

На правах рукописи

**Ндайикунда Жувен**

**РАЗРАБОТКА И АНАЛИЗ МОДЕЛИ ДИНАМИЧЕСКОГО РАСПРЕДЕЛЕНИЯ  
РЕСУРСА БЕСПРОВОДНЫХ УЗЛОВ ДОСТУПА ПРИ ПЕРЕДАЧЕ  
НЕОДНОРОДНОГО ТРАФИКА IoT**

Специальность 2.2.15—

Системы, сети и устройства телекоммуникаций

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени

кандидата технических наук

Москва, 2022

Работа выполнена в ордене Трудового Красного Знамени федеральном государственном бюджетном образовательном учреждении высшего образования «Московский технический университет связи и информатики» (МТУСИ).

**Научный руководитель:** **Степанов Сергей Николаевич** – доктор технических наук, профессор, зав. кафедрой «Сети связи и системы коммутации» МТУСИ.

**Официальные оппоненты:** **Карташевский Вячеслав Григорьевич** – доктор технических наук, профессор, зав. кафедрой «Информационная безопасность» федерального государственного бюджетного образовательного учреждения высшего образования «Поволжский государственный университет телекоммуникаций и информатики» (ФГБОУ ВО ПГУТИ);

**Гайдамака Юлия Васильевна** – доктор физико-математических наук, профессор, профессор кафедры «Прикладная информатика и теория вероятностей» федерального государственного автономного образовательного учреждения высшего образования «Российский университет дружбы народов» (ФГАОУ ВО РУДН),

**Ведущая организация:** Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» (ФГБОУ ВО СПбГУТ)

Защита состоится 02 июня 2022 г. в 15.00 часов на заседании диссертационного совета по защите докторских и кандидатских диссертаций 55.2.002.01 (Д 219.001.04) при МТУСИ по адресу: 111024, г. Москва, ул. Авиамоторная, д. 8а, МТУСИ, аудитория А-211.

С диссертацией можно ознакомиться в библиотеке и на сайте МТУСИ:

<http://www.srd-mtuci.ru/images/Dis-Ndayikunda/dis-Ndayikunda.pdf>.

Автореферат разослан «\_\_\_» \_\_\_\_\_ 2022 г.

Ученый секретарь диссертационного совета

55.2.002.01 (Д 219.001.04) д.т.н, доцент \_\_\_\_\_ Максим Валерьевич Терешонок

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы исследования.** Одной из основных тенденций развития телекоммуникаций является необходимость совместного обслуживания выделенным ресурсом потоков информационных сообщений, отличающихся существенным разнообразием в требованиях к ресурсу и качеству обслуживания. Часто источниками информационных сообщений являются устройства телеметрии, видеокамеры и т.п. Подобные устройства являются элементами сети виртуального оператора, предоставляющего услуги сбора и обработки данных разного рода наблюдений. Эта деятельность регулируется положениями концепции интернета Вещей.

Нередко виртуальные сети разворачиваются в местах, где ограничено или вообще не имеется возможности применения фиксированной проводной связи. Это вынуждает виртуального оператора использовать ресурс беспроводных сетей для обслуживания возникающих информационных потоков. В силу известных причин<sup>1</sup>, этот ресурс ограничен и должен использоваться с максимальной эффективностью.

Чтобы добиться этого результата, нужно решить следующие две задачи. Во-первых, построить модель формирования и обслуживания сессий связи, которая более точно отражает реалии работы действующих узлов беспроводного доступа. Во-вторых, предложить и исследовать сценарии распределения ресурса между поступающими потоками информационных сообщений, которые бы позволили создать условия для их дифференцированного обслуживания. Иначе, как показали численные эксперименты, с ростом нагрузки на канал происходит неконтролируемое перераспределение ресурса в пользу потоков сессий с относительно малыми требованиями к скорости передачи. Этот результат может нарушить принятое соглашение об обслуживании. Решение сформулированных задач позволит находить предпочтительные соотношения между параметрами потоков запросов на информационное обслуживание и характеристиками пропускной способности мультисервисного узла доступа, обеспечивающие гарантированное качество обслуживания клиентов. Именно эти вопросы рассматривались в диссертационной работе, что говорит об актуальности выбранной тематики.

**Степень разработанности темы.** Поставленная задача решалась на базе моделей и методов теории телетрафика, а также возможностей, заложенных в механизмы управления процессом обслуживания сессий связи в современных беспроводных мультисервисных узлах доступа. Различным аспектам решения данной задачи посвящены работы российских и зарубежных

---

<sup>1</sup>В их число входят ограничения физического плана, а также действия регулятора, направленные на создание конкуренции.

авторов. В их числе: Г.П. Башарин, В.М. Вишнеvский, Ю.В. Гайдамака, В.Г. Карташевский, А.Е. Кучеряvый, Е.А. Кучеряvый, В.А. Наумов, А.П. Пшеничников, К.Е. Самуйлов, С.Н. Степанов, М.С. Степанов, И.И. Цитович, и др., а также – T.Bonald, F.P. Kelly, V.B. Iversen, K.W. Ross, J. Virtamo и др. Отдельные вопросы построения и исследования моделей распределения ресурса в беспроводных узлах доступа рассматривались в диссертационных работах: С.Д. Андреева, В.О. Бегишева, Е.А. Кучеряvого, К.А. Агеева и др. Анализ публикаций и выполненных диссертационных исследований показал, что в большинстве теоретических работ либо изучалось действие какого-то одного фактора на процесс распределения ресурса узла доступа (например, зависимость требования к ресурсу от типа сервиса, ограничение доступа, резервирование ресурса и т.д.), либо процесс распределения ресурса рассматривался с избыточной детальностью, что в итоге затрудняло использование построенной математической модели. Задача построения модели, которая, с одной стороны, отражала основные реалии распределения ресурса, а с другой — могла бы использоваться в практических приложениях не рассматривалась, что и определило направление исследований, выполненных в диссертации.

**Цели и задачи работы.** Целью исследования является разработка и анализ процедуры динамического распределения ресурса беспроводного узла доступа, направленной на создание условий по дифференцированному обслуживанию неоднородного трафика и повышению эффективности использования ресурса передачи информации. Для достижения указанной цели необходимо решить следующие частные научные задачи: разработать модель динамического распределения ресурса беспроводного узла доступа при обслуживании неоднородного трафика при наличии ограничения по доступу; определить характеристики качества обслуживания поступающих сессий связи; построить алгоритмы оценки характеристик; сформулировать рекомендации по эффективному распределению ресурса между поступающими потоками разнородного трафика.

#### **Научная новизна.**

1. Построена и исследована обобщенная модель обслуживания неоднородного трафика в беспроводном узле доступа, которая в отличие от известных моделей позволила учесть совместное влияние основных значимых факторов, определяющих совместное обслуживание трафика реального времени и эластичных данных. Среди них: наличие приоритета у трафика реального времени; использование дисциплины Processor Sharing при передаче эластичного трафика; ограничение по доступу для всех видов трафика, зависящее от общего уровня занятости ресурса.

2. Получены выражения для оценки характеристик качества обслуживания заявок через значения входных параметров и стационарных вероятностей обобщенной модели беспроводного узла доступа. В отличие от более ранних исследований, полученные выражения позволяют анализировать действие разного рода процедур, направленных на повышение эффективности использования ресурса передачи узлов доступа и создание условий по дифференцированному обслуживанию потоков неоднородного трафика, основанных на ограничении доступа, зависящего от общего уровня занятости ресурса.
3. Построена система уравнений статистического равновесия, связывающая значения стационарных вероятностей модели и разработан алгоритм ее решения. В отличие от известных реализаций других стандартных методов разработанный алгоритм позволяет вести оценку характеристик для моделей с числом состояний в несколько миллионов, что достаточно для исследования условий по дифференцированному обслуживанию поступающих потоков неоднородного трафика для большинства практических приложений.

**Теоретическая и практическая значимость работы.** Теоретическая значимость работы заключается в построении и исследовании обобщенной модели обслуживания неоднородного трафика в беспроводном узле доступа, которая позволила учесть совместное влияние основных значимых факторов, определяющих совместное обслуживание трафика реального времени и эластичных данных, а также в разработке алгоритмов расчета характеристик подобных моделей. Получены программные реализации построенных в диссертации алгоритмов. Разработанный инструментарий рекомендуется использовать для создания условий по дифференцированному обслуживанию гетерогенного трафика в беспроводных узлах доступа и теоретическом обосновании действий администрации, направленных на повышение эффективности использования ресурса передачи. Результаты диссертации использованы в учебном процессе на кафедре «Сети связи и системы коммутации» МТУСИ. Реализация результатов работы подтверждена соответствующим актом.

**Методы исследования.** Для решения поставленной задачи применялись методы теории телетрафика, теории вероятностей и вычислительной математики.

**Основные положения, выносимые на защиту:**

1. Построенная обобщенная модель обслуживания неоднородного трафика в беспроводном узле доступа позволяет учесть совместное влияние основных значимых факторов, определяющих совместное обслуживание трафика реального времени и эластичных данных.

Среди них: наличие приоритета у трафика реального времени; использование дисциплины Processor Sharing при передачи эластичного трафика; ограничение по доступу для всех видов трафика, зависящее от общего уровня занятости ресурса.

2. Для оценки значений характеристик качества совместного обслуживания сессий трафика реального времени и эластичного трафика данных в рамках построенной модели беспроводного узла доступа рекомендуется использовать метод, основанный на решении системы уравнений равновесия итерационным алгоритмом Гаусса-Зейделя. Этот подход позволяет рассчитать характеристики для моделей с числом состояний в несколько миллионов, что достаточно для исследования условий по дифференцированному обслуживанию поступающих потоков неоднородного трафика для большинства практических приложений.
3. Разработанная модель и алгоритмы оценки ее характеристик позволяют анализировать действие разного рода процедур, направленных на повышение эффективности использования ресурса передачи узлов доступа и создание условий по дифференцированному обслуживанию потоков неоднородного трафика, основанных на ограничении доступа, зависящего от общего уровня занятости ресурса. Среди них динамичный слайсинг, когда распределение выделенного объема ресурса осуществляется на динамической основе и зависит от его загрузки. Для ограничения доступа сессий здесь предлагается использовать процедуру резервирования, основанную на фильтрации поступающих сессий с использованием функции внутренней блокировки. Другой сценарий — статичный слайсинг. Для данного сценария имеющийся ресурс делится между поступающими потоками в определенной пропорции, зависящей от требований сессий связи к показателям качества обслуживания.
4. Выполненное численное исследование показало, что использование динамического слайсинга позволяет на 5–20% уменьшить требование к объему ресурса, обеспечивающего требуемый уровень потерь сессий, по сравнению с применением для этих же целей статичного слайсинга. Наибольший эффект применение предложенной версии динамического слайсинга приносит в ситуации обслуживания эластичного трафика данных с использованием дисциплины Processor Sharing.

**Степень достоверности и апробация результатов.** Полученные теоретические результаты обоснованы доказательствами с использованием математических методов теории телетрафика, подтверждены численными экспериментами. Достоверность положений и выводов диссертации

подтверждается апробацией работы, основные результаты которой обсуждались и докладывались на международной научно-технической конференции «Технологии информационного общества» (Москва, 2019 — 2021 гг.), на отраслевой научно-технической конференции «Телекоммуникационные и вычислительные системы» (Москва, 2019 гг.), на международной научной конференции «Systems of Signals Generating and Processing in the Field of on Board Communications» (Москва, 2020 — 2021 гг.), на международной научной конференции «Conference of Open Innovation Association, FRUCT» (Москва, 2019 гг.), на международной научной конференции «Distributed Computer and Communication Networks: Control, Computation, Communications» (Москва, 2020 — 2021 гг.). По материалам диссертации опубликованы 14 работ, в том числе 3 — в изданиях, включенных в список ВАК РФ и 4 в изданиях, входящих международную базу цитирования SCOPUS.

**Объем и структура работы.** Диссертация состоит из введения, четырех разделов, заключения, списка литературы и приложения. Основная часть (без приложения) изложена на 134 страницах машинописного текста, содержит 46 рисунков и 10 таблиц; список литературы состоит из 111 наименований. Приложение изложено на 1 странице машинописного текста.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

**Во введении** обоснованы выбор темы диссертации, ее актуальность, научная новизна, перечислены цели и основные задачи исследования, сформулированы основные положения, выносимые на защиту.

**В первом разделе** диссертации анализируется использование беспроводных узлов доступа для обслуживания трафика современных коммуникационных приложений. Показана необходимость применения в узлах процедур распределения ресурса, позволяющих создать условия по дифференцированному совместному обслуживанию потоков сессий трафика реального времени и сессий передачи данных. Для научно обоснованной оценки требуемого объема ресурса передачи информации и формулировки процедур его эффективного распределения необходимо построить модель узла доступа, учитывающую перечисленные выше особенности формирования и обслуживания сессий связи, а также показать возможность применения модели для решения практических задач, возникающих в процессе проектирования и эксплуатации беспроводных узлов доступа. Несмотря на актуальность проблематики, анализ выполненных исследований показал

недостаточность полученных ранее результатов для решения рассмотренных выше задач. Первая глава заканчивается постановкой задачи исследования.

**Второй раздел диссертации** посвящен построению и исследованию обобщенной модели распределения ресурса в беспроводном узле доступа. Анализируемая система связи представляет собой изолированную соту сети стандарта LTE, часть ресурса которой (выделенный слайс) арендуется оператором систем наблюдения для передачи информационных потоков, иницируемых видеокамерами и разного рода датчиками. Предполагается, что информационные потоки для части видеокамер (сессии связи) передаются в режиме реального времени, т.е. требуют для своего обслуживания фиксированную скорость передачи на все время соединения. Совместно с трафиком реального времени обслуживаются запросы на передачу эластичных данных. К ним относятся, например, видеокамеры с записью видеоконтента в буфер, а также датчики, генерирующие файлы измерений большого объема, например, фотографии высокого качества. Соответствующие сессии связи обладают свойством эластичности, т.е. могут менять скорость передачи в некоторых пределах без потери качества предоставления услуги. Предполагается, что каждая отдельная видеокамера функционирует как обычный пользователь услуг сети LTE, т.е. напрямую соединяется с базовой станцией. Функциональная модель изолированной соты сети LTE, часть ресурса которой арендуется оператором систем наблюдения показана на рисунке 1.

Пусть:  $C$  — скорость передачи информации, обеспечиваемая всем имеющимся ресурсом, выраженная в битах в секунду,  $r$  — скорость передачи информации, обеспечиваемая одной виртуальной канальной единицей (к.е.), а  $v$  — общее число к.е. В модели рассматривается процесс поступления и обслуживания  $n$  потоков запросов на пересылку трафика сессий реального времени и один поток заявок на пересылку трафика сессий эластичных данных. Запросы  $k$ -го потока поступают по пуассоновскому закону с интенсивностью  $\lambda_k$ , для обслуживания одной сессии требуется  $b_k$  единиц ресурса, а длительность сессии имеет экспоненциальное распределение с параметром  $\mu_k$ , где  $k = 1, 2, \dots, n$ . Если в момент поступления запроса  $k$ -го потока нет достаточного числа свободных канальных единиц, но это количество может быть получено в результате уменьшения скорости передачи находящихся на обслуживании сессий эластичных данных, то соответствующее изменение скоростей выполняется и поступивший запрос принимается на обслуживание.





Рисунок 1 — Функциональная модель узла доступа

Пусть в момент поступления запроса  $i_r$  — число единиц ресурса, занятых на обслуживание трафика реального времени, а  $d$  — число обслуживаемых сессий на передачу эластичных данных. Запрос  $k$ -го потока принимается к обслуживанию, если выполняется условие  $i_r + d + b_k \leq v$ . Выполнение данного условия означает, что все заявки на пересылку файлов уже получают для своего обслуживания один канал и дальнейшее уменьшение скорости передачи эластичных данных невозможно. Таким образом, в ситуации  $i_r + d + b_k > v$  поступивший запрос получает отказ и не возобновляется. Аналогично, поступление запросов на передачу эластичных данных подчиняется пуассоновскому закону с интенсивностью  $\lambda_d$ . Поступивший запрос принимается к обслуживанию, если выполняется условие  $i_r + d + 1 \leq v$ . Объём пересылаемого файла, выраженный в битах, имеет экспоненциальное распределение со средним значением  $F$ . Обозначим через  $\mu_d = r/F$  параметр экспоненциального распределения времени передачи файла с использованием передаточных возможностей одной канальной единицы.



определённым на конечном пространстве состояний. В него входят вектора  $(i_1, \dots, i_n, d)$  с компонентами  $i_1, \dots, i_n, d$ , принимающими значения:

$$\begin{aligned} i_1 &= 0, 1, \dots, \left\lfloor \frac{v}{b_1} \right\rfloor; & i_2 &= 0, 1, \dots, \left\lfloor \frac{v - i_1 b_1}{b_2} \right\rfloor; \\ & \dots & & \dots \\ i_n &= 0, 1, \dots, \left\lfloor \frac{v - i_1 b_1 - \dots - i_{n-1} b_{n-1}}{b_n} \right\rfloor; \\ d &= 0, 1, \dots, v - i_1 b_1 - \dots - i_n b_n. \end{aligned}$$

Множество таких векторов составляет пространство теоретически возможных состояний  $\Omega$ . Пространство реально используемых состояний  $S \subset \Omega$  и определяется выбором  $\varphi_k(i)$ . Пусть  $p(i_1, \dots, i_n, d)$  — стационарная вероятность состояния  $(i_1, \dots, i_n, d) \in S$ . Она имеет интерпретацию доли времени пребывания выделенного ресурса в состоянии  $(i_1, \dots, i_n, d)$  и может использоваться для оценки характеристик совместного обслуживания сессий связи оператора систем наблюдения.

Построенный марковский процесс дает возможность определить показатели качества обслуживания поступающих сессий через  $p(i_1, \dots, i_n, d)$ . Для вычисления  $p(i_1, \dots, i_n, d)$  необходимо сформировать и решить систему уравнений статистического равновесия, связывающую  $p(i_1, \dots, i_n, d)$ . По мнению экспертов наиболее эффективным является подход, основанный на использовании итерационных алгоритмов Гаусса-Зейделя или Якоби. Для их реализации систему уравнений равновесия необходимо записать в виде одного соотношения с коэффициентами при неизвестных вероятностях  $p(i_1, \dots, i_n, d)$ , зависящими от компонент состояния  $(i_1, \dots, i_n, d)$ . Данное соотношение имеет вид:

$$\begin{aligned} P(i_1, \dots, i_n, d) & \left\{ \sum_{k=1}^n \left( \lambda_k (1 - \varphi_k(i_r + d)) + i_k \mu_k I(i_k > 0) \right) + \right. & (1) \\ & \left. + \lambda_d (1 - \varphi_d(i_r + d)) + (v - i_r) \mu_d I(d > 0) \right\} = \\ & = \sum_{k=1}^n P(i_1, \dots, i_k - 1, \dots, i_n, d) \lambda_k (1 - \varphi_k(i_r + d - b_k)) I(i_k > 0) + \\ & + P(i_1, \dots, i_n, d - 1) \lambda_d (1 - \varphi_d(i_r + d - 1)) I(d > 0) + \end{aligned}$$

$$\begin{aligned}
& + \sum_{k=1}^n P(i_1, \dots, i_k + 1, \dots, i_n, d) (i_k + 1) \mu_k I(i_r + d + b_k \leq v) + \\
& + P(i_1, \dots, i_n, d + 1) (v - i_r) \mu_d I(i_r + d + 1 \leq v).
\end{aligned}$$

Здесь  $I(\cdot)$  — индикаторная функция. Её значение равно единице, если сформулированное условие выполняется, и равно нулю в противоположном случае. Для  $P(i_1, \dots, i_n, d)$  выполнено условие нормировки.

Введем показатели обслуживания  $k$ -го потока трафика реального времени. Пусть:  $\pi_k$  — доля потерянных запросов;  $m_k$  — среднее число занятых единиц ресурса;  $y_k$  — среднее число обслуживаемых запросов. Тогда:

$$\begin{aligned}
\pi_k &= \sum_{(i_1, \dots, i_n, d) \in S} p(i_1, \dots, i_n, d) \varphi_k(i_r + d); \\
m_k &= \sum_{(i_1, \dots, i_n, d) \in S} p(i_1, \dots, i_n, d) i_k b_k; \quad y_k = m_k / b_k.
\end{aligned}$$

Определим показатели обслуживания эластичного трафика. Пусть:  $\pi_d$  — доля потерянных запросов;  $y_d$  — среднее число запросов, находящихся на обслуживании;  $I_d$  — интенсивность окончания обслуживания сессий;  $k_d$  — среднее число единиц ресурса, используемых на обслуживание одной сессии;  $T_d$  — среднее время обслуживания сессии. Тогда:

$$\begin{aligned}
\pi_d &= \sum_{(i_1, \dots, i_n, d) \in S} p(i_1, \dots, i_n, d) \varphi_d(i_r + d); \\
y_d &= \sum_{(i_1, \dots, i_n, d) \in S} p(i_1, \dots, i_n, d) d; \\
I_d &= \sum_{(i_1, \dots, i_n, d) \in S | d > 0} p(i_1, \dots, i_n, d) (v - i_r) \mu_d; \\
k_d &= \frac{I_d}{y_d \mu_d}; \quad T_d = \frac{y_d}{\lambda_d (1 - \pi_d)}.
\end{aligned}$$

Получены соотношения между характеристиками обслуживания сессий, которые имеют характер законов сохранения интенсивностей поступающих и обслуженных системой потоков заявок. Найденные соотношения можно использовать для вычисления значений характеристик и косвенной оценки сходимости итерационного метода решения системы уравнений равновесия. Указанные соотношения имеют вид

$$\lambda_k = \lambda_k \pi_k + y_k \mu_k, \quad k = 1, 2, \dots, n; \quad \lambda_d = \lambda_d \pi_d + I_d.$$

Приведенные соотношения позволяют установить альтернативные формулы, которые упрощают оценку введенных характеристик

$$T_d = \frac{1}{\mu_d} \cdot \frac{1}{k_d}.$$

Таким образом, среднее время передачи файла равно среднему времени передачи файла одним каналом, деленному на среднее число каналов, составляющим один макроканал.

**Третий раздел диссертации** посвящен разработке и анализу алгоритмов оценки характеристик качества совместного обслуживания трафика реального времени и эластичного трафика данных с резервированием. Анализ численных методов расчета характеристик совместного обслуживания сессий связи в мультисервисных узлах доступа показал, что наиболее эффективным способом их точного вычисления являются процедуры, основанные на формировании и последующем решении системы линейных уравнений статистического равновесия с помощью итерационного метода Гаусса-Зейделя. Выбор алгоритма основан на использовании свойств матрицы системы уравнений равновесия, среди которых важнейшими являются большое число неизвестных и большое количество нулевых элементов, а также наличие рекуррентных зависимостей для оценки коэффициентов. Приведена формулировка основных шагов алгоритма для мультисервисного узла доступа. Среди них: выбор начальных значений итерационного цикла; рекурсивная формула, связывающая величины последовательных приближений; анализ условий сходимости и критерий останова итерационного процесса.

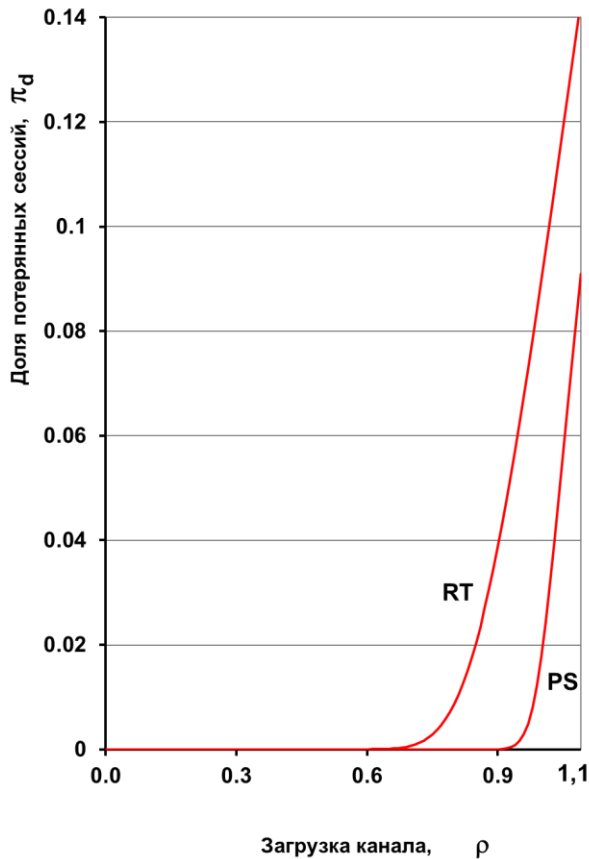
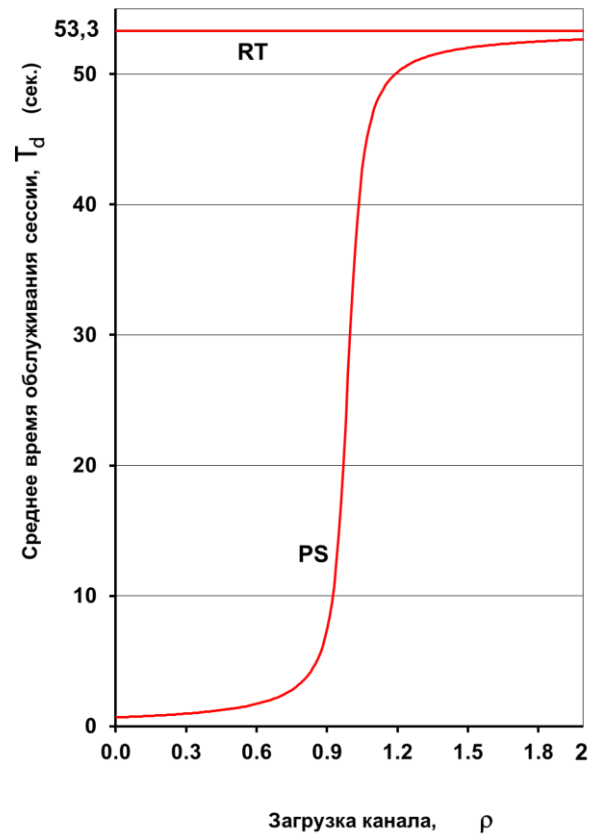
В целях повышения эффективности реализации итерационного алгоритма Гаусса-Зейделя получено выражение для формирования системы уравнений статистического равновесия в виде цикла по целочисленным компонентам состояния. Полученное соотношение дает возможность определять коэффициента матрицы системы уравнений непосредственно в процессе применения итерационного алгоритма, а не хранить их в памяти компьютера. Этот результат существенно упрощает реализацию итерационного метода и позволяет увеличить число состояний в исследуемой модели мультисервисного узла доступа до нескольких миллионов.

Построена вычислительная процедура оценки характеристик обслуживания трафика сервисов реального времени на отдельном слайсе с использованием возможностей ограниченного доступа. Сформулированы определения характеристик обслуживания заявок и рассмотрены алгоритмы их точной и приближенной оценки. Точный расчет характеристик выполняется с использованием итерационных методов. Приведены рекурсивные выражения для оценки последовательных значений стационарных вероятностей. В основу приближенного алгоритма положено

предположение о возможности выполнения соотношений детального баланса. В результате для оценки характеристик используются рекурсивные формулы, связывающие вероятности отдельных макросостояний с фиксированным числом занятых каналов. Численно исследована точность приближенного алгоритма. Показано, что относительная погрешность вычисления характеристик лежит в пределах нескольких процентов для большинства практически интересных случаев.

Построена вычислительная процедура оценки характеристик обслуживания эластичного трафика данных на отдельном слайсе. Сформулированы определения характеристик обслуживания сессий передачи эластичного трафика. В рассматриваемом случае для оценки показателей обслуживания сессий можно использовать явные формулы и рекурсивные алгоритмы. Приведены соответствующие расчетные выражения через значения входных параметров модели. Полученные результаты использовались при численном исследовании эффективности применения процедуры processor Sharing (PS) при обслуживании сессий передачи эластичного трафика. Предположим, что базовая станция может обслужить одновременно максимум 80 абонентов, предоставив каждому пользователю ее услуг скорость не ниже 1 Мбит/с. Из перечисленных данных следуют величины параметров модели передачи эластичного трафика, используемые при проведении вычислений. Предположим, что:  $\nu = 80$ , минимальная скорость одного виртуального канала передачи информации  $r = 1$  Мбит/с, средний объем передаваемого файла  $F$  примем равным 10 Мбайт или 80 Мбит. Величина  $\mu_d$  определяется из соотношения  $\frac{1}{\mu_d} = \frac{80}{1} = 80$  с. Это максимальное среднее время передачи файла с использованием передаточных возможностей одного канала.

Исследована зависимость характеристик передачи эластичного трафика  $\pi_d$  и  $T_d$  от изменения  $\rho = \frac{\beta}{C} = \frac{\lambda}{\nu\mu_d}$  потенциальной нагрузки на один канал и проведено сравнение значений характеристик с их аналогами, рассчитанными при условии передачи файлов по правилам обслуживания трафика реального времени, т.е. при использовании одного канала для передачи одного файла. В этом случае значения характеристик рассчитываются с использованием мультисервисной модели Эрланга. На рисунках 3 и 4, соответственно, показана зависимость потерь сессий и среднего времени передачи файла от  $\rho$ . Символ PS на кривой означает использование дисциплины PS, а символ RT означает передачу файла по правилам трафика реального времени.

Рисунок 3 — Зависимость потерь сессий от  $\rho$ Рисунок 4 — Зависимость среднего времени обслуживания сессии от  $\rho$ 

Приведенные данные показывают эффективность дисциплины PS: в области малых потерь (до 10%) доля потерь сессий и среднее время передачи файла могут уменьшиться более чем в десять раз.

**Четвертый раздел диссертации** посвящен анализу особенностей совместного обслуживания информационных потоков с ярко выраженной неоднородностью требований заявок к ресурсу передачи. Проведенное исследование показало, что с ростом нагрузки на канал происходит неконтролируемое перераспределение ресурса в пользу потоков сессий с относительно малыми требованиями к скорости передачи. Этот результат может нарушить принятое соглашение об обслуживании. Избавиться от перечисленных трудностей можно создав условия по дифференцированному обслуживанию входящих информационных потоков. Рассмотрены особенности реализации двух сценариев:

- Статичный слайсинг (Static Slicing — SS). Для данного сценария имеющийся ресурс делится между поступающими потоками в определенной пропорции, зависящей от требований сессий связи к показателям качества обслуживания. Выделение определенного объема ресурса, который носит название «слайс», выполняется для группы информационных потоков с примерно одинаковыми требованиями к скорости передачи, либо для одного потока, если отмеченное объединение потоков невозможно.
- Динамичный слайсинг (Dynamic Slicing — DS). В рассматриваемом сценарии распределение выделенного объема ресурса осуществляется на динамической основе и зависит от его загрузки. До определенного уровня занятости ресурса он используется всеми поступающими потоками сессий связи. С увеличением загрузки ресурса часть поступающих потоков заявок получает отказ, создавая тем самым приоритет в использовании ресурса у выделенной группы информационных потоков. Для ограничения доступа сессий предлагается использовать процедуру резервирования, основанную на фильтрации поступающих сессий с использованием функции внутренней блокировки.

Показано, что эффективность реализации каждого из предложенных сценариев можно исследовать с помощью комплекса моделей совместного обслуживания неоднородного трафика и алгоритмов оценки их вероятностных характеристик, введенных и исследованных во втором и третьем разделах диссертации. Разработанные алгоритмы отличаются высокой эффективностью реализации и могут применяться для всех практически интересных значений входных параметров. С использованием разработанных алгоритмов проведено сравнение предложенных сценариев создания условий по дифференцированному обслуживанию неоднородного трафика. Сравнение проводилось по результатам решения следующих двух задач:

- Для заданного объема ресурса  $\nu$ , выраженного в к.е., найти разделение ресурса на слайсы с тем, чтобы поступающие потоки сессий обслуживались с одинаковыми характеристиками, выраженными в значениях доли потерянных сессий для трафика реального времени и данных.
- Найти минимальный объем ресурса  $\nu$ , выраженный в к.е., и разделение ресурса на слайсы с тем, чтобы поступающие потоки сессий обслуживались с требуемыми  $\pi$  одинаковыми значениями характеристик, выраженными в значениях доли потерянных сессий для трафика реального времени и данных.



Перечисленные задачи решались для двух моделей генерации и обслуживания гетерогенного трафика, встречающихся в практических приложениях. представляющего из себя смесь «тяжелого» трафика видеоконтента и «легкого» трафика данных. В первой модели трафик данных представляет из себя сессии передачи видеоконтента с низким качеством, требующим относительно невысокую скорость передачи. Он обслуживается по правилам трафика реального времени. Каждый файл передается с использованием возможностей одной канальной единицы.

Во второй модели трафик данных обладает эластичными свойствами, например, представляя из себя файлы, получающиеся после записи видеоконтента в буфер. Он обслуживается по правилам эластичного трафика. Минимальный объем используемого ресурса составляет одну канальную единицу.

Выполненное численное исследование показало, что использование динамического слайсинга позволяет на 5 – 50% уменьшить потери при дифференцированном обслуживании, направленном на выравнивание потерь сессий на фиксированном объеме ресурса, и на 5 – 20% уменьшить требование к объему ресурса, обеспечивающего требуемый уровень потерь сессий, по сравнению с применением для этих же целей статичного слайсинга. Наибольший эффект применение предложенной версии динамического слайсинга приносит в ситуации обслуживания эластичного трафика данных с использованием дисциплины Processor Sharing.

## **ЗАКЛЮЧЕНИЕ**

Основные результаты работы состоят в следующем.

1. Построена и исследована обобщенная модель обслуживания неоднородного трафика в беспроводном узле доступа, которая в отличии от известных моделей позволила учесть совместное влияние основных значимых факторов, определяющих совместное обслуживание трафика реального времени и эластичных данных. Среди них: наличие приоритета у трафика реального времени; использование дисциплины Processor Sharing при передаче эластичного трафика; ограничение по доступу для всех видов трафика, зависящее от общего уровня занятости ресурса.
2. С использованием модели получены выражения для оценки характеристик качества обслуживания заявок через значения входных параметров и стационарных вероятностей обобщенной модели беспроводного узла доступа. Среди них для каждого типа трафика: доли потерянных сессий, средний объем занятого ресурса, среднее время доставки

сообщения, средний объем ресурса, используемый каждой сессией и т.д. Полученные выражения позволяют анализировать действие разного рода процедур, направленных на повышение эффективности использования ресурса передачи узлов доступа и создание условий по дифференцированному обслуживанию потоков неоднородного трафика, основанных на ограничении доступа, зависящего от общего уровня занятости ресурса.

3. Получено алгебраическое представление системы уравнений равновесия исследуемой модели беспроводного узла доступа в виде, удобном для последующей реализации метода Гаусса-Зейделя. Найденное выражение дает возможность записать все уравнения системы в виде одного соотношения с коэффициентами, вычисляемыми с помощью рекуррентных формул, зависящих от компонент состояния модели. Это значительно упрощает реализацию метода и дает возможность увеличить число состояний в исследуемой модели до нескольких миллионов.
4. Получены соотношения между характеристиками обслуживания сессий, которые имеют характер законов сохранения интенсивностей поступающих и обслуженных системой потоков заявок. Найденные соотношения можно использовать для вычисления значений характеристик и косвенной оценки сходимости итерационного метода решения системы уравнений равновесия.
5. Разработанная модель и алгоритмы оценки ее характеристик позволяют анализировать действие разного рода процедур, направленных на повышение эффективности использования ресурса передачи узлов доступа и создание условий по дифференцированному обслуживанию потоков неоднородного трафика, основанных на ограничении доступа, зависящего от общего уровня занятости ресурса. Среди них динамичный слайсинг, когда распределение выделенного объема ресурса осуществляется на динамической основе и зависит от его загрузки. Для ограничения доступа сессий здесь предлагается использоваться процедуру резервирования, основанную на фильтрации поступающих сессий с использованием функции внутренней блокировки. Другой сценарий — статичный слайсинг. Для данного сценария имеющийся ресурс делится между поступающими потоками в определенной пропорции, зависящей от требований сессий связи к показателям качества обслуживания.
6. Выполненное численное исследование показало, что использование динамического слайсинга позволяет на 5 – 20% уменьшить требование к объему ресурса, обеспечивающего требуемый уровень потерь сессий, по сравнению с применением для этих же целей

статичного слайсинга. Наибольший эффект применение предложенной версии динамического слайсинга приносит в ситуации обслуживания эластичного трафика данных с использованием дисциплины Processor Sharing.

Таким образом, в результате проведенных в диссертационной работе исследований построена и проанализирована процедура динамического распределения ресурса беспроводного узла доступа, позволяющая создать условия по дифференцированному обслуживанию неоднородного трафика современных коммуникационных приложений и повысить эффективность использования ресурса передачи информации. Тем самым, цель диссертационного исследования достигнута.

### **СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ АВТОРОМ ПО ТЕМЕ ДИССЕРТАЦИИ.**

#### **Публикации в изданиях, рекомендованных ВАК при Минобрнауки Российской Федерации**

1. Stepanov S.N. The construction and analysis of generalized model of resource sharing for LTE technology with functionality of NB-IoT / S. N. Stepanov, M.S. Stepanov, E.E. Malikova, A. Tsogbadrakh, **Ju. Ndayikunda** // T-Comm. — 2018. — Vol. 12. — No.12. — P. 71-77.
2. Andrabi U.M. Cellular network resource distribution methods for the joint servicing of real-time multiservice traffic and grouped IoT traffic / U.M. Andrabi, S.N. Stepanov, **J. Ndayikunda**, M.G. Kanishcheva // T-Comm. — 2020. — Vol. 14. — No.10. — P. 61-69.
3. Stepanov S.N. Resource allocation model for LTE technology with functionality of NB-IoT and reservation / S.N. Stepanov, **J. Ndayikunda**, M. G. Kanishcheva // T-Comm. — 2021. — Vol. 15. — No.11. — P. 69-76.

#### **Публикации в изданиях, индексируемых в международных наукометрических базах, в т.ч.**

##### **Scopus и WoS.**

4. Stepanov S. Resource Allocation and Sharing for Transmission of Batched NB-IoT Traffic over 3GPP LTE / S. Stepanov, M. Stepanov, A. Tsogbadrakh, **J. Ndayikunda**, U. Andrabi // Conference of Open Innovation Association, FRUCT. — 2019. — P. 422-429.
5. Stepanov S. N. Reservation Based Joint Servicing of Real Time and Batched Traffic in Inter Satellite Link / S. N. Stepanov, U. M. Andrabi, M. S. Stepanov, **J. Ndayikunda** // 2020 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russia. — 2020. — P. 1-5.

6. Muzata. A. R. The Modeling of Elastic Traffic Transmisson by the Mobile Network with NB-IoT Functionality. / A. R. Muzata, V. A. Pershina, M. S. Stepanov, F. Ndimumahoro, **J. Ndayikunda**. // 2021 Systems of Signals Generating and Processing in the Field of on Board Communications. —2021. — P. 1-7.
7. Stepanov S.N. The Analysis of Resource Sharing for Heterogenous Traffic Streams over 3GPP LTE with NB-IoT Functionality / S.N. Stepanov, M.S. Stepanov, U. Andrabi, **J. Ndayikunda**. // In: Vishnevskiy V., Samouylov K. (eds) DCCN 2020. Lecture Notes in Computer Science (LNCS). —Vol. 12563. — P. 422-435.

### Другие статьи и материалы конференций

8. Степанов С.Н. Построение модели динамического распределения радиоресурсов LTE в гетерогенных сетях с трафиком NB-IoT / С.Н. Степанов, **Ж. Ндайикунда** // Технологии информационного общества. Сборник трудов XIII Международной отраслевой научно-технической конференции «Технологии информационного общества». (20-21 марта 2019 г. Москва, МТУСИ). В 2-х томах. —М.: ИД Медиа Пабlishер. — 2019. — Том.1. — С. 136-138.
9. **Ндайикунда Ж.** Сценарий распределения ресурсов в технологии LTE / Ж. Ндайикунда // Труды международной научно-технической конференции «Телекоммуникационные и вычислительные системы –2019» — М.: Горячая линия. — Телеком. — 2019. — С. 84-87.
10. **Ндайикунда Ж.** Оценка качества обслуживания в сетях LTE с ограниченным числом пользователей / Ж. Ндайикунда // Технологии информационного общества. Сборник трудов XIV Международной отраслевой научно-технической конференции «Технологии информационного общества». (18-19 марта 2020 г. Москва, МТУСИ). — М.: ИД Медиа Пабlishер. — 2020. — С. 98-100.
11. **Ндайикунда Ж.** Ресурс базовой станции LTE с функциональностью NB-IoT / Ж. Ндайикунда // Технологии информационного общества. Сборник трудов XIV Международной отраслевой научно-технической конференции «Технологии информационного общества». (18-19 марта 2020 г. Москва, МТУСИ). — М.: ИД Медиа Пабlishер. — 2020. — С.101-103.
12. **Ндайикунда Ж.** Построение модели совместного обслуживания разнородных устройств в гетерогенных сетях LTE / Ж. Ндайикунда // Технологии информационного общества. Сборник трудов XV Международной отраслевой научно-технической конференции

- «Технологии информационного общества». (3-4 марта 2021 г. Москва, МТУСИ). — М.: МТУСИ. — 2021. — С. 67-69.
13. Stepanov S. The Modeling of Resource Sharing for Heterogenous Data Streams over 3GPP LTE with NB-IoT Functionality / S. Stepanov, M. Stepanov, U. Andrabi, **J. Ndayikunda** // The Proc. of international conference Distributed computer and communication networks: control, computation, communications . — 2020. — P. 164-172.
14. Stepanov M.S. Enhancing the Resource Sharing Capabilities of a Network by Deploying Network Slicing Procedure / M.S. Stepanov, S.N. Stepanov, Umer Andrabi, D.S. Petrov, **Juvent Ndayikunda** // The Proc. of international conference Distributed Computer and Communication Networks: Control, Computation, Communications. — 2021.— P. 243-251.