

На правах рукописи



Янковский Никита Андреевич

**МОДЕЛИ И МЕТОДЫ ДИНАМИЧЕСКОГО РАСПРЕДЕЛЕНИЯ  
РЕСУРСОВ В СЕТЯХ 5G**

2.2.15 – Системы, сети и устройства телекоммуникаций

Автореферат

диссертации на соискание ученой степени

кандидата технических наук

Санкт-Петербург – 2026

Работа выполнена на кафедре прикладной информатики федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения».

**Научный руководитель:** **Татарникова Татьяна Михайловна**  
доктор технических наук, профессор

**Официальные оппоненты:** **Канаев Андрей Константинович**  
доктор технических наук, профессор, профессор кафедры «Электрическая связь» Федерального государственного бюджетного образовательного учреждения высшего образования «Петербургский государственный университет путей сообщения Императора Александра I», г. Санкт-Петербург;  
**Викулов Антон Сергеевич**  
кандидат технических наук, руководитель направления беспроводных решений, отдел беспроводных решений ООО «Лаборатория Кьютэк», г. Москва;

**Ведущая организация:** Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого», 195251, г. Санкт-Петербург, вн. тер. г. муниципальный округ Академическое, ул. Политехническая, д. 29 литера Б

Защита состоится «18» июня 2026 г., в 14:00 часов на заседании диссертационного совета 24.2.384.01 при федеральном государственном автономном образовательном учреждении высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения» по адресу: 190000, Санкт-Петербург, ул. Большая Морская, д. 67, лит. А. С диссертацией можно ознакомиться в библиотеке федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения» по адресу: 190000, Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.

Автореферат и текст диссертации размещены на сайте университета [https://dissov.guap.ru/defense/yankovsky\\_na](https://dissov.guap.ru/defense/yankovsky_na).

Автореферат разослан «27» апреля 2026 г.

Ученый секретарь  
Диссертационного совета 24.2.384.01  
кандидат технических наук, доцент



А.М. Сергеев

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы исследования.** В Стратегии развития отрасли связи Российской Федерации на период до 2035 г., подготовленный Министерством цифрового развития России и утвержденной распоряжением правительства РФ от 23.11.2023 г. представлены 19 стратегических целей в области связи, среди которых обеспечение малочисленных населенных пунктов и труднодоступных районов услугами связи 5G на отечественном оборудовании к 2030 году и создание гибридной сети связи на основе отечественной низкоорбитальной спутниковой группировки, сегментов мобильной и фиксированной связи к 2035 году. Эта сеть, по задумке Минцифры, покрывает всю территорию страны, включая федеральные и региональные автомобильные дороги, Арктическую зону и трассу Северного морского пути, а также обеспечит технологические возможности управления беспилотными авиационными системами и глобальное покрытие с возможностью экспорта услуг связи в зарубежные страны.

Сценарии использования сотовых сетей связи пятого поколения или сетей 5G: eMBB (enhanced Mobile Broadband) – сверхширокополосная мобильная связь, URLLC (Ultra-Reliable Low Latency Communication) – сверхнадежная связь с низкими задержками и mMTC (Massive Machine-Type Communications) – массовая межмашинная связь направлены на поддержку большого количества одновременно подключенных устройств, автономность, низкую задержку, эффективное распределение ресурсов, механизмов резервирования и приоритетное планирования данных. Для реализации гибридной сети связи предстоит преодолеть ряд противоречий, связанных прежде всего с гетерогенностью трафика и, соответственно, разных требований к качеству его обслуживания (Quality of Service, QoS).

QoS в сетях 5G представляет собой набор механизмов для управления и приоритизации разного типа трафика таким образом, чтобы обеспечить соответствующий уровень производительности и надежности. Каждый сценарий имеет связанный профиль QoS, который указывает требования к задержке, надежности и пропускной способности. Например, к URLLC трафику, характерному для приложений промышленной автоматизации, здравоохранения, автономных автомобилей предъявляются строгие требования QoS: надежность передачи – не менее 99,999% с задержкой не более 1 мс. В то же время для mMTC трафика, генерируемого такими устройствами как счетчики воды, газа, электричества, контроллеры уличного освещения, датчики парковочного места, различные сенсоры дыма/огня, «умные» мусорные баки и прочих устройств интернета вещей не важны высокая скорость и сверхнизкие задержки, но очень важны автономность и огромное число подключений в сети. Для eMBB трафика, например потокового видео или игр виртуальной/дополненной реальности важна массовость подключений при малом объеме данных.

Для разрешения сложившихся противоречий необходимы новые модели и методы, обеспечивающие обслуживание гетерогенного трафика в сетях 5G согласно требованиям QoS.

**Степень разработанности темы диссертации.** Научные работы в области моделирования гетерогенных телекоммуникационных сетей, в том числе сотовых сетей связи, посвящены особенностям применения теории телетрафика, случайных процессов, теории очередей, имитационного моделирования при оценивании показателей QoS, распределении ресурсов сети, организации множественного доступа к разделяемому ресурсу.

Определяющий вклад в развитие математических моделей гетерогенных телекоммуникационных сетей внесли российские и зарубежные ученые, такие как Г.П. Башарин, В.М. Вишневский, А.Н. Дудин, В.Г. Карташевский, О.И. Кутузов, Б.С. Лившиц, Е.В. Морозов, А.Н. Назаров, Ю.Н. Орлов, А.П. Пшеничников, С.Н. Степанов, Н.А. Соколов, Т.М. Татарникова, Я.В. Фидлин, О.И. Шелухин, М.А. Шнепс, А.Д. Харкевич, Г.Г. Яновский, V.-V. Iversen, M. Gerla, F. Kelly, P.J. Kuhn, L. Юетгоск, K.W. Ross, W. Whitt и др. Предложенные в диссертации модели и методы распределения ресурсов сетей 5G в условиях гетерогенного трафика и высокой плотности устройств основаны на результатах, полученных этими учеными.

Исследованию методов доступа в телекоммуникационных сетях посвящены работы А.С. Аджемова, Б.С. Гольдштейна, Е.А. Крука, А.Е. Кучерявого, К.Е. Самуйлова, В.О. Тихвинского, А.М. Тюрликова, J. Andrews, N. Himayat, F. Vaccelli, S. Rangan, T. Rappoport, S. Singh, и др. Предложенный в диссертации метод распределения ресурсов в восходящем канале уточняет результаты вышеперечисленных ученых применительно к специфике обслуживания данных малого объема, динамического изменения числа используемых преамбул и приоритизации гетерогенного трафика.

Работы таких ученых как А.В. Росляков, А.И. Парамонов, Р.В. Киричек, M. Dohler, H. Yanikomeroglu, J. Nosek и др. посвящены одной из актуальных проблем в последние годы – анализу эффективности функционирования интернета вещей, являющегося источником mMTC трафика. Предложенные в работе модели сосуществования URLLC и eMBB трафика в нисходящем канале развивают идеи этих ученых при оценивании качества обслуживания устройств сетей 5G.

**Объект и предмет исследования.** Объектом исследования являются сети 5G.

Предметом исследования – применение моделей и методов распределения ресурсов в сетях 5G.

**Цель и задачи исследования.** Цель работы состоит в повышении эффективности распределения ресурсов в сотовых сетях 5G для удовлетворения показателям QoS.

Цель работы достигается последовательным решением следующих задач:

1. Разработка математических моделей восходящего канала и нисходящего канала сетей 5G.
2. Разработка метода доступа устройств M2M к разделяемому ресурсу в восходящем канале сетей 5G.
3. Разработка метода мультиплексирования битовых потоков URLLC и eMBB трафика в нисходящем канале сетей 5G
4. Разработка модели динамической ассоциации устройств с базовой станцией в сетях 5G.

**Научная новизна диссертационной работы состоит в следующем:**

1. Математические модели восходящего канала и нисходящего канала отличаются от известных учетом сценариев использования сетей 5G, что позволяет организовать эффективное совместное обслуживание гетерогенного трафика, удовлетворяющее QoS.

2. Метод доступа устройств M2M к ресурсам сетей 5G, отличается от известных учетом специфики обслуживания данных малого объема, динамическим изменением числа используемых преамбул, а также системой приоритизации трафика, что позволяет снизить задержку и улучшить энергоэффективность за счет сокращения повторных передач.

3. Метод мультиплексирования битовых потоков URLLC и eMBB трафика в нисходящем канале сетей 5G отличается от известных возможностью адаптивного изменения схемы распределения ресурсов между различными видами трафика, что позволяет для некоторых вариантов снизить среднюю задержку.

4. Математическая модель динамической ассоциации устройств с базовой станцией, отличается от известных учетом особенностей трафика устройств M2M, расположенных на границе соты, позволяющие снизить среднюю задержку передачи сообщений по нисходящему каналу связи.

**Теоретическая значимость** результатов диссертации заключается в предложении нового комплекса взаимосвязанных моделей и методов управления трафиком, учитывающих особенности современных сотовых сетей связи. Предложенные в работе модели и методы могут быть использованы при теоретическом анализе сотовых сетей связи пятого поколения.

**Практическая значимость** диссертационной работы заключается во внедрении моделей и методов эффективного управления гетерогенным трафиком сетей 5G.

**Положения, выносимые на защиту.**

1. Метод динамического изменения числа преамбул на основе комбинирования методов машинного обучения: регрессии опорных векторов SVR и рекуррентных нейронных сетей RNN позволяет снизить среднюю задержку передачи сообщений по каналу со случайным доступом.

2. Модель нисходящего канала и метод адаптивного мультиплексирования потоков eMBB и URLLC на основе комбинированного использования LDPC-кодов и решетчатой модуляции (TCM), позволяют обеспечить совместную передачу гетерогенного трафика в общем частотно-временном ресурсе с выполнением строгих требований URLLC по задержке (<1 мс) и надежности (>99.999%) при минимальной деградации пропускной способности eMBB-трафика.

3. Модель и метод динамической ассоциации пользовательских устройств с базовыми станциями в гетерогенной сети 5G, основанные на прогнозировании нагрузки с помощью рекуррентных нейронных сетей (LSTM) позволяют снизить среднюю задержку передачи данных внутри соты по сравнению со статической ассоциацией.

**Методы исследования, которые использовались в диссертационной работе:** теория информации, теория связи, теория случайных процессов, теория вероятностей, теория Марковских цепей, методы математической оптимизации и методы имитационного моделирования.

**Степень достоверности и апробация результатов.** Результаты диссертационной работы получены на основе общепринятых математических методов, согласуются с ранее известными результатами и подтверждены результатами имитационного моделирования, публикациями в высокорейтинговых журналах и выступлениями на российских и международных конференциях.

Основные положения диссертационной работы были представлены и обсуждались на следующих конгрессах, конференциях и семинарах: международная научная конференция «Обработка, передача и защита информации в компьютерных системах», 11–15 апреля 2022; международная конференция «Волновая электроника и инфокоммуникационные системы», 2020, 2021, 2023 гг.; международная научно-практической конференция «Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем», Москва, 2023, 2024 гг.; Санкт-Петербургская международная конференция «Региональная информатика и информационная безопасность», Санкт-Петербург, 25–27 октября 2023 год.

**Соответствие работы паспорту научной специальности.** Диссертационная работа соответствует паспорту научной специальности 2.2.15 «Системы, сети и устройства телекоммуникаций» (пунктам 2, 6, 8 и 18).

**Личный вклад автора.** Все результаты, представленные в тексте диссертационной работы, получены автором лично.

**Публикации.** Материалы, отражающие основные результаты диссертации, опубликованы в 17 научных работах: 6 работ – в журналах из перечня ВАК, в том числе 1 из них без соавторов; 5 – в изданиях, индексируемых Scopus, и 5 – в сборниках конференций, индексируемых РИНЦ. По теме диссертации получено 1 свидетельство о государственной регистрации программы для ЭВМ.

## ОБЪЕМ И СТРУКТУРА РАБОТЫ

**Во введении** приводится обоснование актуальности выбранной темы диссертационного исследования и степень её разработанности, сформулированы цель, задачи и положения, выносимые на защиту, описывается новизна, теоретическая и практическая значимость работы.

**Первая глава** диссертации посвящена анализу архитектурных особенностей и проблем распределения ресурсов в сотовых сетях связи пятого поколения.

Выполнен анализ противоречивых требований к обслуживанию гетерогенного трафика при совместной работе сценариев использования сетей 5G. Показано, что для обеспечения QoS популярность получила технология динамического распределения ресурсов – динамическое распределение спектра и адаптивная модуляция и кодирование для оптимизации использования ресурсов на основе требований разных сценариев в реальном времени.

Особое внимание уделено передаче трафика по технологии M2M – "устройство-устройство", которая отличается от классической технологии H2H – "человек-человек", прежде всего высокой плотностью устройств, (на порядок выше, чем в H2H), преобладанием трафика в восходящем канале, разнородными требованиями к QoS и низким энергопотреблением. Эти обстоятельства обусловили рассмотрение задачи сосуществования трафика URLLC и mMTC в восходящем канале связи.

Сценарий URLLC использования сетей 5G характеризуется жесткими требованиями к задержкам и надежности передачи данных, что требует немедленного выделения ресурсов. В то же время сценарий mMTC характеризуется массовостью подключений, но при этом допускает более высокие задержки. Сложность решения задачи обусловлена использованием механизма случайного доступа устройствами mMTC, что приводит к коллизиям и увеличению интерференции. Это в свою очередь создает дефицит ресурсов для сценария URLLC, что усугубляется при высокой плотности устройств. Для решения задачи сосуществования трафика URLLC и mMTC в восходящем канале связи предложено применять методы динамического резервирования ресурсов с выделением мини-слотов для URLLC.

Также рассмотрена задача обеспечения сосуществования eMBB и URLLC трафика в нисходящем канале связи. Сценарий eMBB ориентирован на предоставление услуг с высокой пропускной способностью, но при этом менее чувствителен к задержкам по сравнению со сценарием URLLC. Однако приоритетное выделение ресурсов для URLLC трафика может негативно сказаться на производительности eMBB трафика, особенно в условиях ограниченной полосы пропускания. Для минимизации такого влияния предложено использовать адаптивные алгоритмы планирования ресурсов, основанные на применении мультиплексирования битовых потоков данных. Эти методы способствуют динамическому перераспределению ресурсов между eMBB трафиком и URLLC трафиком, учитывая текущую нагрузку и

требования QoS. При этом применяется гибкое разделение полосы частот, при котором часть ресурсов, выделенных для eMBB трафика, временно заимствуется для обслуживания URLLC трафика.

Показано, что помимо задач, связанных с обеспечением обслуживания гетерогенного трафика, в сетях 5G с высокой плотностью устройств, возникает необходимость динамического перераспределения нагрузки между базовыми станциями. Неравномерное распределение пользователей и устройств может приводить к перегрузке одних базовых станций при недогрузке других, что влечет за собой ухудшение QoS из-за повышенной интерференции и увеличения задержек. Для решения этой проблемы предлагается использовать механизмы координации между базовыми станциями, такие как Cloud-RAN, где управление ресурсами осуществляется централизованно через облачную платформу. Это позволяет оптимизировать распределение нагрузки за счет динамического переключения пользователей между станциями (Handover) на основе прогнозируемой и текущей нагрузки.

**Вторая глава** диссертации посвящена решению задачи сосуществования трафика URLLC и mMTC в восходящем канале связи.

Проблема случайного доступа рассматривается как основное узкое место в модели восходящего канала. Аргументируется, что высокая плотность устройств, генерирующих короткие частые пакеты данных, создает значительную нагрузку на сеть, приводя к коллизиям, задержкам и потерям данных.

Модель восходящего канала, рассматриваемая в работе, представляется системой массового обслуживания (СМО) типа  $G|G|1$  с приоритетным обслуживанием (рисунок 1).

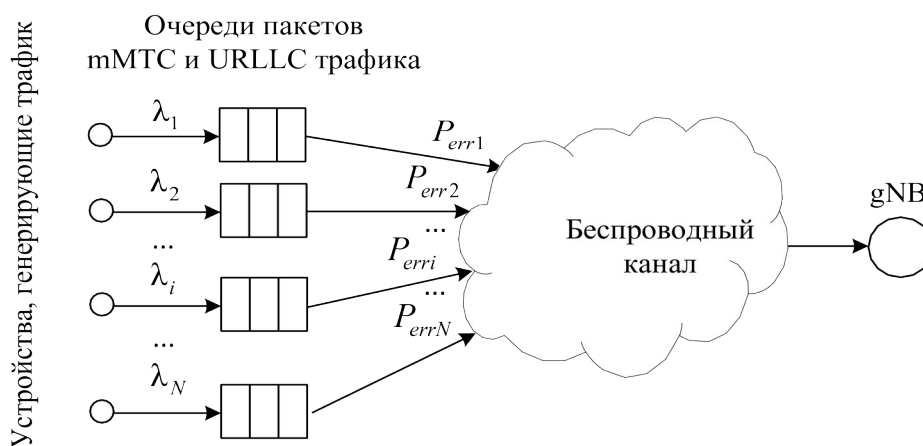


Рисунок 1 – Модель восходящего канала связи

Математическое моделирование восходящего канала сети 5G выполнено имитационным методом. Модель представляет собой замкнутую систему, которая адекватно отражает основные аспекты работы восходящего канала сети 5G: конкуренцию за ресурсы, гетерогенность трафика, неидеальность канала и интеллектуальное управление.

Модель – дискретно-событийная, построена по схеме процессов. Все время взаимодействия  $N$  устройств с базовой станцией (generalized Node B, gNB) сети 5G разбито на фреймы одинаковой длины, поэтому системное время в модели задано последовательностью фреймов.

В модели одновременно развиваются несколько процессов:

*Пользовательское устройство* соты, которое задается интенсивностью  $\lambda_i$ ,  $i = \overline{1, N}$  высокоприоритетного URLLC трафика, либо низкоприоритетного mMTC трафика. Распределение интервалов времени между поступлениями пакетов задано произвольным, изменение значения  $\lambda_i$  выполняется согласно модели burst-трафика.

*Очередь*, в которой находятся необслуженные пакеты каждого пользовательского устройства. Каждое из устройств имеет свою неограниченную очередь сообщений.

*Базовая станция*, которая управляет ресурсами, динамически перераспределяя преамбулы на основе наблюдаемой статистики. В каждом новом цикле имитации работы восходящего канала выполняются следующие процедуры:

- оценивается текущее число активных устройств каждого типа, используя модель машинного обучения SVR;

- прогнозируется интенсивность поступления новых сообщений на следующий фрейм с помощью рекуррентной нейронной сети LSTM;

- на основе прогнозных оценок вычисляется новое оптимальное распределение преамбул  $m_0$  – для URLLC трафика и  $m_1$  – для mMTC трафика на следующий фрейм с обеспечением минимальных гарантий для каждого класса трафика;

- преамбулы  $m_0$  и  $m_1$  отправляются широковещательной рассылкой.

*Беспроводной канал*, который задается вероятностью ошибки  $P_{err}$  для каждого устройства. Значение  $P_{err}$  определяется моделированием полной группы несовместных случайных событий на интервале  $[0.1, 0.01]$ .

Показано, что традиционная четырехэтапная процедура доступа в сотовых сетях 4G, где устройства последовательно передают преамбулу, запрос ресурсов, данные и подтверждение (рисунок 2,а), становится неэффективной при массовой активации датчиков.

Коллизии преамбул в перегруженных слотах приводят к каскадным повторным передачам по схеме  $K$ -повторений, что приводит к увеличению задержки до 10-100 мс и сокращению срока службы батарей. В качестве альтернативы исследуется двухэтапной процедуры, объединяющей преамбулу и данные в одном сообщении (рисунок 2,б).

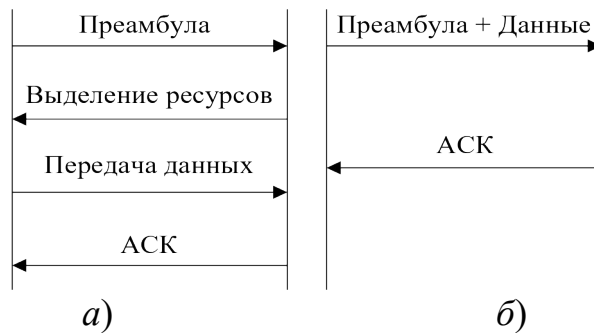


Рисунок 2 – Процедуры доступа к разделяемому ресурсу:  
 а – четырехэтапная; б – двухэтапная

Показано, что для уменьшения задержки передачи пакета необходимо рассматривать две задачи – уменьшение вероятности конфликтов при выборе преамбулы и уменьшения вероятности ошибки при передаче данных.

Время передачи складывается из времени, затрачиваемого на передачу преамбулы  $T_{pr}$ , времени передачи пакетов  $T_{pac}$  и времени обработки полученного пакета или обратной связи  $T_{dp}$ . Если пакет успешно декодирован, то gNB отправляет положительную квитанцию АСК, в противном случае – отрицательную квитанцию NACK. После получения и обработки обратной связи устройство может решить, следует ли выполнять повторную передачу.

Предложено решать проблему возникновения коллизий при помощи динамического управления ресурсами с применением ML-прогнозирования. Предлагаемый метод динамического изменения числа преамбул приведен на рисунке 3 в виде UML-диаграммы. Суть метода заключается, во-первых, в сборе статистики по состоянию слотов:  $S$  – количество слотов, в которых произошла успешная передача данных,  $E$  – количество пустых слотов и  $C$  – количество слотов, в которых произошла коллизия, и во-вторых, прогнозированию числа активных устройств  $u$  методом опорных векторов (SVR-регрессии) для каждого из видов трафика –  $u_0$  для URLLC и  $u_1$  для mMTC.

Для предсказания количества новых запросов  $w$  применяются рекуррентные нейронные сети (RNN-сети), обрабатывающие временные ряды трафика. На основе предсказанных значений  $w_0$  для трафика URLLC и  $w_1$  для трафика mMTC, система динамически распределяет число преамбул  $m_0$  и  $m_1$  для трафика URLLC и mMTC соответственно, с гарантией минимальной квоты для последнего, то есть  $m_{1,t+1}=(L-m_{0,t+1})$ , где  $L$  – общее число преамбул.

Результаты прогнозирования числа активных устройств характеризуются коэффициентом детерминации  $R^2 = 0.998$ , средняя ошибка точности прогноза новых сообщений WAPE (Weighted Average Percent Error) составила в среднем 18%.

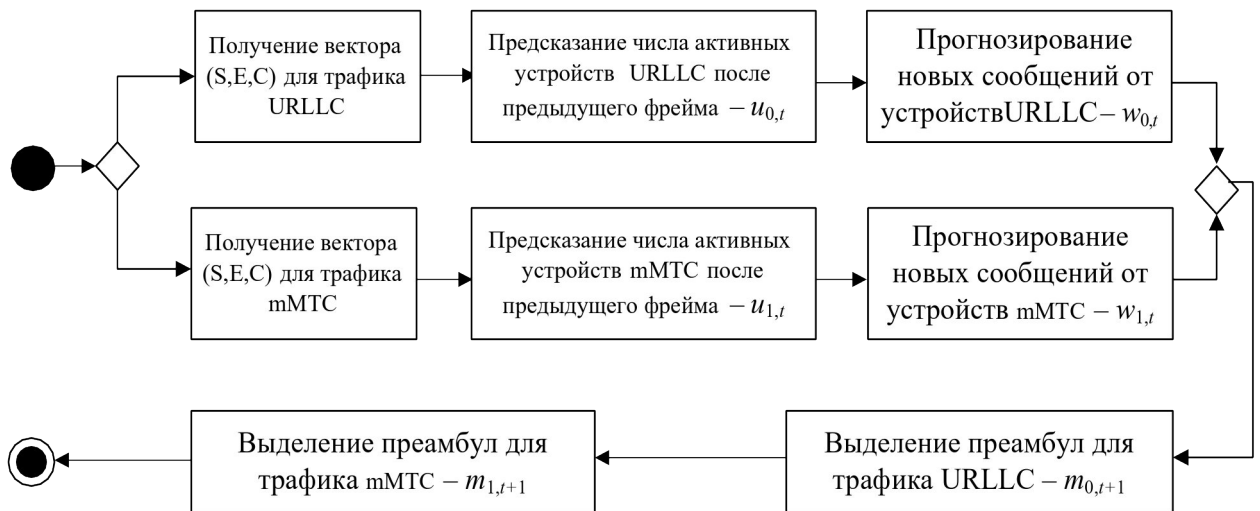


Рисунок 3 – UML-диаграмма метода динамического изменения числа преамбул

Также показано, что использование линейного сетевого кодирования может обеспечить значительное повышение надежности по сравнению со схемой  $K$ -повторений при той же спектральной эффективности, но время передачи при этом увеличивается, что не соответствует требованиям QoS для трафика URLLC. Предложено использовать алгоритм скользящего сетевого кодирования, позволяющего декодировать пакеты «на лету». Обобщенная схема скользящего сетевого кодирования с параметрами  $(K, D, M)$ , где  $K$  – число пакетов в одной передаче,  $D$  – задержка кодирования/декодирования  $i$ -го блока сообщения,  $M$  – число блоков сообщения приведена на рисунке 4.

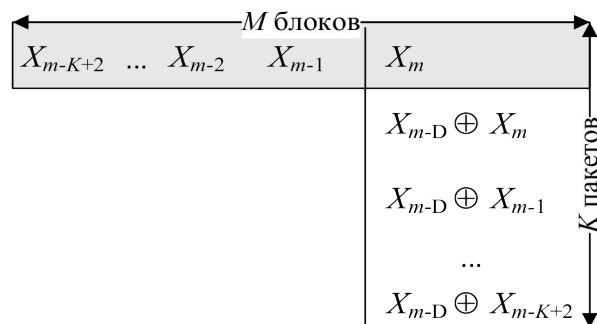


Рисунок 4 – Обобщенная схема скользящего сетевого кодирования

Результаты моделирования (рисунок 5) демонстрируют эффективность использования алгоритма скользящего сетевого кодирования в сравнении со схемой  $K$ -повторений по средней задержке  $E[S]$  и по вероятности  $P$  неудачной передачи после  $K$  повторений.

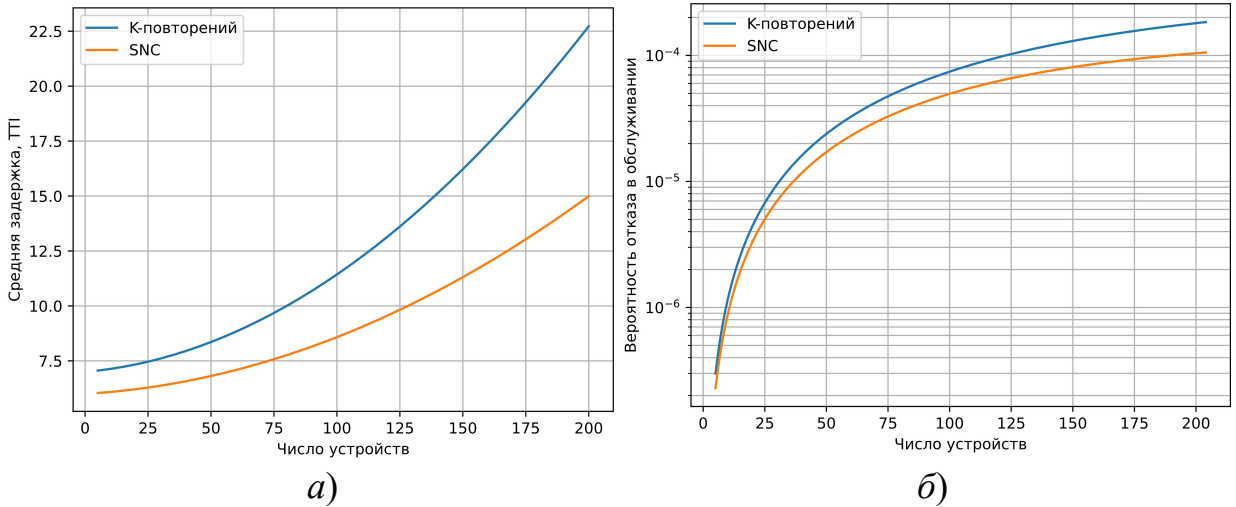


Рисунок 5 – Сравнение алгоритма скользящего сетевого кодирования со схемой  $K$ -повторений:  $a$  – зависимость  $E[S]$  от числа устройств;  $b$  – зависимость  $P$  после  $K$  повторений

Также показано, что совместное применение скользящего сетевого кодирования и динамического распределения преамбул ( $A^*$ ) позволяет добиться снижения средней задержки передачи сообщения как для URLLC трафика, так и для mMTC трафика по сравнению с известными методами ( $M^*$ ). На рисунке 6 приведены результаты имитационного моделирования передачи M2M трафика по восходящему каналу, из которого видно, что при 100-200 устройствах в сети задержка высокоприоритетного трафика снижается на 15-30%.

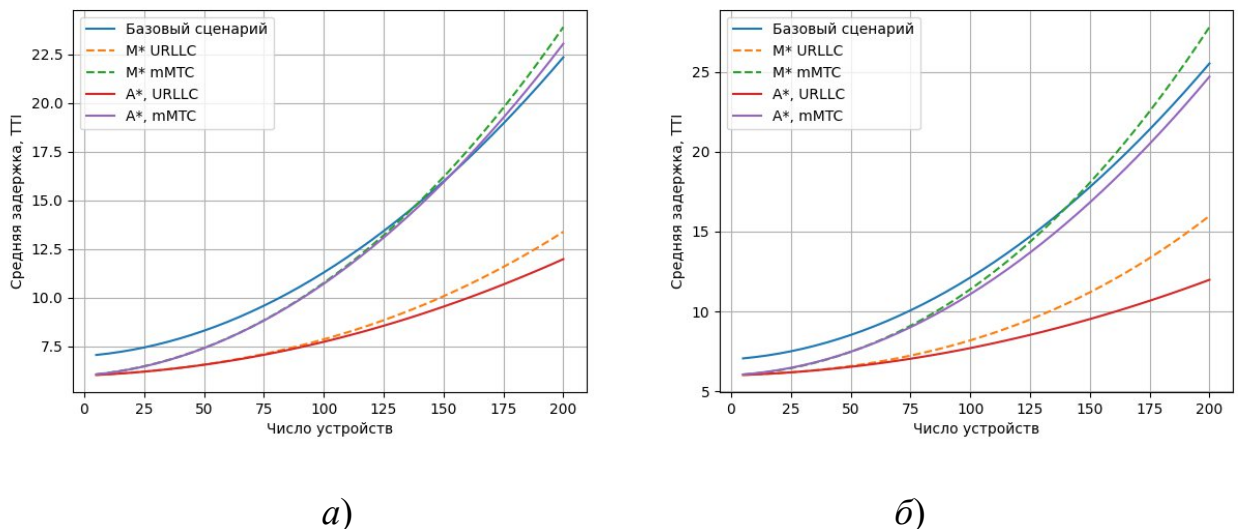


Рисунок 6 – График зависимости средней задержки для каждого вида трафика в восходящем канале от числа устройств в системах:  $a$  –  $M|G|1$ ;  $b$  –  $LN|G|1$

**Третья глава** диссертации посвящена решению задачи совместной передачи гетерогенного трафика eMBB и URLLC в нисходящем канале связи.

Проблема мультиплексирования рассматривается как основное узкое место в модели нисходящего канала. Аргументируется, что традиционные методы мультиплексирования, такие как статическое разделение ресурсов или примитивное наложение потоков приводят к значительному снижению показателей QoS для одного или обоих типов трафика.

Модель нисходящего канала, рассматриваемая в работе, представляется системой массового обслуживания типа  $G|G|1$  с приоритетным обслуживанием (рисунок 7).

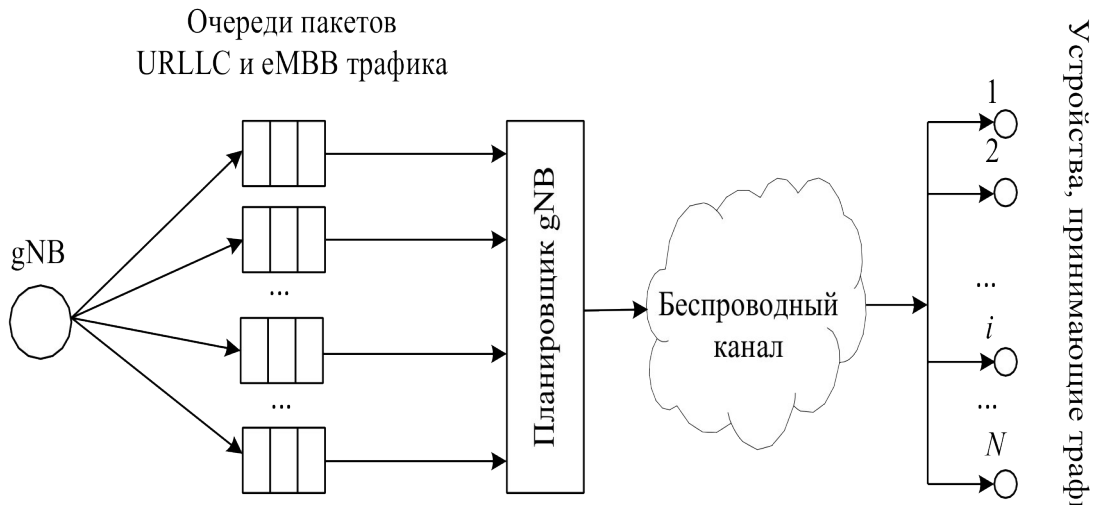


Рисунок 7 – Модель нисходящего канала связи

Математическое моделирование нисходящего канала сети 5G выполнено имитационным методом. Модель – дискретно-событийная, построена по схеме событий, где каждое новое событие начиналось и заканчивалось субфреймом длительностью 1 мс, что соответствует нумерологии 3 в 5G NR глобального стандарта для эфирного интерфейса сетей 5G. Центральное место в модели отведено реализации алгоритмов мультиплексирования, основанного на комбинации LDPC-кодирования, решетчатого кодирования и адаптивной модуляции.

Имитационная модель нисходящего канала состоит из следующих объектов:

*Базовая станция gNB*, которая обрабатывает два типа сообщений от множества пользовательских устройств. В каждом новом цикле имитации работы исходящего канала выполняются следующие процедуры:

- для каждого  $i$ -го устройства генерируются два независимых потока заявок: eMBB трафик с интенсивностью  $\lambda_{i,e}$  и URLLC трафик с интенсивностью  $\lambda_{i,u}$ . Длина  $j$ -го eMBB сообщения  $\bar{V}_{ij}$  задается Бета-распределением вероятностей с математическим ожиданием 1024 байта из интервалом [512, 1536], что соответствует реальной природе объемного

трафика, такого как видео или файлы. Размеры всех URLLC сообщений предполагались фиксированными и равными 32 байтами;

- каждое eMBB сообщение и URLLC сообщение снабжается контрольной суммой (CRC-8);

- каждое eMBB сообщение и URLLC сообщение независимо кодируется LDPC-кодами с разными скоростями: высокая скорость (8/9) для eMBB для максимизации пропускной способности и низкая (1/2) для URLLC для обеспечения сверхнадежности;

- в начале каждого субфрейма планировщик gNB, основываясь на текущем состоянии системы: заполненности очередей, приоритетах и состоянии канала, принимает решение о распределении ограниченного частотно-временного ресурса (ресурсных блоков) между активными устройствами. Планировщик выделяет ресурсы отдельно для eMBB сообщений и URLLC сообщений и определяет оптимальную схему мультиплексирования.

- закодированные битовые последовательности двух типов трафика поступают в блок адаптивного мультиплексирования. На основе выбранной схемы (1, 2 или 3 бита URLLC на OFDM-символ) производится их объединение с помощью перемежения;

- результирующий составной битовый поток подвергается решетчатому кодированию с выбором размерности модуляции, необходимой для передачи символов  $(l + r)$ , где  $l$  – длина сообщения,  $r$  – избыточность применяемого сверточного кода;

- сформированные OFDM-символы размещаются в ресурсных блоках согласно решению планировщика и отправляются на пользовательские устройства.

*Очередь*, которая выделена на gNB для каждого пользовательского устройства. Очередь имеет неограниченный объем для каждого класса трафика. Это допущение моделирует сценарий с высокой нагрузкой, где потери происходят не из-за переполнения буфера, а из-за нарушения задержек.

*Пользовательское устройство* соты, которое выполняет функцию приемника.

В каждом новом цикле имитации работы восходящего канала выполняются следующие процедуры:

- на вход декодера, который представляет собой мягкий выход алгоритма Витерби поступает объединенный битовый поток данных;

- сообщение квантуется в соответствии со следующим правилом: сохраняются наиболее значимые биты  $\log_2(Q)$  LLR, где  $Q$  – количество уровней квантования, LLR – Log Likelihood Ratio – статистическая мера, основанная на функции логарифмического правдоподобия.

- кодовые слова каждого класса трафика передаются деинтерлейверу для разъединения бит сообщений.

- разьединенные сообщения проходят этап LDPC-декодера;
- проверяется контрольная сумма и принимается решение о корректности полученного сообщения.

Результаты моделирования показали, что предложенный метод позволяет достичь значительного снижения средней задержки для URLLC трафика (рисунок 8,а) при минимальном воздействии на eMBB трафик (рисунок 8,б).

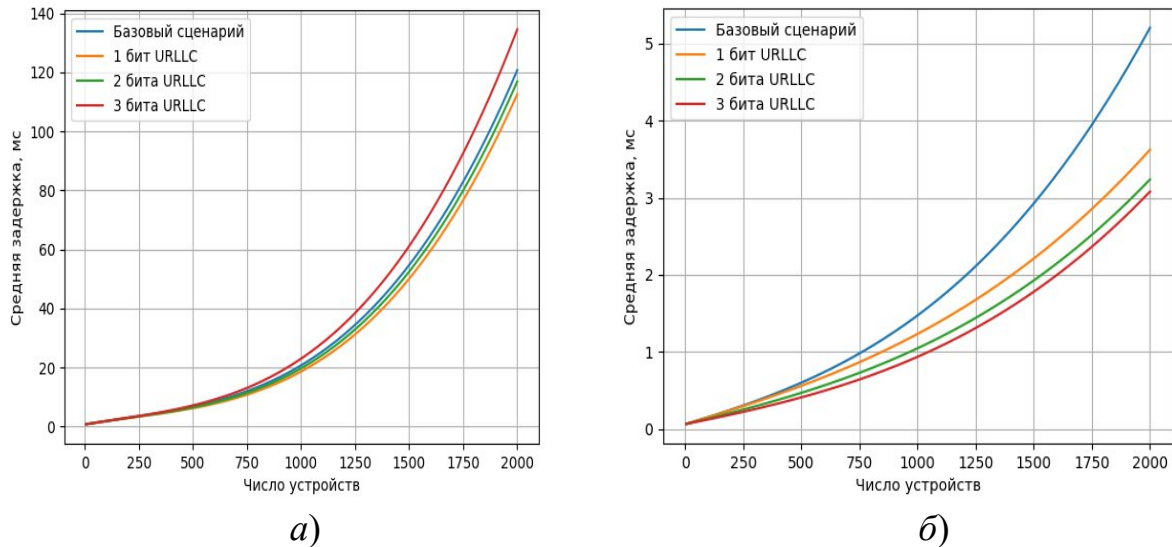


Рисунок 8 – График зависимости средней задержки от количества устройств в нисходящем канале: *а* – для eMBB трафика; *б* – для URLLC трафика

**Четвертая глава** диссертационной работы посвящена решению задачи оптимизации ассоциации пользовательских устройств с базовыми станциями gNB в условиях высокой плотности подключений и гетерогенности трафика.

Показано, что традиционные методы, основанные на максимизации уровня принимаемого сигнала, часто приводят к неравномерному распределению нагрузки между gNB, что вызывает перегрузку одних и недогрузку других базовых станций, ухудшая качество обслуживания QoS и увеличивая задержки.

В качестве альтернативы предложен метод динамической ассоциации, основанный на прогнозировании трафика с использованием рекуррентных нейронных сетей с долгой краткосрочной памятью. Прогнозные оценки объема трафика строятся для каждого нового раунда подключения устройств к gNB. Такой подход позволяет учитывать не только текущее состояние сети, но и пространственно-временные закономерности поступления трафика.

Теоретической основой метода является асимптотическая нижняя граница средней задержки, вывод которой с доказательством теоремы приведен в работе.

Пусть критерий эффективности работы  $m$ -й соты,  $m = \overline{1, M}$  :

$$B_m = \sum_{i=1}^M \left[ \lim_{T \rightarrow \infty} \frac{1}{H_i^T} \sum_{j=1}^{H_i^T} b_{ij} \right], \quad (1)$$

где  $H_i^T$  – число сегментов, отправленных пользователем  $i$  за время  $T$ ;  
 $b_{ij}$  – длительность отправки сегмента  $j$  пользователю  $i$ .

$$B_m = \begin{cases} 0, & \text{если } \sum_{i \in U_m} \frac{K_i^m}{\tau_i} \leq 1, \\ \sum_{i \in U_m} \max \left( \sqrt{K_i^m \mu_m} - \tau_i, 0 \right) & \text{иначе,} \end{cases} \quad (2)$$

где  $K_i^m = \frac{\bar{V}_i}{C_i}$ ,  $\bar{V}_i$  – средний размер передаваемого сообщения,  $C_i$  –  
максимально достижимая скорость для  $i$ -го устройства;

$\mu_{0m}$  – решение уравнения  $\sum_i \frac{K_i^m}{\max \left( \sqrt{K_i^m \mu_{0m}} - \tau_i \right)} = 1$ ;

$\tau_i$  – интервал времени, за который отправлено  $H$  пакетов;

Для сценария перегрузки:

$$b_i^m = \lim_{T \rightarrow \infty} \frac{1}{H_i^T} \sum_{j=1}^{H_i^T} b_{ij} > 0, \quad (3)$$

тогда  $\mu_{0m} = \left( \sum_i \sqrt{K_i^m} \right)^2$ .

Следовательно,

$$B_m = \sum_i \sqrt{K_i^m \mu_{0m}} - \sum_i \tau_i = \sqrt{\mu_{0m}} \sum_i \sqrt{K_i^m} - \sum_i \tau_i = \left( \sum_i \sqrt{K_i^m} \right)^2 - \sum_i \tau_i. \quad (4)$$

И средняя задержка по всем сотам может быть определена

$$B = \frac{1}{N} \sum_{m=1}^M B_m = \frac{1}{N} \sum_{m=1}^M \left( \sum_i \sqrt{K_i^m} \right)^2 - \frac{1}{N} \sum_{i=1}^N \tau_i. \quad (5)$$

Алгоритм реализации метода динамической ассоциации включает следующие основные шаги:

1. Сбор данных: пассивный мониторинг трафика на gNB в  $i$ -м раунде подключения устройств к gNB, включая количество ресурсных блоков, размер транспортных блоков и схемы модуляции.

2. Преобразование данных: формирование вектора признаков для количества устройств в  $(i+1)$  раунде подключения устройств к gNB.

3. Прогнозирование трафика: использование LSTM-сети для предсказания нагрузки на gNB в  $(i+1)$  раунде на основе вектора признаков  $i$ -го раунда подключения устройств к gNB.

4. Перераспределение пользователей: динамическое изменение ассоциации устройств с gNB на основе прогнозных оценок объема трафика.

Для оценки эффективности метода проведено имитационное моделирование.

Сравнение метода динамической ассоциации проводилось с традиционными методами ассоциации на основе RSSI при различных алгоритмах планирования ресурсов Round Robin и Maximum Throughput:

- жадным алгоритм, при котором оптимизация ассоциации пользовательских устройств с gNB выполняется на основе локально оптимального выбора без гарантии оптимального итогового решения;

- методом Гаусса-Зейделя, при котором многомерной оптимизации сводится к многократному использованию метода одномерной оптимизации

Сравнительные результаты приведены на рисунке 9.

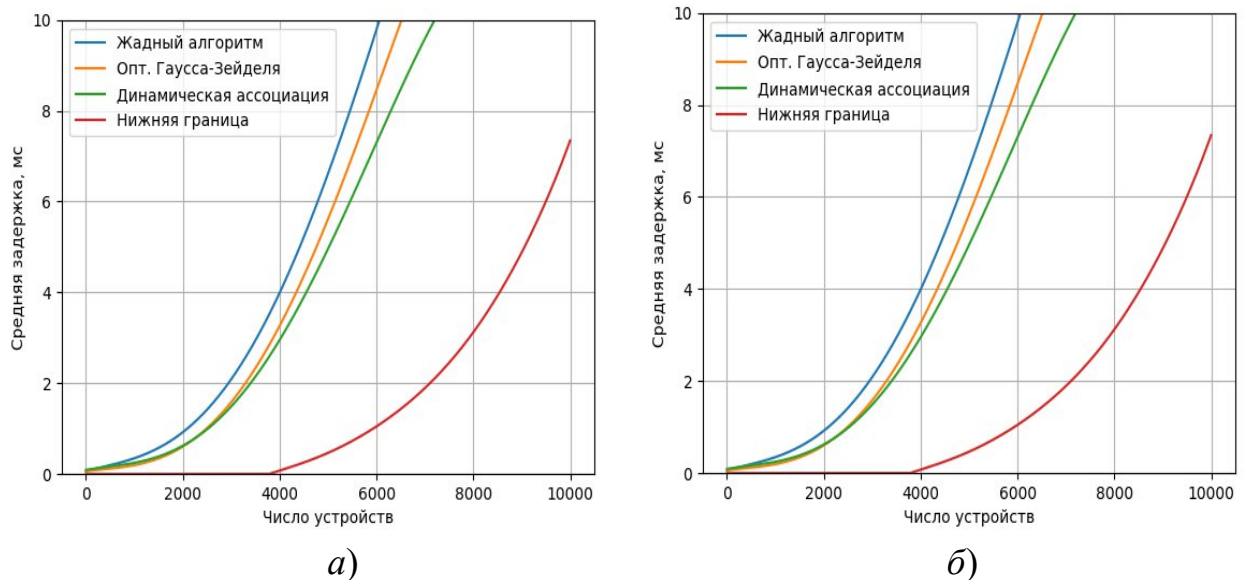


Рисунок 9 – Зависимость средней задержки в нисходящем канале для алгоритма планирования: *а* – Round Robin; *б* – Maximum Throughput

Анализ результатов сравнения показал, что динамическая ассоциация позволяет значительно снизить среднюю задержку передачи данных по сравнению со статическими схемами распределения, особенно в условиях высокой нагрузки.

## ЗАКЛЮЧЕНИЕ

1. Разработана модель восходящего канала как система массового обслуживания типа  $G|G|1$  с приоритетным обслуживанием URLLC трафика, для которой:

- предложен метод скользящего сетевого кодирования, показавший существенное преимущество по сравнению с традиционными схемами  $K$ -

повторений. Метод позволяет значительно снизить вероятность ошибок и среднюю задержку передачи сообщений за счет совмещения преимуществ сетевого кодирования и механизма повторных передач. Важной особенностью метода является возможность декодирования данных «на лету», что критически важно для сценария URLLC. Результаты имитационного моделирования подтвердили способность метода обеспечивать надежность на уровне 99.999% при одновременном снижении задержки;

– алгоритм динамического изменения числа преамбул на основе методов машинного обучения: регрессии опорных векторов SVR и рекуррентных нейронных сетей RNN. Этот подход позволяет прогнозировать количество активных устройств и интенсивность поступающего трафика, для динамического перераспределения ресурсов случайного доступа (преамбул) между высокоприоритетным URLLC и низкоприоритетным mMTC трафиком. Это способствует минимизации коллизий, снижению задержки и эффективному использованию доступной полосы пропускания.

2. Разработана модель нисходящего канала как система массового обслуживания типа  $G|G|1$  с приоритетным обслуживанием URLLC трафика, для которой:

– предложен метод адаптивного мультиплексирования гетерогенного трафика eMBB и URLLC. В основе метода лежит использование решетчатого кодирования в сочетании с LDPC-кодами, что позволяет интегрировать трафик с различными требованиями в единый ресурсный блок без значительной деградации пропускной способности eMBB. Результаты имитационного моделирования показали, что метод позволяет обеспечить сверхнизкую задержку передачи URLLC-сообщений менее 1 мс при высокой вероятности успешной доставки более 99.999% согласно требованиям QoS.

3. Разработан метод динамической ассоциации пользователей с базовыми станциями, основанный на прогнозировании нагрузки с использованием сетей LSTM. В отличие от традиционных методов, ориентированных на максимальный уровень сигнала RSSI, предложенное решение позволяет балансировать нагрузку между сотами, учитывая пространственно-временные характеристики трафика. Теоретической основой метода является выведенная в работе асимптотическая нижняя граница средней задержки. Результаты моделирования показали, что динамическая ассоциация позволяет значительно снизить среднюю задержку передачи данных по сравнению со статическими схемами распределения.

## ПУБЛИКАЦИИ АВТОРА ПО ТЕМЕ ИССЛЕДОВАНИЯ

### Публикации в рецензируемых изданиях из перечня ВАК:

1. Янковский, Н. А. О мультиплексировании потоков данных с использованием решетчатого кодирования в централизованных беспроводных сетях / Н. А. Янковский, И. А. Пастушок // Информационно-управляющие системы. 2021. № 2. С. 52-59.

2. Янковский, Н. А. Использование сетевого кодирования для массовых межмашинных коммуникаций / Н. А. Янковский // Успехи современной радиоэлектроники. – 2023. – Т. 77, № 8. – С. 50-57.

3. Янковский, Н. А. Оценка эффективности алгоритмов ассоциации и планирования для передачи чувствительного к задержке трафика по нисходящему каналу беспроводных сетей / Н. А. Янковский, Т. М. Татарникова // Успехи современной радиоэлектроники. – 2024. – Т. 78, № 8. – С. 33-38.

4. Yankovskii, N. A. Impact of correlated arrival and service flows for radio resource management in centralized wireless networks / N. A. Yankovskii, T. M. Tatarnikova // T-Comm. – 2024. – Vol. 18, No. 5. – P. 42-48.

5. Янковский Н. А. Динамическое управление ресурсами в восходящем канале 5G / Н. А. Янковский, Т. М. Татарникова // Информация и Космос. – 2025. - №4. – С. 38-44.

6. Татарникова, Т. М. Динамическое распределение ресурсов в восходящем канале сети 5G / Т. М. Татарникова, Ю. А. Ямщиков, Н. А. Янковский // Морская радиоэлектроника. – 2025. – № 4(94). – С. 22-28.

### Свидетельства о результатах интеллектуальной деятельности:

1. Свидетельство о государственной регистрации программы для ЭВМ № 2023684991 Российская Федерация. Программа расчёта и анализа задержки в восходящем канале межмашинных коммуникаций: № 2023684553: заявл. 14.11.2023: опубл. 21.11.2023 / Н. А. Янковский.

### Публикации изданиях, индексируемых в Web of Science и Scopus:

1. Pastushok, I. A. Bit Stream Multiplexing in 5G networks / I. A. Pastushok, N. A. Boikov, N. A. Yankovskii // 2020 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), St. Petersburg, Russia, 2020. P. 9131511.

2. Pastushok, I. A. Bitstreams multiplexing with trellis-coded modulation and the fixed point ldpc decoding procedure for centralised wireless networks/ I. A. Pastushok, N. A. Boikov, N. A. Yankovskii // Lecture Notes in Computer Science. – 2020. – Vol. 12525 LNCS. – P. 381-392.

3. Pastushok, I. A., On dependent distributions of appearance and servicing of claims in the M/M/1 / I. A. Pastushok, N. A. Yankovskii // 2021 Wave Electronics

and its Application in Information and Telecommunication Systems, WECONF 2021 - Conference Proceedings, Saint-Petersburg, 31 May 2021 - 04 June 2021. – Saint-Petersburg, 2021. – Vol. 4. – P. 9470500.

4. Enhanced Algorithm for Dynamic Changing of Frame Size in Radio Frequency Identification Systems / N. A. Yankovskii, A. Y. Kolesnikova, K. V. Afanasev, N. V. Stepanov // Wave Electronics and Its Application in Information and Telecommunication Systems. – 2022. – Vol. 5, No. 1. – P. 532-535.

5. Tatarnikova, T., Difficulties in Implementing a Heavy-tailed Distribution Using the Example of the Pareto Distribution in Modeling Network Traffic/ T. M. Tatarnikova, N. A. Yankovskii // Wave Electronics and Its Application in Information and Telecommunication Systems. – 2023. – Vol. 6, No. 1. – P. 430-433.

### **Публикации в других изданиях и конференциях:**

1. Янковский, Н. А. Сетевое кодирование в межмашинных коммуникациях / Н. А. Янковский // Волновая электроника и инфокоммуникационные системы: Материалы XXVI Международной научной конференции. В 3-х частях, Санкт-Петербург, 29 мая – 02 июня 2023 года. – Санкт-Петербург: ГУАП, 2023. – С. 137-141.

2. Янковский, Н. А. Использование методов машинного обучения для предсказания числа меток в RFID системах / А. А. Цыганкова, Н. А. Янковский // Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем ("Радиоинфоком-2024"): Сборник научных статей по материалам VIII Международной научно-практической конференции, Москва, 18–22 ноября 2024 года. – Москва: МИРЭА - Российский технологический университет, 2024. – С. 239-243.

3. Алгоритм оптимизации ассоциации пользователей к базовым станциям / Д. А. Русановская, Н. А. Янковский, А. Д. Каркина, В. А. Ушаков // Научному прогрессу – творчество молодых. – 2024. – № 1. – С. 323-326.

4. Использование скользящего сетевого кодирования в восходящем канале межмашинных коммуникаций / Н. А. Янковский, П. А. Гондунова, Д. М. Иванов, Г. Е. Есипов // Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем ("Радиоинфоком-2023"): Сборник научных статей по материалам VII Международной научно-практической конференции, Москва, 20–24 ноября 2023 года. – Москва: МИРЭА-Российский технологический университет, 2023. – С. 143-147.

5. Витвинов, В. К. Программно-аппаратный комплекс для исследования эффективности использования QUIC-протокола в системах интернета вещей / В. К. Витвинов, Н. А. Янковский // Сборник докладов Второй Международной научной конференции «Обработка, передача и защита информации в компьютерных системах». Санкт-Петербург, 11-15 апреля 2022 года. – С. 178-182.