

На правах рукописи

До Фук Хао

**РАЗРАБОТКА МОДЕЛЕЙ И МЕТОДОВ МИНИМИЗАЦИИ СЕТЕВОЙ
ЗАДЕРЖКИ НА ФРАГМЕНТЕ ГИБРИДНОЙ СЕТИ СВЯЗИ С
ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА**

2.2.15. Системы, сети и устройства телекоммуникаций

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Санкт-Петербург – 2025

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» на кафедре программной инженерии и вычислительной техники.

Научный руководитель: кандидат технических наук
Березкин Александр Александрович

Официальные оппоненты: **Канаев Андрей Константинович**,
доктор технических наук, профессор
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Петербургский государственный университет путей
связи Императора Александра I», кафедра
электрической связи, профессор

Карташевский Игорь Вячеславович,
доктор технических наук, доцент,
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Поволжский государственный университет
телекоммуникаций и информатики», Научно-
исследовательская лаборатория «Инновационные
проекты», заведующий лабораторией

Ведущая организация: Федеральное государственное бюджетное
учреждение науки «Институт проблем управления
им. В.А. Трапезникова Российской академии наук»
(ИПУ РАН), г. Москва

Защита диссертации состоится 25 декабря 2025 года в 16.00 на заседании диссертационного совета 55.2.004.01, созданном на базе федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича», по адресу: Санкт-Петербург, пр. Большевиков, д. 22, корп. 1, ауд. 554/1.

С диссертацией можно ознакомиться в библиотеке СПбГУТ по адресу: Санкт-Петербург, пр. Большевиков, д. 22, корп. 1 и на сайте www.sut.ru.

Автореферат разослан 24 октября 2025 года.

Ученый секретарь
диссертационного совета 55.2.004.01,
д-р техн. наук, доцент

А.Г. Владыко

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Современный этап развития телекоммуникаций характеризуется экспоненциальным ростом объёмов данных и повышенными требованиями к обеспечению глобального и качественного доступа к информационным ресурсам. В этом контексте создание гибридных орбитально-наземных сетей связи является одним из ключевых направлений, что находит отражение в «Стратегии развития отрасли связи Российской Федерации на период до 2035 года», утвержденный распоряжением правительства РФ 3339-р от 24 ноября 2023 года. Такие сети призваны обеспечить надежное покрытие на всей территории страны, включая труднодоступные и стратегически важные регионы, такие как Арктическая зона.

Однако, неотъемлемой особенностью спутникового сегмента таких гибридных сетей является значительная сетевая задержка. Этот фактор становится основным ограничением, которое существенно снижает качество обслуживания (QoS) для чувствительных ко времени приложений, таких как видеоконференцсвязь, передача телеметрических данных и сервисы реального времени. Сложность управления такими сетями усугубляется изменчивостью параметров каналов связи, что делает задачу минимизации задержки нетривиальной и требующей новых подходов.

В этих условиях применение технологий искусственного интеллекта (ИИ) открывает перспективные возможности для адаптивного управления сетевыми ресурсами. Методы ИИ позволяют в реальном времени анализировать состояние сети, прогнозировать нагрузку и оптимизировать передачу данных, тем самым смягчая негативное влияние задержек. Таким образом, разработка моделей и методов минимизации сетевой задержки с использованием ИИ является актуальной и своевременной научно-технической задачей, имеющей важное значение для повышения эффективности гибридных сетей связи и реализации стратегических задач в области связи. Актуальность настоящего исследования обусловлена быстрым развитием спутниковых систем и ростом требований к качеству обслуживания в гибридных сетях связи.

Степень разработанности темы. Анализ как отечественной, так и зарубежной научной литературы показывает, что вопросы минимизации сетевых задержек в спутниковых каналах связи уже привлекали внимание учёных, однако исследования в этой области не были проведены в полном объёме и комплексно. В частности, работа Nie (2021) посвящена моделированию и анализу характеристик сетевых задержек в спутниковых каналах при различных условиях, таких как высоты орбит, типы сигналов и погодные условия.

В работе акцентируется внимание на таких факторах, как доплеровский сдвиг, многолучевое распространение, помехи и другие, которые влияют на задержки сигнала. Другие исследования (например, Zong, 2023; Deng, 2021; Хохлов, 2021; Кулиева, 2023) сосредоточены на оптимизации методов передачи данных для повышения эффективности использования пропускной способности спутниковых каналов и совершенствовании технологий управления потоками данных с целью снижения задержек и потерь пакетов. В последние годы наблюдается рост интереса к применению машинного обучения и нейронных сетей для улучшения управления трафиком и снижения задержек в спутниковых каналах связи. Например, в работе Pacheco (2020) представлены алгоритмы

машинного обучения, которые позволяют прогнозировать поступление пакетов на приёмный терминал на основе характеристик потока кадров. Несмотря на существующие методы уменьшения сетевых задержек, их эффективность остаётся недостаточной в условиях растущей нагрузки на коммуникационные сети. Использование технологий искусственного интеллекта открывает возможности для создания более точных и адаптивных моделей, способных быстро реагировать на изменения в сети и оптимизировать передачу данных. Однако комплексные решения, объединяющие методы искусственного интеллекта и современные алгоритмы управления потоками данных, остаются недостаточно разработанными, что подчёркивает необходимость дальнейших исследований в этой области.

Объект исследования - спутниковый канал связи фрагмента гибридной орбитально-наземной сети связи.

Предмет исследования - методы и алгоритмы уменьшения сетевой задержки.

Цель диссертационной работы заключается в разработке и исследовании моделей и методов на основе технологий искусственного интеллекта, направленных на снижение сетевой задержки и повышение эффективности передачи данных в гибридных орбитально-наземных сетях связи.

Для достижения поставленной цели были определены следующие задачи исследования:

1. Провести анализ существующих подходов к снижению сетевой задержки в спутниковых и гибридных сетях связи, выявить их ограничения и определить перспективные направления для применения технологий искусственного интеллекта.
2. Разработать модель прогнозирования трафика и динамической приоритизации пакетов на основе нейросетевых методов для сокращения задержки в очереди.
3. Разработать методы многокритериальной оптимизации распределения сетевых ресурсов в многолучевых спутниковых системах с использованием гибридного подхода, сочетающего графовые нейронные сети и эволюционные алгоритмы.
4. Разработать и исследовать метод сжатия данных без потерь на основе пространственно-временных графовых нейронных сетей (ST-GNN) как средства снижения задержки передачи.
5. Провести сравнительный анализ эффективности предложенных моделей и методов посредством имитационного моделирования и экспериментальных исследований, оценивая ключевые показатели качества обслуживания, такие как средняя задержка и коэффициент потерь пакетов.

Научная новизна результатов исследования состоит в следующем:

1. Предложен подход к упреждающему управлению трафиком, который, в отличие от методов, использующих сложные модели, обосновывает достаточность вычислительно эффективной архитектуры MLP для решения задачи управления приоритетами. Другим элементом новизны является аналитически обоснованное правило приоритизации, связывающее прогноз с параметрами управления очередью, что представляет собой альтернативу распространённым эвристическим подходам.
2. Разработан гибридный метод оптимизации ресурсов, в котором графовая нейронная сеть (GNN) применяется в качестве предобработчика для классификации трафика перед этапом многокритериальной оптимизации с помощью адаптированных эволюционных алгоритмов (NSGA-II/SPEA2).

3. Предложен механизм адаптивного сжатия, где ST-GNN используется для вероятностного моделирования данных с явным учётом их графовой (пространственной) структуры. Такой подход позволяет кодеку динамически изменять стратегию сжатия в зависимости от локальных пространственно-временных корреляций, что является развитием классических методов, которые не учитывают топологию сети.

Теоретическая значимость работы заключается во вкладе в развитие теоретических основ управления гибридными сетями по трем ключевым направлениям:

1. Для теории систем массового обслуживания: предложены модели упреждающего управления, основанные на связке «прогноз нагрузки → динамическая приоритизация», и подходы к их аналитической оценке.

2. Для теории оптимизации сетей: предложена концептуальная модель, устанавливающая формальную связь между методами обработки графовых сигналов, эволюционной оптимизацией и теорией массового обслуживания для решения задач распределения ресурсов.

3. Для теории информации и кодирования: расширен класс моделей источников информации за счет введения графовой компоненты для описания данных с пространственно-временной структурой, что закладывает основу для разработки кодеков с совместной оптимизацией сжатия и маршрутизации.

Практическая значимость работы подтверждается следующими результатами:

1. Разработан вычислительно эффективный метод прогнозирования и приоритизации трафика (на основе MLP), который в ходе натурных испытаний на опытном GEO-канале продемонстрировал возможность снижения средней задержки до 19,6% и потерь пакетов до 19,42%.

2. Обоснована высокая потенциальная эффективность гибридного подхода к оптимизации ресурсов (на основе GNN и эволюционных алгоритмов), подтвержденная результатами имитационного моделирования (снижение времени ожидания в очереди составило до 35,1%), что доказывает целесообразность дальнейших разработок в данном направлении.

3. Разработан и успешно реализован аппаратный прототип кодека сжатия данных без потерь (на основе ST-GNN), продемонстрировавший в ходе натурных экспериментов высокую устойчивость к деградации спутникового канала (рост времени передачи на 11% против 46% у GZIP), что подтверждает его практическую жизнеспособность.

Методы исследования. Для решения поставленных в диссертации задач использовался комплекс методов, включающий теорию вероятностей и математическую статистику, теорию графов, теорию массового обслуживания, теорию информации и теорию оптимизации. Широко применялись методы машинного обучения, в частности, архитектуры глубоких нейронных сетей, а также эволюционные вычисления для решения задач многокритериальной оптимизации.

Положения, выносимые на защиту:

1. Модель прогнозирования и динамической приоритизации трафика на основе многослойного персептрона (MLP) для снижения задержки и потерь пакетов в спутниковом сегменте гибридной орбитально-наземной сети.

2. Методы оптимизации распределения сетевых ресурсов на основе графовых нейронных сетей (GNN) и эволюционных алгоритмов (NSGA-II/SPEA2) для минимизации задержки в спутниковой сегменте гибридной орбитально-наземной сети связи.

3. Метод сжатия данных без потерь на основе пространственно-временной графовой нейронной сети (ST-GNN) для снижения сетевой задержки в спутниковом канале связи.

Степень достоверности результатов. Достоверность полученных автором научных и практических результатов диссертационной работы подтверждается корректным использованием математического и имитационного моделирования, соответствием расчетов с результатами экспериментальных исследований, проведенных лично автором. Основные теоретические и практические результаты работы, в том числе разработанные имитационные модели, использованы в учебном процессе кафедры сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича при чтении лекций и проведении практических занятий и лабораторных работ, а также при проведении испытаний в научно-исследовательской и испытательной лаборатории инновационных инфокоммуникаций ПАО «Ростелеком», что подтверждается соответствующими актами внедрения.

Апробация результатов исследования. Основные результаты диссертационной работы докладывались и обсуждались на 27st International Conference on Distributed Computer and Communication Networks, DCCN 2024 (Москва, 2024); 8 и 9-я международные конференции «Информация, управление и коммуникационные технологии» (ICCT) (Владикавказ, 2024; Гомель, 2025); The 25th и 26th International Conference on Advanced Communication Technology, ICACT 2023 и 2024 (Корея, 2023, 2024); The 1st и 2nd International Conference on Advanced Computing & Next-Generation Communication (ICACNGC 2022, 2023); на 79-й конференциях Санкт-Петербургского отделения Общероссийской общественной организации «Российское научно-техническое общество радиотехники, электроники и связи им. А. С. Попова» (СПб НТОРЭС) (Санкт-Петербург, 2024), Международной научно-технической и научно-методической конференции «Актуальные проблемы инфотелекоммуникаций в науке и образовании» СПбГУТ (Санкт-Петербург, 2023, 2024), ПКМ; Подготовка профессиональных кадров в магистратуре для цифровой экономики (ПКМ, 2023);

Публикации по теме диссертации. Основные результаты диссертации изложены в 18 опубликованных работах, в том числе в 6 работах, опубликованных в журналах из перечня ВАК Министерства образования и науки Российской Федерации; в 5 работах, опубликованных в трудах, индексируемых Scopus и Web of Science, и 7 публикациях в других изданиях и материалах конференций.

Соответствие специальности. Диссертационная работа соответствует пунктам п. 1, 2, 3, 18, 19 паспорта специальности 2.2.15. Системы, сети и устройства телекоммуникаций.

Личный вклад автора. Результаты теоретических и экспериментальных исследований получены автором самостоятельно. В работах, опубликованных в соавторстве, соискателю принадлежит основная роль при постановке и решении задач, а также обобщении полученных результатов

Структура и объем диссертации. Диссертация состоит из введения, 5 глав, выводов, списка используемых источников и двух приложений. Полный объем диссертации составляет 168 страниц, из них к основной части относится 102 страниц. Работа содержит 35 рисунка, 15 таблиц и список из 110 литературных источников.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы, определены цели, задачи, новизна и практическая значимость работы, а также область применения и положения на защиту. Глава 1 описывает проблему задержек в спутниковых сетях, раскрывает особенности QoS и подчёркивает важность применения ИИ для минимизации задержек. Глава 2 анализирует существующие методы снижения задержек, включая традиционные и ИИ-решения, с акцентом на их преимущества и ограничения в гибридных сетях. Глава 3 диссертации посвящена разработке моделей для минимизации сетевой задержки в спутниковых каналах связи с использованием технологий искусственного интеллекта (ИИ). Введение (раздел 3.1) подчёркивает актуальность применения ИИ для решения проблемы сетевой задержки. Раздел 3.2 рассматривает сложные процессы передачи данных через спутниковые каналы и необходимость их оптимизации. Основное внимание уделяется методике организации передачи трафика с помощью ИИ (раздел 3.3), где представлена модель классификации и приоритезации сетевого трафика с использованием нейронной сети MLP, что позволяет улучшить QoS за счет оптимального распределения ресурсов.

На рисунке 1 показаны ключевые этапы передачи трафика в предлагаемом решении. Трафик проходит через этап «Определение типа трафика на основе ИИ», где определяется его тип и приоритет. Затем трафик сортируется по приоритету в очереди, оптимизируя распределение ресурсов и повышая QoS в спутниковой связи.

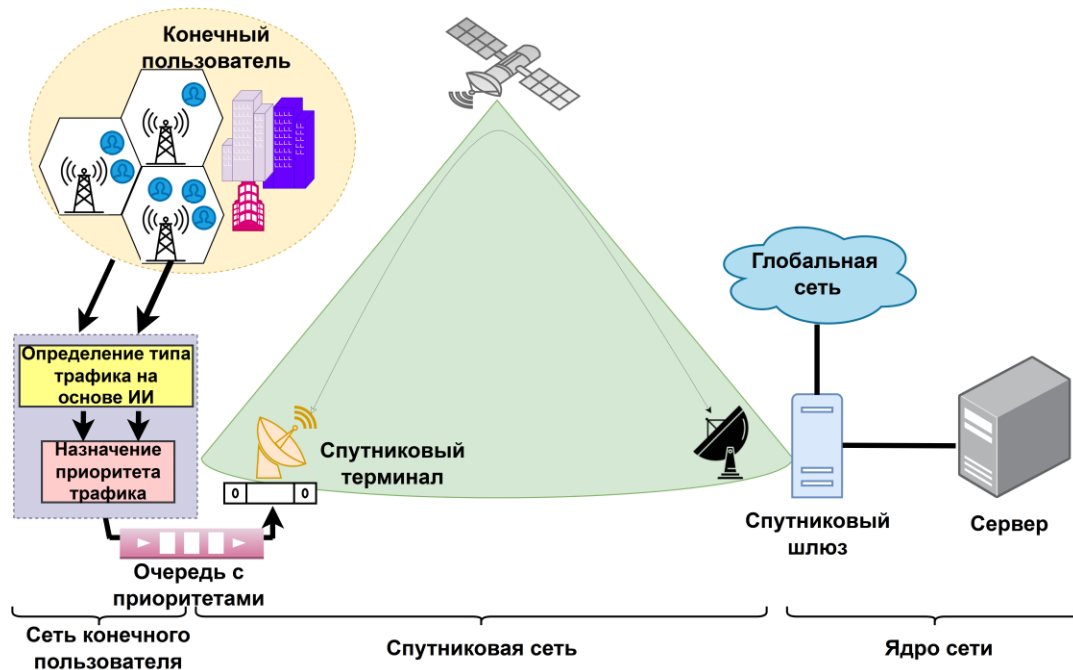


Рисунок 1 – Предлагаемая организация гибридной спутниковой сети связи

На рисунке 2 представлен процесс определения типа трафика и назначения приоритета для пользователей. Модель ИИ на основе MLP классифицирует четыре типа трафика для повышения QoS в спутниковой сети. Приоритеты распределены следующим образом: 1 — сетевое управление, 2 — критические коммуникации (тактильный интернет, AR), 3 — мультимедийные конференции (голос, видео), 4 — потоковое вещание (аудио и

видео), 5 — критический трафик (важные приложения), 6 — приоритетный трафик (менее критичные приложения), 7 — фоновые задачи (передача файлов).

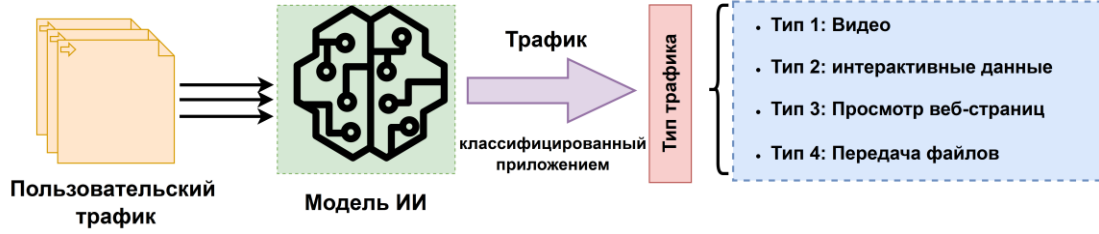


Рисунок 2 – Нейросетевая модель определения типа и назначения приоритетов сетевому трафику

На рисунке 3 представлено распределение уровней приоритета трафика для оптимизации спутниковой связи. После сортировки трафика следующим шагом является обеспечение приоритетной передачи через спутниковые сети. В **разделе 3.4** обсуждается обучение модели ИИ для классификации трафика на четыре типа: файлы, потоковое видео, веб-страницы и данные приложений в реальном времени, используя набор данных в формате PCAP. Рисунок 4 демонстрирует результаты 150 экспериментов с точностью модели MLP, которая превышает 90% при 87 итерациях.



Рисунок 3 – Присвоение уровней приоритета для сетевого трафика

Математическое представление модели передачи приоритезированного трафика описано в **разделе 3.5**, где детально объясняются этапы обработки трафика и присвоение уровней приоритета для эффективного распределения ресурсов и снижения задержек в ограниченных спутниковых каналах связи.

В математических представлениях используются следующие обозначения: T_i - тип трафика (где $i = 1, 2, \dots, m$); p_j - уровень приоритета (где $j = 1, 2, \dots, n$); x_i - входные параметры типа T_i ; y_i - прогнозируемый тип трафика из модели MLP; r_i - эффективная скорость передачи пакетов типа T_i ; l_i - допустимая задержка для T_i ; u_i - уровень срочности (целочисленное значение от 0 до i); B - доступная полоса пропускания (пакеты/с); D - возникшая задержка; L - коэффициент потери пакетов (%).

$$y_i = f_{MLP}(x_i) \quad (1)$$

где f_{MLP} - обученная модель многослойного персептрона, предназначенная для предсказания типа трафика на основе входных признаков x_i .

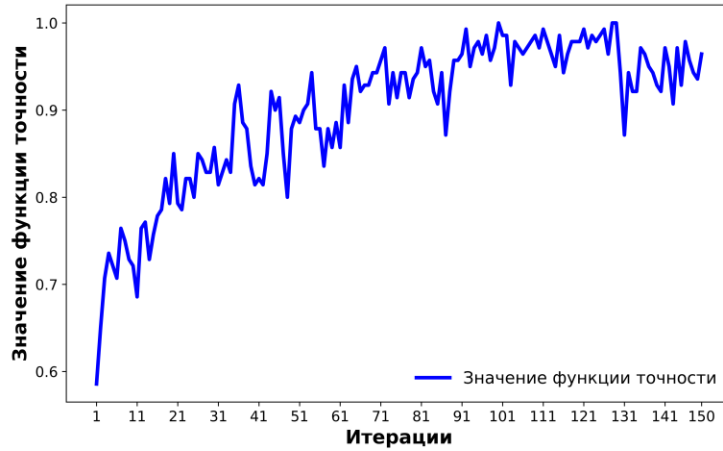


Рисунок 4 – Классификатор точности для обучающих данных модели на основе ИИ

Уровень приоритета для каждого типа трафика рассчитывается по следующей формуле:

$$p_i = w_1 f_1(r_i, B) + w_2 f_2(d_i, D) + w_3 f_3(l_i, L) + w_4 u_i + w_5 y_i \quad (2)$$

где веса $w_1, \dots, w_5$ соответствуют относительному влиянию каждого компонента (требований к скорости передачи данных, величине задержки, величине потерь, уровню срочности и определяемому типу трафика) при определении уровня приоритета для передаваемого трафика.

Функции отображения f_1, f_2 и f_3 определяются следующим образом:

$$f_1(r_i, B) = \begin{cases} \frac{r_i}{B}, & \text{если } r_i \leq r_{th}; \\ 1 + \xi \frac{(r_i - r_{th})}{(B - r_{th})}, & \text{если } r_i > r_{th}; \end{cases} \quad (3)$$

$$f_2(d_i, D) = \frac{1}{1 + e^{\alpha(D - d_i)}};$$

$$f_3(l_i, L) = 1 - e^{-\beta(L - l_i)};$$

где r_{th} - выделяемая полоса пропускания для передачи трафика; $f_1(r_i, B)$ моделирует требования к скорости передачи данных и соотношение с доступной полосой; $f_2(d_i, D)$ представляет зависимость между допустимой и текущей задержкой в сети; $f_3(l_i, L)$ моделирует соотношение допустимых потерь к текущему уровню потерь пакетов.

Параметры α, β и ξ должны быть выбраны на основе характеристик передачи трафика в спутниковых каналах и желаемого поведения системы. Значение α определяет спад допустимой задержки передачи пакетов при увеличении сетевой задержки; более высокие значения α приводят к более выраженному снижению. β определяет снижение допустимой потери пакетов при увеличении уровня потерь, и высокие значения β ускоряют это снижение. ξ формирует нелинейную зависимость между требуемой скоростью передачи данных и доступной полосой. Выбор ξ зависит от требований приложения к скорости передачи при ограниченной полосе.

Средняя задержка D_{avg} вычисляется как средняя задержка передачи пакетов, вычисляемая по формуле:

$$D_{avg} = \frac{1}{N} \sum_{i=1}^N D_i, \quad (4)$$

где D_i — задержка i -го пакета; N - общее количество пакетов.

Для оценки эффективности предлагаемого решения по приоритизации трафика рассмотрены следующие модели:

1. **FCFS** - обслуживает пакеты по порядку поступления;
2. **PQ** - использует статическое присвоение приоритетов;
3. **WFQ** - применяет пропорциональные веса для разных типов трафика;
4. **RED** - управляет очередью путем предварительного уничтожения пакетов;
5. **RR** - обрабатывает пакеты циклически.

Для предложенной схемы приоритизации примем скорость обслуживания μ_i для каждого типа трафика. Уровень приоритета p_i определяет то, как корректируется μ_i : более высокий приоритет получает более высокую скорость обслуживания, обеспечивая меньшие задержки для трафика с высоким приоритетом и наоборот для низкоприоритетного трафика. Скорость обслуживания μ_i регулируется в соответствии с выражением:

$$\mu_i = B(1 + p_i AF), \quad (5)$$

где B - доступная полоса пропускания в сети (пакетов/с); AF (*Adjustment Factor*) - коэффициент, используемый при приоритизации трафика для изменения скорости обслуживания исходя из уровня приоритета.

После определения скорости передачи μ_i на основе уровня приоритета p_i (5) можно рассчитать задержку D_i каждого пакета трафика типа i , используя формулу модели массового обслуживания М/М/1:

$$D_i = \frac{1}{\mu_i - \lambda}, \quad (6)$$

где λ скорость поступления пакетов; μ_i скорость обслуживания для трафика типа i . Более высокая скорость обслуживания μ_i приводит к меньшей задержке D_i , в то время как более низкая скорость обслуживания приводит к более высокой задержке.

Пусть $\mu_{proposed}^i$ - скорость обслуживания для трафика типа i в предложенной модели, μ_{other}^i - скорость обслуживания для трафика типа i в схеме обслуживания FCFS, а $D_{proposed}^i$ и D_{other}^i - задержки для трафика типа i в предложенной модели и модели FCFS соответственно. Учитывая, что более высокая скорость обслуживания приводит к меньшей задержке, имеем:

$$\mu_{proposed}^i > \mu_{other}^i \Rightarrow D_{proposed}^i < D_{other}^i. \quad (7)$$

Поэтому если задержки $D_{proposed}^i$ для отдельных типов трафика в целом ниже, чем D_{other}^i то средняя задержка $D_{proposed}^{avg}$ предлагаемой схемы также будет ниже D_{other}^{avg} :

$$\sum_{i=1}^N D_{proposed}^i < \sum_{i=1}^N D_{other}^i \Rightarrow D_{proposed}^{avg} < D_{other}^{avg}. \quad (8)$$

Пусть $\rho_{proposed}^i$ и $\mu_{proposed}^i$ коэффициент загрузки и скорость обслуживания для трафика типа i в предложенной модели, а ρ_{other}^i и μ_{other}^i для других типов трафика. Коэффициент загрузки канала ρ отражает долю времени, когда оборудование занято передачей пакетов.

При полной загрузке ($\rho \rightarrow 1$) очередь растет, увеличивая вероятность потери пакетов из-за переполнения буфера. Для сравнения эффективности предложенного метода с базовыми подходами используется модель М/М/1/К, которая аналогична М/М/1, но с

конечной длиной очереди K . Коэффициент загрузки для трафика типа i , ρ_i определяется формулой:

$$\rho_i = \frac{\lambda}{\mu_i}, \quad (9)$$

где λ - скорость поступления пакетов.

Учитывая, что более высокая скорость обслуживания μ_i приводит к более низкому коэффициенту загрузки ρ_i , имеем:

$$\mu_{proposed}^i > \mu_{other}^i \Rightarrow \rho_{proposed}^i < \rho_{other}^i \quad (10)$$

В очереди М/М/1/К коэффициент потери пакетов L_{rate} определяется следующим образом:

$$L_{rate} = \frac{(1-\rho)\rho^K}{1-\rho^{K+1}} \quad (11)$$

Более низкий коэффициент загрузки подразумевает более низкий коэффициент потери пакетов:

$$\rho_{proposed}^i < \rho_{other}^i \Rightarrow L_{rate, proposed}^i < L_{rate, other}^i \quad (12)$$

Средний коэффициент потери пакетов $L_{rate, avg}$ для каждого типа трафика рассчитывается следующим образом:

$$L_{rate}^{avg} = \frac{1}{N} \sum_{i=1}^N L_{rate}^i \quad (13)$$

Из формулы (13) следует, что если коэффициент потери пакетов $L_{rate, proposed}^i$ для отдельных типов трафика в целом будет ниже, чем $L_{rate, other}^i$ то средний коэффициент потери пакетов $L_{rate, proposed}^{avg}$ для конкретного типа трафика также будет ниже, чем $L_{rate, other}^{avg}$:

$$\sum_{i=1}^N L_{rate, proposed}^i < \sum_{i=1}^N L_{rate, other}^i \quad (14)$$

$$\Rightarrow L_{rate, proposed}^{avg} < L_{rate, other}^{avg} \quad (15)$$

Раздел 3.6 посвящён анализу задержек при различных уровнях приоритета. Используя параметры $\lambda = 5$ пак./с, $\mu = 12$ пак./с и $AF = 0.2$, модель тестируется на 20000 пакетах с четырьмя типами трафика. На рисунке 5 продемонстрировано, что приоритизация снижает среднюю задержку на 2,8 – 10,3% в зависимости от типа трафика. Рисунки 6-9 подтверждают эффективность предложенного метода приоритизации, где для видеотрафика задержка снизилась на 10,18% для интерактивного трафика на 7,86% веб трафика на 5,75 и при передаче файлов на 2,8%.

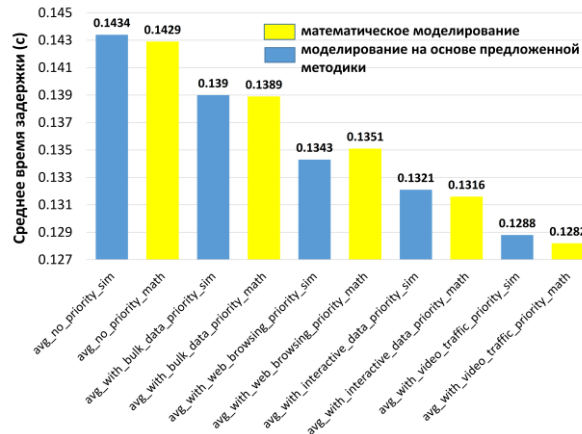


Рисунок 5 – Расхождение вычисленного среднего времени задержки передачи пакетов между математическим и имитационным моделированием

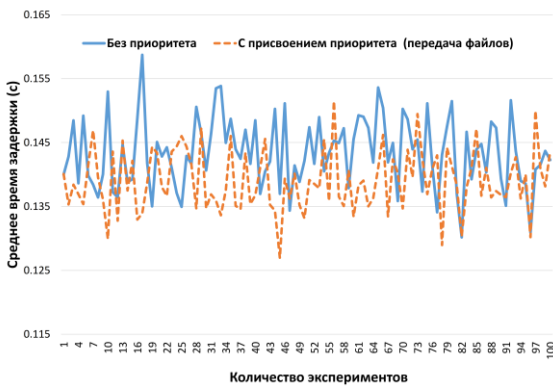


Рисунок 6 – Среднее время задержки пакетов при передаче файлов

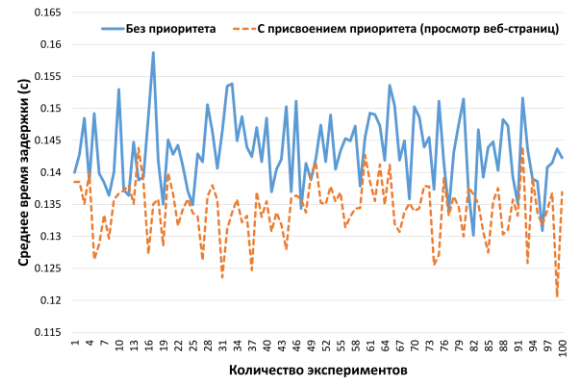


Рисунок 7 – Среднее время задержки пакетов при просмотре веб-страниц

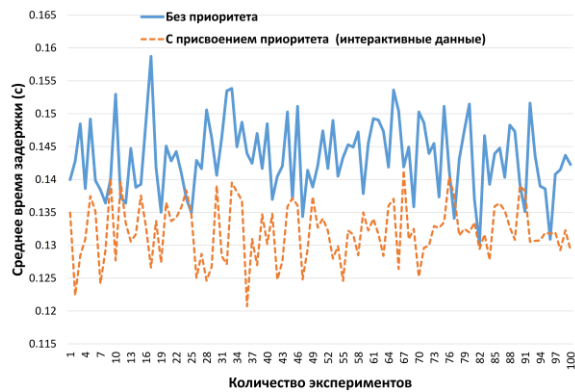


Рисунок 8 – Среднее время задержки пакетов при передаче интерактивных данных

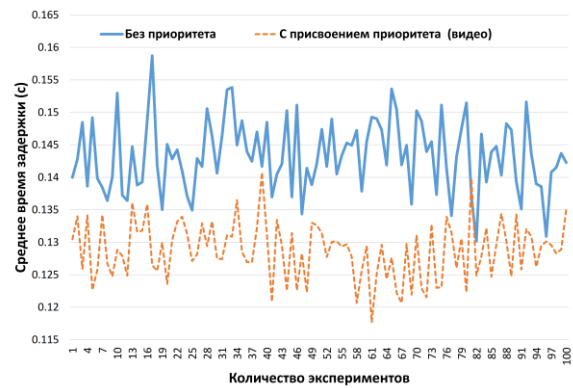


Рисунок 9 – Среднее время задержки пакетов при передаче видеотрафика

В разделе 3.7 сравниваются коэффициенты потерь пакетов в модели M/M/1/K при использовании предложенного метода приоритизации и метода FCFS при использовании. Рисунок 10 показывает, что результаты моделирования на основе предложенной методики, достаточно близки к математическим расчётам (например, 1,75% против 1,78% без приоритета). При введении приоритетов расхождение между расчётами и моделированием не превышает 0,03%, что подтверждает точность модели.

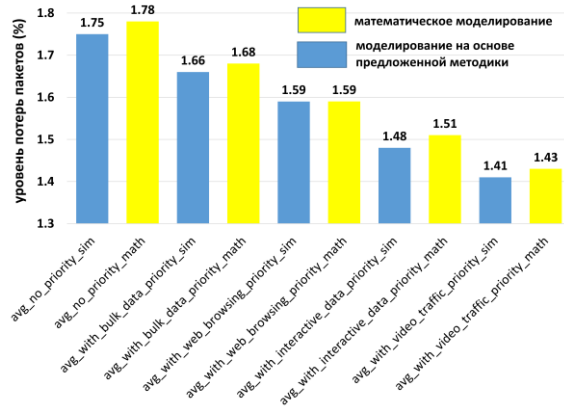


Рисунок 10 – Расхождение вычисленного коэффициента потери пакетов между математическим и имитационным моделированием

Глава 4 посвящена разработке методов ИИ для минимизации сетевой задержки в спутниковом сегменте гибридных спутниковых сетях. Основной фокус сделан на **оптимизации распределения ресурсов** в многолучевых системах (MBS) с учётом ограничений по спектру, мощности и QoS.

В **разделе 4.1** описывается задача распределения ресурсов, включая моделирование канала (**рис. 11**), расчёт скорости передачи и качества сигнала.

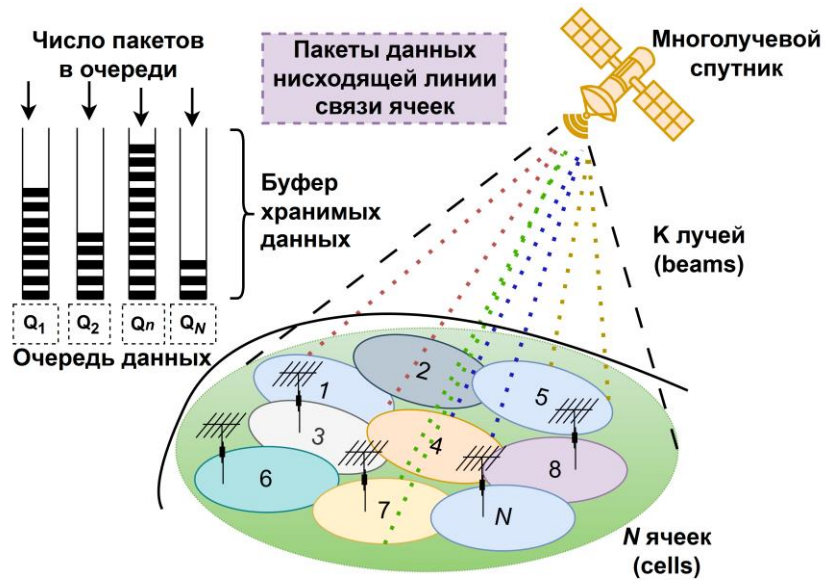


Рисунок 11 – Многолучевая система

Формулируется оптимизационная модель, минимизирующая задержку, потери пакетов и энергопотребление при соблюдении системных ограничений. Также приводится уравнение бюджета спутникового канала связи для оценки мощности сигнала на приёмной стороне.

$$P_r = \frac{P_t G_t G_r}{L_{fs} L_m}, \quad (16)$$

где P_r - приемная мощность, P_t - мощность передатчика, G_t - усиление антенны передатчика, G_r - усиление антенны приемника, L_{fs} - потери на свободном пространстве,

L_m - различные потери (атмосферное затухание, потери на оборудовании и др.). Усиление антенн рассчитывается по формуле:

$$G = \eta \left(\frac{\pi D}{\lambda} \right)^2, \quad (17)$$

где η - эффективность антенны, D - диаметр антенны, λ - длина волны сигнала. Потери в свободном пространстве, которые представляют собой затухание сигнала из-за расстояния, определяются выражением $L_{fs} = \left(\frac{4\pi d}{\lambda} \right)^2$, где d - расстояние распространения сигнала.

Достижимая скорость передачи данных в канале связи зависит от SNR. В спутниковом канале мощность шума (N_p) рассчитывается как $N_p = N_0 B^n$, где N_0 - спектральная плотность мощности шума, а B^n - выделенная полоса частот для ячейки n . Спектральная плотность определяется как $N_0 = k_b T_{sys}$, где k_b - постоянная Больцмана, а T_{sys} - температура шума системы. Учитывая помехи от других ячеек (I_{cci}), отношение сигнал/помехи плюс шум (SINR) для ячейки n определяется как:

$$\text{SINR}^n = \frac{P_r^n}{N_p + I_{cci}^n}. \quad (18)$$

Применяя теорему Шеннона, можно определить теоретически максимальную скорость передачи данных в ячейке n :

$$R^n = B^n \log_2 (1 + \text{SINR}^n). \quad (19)$$

BIP (паттерн освещения лучей) определяет набор ячеек, освещенных лучами спутника в определенный временной интервал. Эта схема влияет на распределение ресурсов, таких как мощность и пропускная способность, и определяет потенциал одновременного обслуживания пользователей в разных географических зонах. Математически BIP во временном интервале t можно представить в виде вектора:

$$X_t = \{x_t^1, x_t^2, \dots, x_t^N \mid x_t^n = 0,1\}, \quad (20)$$

где N - общее количество ячеек; x_t^n - двоичная переменная, указывающая статус освещения ячейки n в временном интервале t ; $x_t^n = 1$ - ячейка n освещена в временном интервале t ; $x_t^n = 0$ - ячейка n не освещена в временном интервале t .

Основные цели предлагаемого метода оптимизации заключаются в минимизации задержки, потерь пакетов и потребления энергии при соблюдении ограничений системы. Мы сосредоточены на оптимизации трех ключевых показателей производительности: задержка передачи в системе d_{sys} , представляющая среднее время, которое данные пакеты проводят в системе, включая задержки в очереди и время передачи; коэффициент потерь данных lr_{sys} , который количественно выражает долю потерянных пакетов из-за переполнения буфера или ошибок передачи и нагрузка по потреблению энергии Pl_{sys} , отражающая общее потребление энергии спутниковой системой, включая мощность передачи и другие бортовые системы.

Целевая функция G для задачи минимизации имеет следующий вид:

$$G = \min (\beta_1 d_{sys} + \beta_2 lr_{sys} + \beta_3 Pl_{sys}) \quad (21)$$

где d_{sys} - задержка передачи системы, lr_{sys} - коэффициент потерь пакетов данных, Pl_{sys} - нагрузка по потреблению энергии, $\beta_1, \beta_2, \beta_3$ - весовые коэффициенты, представляющие относительную важность каждой цели, определяются в соответствии с эволюционным алгоритмом, описанным в разделе 4.3.

В задаче оптимизации используются переменные, такие как x_t^n, P_t^k и B^n , представляющие статус освещения ячеек, распределение мощности по лучам и распределение полосы пропускания. Цель состоит в минимизации задержки, потерь пакетов и потребления энергии, соблюдая системные ограничения и оптимизируя использование ресурсов. Ограничения модели системы описаны следующим образом: C1 - ограничение активации луча: $\sum_{i=1}^N x_t^i = K$ и $x_t^n \in \{0,1\}$; C2 - ограничение по мощности: $\sum_{k=1}^K P_t^k \leq P_{tot}$; C3 - ограничение мощности индивидуального луча: $P_{min} \leq P_t^k \leq P_{max}$ для всех k ; C4 - ограничение полосы пропускания: $B^n \leq B_{tot}$ для всех n . Каждая ячейка оснащена очередью, моделируемой как буфер с принципом «первый пришёл, первый вышел» (FIFO), где входящие пакеты данных хранятся перед передачей. Модель очереди позволяет анализировать длину очереди Φ_t^n (количество пакетов данных в ячейке n в временном интервале t), задержку пакета L_t время, которое пакет проводит в очереди перед передачей, предыдущую длину очереди Φ_{t-1}^n (количество пакетов из предыдущего интервала), новые поступления ω_t^n (количество новых пакетов в текущем интервале), а также скорость обслуживания, определяемую выделенной пропускной способностью и условиями канала.

Длина очереди изменяется согласно следующему уравнению:

$$\Phi_t^n = \Phi_{t-1}^n - \frac{x_{t-1}^n P_{t-1}^k R^n}{B^n} + \omega_t^n, \quad (22)$$

где $\frac{x_{t-1}^n P_{t-1}^k R^n}{B^n}$ представляет количество пакетов, обслуженных в предыдущем временном интервале, с учетом состояния освещенности ячейки, выделенной мощности, достижимой скорости передачи данных и пропускной способности.

Рассмотрим систему с дискретным временем. Задержку i -го пакета в такой системе определим следующим образом:

$$L_t = \begin{bmatrix} l_t^1 \\ l_t^2 \\ \vdots \\ l_t^N \end{bmatrix} = \begin{bmatrix} t_s & 2t_s & \cdots & t_{th} \\ t_s & 2t_s & \cdots & t_{th} \\ \vdots & \vdots & \ddots & \vdots \\ t_s & 2t_s & \cdots & t_{th} \end{bmatrix}, \quad (23)$$

где l_t^n - время задержки пакета в ячейке n в интервале t относительно текущего интервала; t_s - продолжительность одного временного интервала; t_{th} - максимальное время порога очереди, представляющее максимально допустимую задержку, после которой пакет считается утерянным.

Индивидуальные показатели рассчитываются следующим образом:

- задержка передачи системы: $d_{sys} = \sum_{i=1}^N \sum_{j=1}^{t_{th}} \phi_{t,j}^i l_j^i$,
- коэффициент потерь пакетов данных: $lr_{sys} = \frac{\sum_{i=1}^N \phi_{t,t_{th}}^i}{\sum_{i=1}^N \sum_{j=1}^{t_{th}} \phi_{t,j}^i}$,

- нагрузка по потреблению энергии: $Pl_{sys} = \frac{\sum_{i=1}^N x_t^i P_t^i}{P_{tot}}$.

(21) представляет собой сложную функцию оптимизации, отражающую задачу учета множества конфликтующих целей: экономия энергии, улучшение задержек и минимизация потерь пакетов в спутниковых сетях. Для решения этой многокритериальной задачи подходят генетические алгоритмы, такие как NSGA-II и SPEA2, которые помогают эффективно исследовать компромиссы и находить оптимальные решения по Парето, предоставляя сбалансированные варианты для принятия решений.

Раздел 4.3 рассматривает применение генетических алгоритмов для многокритериальной оптимизации распределения ресурсов. Представлены две стратегии: NSGA-II (раздел 4.3.1) и SPEA2 (раздел 4.3.2). Таблица 1 содержит псевдокод NSGA-II, использующего недоминирующую сортировку для эволюции популяции решений к Парето-фронту. Таблица 2 иллюстрирует алгоритм SPEA2, основанный на механизмах естественного отбора, направленных на нахождение сбалансированных решений между конфликтующими целями. Оба подхода применяются для эффективной оптимизации параметров в многолучевых спутниковых сетях.

Таблица 1. Генетический алгоритм с недоминантной сортировкой II

Algorithm 1: NSGA-II (Генетический алгоритм с недоминантной сортировкой II)

```

1: Инициализировать популяцию  $P$  случайными решениями
2: Оценить целевые функции для каждого решения в  $P$ 
3: Инициализировать пустые множества  $F_1, F_2, F_k$  для хранения фронтов
4: Инициализировать пустое множество  $F$ 
5: Установить счётчик поколений  $t = 0$ 
6: while условие завершения не выполнено do
7:   Создать потомков  $Q$ , выполняя генетические операции (скрещивание и мутация) на  $P$ 
8:   Объединить популяции:  $R = P \cup Q$ 
9:   Выполнить недоминантную сортировку на  $R$ , чтобы создать фронты  $F_1, F_2, F_k$ 
10:  Инициализировать  $P'$  как пустое множество
11:  Установить  $i = 1$ 
12:  while  $|P'| + |F_i| \leq N$  do
13:    Присвоить ранг  $i$  индивидам из  $F_i$ 
14:    Добавить  $F_i$  в  $P'$ 
15:    Увеличить  $i$ 
16:  end while
17:  Отсортировать оставшихся индивидов в  $F_i$  на основе дистанции загруженности
18:  Добавить индивидов из  $F_i$  с наибольшей дистанцией загруженности в  $P'$ , пока  $|P'| = N$ 
19:  Обновить  $P$ , выбрав первых  $N$  индивидов из  $P'$ 
20:  Увеличить счётчик поколений:  $t = t + 1$ 
21: end while
22: return Аппроксимацию фронта Парето  $P$ 

```

Раздел 4.4 включает результаты двух экспериментов: первый (4.4.1) моделирует сценарий распределения ресурсов, второй (4.4.2) анализирует производительность в терминах задержки и потерь. Рисунки 12a и 12b показывают обратную зависимость между

энергопотреблением и ключевыми метриками: снижение потребления ведёт к увеличению задержки и потерь. NSGA-II демонстрирует лучшие показатели по задержке (51,6–53,3 мс) и потерям (0,31%–0,36%), тогда как SPEA2 оказывается более энергоэффективным.

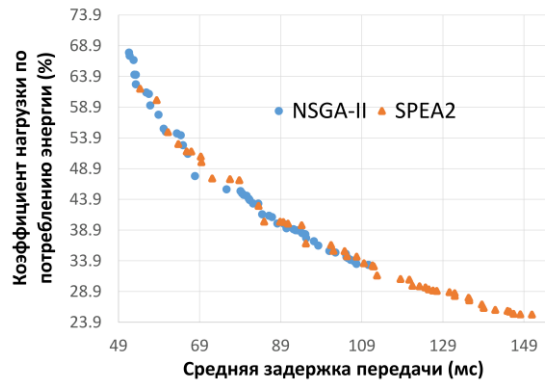


Рисунок 12а – Взаимосвязь между потреблением энергии спутниковой системы и средней задержкой передачи

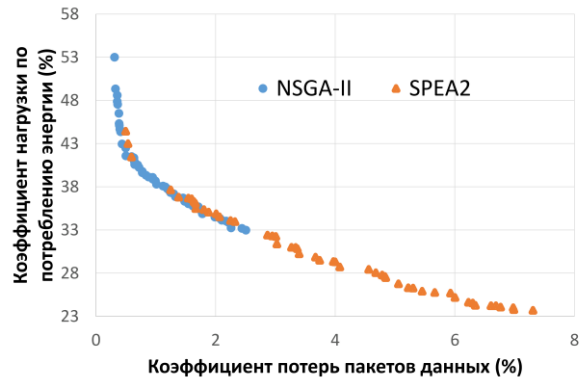


Рисунок 12б – Взаимосвязь между потреблением энергии спутниковой системы и коэффициентом потерь данных

Таблица 2. Алгоритм эволюции Strength Pareto 2 (SPEA-2)

Algorithm 2: Алгоритм эволюции Strength Pareto 2 (SPEA-2)

- 1: Инициализировать популяцию P
 - 2: Инициализировать архив A
 - 3: **repeat**
 - 4: Рассчитать сырые значения приспособленности и силу для каждого решения в P
 - 5: Рассчитать оценку плотности для каждого решения в P
 - 6: Обновить архив A не доминирующими решениями из P
 - 7: Выбрать индивидов из $P \cup A$ для следующего поколения
 - 8: Заменить P выбранными индивидами
 - 9: **until** условие завершения не выполнено
-

Рисунки 13а и 13б подтверждают, что с ростом пользовательского трафика (до 2000 Мбит/с) увеличиваются и задержка, и потери. Однако NSGA-II сохраняет преимущество над SPEA2: при 2000 Мбит/с задержка составляет 51,58 мс против 54,33 мс, а потери — 0,092% против 0,123%. При этом время выполнения у SPEA2 на 26% выше, что указывает на более высокую вычислительную сложность.

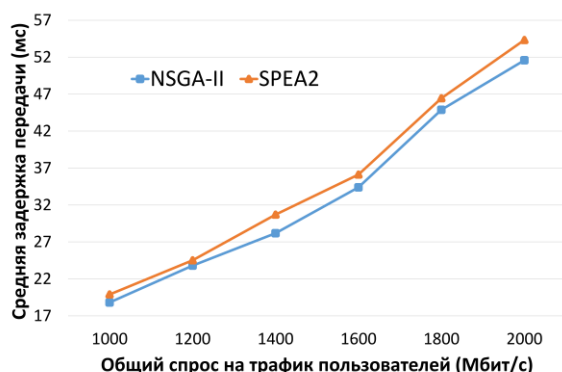


Рисунок 13а – Средняя задержка передачи при различных общих требованиях к пользовательскому трафику

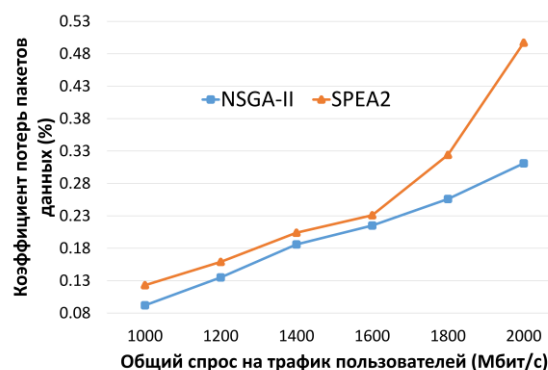


Рисунок 13б – Коэффициент потерь данных при различных общих требованиях к пользовательскому трафику

Раздел 4.5 посвящён симуляции системы с приоритетной очередью (М/М/1) в среде SimPy. Рисунок 14 показывает, что приоритизация значительно снижает среднее время ожидания пакетов: для OFDM — на 35,1%, для QAM — на 22,5%, и для MIMO — на 15%. Это подтверждает эффективность предложенного метода в условиях гибридной сети связи при различной нагрузке.

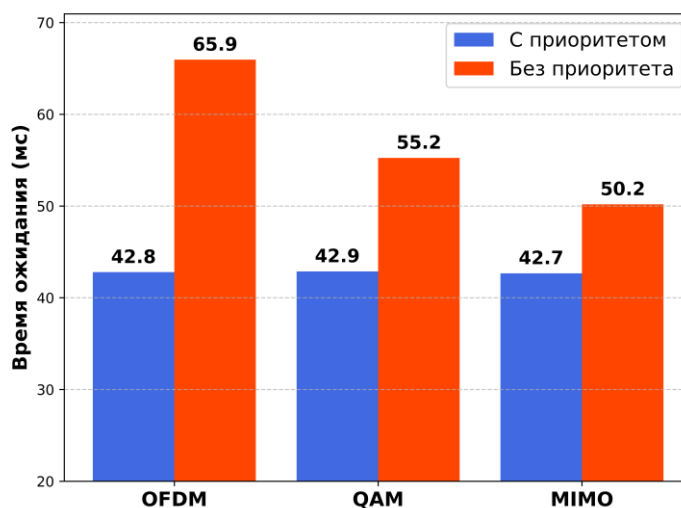


Рисунок 14 – Сравнение времени ожидания в очереди при использовании разных технологий

В **главе 5** рассматриваются методы сжатия данных без потерь на основе пространственно-временных графовых нейросетей (ST-GNN), направленные на снижение сетевой задержки в спутниковом канале связи. В **разделах 5.1–5.4** представлены актуальность задачи, обзор существующих методов, теоретические основы ST-GNN, а также описание архитектуры компрессора, предсказателя и декомпрессора. Основу составляет применение моделей LSTM и GRU для повышения точности прогнозирования трафика и эффективности сжатия в условиях ограниченной пропускной способности.

Рисунок 15 иллюстрирует общий процесс сжатия в спутниковом канале связи, включая стадии предсказания, кодирования и восстановления.



Рисунок 15 – Основной поток сжатия и передачи данных

Рисунок 16 раскрывает структуру ИИ-компрессора, который принимает 5-минутные сегменты трафика и формирует вероятностные распределения с использованием скользящего окна.

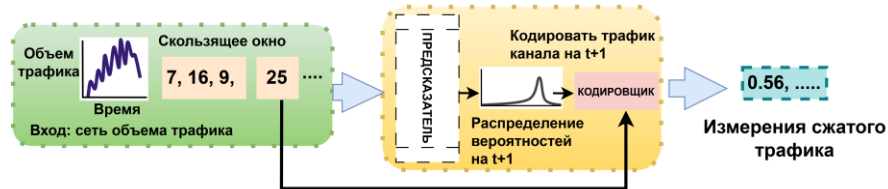


Рисунок 16 – Обзор модуля компрессора на основе искусственного интеллекта с его входными и выходными данными.

Рисунок 17 демонстрирует механизм предсказания с помощью ST-GNN: трафик кодируется через MLP, передаётся по связям графа и агрегируется для получения скрытых состояний, используемых в RNN для генерации последующих прогнозов.

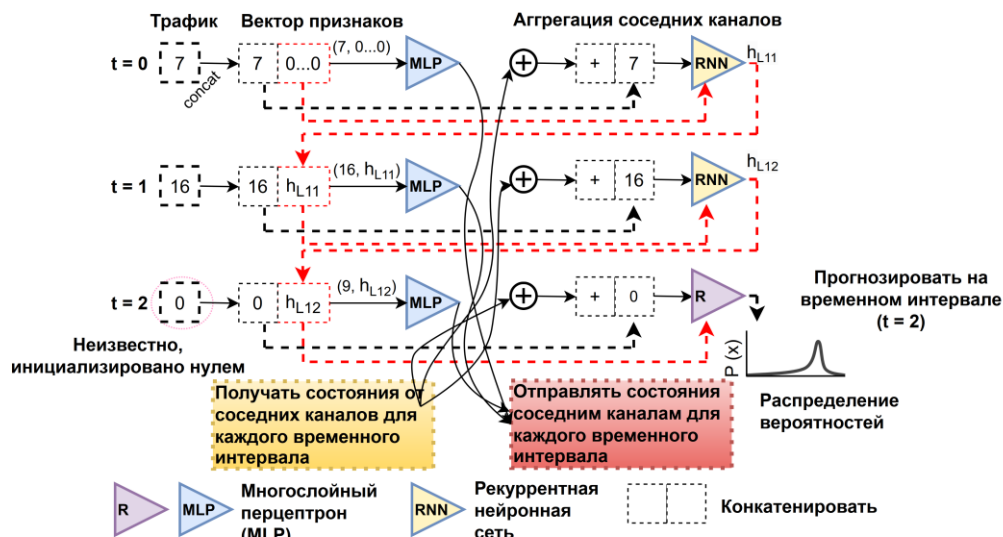


Рисунок 17 – Обзор модуля компрессора на основе ИИ с его входами и выходами

Рисунок 18 описывает весь цикл компрессии и декомпрессии. На этапе сжатия ST-GNN предсказывает распределения вероятностей, которые затем используются для эффективного кодирования трафика. При декомпрессии тот же механизм восстанавливает данные поэтапно, обеспечивая согласованность. Такой подход позволяет свести зависимость от канала до минимума за счёт переноса вычислений на локальные устройства.

В разделе 5.5 представлены экспериментальные результаты двух экспериментов, тестирующих методы сжатия и их влияние на сетевые задержки. Эксперимент 1 (5.5.1) оценивает эффективность сжатия данных, а эксперимент 2 (5.5.2) анализирует влияние методов сжатия на сетевые параметры.

Таблица 3 показывает сравнительные показатели производительности для моделей сжатия сетевого трафика на основе архитектур GRU и LSTM, с использованием предсказателя (рис. 17). Размер модели составляет 230 килобайт для GRU и 290 килобайт для LSTM, что связано с архитектурными различиями между ними. Среднее время кодирования и декодирования также указывает на вычислительную эффективность: GRU требует около 0,946 секунд для кодирования и 0,482 секунд для декодирования, в то время как LSTM около 0,933 секунд и 0,496 секунд соответственно.

В данном эксперименте используются параметры для оценки времени задержки в спутниковом канале связи. Ключевые характеристики включают высоту орбиты 1560 км, частоту нисходящего канала 21.5 ГГц, общую полосу пропускания 500 МГц, мощность спутника 34 дБВт и спектральную плотность мощности шума -187.6 дБм/Гц. Также рассматриваются усиление антенны передатчика (33.3 дБи), усиление антенны приемника (41.3 дБи) и потери на пути 202.1 дБ. Эти параметры играют важную роль в анализе и оптимизации производительности спутниковых систем.

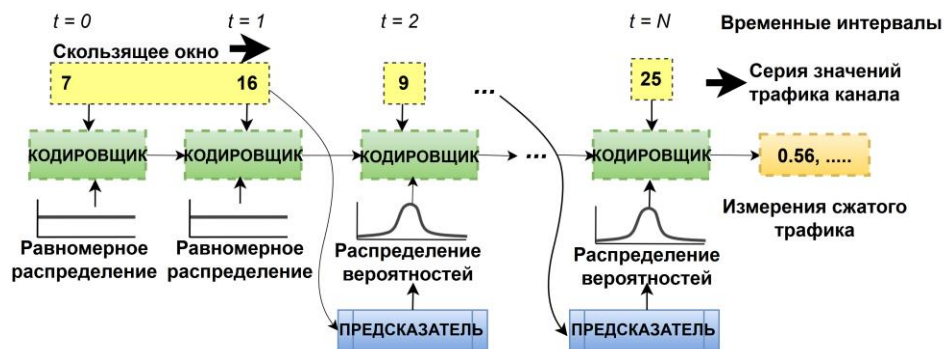


Рисунок 18 – Обзор модуля сжатия на основе ИИ с его входами и выходами

В исследовании оценивается эффективность модели ST-GNN для сжатия данных с целью снижения сетевой задержки, используя как теоретический расчёт битрейта (2335 Мбит/с, по Таблице 4), так и эмпирические данные трафика (≈ 6750 МБ каждые 5 минут).

Таблица 3. Сравнительные показатели производительности для
двух различных конфигураций ПРЕДСКАЗАТЕЛЯ

Предсказатель	Размер модели (кб)	Среднее время кодирования (с)	Среднее время декодирования (с)	Коэффициент сжатия
LSTM	290	0.933	0.496	2.426
GRU	230	0.946	0.482	3.397

Размер до и после сжатия обозначаются как L_1 и L_2 , а задержка рассчитывается как:

- До сжатия: $t_{\text{до}} = L_1 / \text{bitrate}$
- После: $t_{\text{после}} = L_2 + t_{\text{код}} + t_{\text{декод}}$

Снижение задержки рассчитывается по формуле:

$$\Delta t = \frac{t_{\text{до}} - t_{\text{после}}}{t_{\text{до}}} \quad (24)$$

Результаты показывают уменьшение времени передачи на 9,35% (LSTM) и 21,1% (GRU). Таблица 4 демонстрирует зависимости эффективности от объёма и битрейта при использовании LSTM.

Таблица 4. Процентное снижение времени задержки (%) в зависимости от размеров объёма трафика и битовых скоростей для передачи данных (используя LSTM в качестве предсказателя)

Объем трафика для каждого интервала (Мб) / Скорость передачи данных (Мбит/с).	2000	2100	2200	2300	2400	2500
6500	11.18	12.64	10.44	8.25	6.05	3.85
6600	12.71	13.34	11.18	9.01	6.85	4.68
6700	14.15	14.02	11.89	9.75	7.62	5.49
6800	15.5	14.68	12.58	10.48	8.38	6.28
6900	16.78	15.31	13.24	11.18	9.11	7.04
7000	17.98	15.94	13.9	11.86	9.82	7.78

В таблице 5 предоставлено процентное снижение времени задержки, для различных размеров объемов трафика и битовых скоростей для передачи данных, используя GRU в качестве модели предсказания.

Таблица 5. Перцентное снижение времени задержки (%) в зависимости от размеров объема трафика и битовых скоростей при передаче данных (используя GRU в качестве предсказателя)

Объем трафика для каждого интервала (Мб) / Скорость передачи данных (Мбит/с).	2000	2100	2200	2300	2400	2500
6500	26.62	24.43	22.23	20.03	17.84	15.64
6600	27.29	25.13	22.96	20.8	18.64	16.47
6700	27.94	25.8	23.67	21.54	19.41	17.28
6800	28.56	26.46	24.36	22.26	20.16	18.06
6900	29.17	27.1	25.03	22.96	20.89	18.82
7000	29.76	27.72	25.68	23.64	21.6	19.56

Раздел 5.6 посвящён практической валидации предложенного метода ST-GNN в условиях реальной спутниковой среды с использованием физического GEO-модема и периферийных серверов с GPU. Целью является подтверждение устойчивости результатов, полученных в симуляции, в рамках аппаратного эксперимента.

Эксперименты проводились с использованием модема Gilat SkyEdge CPE, с уровнем приёма (4,6 дБ) и передачи (65,3 дБм), что зафиксировано на рисунке 19.

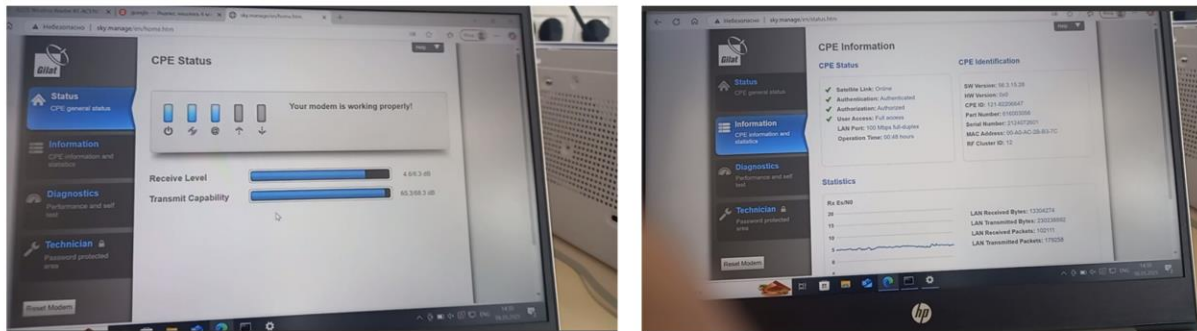


Рис.19. Интерфейс GEO-модема (а) панель «Статус» с отображением текущих уровней приёма (RX) и передачи (TX); (б) панель «Информация» с состоянием канала, идентификаторами устройства и статистикой трафика.

Панель модема показывает стабильное соединение, полнодуплексную работу по Ethernet на скорости 100 Мбит/с, и загрузку трафика до 230 МБ, что служит эталонной отправной точкой для анализа сжатия. На рисунке 20 представлены два сервера на базе GPU NVIDIA RTX 4080, выполняющие роль кодировщика и декодировщика соответственно. Оба устройства работают под управлением Ubuntu 22.04 с поддержкой CUDA 12 и PyTorch 2.3, обеспечивая прямой прозрачный маршрут между локальной сетью и модемом. Рисунок 21 отображает результаты живых ICMP-тестов (пинг до DNS Google), которые показали среднее RTT около 841 мс, с диапазоном 700–1079 мс, что подтверждает реалистичный бюджет задержки для GEO-соединений. Рисунок 22 демонстрирует сам модем, применённый в лабораторных испытаниях.



Рис. 20. Два сервера на периферии с GPU NVIDIA RTX 4080 один кодирущик для восходящего канала, другой декодирущик для нисходящего канала работают под CUDA 12 и PyTorch 2.3, подключены в линию между локальной сетью и GEO-модемом.

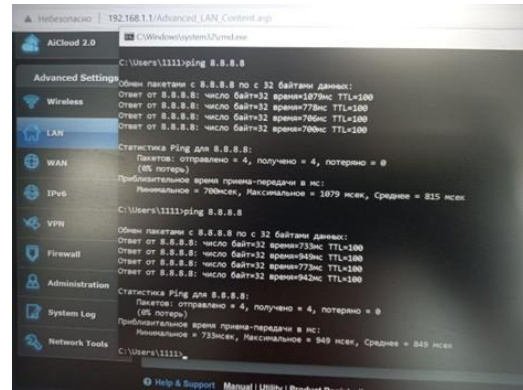


Рис. 21. Окно диагностики роутера с результатами пинга по активной GEO спутниковой ссылке (мин. ≈ 700 мс, макс. ≈ 1079 мс, среднее $\approx 0,8$ с RTT).

Экспериментальные измерения на тестовом наборе «Собранные данные» визуализированы в рисунке 23 круговая диаграмма демонстрирует вклад каждого этапа в общую задержку передачи. В случае ST-GNN почти 98% времени уходит на кодирование и декодирование, тогда как время передачи составляет менее 2%. Напротив, для GZIP около 99,9% времени приходится на эфир, а вычислительная часть почти неощутима.

Эти данные обобщены в таблице 6, где показано процентное распределение компонентов задержки:

- ST-GNN: кодирование 36.2%, передача 1.9%, декодирование 62.0%.
- GZIP: передача 99.9%, вычисления $<0.1\%$.



Рис. 22. GEO-модем, используемый в аппаратной тестовой среде.

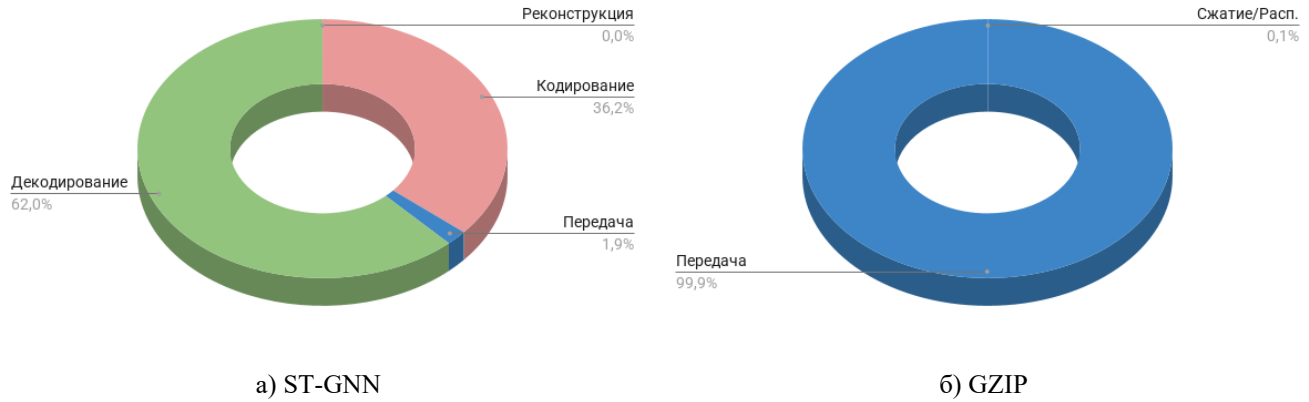


Рис. 23: Вклад компонентов задержки в общую задержку сигнала на основе собранных данных

В таблице 6 представлены два распределения долей общего времени задержки (%) между разными этапами обработки данных для алгоритмов ST-GNN и GZIP (по результатам экспериментов) на основе полного собранного набора данных.

В таблице 7 предоставлено абсолютное время выполнения полного цикла передачи в секундах при двух уровнях RTT 500 мс (моделирование) и 841 мс (реальный).

- При RTT = 500 мс, ST-GNN выполняет цикл за 376.9 с, GZIP за 453.0 с.
- При RTT = 841 мс, ST-GNN требует 419.9 с, GZIP 662.2 с.

Таким образом, при росте RTT производительность GZIP деградирует значительно сильнее (+46%) по сравнению со ST-GNN (+11%), что расширяет относительный выигрыш ST-GNN до 242 секунд или 37%.

Таблица 6: Доля от общего времени (%) и суммарное снижение задержки

Компонент / Метрика	Значение (%)	
	ST-GNN	GZIP
Кодирование/Сжатие	36.16	0.05
Передача	1.87	99.93
Декодирование/Распаковка	61.96	0.02
Реконструкция ST-GNN	0.01	—

Эти данные подтверждают, что в условиях высокой задержки ST-GNN обеспечивает лучшее распределение вычислительной нагрузки, сводя к минимуму влияние передачи по каналу связи, особенно при наличии GPU-ресурсов на шлюзах.

Таблица 7: Абсолютное время выполнения при различных RTT

Метод	Кодирование	Передача	Декодирование	Восстановление	Итого
RTT=500 мс (моделирование)					
ST-GNN	136.29	7.06	233.56	0.025	376.94
GZIP	0.25	452.69	0.07	–	453.01
RTT=841 (реальный)					
ST-GNN	154.39	20.71	244.77	0.028	419.89
GZIP	0.27	661.87	0.07	–	662.21

Аппаратные тесты подтвердили результаты моделирования: даже при практическом RTT GEO-соединения (841 мс) графовый кодек ST-GNN значительно сокращает общее время передачи по сравнению с GZIP.

ЗАКЛЮЧЕНИЕ

В диссертационной работе решена актуальная научно-техническая задача по снижению сетевой задержки на фрагменте гибридной орбитально-наземной сети связи путем разработки и исследования моделей и методов на основе технологий искусственного интеллекта. Основные научные и практические результаты, полученные в ходе исследования, заключаются в следующем:

1. Разработана модель упреждающего управления трафиком, сочетающая вычислительно эффективный многослойный персептрон (MLP) для прогнозирования нагрузки и аналитически обоснованное правило для динамической приоритизации. Эффективность модели подтверждена экспериментально на опытном GEO-канале, где было зафиксировано снижение средней задержки до 19,6% и потерь пакетов до 19,42% по сравнению с базовыми алгоритмами.

2. Предложены методы многокритериальной оптимизации ресурсов, основанные на гибридном подходе, где графовая нейронная сеть (GNN) применяется для предварительной классификации трафика, а адаптированные эволюционные алгоритмы (NSGA-II/SPEA2) — для распределения ресурсов. Результаты имитационного моделирования показали высокий потенциал подхода, продемонстрировав возможность снижения времени ожидания в очереди до 35,1%.

3. Разработан и реализован на основе программно-аппаратного комплекса метод сжатия данных без потерь на основе пространственно-временной графовой нейронной сети (ST-GNN). В ходе натурных экспериментов на реальном спутниковом маршруте продемонстрирована его высокая устойчивость к деградации канала: при росте RTT общее время передачи для предложенного метода увеличилось лишь на 11% против 46% у стандартного алгоритма GZIP, что подтверждает практическую жизнеспособность решения.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в рецензируемых научных изданиях, рекомендованных ВАК

1. Фук Х.До. Анализ методов минимизации сетевой задержки в каналах связи при передаче трафика промышленного интернета вещей / Фук Х.До., Березкин А.А., Парфенов Д.Д., Кукунин Д.С., Киричек Р.В. // Электросвязь. – 2023. – № 1. – С. 8-16.
2. Ф. Х. До. Графовые нейронные сети для классификации трафика в каналах спутниковой связи: сравнительный анализ / До Ф. Х., Ле Ч. Д., Берёзкин А.А., Киричек Р.В. // Труды учебных заведений связи. – 2023. – Т. 9. – №. 3. – С. 14-27.
3. Ф. Х. До. Методика приоритизации трафика в спутниковых каналах связи на базе искусственного интеллекта. Часть I. Анализ задержки передачи трафика / До Ф. Х., Березкин А.А., Бородко А.В., Киричек Р.В. // Электросвязь. – 2024. – № 6. – С. 37-46.
4. Ф. Х. До. Методика приоритизации трафика в спутниковых каналах связи на базе искусственного интеллекта. Часть II. Анализ потерь пакетов. / До Ф. Х., Березкин А.А., Бородко А.В., Киричек Р.В. // Электросвязь. – 2024. – № 7. – С. 8-13.
5. Ф. Х. До. Методика тестирования методов минимизации сетевой задержки на фрагменте гибридной сети связи. Часть I. Интеллектуальная приоритезация и сокращение задержек в очередях // Электросвязь. – 2025. – № 7. – С. 62-72.
6. Ф. Х. До. Методика тестирования методов минимизации сетевой задержки на фрагменте гибридной сети связи. Часть II. Сжатие ST-GNN и снижение задержки передачи // Электросвязь. – 2025. – № 8. – С. 17-26.

Публикации в изданиях, входящих в международные базы цитирования

7. Do P.H., Le T.D., Berezkin A., Kirichek R. Optimizing Resource Allocation for Multi-beam Satellites Using Genetic Algorithm Variations // Distributed Computer and Communication Networks: 27th International Conference; Vishnevsky, V.M., Samouylov, K.E., Kozyrev, D.V. (eds). DCCN 2024, Lecture Notes in Computer Science, Vol. 15460. Springer, Cham. – Pp. 16-29.
8. Phuc Hao Do, Le D.T., Berezkin A., Kirichek R. ST-GNN-Powered Compression for Latency Reduction in Hybrid Satellite-Terrestrial Networks. Conference: 2024 8th International Conference on Information, Control, and Communication Technologies (ICCT). – Pp. 1-10.
9. Do P. H. et al. A Horizontal Federated Learning Approach to IoT Malware Traffic Detection: An Empirical Evaluation with N-BaIoT Dataset // Conference: 2024 26th International Conference on Advanced Communications Technology (ICACT). – Pp. 1494-1506.
10. Do P. H. et al. A Horizontal Federated-Learning Model for Detecting Abnormal Traffic Generated by Malware in IoT Networks // Conference: 2023 25th International Conference on Advanced Communication Technology (ICACT). – Pp. 28-36.
11. Phuc Hao Do, Le D.T., Berezkin A., Kirichek R. Optimizing Federated Learning Performance in GEO Satellite Networks using Successive Convex Approximation. 9th International Conference on Information, Control, and Communication Technologies (ICCT). – IEEE, 2025.

Публикации в других изданиях и сборниках научных трудов и конференций

12. Х. Ф. До. Обзор решений на основе искусственного интеллекта для минимизации задержки в спутниковых сетях связи / А.А. Березкин, Х. Ф. До, Р.В. Киричек // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2023): Сборник научных статей. XII Международная научно-техническая и научно-методическая конференция. В 4 т., Санкт-Петербург, 28 февраля – 01 марта 2023 года. Том 1. – Санкт-Петербург: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2023. – С. 167-171.

13. Ф. Х. До. Исследование предиктивной аналитики на базе микроконтроллера с применением методов TinyML / Ф. Х. До, Ч. Д. Ле, М. А. Зуйков [и др.] // Инновационное приборостроение. – 2023. – Т. 2. – № 2. – С. 64-75.

14. Phuc Hao Do, Le D.T., Berezkin A., Kirichek R. Advancing Satellite Communications: Multi-Objective Optimization with Genetic Algorithms // Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь (DCCN-2024) = Distributed computer and communication networks: control, computation, communications (DCCN-2024): материалы XXVII Международной научной конференции. Россия, Москва, 23–27 сентября 2024 г. / под общ. ред. В. М. Вишневого и К. Е. Самуйлова. – С. 1-6.

15. Х. Ф. До. Анализ методов искусственного интеллекта для улучшения качества обслуживания в низкоорбитальных спутниковых группировках / А.А. Березкин, Х. Ф. До, Р.В. Киричек // Актуальные проблемы инфотелекоммуникаций в науке и образовании: Сборник научных статей XIII Международной научно-технической и научно-методической конференции в 4 т., Санкт-Петербург, 27–28 февраля 2024 года. – Санкт-Петербург: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2024. – С. 96-100.

16. Х. Ф. До. Динамическое распределение ресурсов с использованием искусственного интеллекта в спутниковых каналах связи / А.А. Березкин, Х. Ф. До, Р.В. Киричек // Подготовка профессиональных кадров в магистратуре для цифровой экономики (ПКМ-2023). Всероссийская научно-техническая и научнометодическая конференция магистрантов и их руководителей; Сборник лучших докладов: в 2 т. Т.2. / сост. Н. Н. Иванов. – СПбГУТ.: – Санкт-Петербург, 2024. – С. 234-239.

17. До Х. Ф., Зуйков М.А., Березкин А.А., Киричек Р.В. Анализ методов уменьшения сетевой задержки в гибридных орбитально-наземных сетях связи. // 79-я Научно-техническая конференция СПб НТО РЭС им. А.С. Попова, посвященная Дню радио – Санкт-Петербург: Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина), 2024. – С. 143-146.

18. Phuc Hao Do, Le D.T., Berezkin A., Kirichek R. Reducing latency in hybrid satellite-terrestrial networks through lossless network traffic compression using spatio-temporal graph neural networks. // Информационные технологии и технические средства управления (ICCT-2024) = International Conference on Information, Control, and Communication Technologies (ICCT-2024): материалы VIII Междунар. научн. конф., 1–5 октября 2024 г., Владикавказ / под общ. ред. Е.А. Барабановой, К.А. Вытовтова; Ин-т проблем упр. им. В.А. Трапезникова Рос. акад. наук Минобрнауки РФ. – С. 117-119.

Подписано в печать 22.10.2025. Формат 60×84 1/16.
1,0 а.л. Тираж 100 экз.

Отпечатано в СПбГУТ, 193232, Санкт-Петербург, пр. Большевиков, д. 22, корп. 1