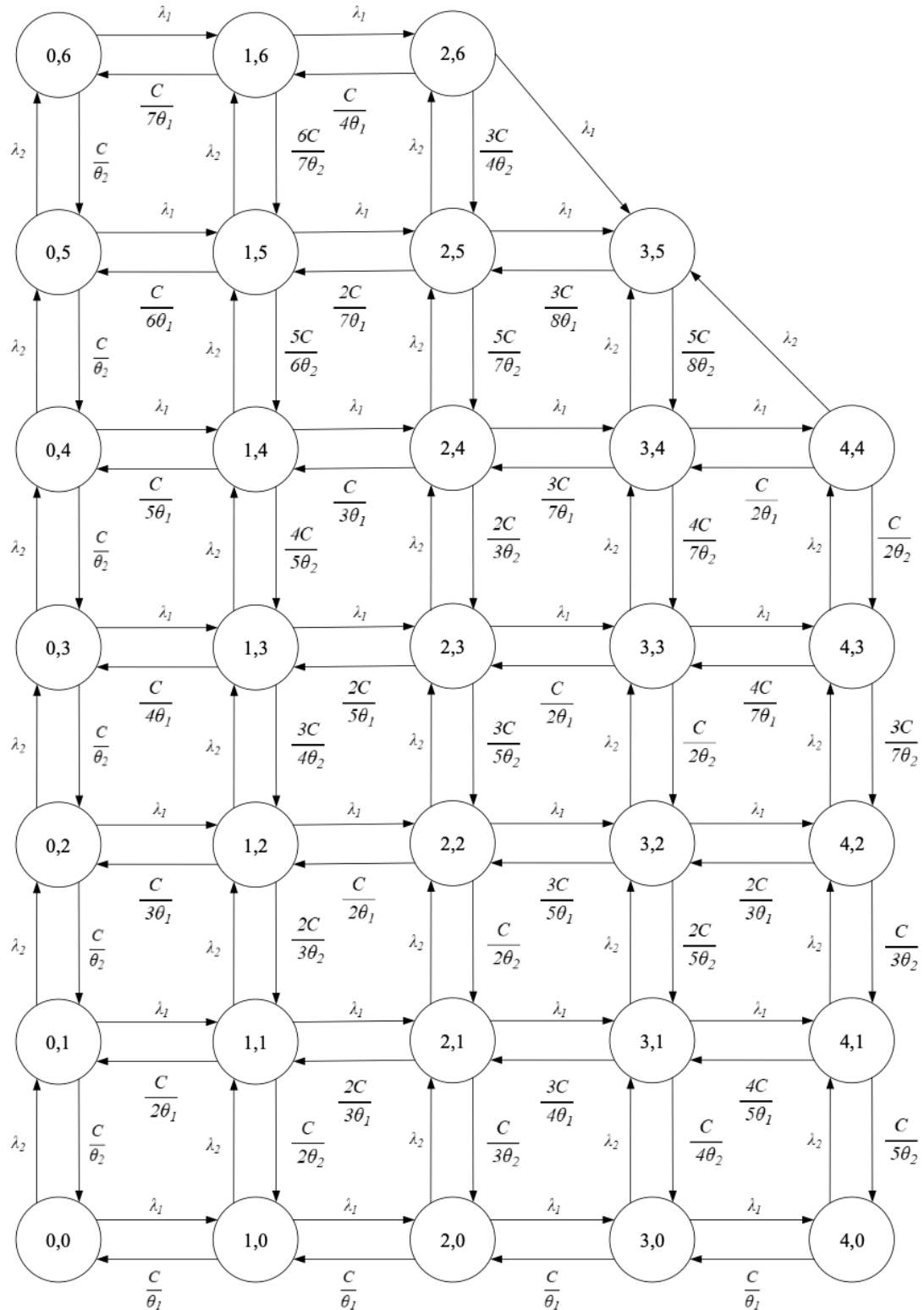


Рисунок 6. Матрица интенсивностей перехода I конфигурации.

Матрица интенсивности перехода состояний для конфигурации II представлена на рисунке 7.

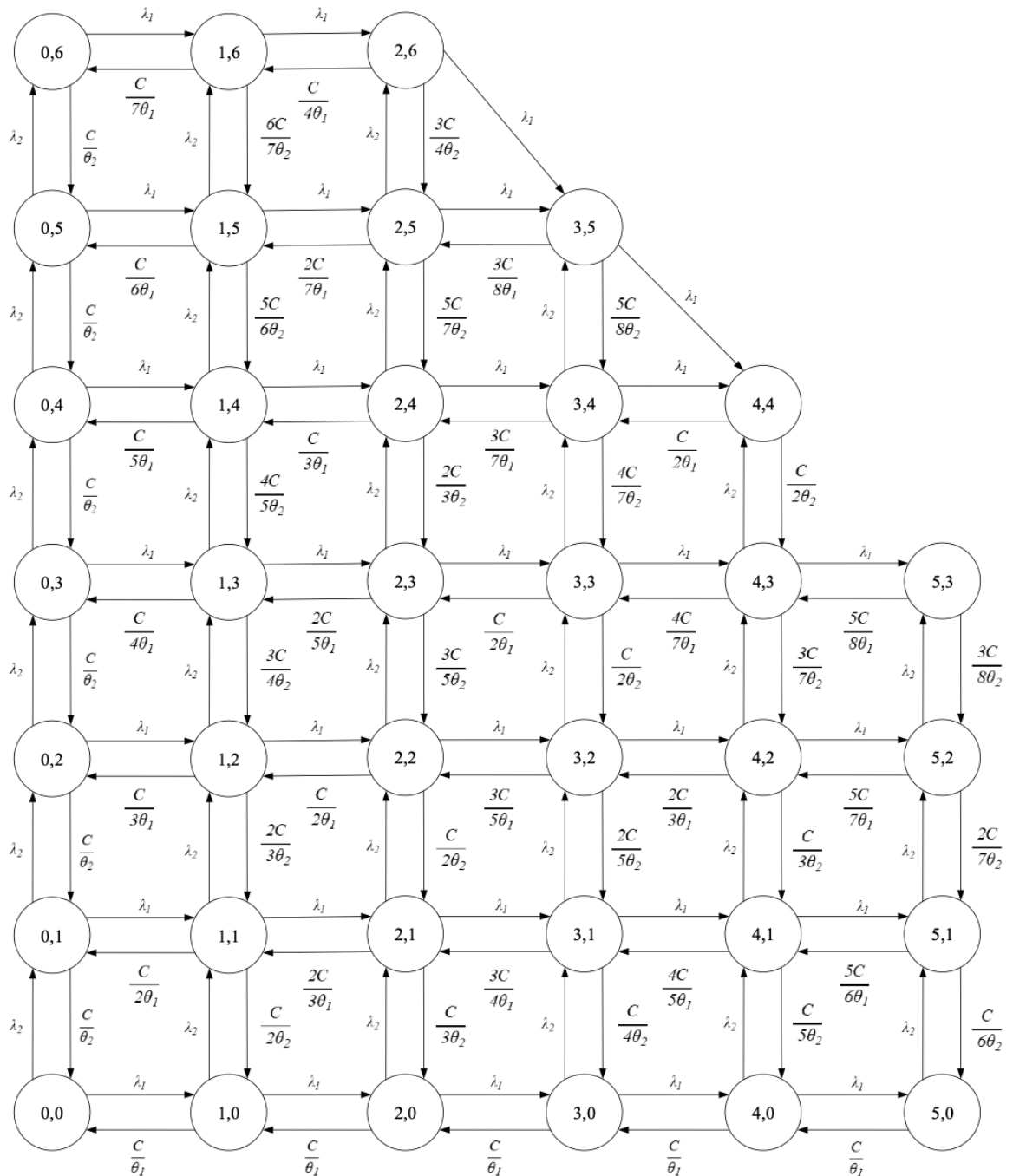


Рисунок 7. Матрица интенсивностей перехода II конфигурации.

Определим пространства состояний. Заявки первого типа блокируются в шести состояниях: (4,0), (4,1), (4,2), (4,3), (4,4). Заявки второго типа блокируются в шести состояниях: (0,6), (1,6), (2,6), (3,5), (4,4), (5,3). Заявки

первого типа не вытесняются заявками второго типа. Заявки второго типа вытесняются заявками первого типа в состоянии (2,6), (3,5).

Заключение. В работе была рассмотрена модель схема доступа к радиоресурсам сети, содержащая ряд решений по увеличению эффективности обслуживания средствами конфигурации логической части сети. Эластичный трафик позволил обслуживать заявки более эффективно в случаях, когда имеются свободные ресурсы в сети. Нарезка сети на слайсы с выделением гарантированных частей для разного типа заявок с сохранением общедоступных ресурсов обеспечивает гарантированную минимальную скорость обслуживания и расставляет приоритет обслуживания заявок.

В дальнейшем следует произвести анализ модели методом имитационного моделирования. Результаты моделирования помогут сделать вывод о целесообразности применимости модели в реальных сценариях использования.

Литература.

1. Башарин Г. П., Самуйлов К. Е., Яркина Н. В., Гудкова И. А. Новый этап развития математической теории телетрафика. М.: Автоматика и телемеханика, 2009. № 12. С. 16-28.

2. Савич В.Н., Дымова П.И., Поляков Н.А., Власкина А.В., Гудкова И.А. К анализу системы массового обслуживания с двумя очередями и нетерпеливым эластичным трафиком с минимальной скоростью // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем. М.: РУДН, 2019. С. 103-107.

3. Haenggi M., Andrews J. G., Baccelli F., Dousse O., Franceschetti M., Stochastic geometry and random graphs for the analysis and design of wireless networks. IEEE journal on selected areas in communications №27 (7), 2009. pp. 1029–1046. URL: ieeexplore.ieee.org/abstract/document/5226957.

4. Самуйлов К.Е., Маркова Е.В. К приближенному анализу вероятности блокировки для модели схемы доступа с индивидуальными потолками скорости передачи эластичного трафика в сети LTE. IX Международная научно-практическая конференция «Современные информационные технологии и ИТ-образование»: Сб. трудов (Ноябрь 14–16, 2014, Москва). С. 417–425.

5. Begishev V., Moltchanov D., Sopin E., Samuylov A., Andreev S., Koucheryavy Y., Samouylov K.. Quantifying the Impact of Guard Capacity on Session Continuity in 3GPP New Radio Systems. IEEE Transactions on Vehicular Technology №12, 2019. p. 1. URL: ieeexplore.ieee.org/document/8878110.

6. Popovski P., Trillingsgaard K. F., Simeone O., Durisi G., 5G Wireless Network Slicing for eMBB, URLLC, and mMTC: A Communication-Theoretic View. 2018. IEEE Access, №6, pp. 55765-55779. URL: ieeexplore.ieee.org/document/8476595.

7. Яркина Н. В., Гайдамака Ю. В., Самуйлов К. Е. О задаче разделения ресурсов при нарезке сети 5G. Материалы международной научно-технической конференции «Инновационные технологии в телекоммуникациях», 2019. С. 74–76.

8. Гарибян А.А., Волынец А.С., Самуйлов К.Е. Модель обслуживания гетерогенного трафика для анализа технологии нарезки беспроводной сети LTE. Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем, М.: РУДН, 2019. С. 31-24.

9. Гончаренко С.Н., Радимов И.Р. Моделирование вероятностных характеристик блокировки запросов на предоставление доступа к радиоресурсам беспроводной сети // Инженерный вестник Дона, 2024. №3. URL: ivdon.ru/ru/magazine/archive/n3y2024/9100.

10. Радимов И.Р. Сценарное моделирование схемы доступа к радиоресурсам беспроводной сети с двумя слайсами // Современная наука:

актуальные проблемы теории и практики. Серия: Естественные и технические науки, 2023. № 12. С. 114-120.

References

1. Basharin G. P., Samujlov K. E., Yarkina N. V., Gudkova I. A. Avtomatika i telemexanika [Automation and telemechanics], Avtomatika i telemexanika. Moskva, 2009. № 12. pp. 16-28.

2. Savich V.N., Dy`mova P.I., Polyakov N.A., Vlaskina A.V., Gudkova I.A. Informacionno-telekommunikacionny`e texnologii i matematicheskoe modelirovanie vy`sokotexnologichny`x system [Information and telecommunication technologies and mathematical modeling of high-tech systems]. Moskva: RUDN, 2019. pp. 103-107.

3. Haenggi M., Andrews J. G., Baccelli F., Dousse O., Franceschetti M., IEEE journal on selected areas in communications №27 (7), 2009. pp. 1029–1046. URL: ieeexplore.ieee.org/abstract/document/5226957.

4. Samujlov K.E., Markova E.V. IX Mezhdunarodnaya nauchno-prakticheskaya konferenciya «Sovremenny`e informacionny`e texnologii i IT-obrazovanie»: Sb. trudov [IX International Scientific and Practical Conference "Modern Information Technologies and IT Education": Proceedings], Moskva, 2014. pp. 417–425.

5. Begishev V., Moltchanov D., Sopin E., Samuylov A., Andreev S., Koucheryavy Y., Samouylov K., IEEE Transactions on Vehicular Technology №12, 2019. p. 1. URL: ieeexplore.ieee.org/document/8878110.

6. Popovski P., Trillingsgaard K. F., Simeone O., Durisi G., IEEE Access, 2018. №6, pp. 55765-55779. URL: ieeexplore.ieee.org/document/8476595.

7. Yarkina N. V., Gajdamaka Yu. V., Samujlov K. E. Materialy` mezhdunarodnoj nauchno-texnicheskoj konferencii «Innovacionny`e texnologii v telekommunikacijax», 2019. pp. 74–76.



8. Garibyan A.A., Voly`necz A.S., Samujlov K.E. Informacionno-telekommunikacionny`e texnologii i matematicheskoe modelirovanie vy`sokotexnologichny`x system [Information and telecommunication technologies and mathematical modeling of high-tech systems], Moskva. RUDN. 2019. pp. 31-24.

9. Goncharenko S.N., Radimov I.R. Inzhenernyj vestnik Dona, 2024, №3. URL: ivdon.ru/ru/magazine/archive/n3y2024/9100.

10. Radimov I.R. Sovremennaya nauka: aktual`ny`e problemy` teorii i praktiki. Seriya: Estestvenny`e i texnicheskie nauki, 2023, № 12. pp. 114-120.

Дата поступления: 16.03.2024

Дата публикации: 26.04.2024