

**Федеральное государственное автономное образовательное учреждение
высшего образования
«РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ
ИМЕНИ ПАТРИСА ЛУМУМБЫ»**

На правах рукописи

Кочеткова Ирина Андреевна

**МОДЕЛИ ДЛЯ АНАЛИЗА ЭФФЕКТИВНОСТИ
ПРИОРИТЕТНОГО ДОСТУПА В БЕСПРОВОДНЫХ СЕТЯХ
С ЭЛАСТИЧНЫМ И ПОТОКОВЫМ ТРАФИКОМ**

Специальность 1.2.3 – Теоретическая информатика, кибернетика
(по физико-математическим наукам)

Диссертация

на соискание ученой степени
доктора физико-математических наук

Научный консультант
доктор технических наук, профессор
Самуйлов Константин Евгеньевич

Москва – 2025

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	7
ГЛАВА 1 ПРИОРИТЕТНОЕ ОБСЛУЖИВАНИЕ ТРАФИКА В МУЛЬТИСЕРВИСНЫХ БЕСПРОВОДНЫХ СЕТЯХ	22
1.1 Приоритизация услуг	22
1.1.1 Сети новых поколений	22
1.1.2 Классы услуг.....	25
1.1.3 Типы трафика	30
1.2 Модели трафика в беспроводных сетях	32
1.2.1 Модели с одним типом трафика.....	32
1.2.2 Модели с потоковым и эластичным трафиком.....	35
1.2.3 Полнодоступная модель «тройной услуги»	36
1.3 Особенности моделирования радиоресурса	38
1.3.1 Модель требования к ресурсу.....	38
1.3.2 Модель затухания сигнала для анализа случайного требования к ресурсу	42
1.3.3 Ресурсная система массового обслуживания.....	45
1.4 Система нарезки радиоресурсов	48
1.4.1 Особенности нарезки радиоресурсов	48
1.4.2 Нестационарные требования к объему сегментов.....	49
1.4.3 Принципы перераспределения ресурса	53
1.5 Система совместного доступа к радиочастотам	54
1.5.1 Особенности использования нелицензируемых и лицензируемых радиочастот.....	54
1.5.2 Механизм снижения мощности сигнала	56
1.5.3 Замечание к нестационарному режиму	60
1.6 Описание комплекса моделей	62
1.6.1 Нотация приоритетов в символике Башарина-Кендала.....	62
1.6.2 Матричная нотация приоритетов	65
1.6.3 Комплекс моделей приоритетного доступа	69

ГЛАВА 2	МОДЕЛИ С ПОТОКОВЫМ ТРАФИКОМ И ПОСТОЯННЫМ ОБЪЕМОМ РЕСУРСА	75
2.1	Модель сверхнадежной передачи с малой задержкой и широкополосной связи	75
2.1.1	Системная модель	75
2.1.2	Схемы приоритетного доступа	77
2.1.3	Модели математической теории телетрафика	80
2.2	Система с узкополосным и широкополосным трафиком	83
2.2.1	Матрица интенсивностей переходов в блочно-трехдиагональном виде	83
2.2.2	Матричный алгоритм расчета стационарного распределения	87
2.2.3	Алгоритм снижения размерности системы уравнений равновесия	91
2.3	Ресурсная система со снижением мощности	96
2.3.1	Модель требования к ресурсу	96
2.3.2	Вероятности занятия ресурса	98
2.3.3	Условные вероятности выделения ресурса	100
2.4	Анализ показателей эффективности передачи широкополосного трафика	103
2.4.1	Вероятность приостановки обслуживания	103
2.4.2	Средняя скорость передачи трафика	108
2.4.3	Вероятность прерывания обслуживания	110
2.5	Модель доступа для видео по требованию и видеоконференции	111
2.5.1	Системная модель	111
2.5.2	Схемы приоритетного доступа	113
2.5.3	Модели математической теории телетрафика	115
2.6	Система с одноадресным и многоадресным трафиком	117
2.6.1	Рекуррентный алгоритм расчета стационарного распределения	117
2.6.2	Алгоритм снижения размерности системы уравнений равновесия ..	121
2.6.3	Средняя скорость передачи многоадресного трафика	127

ГЛАВА 3	МОДЕЛИ С ЭЛАСТИЧНЫМ ТРАФИКОМ И ПОСТОЯННЫМ ОБЪЕМОМ РЕСУРСА	132
3.1	Модель обслуживания трафика в нелицензируемых радиочастотах	132
3.1.1	Системная модель	132
3.1.2	Схемы приоритетного доступа	133
3.1.3	Модели математической теории телетрафика	135
3.2	Система с минимальной скоростью передачи	136
3.2.1	Мультипликативный вид стационарного распределения	136
3.2.2	Показатели эффективности передачи эластичного трафика	138
3.2.3	Анализ вероятности снижения скорости передачи	139
3.3	Модель нарезки радиоресурсов	143
3.3.1	Системная модель	143
3.3.2	Схемы приоритетного доступа	145
3.3.3	Модели математической теории телетрафика	147
3.4	Система с сигналами для перераспределения ресурса	149
3.4.1	Матрица интенсивностей переходов в блочно-трехдиагональном виде	149
3.4.2	Матричный алгоритм расчета стационарного распределения	152
3.4.3	О влиянии интенсивности потока сигналов на показатели эффективности	153
3.5	Управляемая система перераспределения ресурса	154
3.5.1	Марковский процесс принятия решения	154
3.5.2	Вероятность перераспределения ресурса по сигналу	158
3.5.3	О влиянии весов в функции вознаграждения на политику	160
3.6	Нахождение политики перераспределения ресурса	162
3.6.1	Система уравнений относительно вознаграждения	162
3.6.2	Итерационный алгоритм Ховарда	165
3.6.3	Применение искусственной нейронной сети	168

ГЛАВА 4	МОДЕЛИ С ЭЛАСТИЧНЫМ ТРАФИКОМ И ИЗМЕНЯЕМЫМ ОБЪЕМОМ РЕСУРСА	171
4.1	Модель передачи данных в миллиметровом диапазоне	171
4.1.1	Системная модель	171
4.1.2	Схемы приоритетного доступа	173
4.1.3	Модели математической теории телетрафика	175
4.2	Система с полным отказом ресурса	177
4.2.1	Матрица интенсивностей переходов в блочном виде.....	177
4.2.2	Рекуррентный алгоритм расчета стационарного распределения	179
4.2.3	Анализ вероятности блокировки.....	182
4.3	Система с марковским входящим потоком.....	184
4.3.1	Матрица интенсивностей переходов в блочном виде.....	184
4.3.2	Матричный алгоритм расчета стационарного распределения	188
4.3.3	Анализ вероятности блокировки.....	194
4.4	Анализ вложенного поглощающего случайного процесса.....	196
4.4.1	Построение поглощающего случайного процесса	196
4.4.2	Матрица интенсивностей переходов для пуассоновского потока.....	198
4.4.3	Матрица интенсивностей переходов для марковского потока	202
4.5	Алгоритм расчета моментов времени передачи трафика.....	205
4.5.1	Преобразование Лапласа-Стилтьеса времени передачи трафика.....	205
4.5.2	Матричный алгоритм расчета моментов времени.....	206
4.5.3	Анализ среднего времени передачи.....	207
4.6	Система с частичным отказом ресурса	210
4.6.1	Матрица интенсивностей переходов в блочном виде.....	210
4.6.2	Поглощающий случайный процесс	212
4.6.3	Моменты времени передачи трафика	217

ГЛАВА 5 МОДЕЛИ С ПОТОКОВЫМ ТРАФИКОМ И ИЗМЕНЯЕМЫМ ОБЪЕМОМ РЕСУРСА	220
5.1 Модель совместного использования радиочастот	220
5.1.1 Системная модель	220
5.1.2 Схемы приоритетного доступа	222
5.1.3 Модели математической теории телетрафика	227
5.2 Система с полным отказом ресурса	231
5.2.1 Матричный алгоритм расчета стационарного распределения	231
5.2.2 Алгоритм расчета среднего числа приостановленных сессий	234
5.2.3 Анализ среднего числа приостановленных сессий	239
5.3 Анализ периода занятости системы	240
5.3.1 Предварительные обозначения	240
5.3.2 Преобразование Лапласа-Стилтьеса времени первого достижения пустого состояния	242
5.3.3 Матричный алгоритм расчета преобразования Лапласа-Стилтьеса длительности периода занятости	247
5.4 Система с частичным отказом ресурсом	249
5.4.1 Матрица интенсивностей переходов в блочно-трехдиагональном виде	249
5.4.2 Матричный алгоритм расчета стационарного распределения	256
5.4.3 Анализ показателей эффективности	268
5.5 Управляемая система гистерезисного доступа к ненадежному ресурсу	271
5.5.1 Задача выбора порогов гистерезисного доступа как политики управления	271
5.5.2 Применение генетического алгоритма для нахождения политики ...	273
5.5.3 Анализ функции стоимости использования ненадежного ресурса ...	276
5.6 Ресурсная система с зависимым занятием ресурса	279
5.6.1 Модель требования к ресурсу	279
5.6.2 Вероятности занятия ресурса	282
5.6.3 Условные вероятности выделения ресурса	289
ЗАКЛЮЧЕНИЕ	295
СПИСОК ЛИТЕРАТУРЫ	297

ВВЕДЕНИЕ

Актуальность темы исследования.

Современные сети связи активно эволюционируют: технологии четвертого (4th generation, 4G) и пятого поколений (5G) постепенно дополняются инновационными решениями как в проводной, так и в беспроводной среде. Расширение спектра услуг – от потоковой передачи данных до интернета вещей – формирует новые требования к инфраструктуре: помимо увеличения пропускной способности, критически важными становятся снижение задержек и повышение надежности каналов связи. Эти требования стимулируют разработку принципиально новых архитектур, таких как сети шестого поколения (6G) или сеть 2030, где ключевым элементом является обеспечение качества обслуживания (quality of service, QoS) для разнородных приложений.

Ключевой характеристикой современных услуг является вариативность требований к скорости передачи данных: она может быть гарантированной, адаптивной в пределах максимального значения, или полностью динамической. Для анализа таких сценариев трафик принято разделять на два типа: потоковый и эластичный. В первом случае пользователь задает время предоставления услуги, а скорость передачи остается фиксированной или регулируется в рамках заданных порогов. Объем переданных данных прямо зависит от выбранной скорости. Во втором случае пользователь определяет объем данных, а скорость адаптируется под доступный ресурс сети. Время завершения услуги при этом зависит от текущей пропускной способности канала. Дополнительным критерием классификации служит режим передачи данных: одноадресный «точка – точка», когда данные доставляются одному получателю, и многоадресный «точка – много точек», когда осуществляется одновременная передача одних данных множеству пользователей.

Приоритизация услуг в сетях 5G – это механизм управления распределением ресурсов между разнородными сервисами и пользователями, основанный на классах QoS. В условиях перегрузки сети, когда доступный ресурс не покрывает

потребности всех сессий, система выделяет ресурсы в первую очередь для высокоприоритетных услуг. В штатных режимах работы низкой и нормальной нагрузки приоритеты используются для оптимизации QoS. Важным элементом такой системы являются приоритетные события – внешние или внутренние триггеры, которые изменяют объем ресурса, доступного для определенных классов трафика, и требования к QoS для текущих сессий.

Технология нарезки сети (network slicing), включая распределение радиоресурсов, формирует гибкую архитектуру для управления приоритетным доступом. Каждый изолированный сегмент сети обслуживает трафик определенного класса, однако ресурсы между сегментами могут динамически перераспределяться в зависимости от текущих условий. Контроллер, осуществляющий мониторинг и управление нарезкой, обладает высшим приоритетом при принятии решений.

Ограниченность доступного радиочастотного спектра стимулирует внедрение систем совместного доступа (spectrum sharing), которые делятся на два типа. В лицензируемых диапазонах два оператора делят выделенный частотный ресурс, а трафик владельца лицензии имеет безусловный приоритет. В нелицензируемых диапазонах, например, в частотах Wi-Fi, приоритетным считается трафик сети Wi-Fi, но его обслуживание зависит от алгоритмов, которые предотвращают коллизии с сотовым трафиком. При моделировании таких систем необходимо учитывать иерархию приоритетов: в первом случае гарантии QoS для владельца лицензии, во втором – баланс между справедливым распределением ресурса и соблюдением соглашения о качестве обслуживания.

Моделирование ресурсов в сетях 5G требует учета неоднородности беспроводной среды, которая напрямую влияет на доступную пропускную способность и качество связи. Пользователи вблизи базовой станции в зоне с высоким отношением сигнала к шуму получают максимальную скорость. Удаленные устройства сталкиваются с затуханием сигнала и снижением скорости. Прямая видимость обеспечивает стабильный канал связи, но блокировки пути прямой видимости вызывают потери.

Для моделирования и анализа беспроводных сетей используются методы теории массового обслуживания и математической теории телетрафика. Этот аппарат развивается и адаптируется под новые особенности управления приоритетным доступом и обслуживанием пользователей. В том числе, за счет использования моделей стохастической геометрии. Приоритетные системы массового обслуживания учитывают различные варианты приоритетов. Они позволяют исследовать мультисервисные сценарии с различными политиками доступа к обслуживающим приборам и очереди. В том числе, за счет вытеснения более приоритетными заявками менее приоритетных.

Для учета особенностей приоритизации услуг в беспроводных сетях 4G и 5G в моделях математической теории телетрафика должна учитываться возможность снижения скорости передачи трафика, структура ресурса – постоянный или изменяемый, влияние внешних событий, а также приоритеты для пар классов заявок. В результате возникает важная научная проблема создания комплекса марковских моделей для анализа и расчета показателей эффективности приоритетного доступа и обслуживания эластичного и потокового трафика в беспроводной сети, имеющая важное хозяйственное значение для развития теоретических основ информатики.

Степень разработанности темы исследования.

Значительный вклад в развитие данной тематики внесли следующие российские и зарубежные ученые и исследователи. В области методов анализа показателей эффективности беспроводных сетей: Андреев С.Д., Барабанова Е.А., Бегишев В.О., Вишневский В.М., Вытовтов К.А., Гайдамака Ю.В., Гольдштейн Б.С., Деарт В.Ю., Ибрагимов Б.Г., Киричек Р.В., Крук Е.А., Кулябов Д.С., Кучерявый А.Е., Кучерявый Е.А., Ляхов А.И., Маколкина М.А., Молчанов Д.А., Мутханна А.С.А., Орлов Ю.Н., Парамонов А.И., Пшеничников А.П., Росляков А.В., Самуйлов К.Е., Смелянский Р.Л., Соколов Н.А., Хоров Е.М., Цитович И.И., Яновский Г.Г., Akyildiz I., Araniti G., Caire G., Chatzinotas S., Ghrayeb A., Ksentini A., Popovski P., Taleb T., Verikoukis C., Whitt W. и др.

По моделям математической теории телетрафика с потоковым и эластичным трафиком: Башарин Г.П., Гайдамака Ю.В., Дудин А.Н., Зорин А.В., Рыков В.В., Самуйлов К.Е., Степанов С.Н., Федоткин А.М., Яшков С.Ф., Voxma O., Iversen V.B., Fiems D. Kelly F.P., Ross K.W. и др. По ресурсным системам массового обслуживания: Гайдамака Ю.В., Моисеев А.Н., Моисеева С.П., Наумов В.А., Самуйлов К.Е., Сопин Э.С., Тихоненко О.М. По управляемым системам массового обслуживания: Ефросинин Д.В., Зорин А.В., Рыков В.В., Федоткин М.А., Ховард Р.А.

По системам массового обслуживания с ненадежными приборами и случайной среде: Вишневский В.М., Дудин А.Н., Меликов А.З., Моисеева С.П., Назаров А.А., Нетес В.А., Пауль С.В., Рыков В.В., Goswami V., Jain M., Krishnamoorthy A., Panda G., Sztrik J. и др. По системам массового обслуживания с повторными вызовами и орбитами: Вишневский В.М., Ефросинин Д.В., Меликов А.З., Моисеев А.Н., Морозов Е.В., Назаров А.А., Пауль С.В., Румянцев А.С., Степанов С.Н., Phung-Duc T., Sztrik J., Wang J. и др.

Отметим и других ученых, которые внесли значительный вклад в развитие методов математической теории телетрафика и теории массового обслуживания, включая приоритетные системы: Бронштейн О.И., Гнеденко Б.В., Горцев А.М., Джейсуол Н., Димитров Б.Н., Духовный И.М., Зейфман А.И., Ивницкий В.А., Карташевский В.Г., Клименок В.И., Климов Г.П., Нейман В.И., Печинкин А.В., Пшеничников А.П., Разумчик Р.В., Сатин Я.А., Семенова О.В., Соколов Н.А., Терпугов А.Ф., Тюрликов А.М., Фархадов М.П., Харкевич А.Д., Хинчин А.Я., Хохлов Ю.С., Цициашвили Г.Ш., Шнепс М.А., Bruneel H., Chaudhry M., Pagano M., Walraevens J. и др.

Цель исследования состоит в создании теоретических основ и комплекса марковских моделей для анализа и расчета показателей эффективности приоритетного доступа эластичного и потокового трафика с применением механизмов прерывания и приостановки сессии, снижения скорости передачи и мощности сигнала.

Достижение сформулированной цели достигается путем решения следующих **задач исследования**:

1. Разработка методологии построения и анализа моделей обслуживания мультисервисного трафика в условиях приоритетного доступа.
2. Разработка моделей приоритетного обслуживания нескольких классов эластичного и потокового трафика в условиях снижения объема выделенных ресурсов, прерывания и приостановки обслуживания сессий.
3. Разработка моделей приоритетного обслуживания мультисервисного трафика в среде ненадежных ресурсов и политики перераспределения ресурсов методом сетевой нарезки.
4. Разработка алгоритмов для анализа и расчета показателей эффективности приоритетного доступа в беспроводных сетях, включая модель марковского процесса принятия решений.

Научная новизна результатов исследования состоит в следующем:

1. Комплекс моделей приоритетного доступа эластичного и потокового трафика в беспроводных сетях построен в единой концепции описания схем приоритетного доступа в виде матрицы приоритетов, позволяющей задавать приоритет для каждой пары классов заявок, возможность снижать скорость, учитывать структуру и разные типы ресурса, влияние на систему внешних событий. Ранее в классификации систем с приоритетным обслуживанием по большей части предполагался линейный относительно приоритета порядок классов заявок, акцент был на порядке доступа к ресурсу – относительный или абсолютный, его применении во времени, принципе выбора заявок для вытеснения.
2. Модель доступа узкополосного и широкополосного потокового трафика учитывает задержку на попытки начала и возобновления передачи широкополосного трафика в виде системы с орбитами, а также использует дискриминаторное разделение слота времени между сессиями пропорционально требованиям к скорости передачи трафика в ресурсной системе. Ранее в моделях для анализа сверхнадежной передачи с малой

задержкой и широкополосной связи использовались системы с очередями и равным разделением кадра времени. Модель доступа одноадресного и многоадресного потокового трафика реализует взаимный приоритет – многоадресный трафик прерывает одноадресный трафик, который в свою очередь, снижает скорость многоадресной сессии, а также учитывает два режима мультимедиа – завершение сессии по первому и последнему пользователям. Ранее исследовались системы без приоритета многоадресного трафика и в комбинации с одним из режимов мультимедиа.

3. Модель доступа эластичного трафика с минимальной скоростью передачи реализует приоритет посредством прерывания менее приоритетной сессии. Ранее приоритетное обслуживание нескольких классов эластичного трафика моделировалось с помощью дискриминаторного разделения ресурса между сессиями пропорционально весовым коэффициентам. Модель доступа эластичного трафика с управляемым по сигналам перераспределением объема ресурса учитывает выбор объема перераспределения ресурса при помощи марковского процесса принятия решений. Ранее в системах массового обслуживания для анализа нарезки радиоресурсов перераспределение происходило по внутренним, связанным с сессиями, событиям.
4. Модель доступа эластичного трафика с минимальной скоростью передачи и ненадежным ресурсом реализует приоритет посредством приостановки сессии и размещения ее в конечную очередь для ожидания возобновления обслуживания, а также учитывает марковский входящий поток. Ранее системы с ожиданием эластичного трафика исследовались для надежного ресурса и бесконечной очереди.
5. Для модели доступа потокового трафика к ненадежному ресурсу алгоритм расчета среднего числа приостановленных сессий записан в скалярном виде. Ранее расчет такой характеристики проводился матричным решением системы уравнений равновесия. Модель доступа потокового трафика к

частично ненадежному ресурсу реализует выбор направления нагрузки на ненадежный или надежный ресурс при помощи управляемой системы массового обслуживания. Ранее системы с частичным отказом ресурса исследовались с фиксированной политикой управления. Модель доступа потокового трафика со случайным требованием к объему ненадежного ресурса учитывает зависимое занятие двух типов ресурса – мощности сигнала и слота времени. Ранее ресурсные системы исследовались с надежными приборами и независимым занятием разных ресурсов.

Теоретическая значимость работы. Теоретическая значимость результатов работы обоснована тем, что доказаны теоремы, леммы и утверждения, позволяющие вычислять стационарные распределения вероятностей состояний моделей систем приоритетного доступа в беспроводных сетях, их показатели эффективности, а также осуществлять выбор политик управления работой этих систем; применительно к проблематике диссертации результативно использован комплекс методов математической теории телетрафика, теории массового обслуживания и дискретно-событийного статистического моделирования; изложены различные схемы приоритетного доступа эластичного и потокового трафика, типы ресурса, приоритетные события, механизмы реализации приоритета и т.п.; изучено влияние на эффективность приоритетного доступа повторных попыток возобновить обслуживание, многоадресного режима передачи данных, марковского потока запросов на передачу трафика, полного и частичного отказа ресурса; проведена модернизация существующей классификации моделей математической теории телетрафика и теории массового обслуживания в обозначениях Башарина – Кендалла, отражающая в виде матрицы приоритетов особенности приоритетного доступа в современных беспроводных сетях.

Практическая значимость работы. Значение полученных результатов работы для практики подтверждается тем, что разработаны алгоритмы расчета показателей эффективности приоритетного доступа в беспроводных сетях для сценариев сверхнадежной передачи данных с малой задержкой и критически важных данных, в миллиметровом диапазоне и нелицензируемом диапазоне

радиочастот, при использовании технологии нарезки сети и системы совместного использования радиочастот; созданы модели для определения оптимальных для заданных целевых функций политики управления перераспределением ресурса между сегментами сети и политики управления распределением нагрузки при совместном использовании радиочастот; представлены рекомендации по выбору параметра потока попыток возобновления обслуживания широкополосного трафика, уровней скорости широкополосного трафика и видеоконференции, параметра потока сигналов контроллера нарезки сети.

Методология и методы исследования. Для решения поставленных в работе задач исследования использовались методы математической теории телетрафика, теории массового обслуживания, марковских случайных процессов, марковских процессов принятия решения, стохастической геометрии, теории вероятностей, дискретно-событийного статистического моделирования.

Положения, выносимые на защиту.

1. Комплекс моделей схем приоритетного доступа эластичного и потокового трафика в виде марковских систем массового обслуживания и алгоритмы расчета характеристик случайных величин прерывания сессии, скорости передачи, объема используемого ресурса применимы для анализа сценариев сетей четвертого и пятого поколений с несколькими классами пользователей, ненадежными и совместно используемыми ресурсами.
2. Модель доступа узкополосного и широкополосного потокового трафика и матричный алгоритм расчета стационарного распределения позволяют рассчитать показатели эффективности сверхнадежной передачи с малой задержкой и широкополосной связи – вероятность прерывания и среднюю скорость сессии, а также позволяют настроить параметр потока попыток возобновления обслуживания и уровни скорости широкополосного трафика. Модель доступа одноадресного и многоадресного потокового трафика и рекуррентный алгоритм расчета стационарного распределения позволяют настроить уровни скорости видеоконференции для максимизации ее

среднего значения при ограничениях на вероятность блокировки и вероятность прерывания сессии.

3. Модель доступа эластичного трафика с минимальной скоростью передачи позволяет рассчитать показатели эффективности обслуживания трафика в нелицензируемых радиочастотах – вероятность прерывания сессии, вероятность снижения скорости ниже минимального порога и среднюю скорость. Модель доступа эластичного трафика с управляемым по сигналам перераспределением объема ресурса применима для определения оптимальной политики управления перераспределением ресурса между сегментами сети при нарезке радиоресурсов – объемов ресурса каждого из сегментов – по критерию максимума взвешенной суммы коэффициентов соответствия равному делению ресурса, результативности сигналов контроллера и использования ресурса, а также позволяет настроить параметр потока сигналов контроллера для максимизации взвешенной суммы коэффициентов.
4. Модель доступа эластичного трафика с минимальной скоростью передачи и ненадежным ресурсом, матричные алгоритмы расчета стационарного распределения и символьного вычисления преобразования Лапласа-Стилтьеса времени передачи трафика позволяют рассчитать показатели эффективности обслуживания критически важных данных в миллиметровом диапазоне – моменты времени передачи данных.
5. Модель доступа потокового трафика к ненадежному ресурсу, рекуррентный алгоритм в скалярном виде и матричные алгоритмы для символьного расчета периода занятости системы позволяют рассчитать показатели эффективности совместного использования радиочастот – моменты длительности периода занятости. Модель доступа потокового трафика к частично ненадежному ресурсу применима для определения оптимальной политики управления распределением нагрузки при совместном использовании радиочастот – объемом совместно используемого ресурса и порогами гистерезисного управления – по критерию минимума средней стоимости использования

ресурса и обслуживания сессий. Модель доступа потокового трафика со случайным требованием к объему ненадежного ресурса и рекуррентный алгоритм расчета стационарного распределения позволяют рассчитать показатели эффективности адаптивного управления мощностью сигнала – вероятность прерывания и среднюю скорости сессии.

Степень достоверности результатов работы. Достоверность полученных результатов подтверждается тем, что теория построена на известных методах теории массового обслуживания и математической теории телетрафика, используемых при доказательствах теорем, лемм и утверждений; идея базируется на анализе и обобщении передового опыта в обеспечении приоритетного доступа в беспроводных сетях; установлено качественное совпадение частных случаев разработанных автором моделей с известными моделями теории массового обслуживания и математической теории телетрафика; использован современный метод дискретно-событийного статистического моделирования для верификации результатов работы.

Апробация результатов работы.

По результатам работы автор выступил с пленарными докладами на международных конференциях 5th International Conference on Recent Trends in Information Technology, ICRTIT (2016: Ченнаи, Индия), 2nd International Conference on Future Networks and Distributed Systems, ICFNDS (2018: Амман, Иордания), 23rd International Conference named after A. F. Terpugov on Information Technologies and Mathematical Modelling, ITMM (2024: Карши, Узбекистан).

Основные положения работы были апробированы на следующих международных конференциях – International Conference on Analytical and Stochastic Modelling Techniques and Applications, ASMTA (2016: Кардифф, Великобритания, 2019: Москва), IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, European Conference on Modelling and Simulation, ECMS (2015: Албена, Болгария, 2017: Будапешт, Венгрия), IEEE Global Communications Conference, GLOBECOM (2016: Вашингтон, США), International Conference on Future Networks and Distributed Systems, ICFNDS (2018: Амман,

Иордания), International Conference on Numerical Analysis and Applied Mathematics, ICNAAM (2014, 2018: Родос, Греция), International Congress on Ultra Modern Telecommunications and Control Systems and Workshops, ICUMT (2014: Санкт-Петербург, 2015: Брно, Чехия, 2016: Лиссабон, Португалия, 2017: Мюнхен, Германия, 2018: Москва), International Symposium on Computer and Information Sciences, ISCIS (2016: Краков, Польша), ITU Kaleidoscope Academic Conference, K (2014: Санкт-Петербург), International Conference on Next Generation Teletraffic and Wired/Wireless Advanced Networks and Systems, NEW2AN (2012, 2013, 2015-2020: Санкт-Петербург).

Положения работы обсуждались также на следующих научных мероприятиях – Международной конференции «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь», DCCN (2016, 2019, 2022, 2024: Москва, ИПУ РАН, РУДН), Международной конференции имени А.Ф. Терпугова «Информационные технологии и математическое моделирование», ИТММ (2016: Катунь, 2017: Казань, 2019: Саратов, 2023: Томск, ТГУ), Международной конференции «Современные сетевые технологии», MoNeTec (2014: Москва, ВМК МГУ), Международной отраслевой научно-технической конференции «Технологии информационного общества» (2012-2017: Москва, МТУСИ), Международной научно-практической конференции «Современные информационные технологии и ИТ-образование» (2013, 2014, 2016, 2017, 2019: Москва, ВМК МГУ), Всероссийском совещании по проблемам управления, ВСПУ (2014, 2019, 2024: Москва, ИПУ РАН), Международной научной конференции «Математическое и программное обеспечение информационных, технических и экономических систем» (2018-2020: Томск, ТГУ), Молодежной научной конференции «Задачи современной информатики» (2015: Москва, ФИЦ ИУ РАН), Всероссийском симпозиуме по прикладной и промышленной математике, ВСППМ (2018, 2019: Сочи, ФИЦ ИУ РАН), Международном семинаре по прикладным проблемам в теории вероятностей и математической статистике, АРТР+MS (2017: Реджо-ди-Калабрия, Италия, 2018: Лиссабон, Португалия, ФИЦ ИУ РАН, РУДН), Всероссийской конференции с международным участием «Информационно-

телекоммуникационные технологии и математическое моделирование высокотехнологичных систем», ИТТММ (2014-2024: Москва, РУДН), а также на научном межвузовском семинаре «Современные телекоммуникации и математическая теория телетрафика» РУДН, МТУСИ, ИППИ РАН, ТГУ, ИПМ РАН (2021, 2024), на научном семинаре кафедры автоматизации систем вычислительных комплексов ВМК МГУ в рамках консорциума «Сетевые и облачные технологии» (2025).

За часть научных результатов работы автор удостоен премии Правительства Москвы молодым ученым в номинации «Информационно-коммуникационные технологии» за разработку комплекса вероятностных моделей схем приоритетного управления радиоресурсами беспроводных сетей последующих поколений (2016).

Реализация результатов работы.

Научные результаты работы использованы в следующих 6 научно-исследовательских работах (НИР), выполненных под непосредственным руководством автора:

- Разработка моделей и алгоритмов нарезки радиоресурсов и приоритетного доступа в беспроводной сети 6G, рук. Кочеткова И.А. // проект системы грантовой поддержки научных проектов РУДН № 025319-2-000 (2023-2024);
- Исследование и разработка моделей и интеллектуальных алгоритмов совместного обслуживания трафика с малыми задержками и широкополосного доступа в беспроводных сетях пятого поколения, рук. Гудкова И.А. // грант РФФИ стабильность № 20-37-70079 (2020-2021);
- Вероятностные модели сегментации радиоресурсов беспроводных сетей и методы расчета характеристик обслуживания пользователей, рук. Гудкова И.А. // грант совета по грантам Президента РФ № МК-2588.2019.9 (2019-2020);
- Разработка информационной технологии, математического, алгоритмического и программного обеспечения для моделирования взаимодействия устройств в беспроводных сетях пятого поколения,

рук. Гудкова И.А. // стипендия совета по грантам Президента РФ № СП-2987.2016.5 (2016-2018);

- Исследование и разработка информационно-телекоммуникационной технологии и моделей беспроводных межмашинных взаимодействий для управления городской инфраструктурой, рук. Гудкова И.А. // грант РФФИ мол_а_мос № 15-37-70016 (2016-2017);
- Разработка комплекса марковских моделей для анализа показателей эффективности схем доступа в беспроводных мультисервисных сетях с приоритетным обслуживанием, рук. Гудкова И.А. // грант РФФИ мол_а_мос № 16-37-00421 (2016-2017).

Научные результаты работы использованы также в НИР по другим грантам РФФИ и РНФ, в рамках государственного задания Минобрнауки РФ, выполненных при участии автора.

Научные результаты работы использованы в учебном процессе в РУДН на факультете физико-математических и естественных наук в разработанных и читаемых автором студентам 3 курса бакалавриата направлений подготовки 02.03.01 «Математика и компьютерные науки», 02.03.02 «Фундаментальная информатика и информационные технологии» и 09.03.03 «Прикладная информатика» дисциплинах «Модели мультисервисных сетей с приоритетами» и «Анализ приоритетного доступа в мультисервисных сетях».

Публикации. Основные научные результаты диссертации изложены в 55 опубликованных работах, из них 31 работа опубликована в изданиях, индексируемых в международных базах индексации и цитирования (МБЦ) Web of Science и Scopus, 12 работ в рецензируемых научных изданиях, индексируемых в наукометрической базе данных RSCI, 1 работа в рецензируемом научном издании, рекомендованном перечнем Высшей аттестационной комиссии при Минобрнауки РФ (Перечень ВАК РФ). По теме работы автор также имеет 7 зарегистрированных свидетельств на программу для электронных вычислительных машин, 2 монографии.

Соответствие паспорту специальности. Работа соответствует следующим пунктам паспорта научной специальности 1.2.3 «Теоретическая информатика, кибернетика»:

- п. 9 «Математическая теория исследования операций» в части комплекса моделей приоритетного доступа эластичного и потокового трафика в беспроводных сетях (результат 1), исследования политики управления перераспределением ресурса при нарезке радиоресурсов (результат 3) и политики управления распределением нагрузки между ненадежным и надежным ресурсами при совместном использовании радиочастот (результат 5);
- п. 11 «Распределенные многопользовательские системы» в части моделирования системы доступа виртуальных операторов к распределенному ресурсу базового оператора при нарезке радиоресурсов (результат 3) и моделирования системы доступа пользователей мобильного оператора к распределенному ресурсу при совместном использовании радиочастот (результат 5);
- п. 12 «Модели информационных процессов и структур» в части моделирования процесса предоставления услуг видео по требованию и видеоконференции (результат 2) и процесса передачи трафика в нелицензируемых радиочастотах (результат 3);
- п. 23 «Новые интернет - технологии, включая средства поиска, анализа и фильтрации информации» в части моделирования перераспределения ресурса для технологии нарезки радиоресурсов (результат 3) и моделирования процесса передачи данных в миллиметровом диапазоне (результат 4);
- п. 25 «Методы высоконадежной обработки информации и обеспечения помехоустойчивости информационных коммуникаций для целей передачи, хранения и защиты информации» в части моделирования сверхнадежной передачи информации с малой задержкой (результат 2).

Личный вклад автора. Автор лично получил основные результаты диссертации. Автор лично участвовал во всех этапах исследования, выполнил апробацию результатов исследования, разработал сценарии для проведения численных экспериментов. При непосредственном участии автора обработаны результаты численных экспериментов, выполнена их интерпретация, подготовлены основные публикации по работе.

Объем и структура работы. Текст диссертации включает в себя введение, основную часть из пяти глав и заключение. Каждая глава разделена на 6 разделов, а раздел – на 3 параграфа. Диссертация включает в себя список литературы из 400 библиографических ссылок. Работа изложена на 340 страницах текста, содержит 50 рисунков и 13 таблиц.

Глава 1 ПРИОРИТЕТНОЕ ОБСЛУЖИВАНИЕ ТРАФИКА В МУЛЬТИСЕРВИСНЫХ БЕСПРОВОДНЫХ СЕТЯХ

1.1 ПРИОРИТИЗАЦИЯ УСЛУГ

1.1.1 Сети новых поколений

Сети связи развиваются стремительно и динамично – четвертое поколение (4th generation, 4G) и пятое поколение (5G) – появляются новые проводные и беспроводные технологии, благодаря которым скорости передачи данных измеряются уже не в Мбит/с и Гбит/с, а стремятся к значениям в Тбит/с [1–14]. Приложения требуют не только все более высокой пропускной способности, но и низких задержек передачи данных и высокого уровня надежности. Ожидается, что число подключенных устройств вырастет с 28,5 миллиарда (2020 год) до 100 миллиардов в 2025 году, а трафик увеличится с 33 Збайт (2018 год) до 175 Збайт к 2025 году. Спектр предоставляемых пользователям услуг расширяется, растет уровень использования методов искусственного интеллекта (ИИ, artificial intelligence, AI) и машинного обучения (МО, machine learning, ML). Все эти тенденции приводят к необходимости разработки нового поколения сетей, которые бы соответствовали будущим услугам и приложениям – сетей шестого поколения (6G) или, в терминологии Международного союза электросвязи (МСЭ, International Telecommunication Union, ITU), сетей 2030 (Network 2030) [15–26].

Сеть 6G¹ – это система, включающая в себя различные компоненты и аспекты, которые позволяют организовать интегрированную, высокоавтоматизированную и основанную на интеллектуальных методах разделения инфраструктуру. Она объединяет разнородные коммуникации,

¹ Определение дано в спецификации МСЭ TS «Network 2030 Architecture Framework» в 2020 году.

вычислительные ресурсы, ресурсы для хранения данных и ресурсы для развертывания сетевых сервисов и приложений по всем сегментам сети: проводная и беспроводная связь, ядро сети, сеть граничных вычислений и спутниковая связь, а также их комбинации [27–30]. Пользователь может получить доступ к этой сети из одного или нескольких местоположений.

Ключевой аспект в построении сети 6G – обеспечение требуемого качества обслуживания (quality of service, QoS) пользователей [31–37], включая сверхнизкую задержку передачи данных, сверхбольшой объем данных, энергоэффективность и т.д. [38,39]. Причем по большей части данные требования должны быть выполнены для каждого «деления» «нарезанной» инфраструктуры. Концепция такого деления – нарезка сети (network slicing) – охватывает все сетевые и облачные сегменты: сеть радиодоступа, ядро сети, транспортную сеть, сеть граничных вычислений (edge computing) – и позволяет одновременно развертывать несколько логических, автономных и независимых разделенных сетевых ресурсов и группы сетевых и сервисных функций на единой инфраструктурной платформе [40–42].

Новые сценарии использования сетей 6G, такие, как голографическая коммуникация, тактильный интернет для удаленных операций, цифровые двойники и другие, выдвигают высокие требования к скорости передачи данных, задержке, безопасности и поддержке алгоритмов на базе ИИ. Кроме того, сети должны быть гетерогенными, чтобы обеспечить необходимое качество обслуживания. В связи с этим, существующие технологические и системные решения сетей предыдущих поколений не могут полностью удовлетворить этим требованиям. В настоящее время международные организации по стандартизации, такие как МСЭ и 3GPP (3rd Generation Partnership Project), а также мировое научно-исследовательское сообщество, включая исследовательские центры компаний, детально прорабатывают архитектуру будущих сетей, их сценарии использования, требования, характеристики и технологические решения. Эти усилия направлены на создание сетей 6G, которые будут способны обеспечить высокое качество обслуживания для всех новых сценариев использования.

Важной частью будущей архитектуры является концепция граничных вычислений, а именно мобильных граничных вычислений (multi-access edge computing, MEC) [43]. MEC представляет собой систему, предоставляющую среду информационно-телекоммуникационных сервисов и возможностей облачных вычислений (cloud computing) на границе сети доступа, которая поддерживает одну или несколько технологий доступа и расположена в непосредственной близости к пользователям. Понятие «границы» включает в себя не только границу сети радиодоступа, но и все устройства, начиная от конечного устройства (например, датчики, смартфоны и т.д.) до локальных центров обработки данных провайдеров.

Автор с коллегами провел исследование в области облачных вычислений и рассмотрел две задачи: распределение нагрузки между серверами и миграция виртуальных машин между серверами. В рамках задачи распределения нагрузки [44,45], была разработана управляемая система массового обслуживания (СМО) с двумя группами приборов разной интенсивности обслуживания. В систему поступают два потока заявок, которые при занятых приборах помещаются в очереди. Функционирование системы описывается Марковским процессом принятия решений (МППР, Markov decision process, MDP). Управляющее действие заключается в выборе очереди, когда освобождается один из приборов, при условии, что очереди не пусты. Действие зависит от группы приборов, в которой освободилось место, а также от класса заявки, которая этот прибор покинула. Функция вознаграждения за нахождение в состоянии прямо пропорциональна числу заявок на каждой из групп приборов с учетом класса заявок и заявок в очередях.

В рамках задачи миграции виртуальных машин [46,47], разработана СМО с классами заявок, которые перемещаются между группами приборов. Система состоит из групп приборов, которые описывают сервера. На этих серверах развернуты виртуальные машины, на которых выполняются задачи от пользователей – заявки разных классов. В зависимости от занятости групп приборов, классы заявок могут перемещаться между группами, но только целиком весь класс. Миграция может произойти только в момент поступления новой заявки,

при условии, что для нее окажется недостаточно места в текущей группе. Новая группа выбирается по принципу минимизации свободного места в новой группе приборов при условии перемещения на нее класса заявок.

Отметим, что помимо исследований, связанных непосредственно с процессом передачи данных пользователей, имеет значение и обслуживание сигнальной нагрузки. Использование радиоканала случайного доступа (random access channel, RACH) – это механизм, который позволяет пользовательским устройствам инициировать подключение к сети в первый раз или после периода неактивности, установить соединение для начала передачи данных [48–50]. Радиоканал RACH поддерживает, в том числе массовые подключения для межмашинного взаимодействия (massive machine-type communications, mMTC).

Автор с коллегами провел исследование процедуры установления соединения по радиоканалу RACH [51–53]. Построена цепь Маркова с поглощающим состоянием для расчета среднего времени установления соединения после успешного обмена четырьмя сообщениями между устройством и базовой станцией. Начальное состояние соответствует начальному моменту процедуры, а поглощающее – моменту, когда соединение установлено. Промежуточные состояния описывают вектора с компонентами по числу переданных сообщений процедуры RACH – успешных и неуспешных, состояния этих сообщений.

1.1.2 Классы услуг

За последние 50 лет развития стандартизации в телекоммуникационных сетях, вопрос обеспечения качества обслуживания QoS претерпел значительные изменения. Качество и надежность связи постоянно повышаются. Вначале появилась возможность настройки качества негарантированной доставки «best effort» для разных пользователей и сценариев использования. Затем стало доступно предоставление гарантии качества обслуживания – приоритета – для пользователей, готовых за это заплатить – дифференцированного QoS (differentiated QoS, D-QoS).

Приоритет в передаче данных реализуется с помощью различных методов обработки как на уровне соединений, так и на уровне пакетов. На уровне соединений это может быть использование выделенных линий, виртуальных соединений для асинхронной передачи пакетов (asynchronous transfer mode, ATM) и ретрансляции кадров (frame relay), многопротокольной коммутации по меткам (multiprotocol label switching, MPLS), интегрированных услуг (integrated services, IntServ), протокола резервирования сетевых ресурсов (resource reservation protocol, RSVP). На уровне пакетов приоритет реализуется с помощью заголовка пакета, а именно его поля с типом обслуживания (type of service, ToS). Это может быть дифференцированное обслуживание (differentiated service, DiffServ), идентификаторы классов качества обслуживания для сетей 4G (QoS class identifier, QCI) и 5G (5G QoS identifier, 5QI).

Отметим, что концепция нарезки сети 5G позволяет создавать логические сети поверх одной физической инфраструктуры, а затем настраивать их, изолируя в соответствии с требованиями каждого сегмента сети под разные сценарии использования, арендаторов и виртуальных операторов. Таким образом, с бизнес-точки зрения, концепция нарезки сети является способом организации приоритета для разных клиентов, что похоже на классические механизмы D-QoS.

Приоритизация услуг для сетей 5G определяется в спецификации 3GPP TS 23.501 и фактически задается классом 5QI и приоритетом занятия и удержания ресурса (allocation and retention priority, ARP). Класс 5QI характеризуется семью параметрами: (1) идентификатор 5QI; (2) тип ресурса (resource type); (3) уровень приоритета (priority level); (4) задержка передачи пакета (packet delay budget, PDB); (5) частота ошибок в пакетах (packet error rate, PER); (6) окно усреднения (averaging window); (7) максимальный объем пакета данных (maximum data burst volume, MDBV).

Тип ресурса для услуги определяет, предоставляется ли гарантированная скорость передачи данных (guaranteed bit rate, GBR) или нет (non-GBR). Если да, то ресурс подразделяется на просто GBR и критичный к задержке GBR (delay-critical GBR). Они отличаются разными требованиями к задержке передачи пакета и

частоте ошибок в пакетах, а также максимальный объем пакета данных задается только для delay-critical GBR. Пример – услуги сверхнадежной передачи данных с малой задержкой (ultra-reliable low-latency communications, URLLC) [54–71] и передача критически важных данных [72–77].

Услуги типа GBR предоставляются с гарантированной скоростью (guaranteed flow bit rate, GFBR) и до максимальной скорости (maximum flow bit rate, MFBR), при этом скорость в интервале от GFBR до MFBR устанавливается в соответствии с относительным уровнем приоритета. Для услуг типа non-GBR есть два порога на максимальную скорость: на одну сессию пользователя (session aggregate maximum bit rate, Session-AMBR) и на все сессии пользователя (user equipment AMBR, UE-AMBR). Кроме того, есть порог на максимальную скорость для пользователя по всем услугам GBR и non-GBR, которые предоставляются в сегменте сети (UE per slice-maximum bit rate, UE-Slice-MBR).

Уровень приоритета отражает приоритет занятия ресурса трафиком различных услуг как от одного, так и от разных пользователей. Чем ниже его значение, тем выше приоритет. В случае высокой загрузки сети, когда ресурса недостаточно для обеспечения качества обслуживания для всех пользователей, пользователи с более высоким приоритетом получают больше ресурса, чем пользователи с более низким приоритетом. При низкой и нормальной загрузке сети уровень приоритета используется для распределения ресурса между пользователями. Однако помимо уровня приоритета, планировщик занятия ресурса может также учитывать другие факторы, такие как тип ресурса или условия радиоканала, чтобы обеспечить наилучшее предоставление услуг и пропускную способность сети в целом.

Приоритет занятия и удержания ресурса ARP содержит три параметра: уже упомянутый уровень приоритета (priority level), идентификатор занятия ресурса (pre-emption capability) и идентификатор удержания ресурса (pre-emption vulnerability). Параметр ARP определяет, будет ли принят или отклонен запрос на занятие или увеличение объема уже занятого ресурса при отсутствии свободного ресурса для этой операции. Обычно ARP используется для управления доступом к

услугам типа GBR, а также для выбора потока данных, у которого будет уменьшен занятый им ресурс, в условиях высокой загрузки сети.

Идентификатор занятия ресурса определяет, сможет ли пользователь получить доступ к ресурсам, которые уже используются услугой с более низким приоритетом. Идентификатор удержания ресурса определяет, сможет ли пользователь потерять выделенные ему ресурсы, чтобы оператор мог принять запрос на услугу с более высоким приоритетом.

Приоритизация услуг в соответствии со спецификацией 3GPP TS 23.501 (V18.5.0) представлена в таблице 1.1. Для краткости, в таблице приведены только идентификатор 5QI, приоритет по умолчанию и некоторые примеры услуг.

Таблица 1.1 – Классы услуг согласно спецификации 3GPP TS 23.501 (V18.5.0)

Идент. 5QI	Приоритет	Примеры услуг
Услуги с гарантированной скоростью (GBR)		
1	20	Голосовое общение в реальном времени
2	40	Видео общение в реальном времени
3	30	Игры в реальном времени, приложения V2X (vehicle-to-everything), распределение электроэнергии (среднее напряжение), мониторинг автоматизации процессов
4	50	Видео с буферизацией (не общение)
65	7	Критически важные голосовые данные пользователя по одному нажатию кнопки
66	20	Не критически важные голосовые данные пользователя по одному нажатию кнопки
67	15	Критически важные видео данные пользователя
75	25	Приложения V2X, приложения A2X (aircraft-to-everything)
71-74, 76	56	Данные в реальном времени по нисходящему каналу

Услуги без гарантированной скорости (Non-GBR)		
6	60	Видео с буферизацией, приложения TCP (transmission control protocol) (интернет, почта, чат, обмен файлами по протоколам FTP (file transfer protocol) и P2P (peer-to-peer), прогрессивное видео), загрузка модели AI/ML для распознавания изображений (топология модели)
7	70	Видео и аудио в реальном времени, интерактивные игры, загрузка модели AI/ML для распознавания изображений (весовые коэффициенты модели)
8-10	80, 90	Видео с буферизацией, приложения TCP, услуги по спутниковому каналу со схожими характеристиками
69	5	Сигнальные сообщения при передаче критически важных данных
70	55	Критически важные данные для услуг 5QI 6, 8, 9
79	65	Приложения V2X
80	68	Дополненная реальность с широкополосной связью с низкой задержкой
Услуги с гарантированной скоростью и критичные к задержке (delay-critical GBR)		
82	19	Дискретная автоматизация
83	22	Дискретная автоматизация, приложения V2X (низкий уровень автоматизации: колонны транспортных средств, усовершенствованное вождение – изменение полосы движения)
84	24	Интеллектуальные транспортные системы
85	21	Распределение электроэнергии (высокое напряжение), приложения V2X (дистанционное вождение), разделение выходных данных модели AI/ML (передача данных в восходящем канале по распознаванию изображений)
86	18	Приложения V2X (высокий уровень автоматизации и усовершенствованное вождение – предотвращение столкновений, колонны транспортных средств)
87, 88	25	Интерактивные услуги (отслеживание движения), разделение выходных данных модели AI/ML (передача данных в нисходящем канале по распознаванию изображений)
89, 90	25	Визуальный контент для облачного и граничного рендеринга, для рендеринга по разделенным данным

Объемы трафика для различных услуг могут значительно варьироваться [78,79]. В таблице 1.2 показан пример структуры трафика по классам услуг мобильного оператора Vodafone Portugal в 2018 году. Автор с коллегами провел исследование этого трафика по данным, которыми поделился профессор Университета Лиссабона Luis M. Correia [80]. Большую его часть составляют веб-приложения, приложения для обмена мгновенными сообщениями и игры.

Таблица 1.2 – Структура трафика услуг мобильного оператора

Услуга	Доля	Услуга	Доля
Веб-приложения	23,24%	Безопасность	4,02%
Другие приложения	15,58%	Потоковая передача музыки	1,90%
Приложения для обмена мгновенными сообщениями	13,36%	Сетевые службы	1,53%
Игры	12,54%	Одноранговые приложения	0,66%
Передача файлов	10,27%	Мобильные терминалы	0,29%
Электронная почта	6,32%	Файловые системы	0,02%
Потоковые приложения	5,84%	Транзакции баз данных	0,01%
IP-телефония	4,40%	Устаревшие протоколы	< 0,01%

1.1.3 Типы трафика

Таким образом, услуги можно разделить на два класса: с гарантированной скоростью передачи данных (GBR) и без нее (Non-GBR). Еще один способ классификации услуг основан на ключевом параметре, который определяет пользователь при предоставлении услуги и который влияет на генерируемый трафик. Разница между этими двумя видами услуг заметна, прежде всего, по времени их предоставления. В одном случае, например, при телефонных разговорах, время регулируется пользователем. В другом случае, например, при передаче файлов, время может меняться в зависимости от загрузки сети.

В этом контексте выделяют два типа трафика: потоковый (streaming) и эластичный (elastic). Потоковый трафик возникает при предоставлении услуги, когда пользователь сам определяет время ее использования. Скорость передачи данных при этом может быть постоянной или меняться в соответствии с определенными порогами. Объем передаваемых данных зависит от установленной скорости. Эластичный трафик связан с предоставлением услуги, когда пользователь определяет объем данных для передачи. В этом случае скорость передачи данных меняется динамически в зависимости от количества пользователей. Время предоставления услуги, в свою очередь, зависит от установленной скорости передачи данных.

Существует еще один подход к классификации услуг и типов трафика – по режимам передачи данных. Это деление на два вида: «точка – точка» (point-to-point) и «точка – много точек» (point-to-multipoint). В первом случае трафик называется одноадресным (unicast), а во втором – многоадресным (multicast), или передаваемым в режиме мультивещания [81–93]. Один из вариантов мультивещания – это широковещание (broadcast). В беспроводных сетях мультивещание и широковещание реализуются с помощью системы eMBMS (evolved multimedia broadcast/multicast service) / FeMBMS (further evolved MBMS) в сетях 4G и системы MBS (multicast and broadcast service) в сетях 5G.

Автор с коллегами провел исследование в области мультивещания и рассмотрел две задачи: анализ процедуры восстановления переданных файлов, если они были доставлены ненадежно, и формирование группы мультивещания. В рамках задачи анализа процедуры восстановления файлов [94–96] использовалась СМО с групповым обслуживанием и случайным временем формирования группы. Эта система состоит из одного прибора, который описывает передачу пользователям поврежденных файлов в режиме мультивещания. Группа заявок формируется, когда контроллер опрашивает всех пользователей, чтобы узнать, повреждены ли полученные ими файлы.

В задаче формирования группы мультивещания [97–101] исследовалось, как мобильность пользователей в соте влияет на структуру этих групп. Мобильных пользователей можно описать как сеть массового обслуживания (СеМО), которая закрытая, если количество активных пользователей постоянно, или открытая, если пользователи могут отключаться и становиться «пассивными». Алгоритм направлен на то, чтобы обеспечить справедливую скорость передачи данных в разных группах, например, с помощью индекса Джейна (Jain's fairness index).

1.2 МОДЕЛИ ТРАФИКА В БЕСПРОВОДНЫХ СЕТЯХ

1.2.1 Модели с одним типом трафика

Независимо от того, какие услуги имеют приоритет в мультисервисной сети, у разных видов трафика – потокового, эластичного, одноадресного и многоадресного – есть свои особенности обслуживания. При анализе такой сети с помощью моделей математической теории телетрафика [102–125] и теории массового обслуживания [126–154] эти особенности учитываются в дисциплинах обслуживания. Для одноадресного потокового трафика используется дисциплина «первый пришел – первый обслужился» (first come first served, FCFS), для многоадресного потокового трафика – обслуживание «прозрачных» заявок (transparent, Tr), а для одноадресного эластичного трафика – разделение процессора (processor sharing, PS). Отметим, что многоадресный режим передачи эластичного трафика возможен, но встречается реже. Для краткости будем говорить об одноадресном трафике и многоадресном трафике, не используя слово «потоковый», а также об эластичном трафике, не используя слово «одноадресный».

Для некоторых моделей мультисервисных сетей, где используется сочетание одноадресного, многоадресного и эластичного трафика, стационарное распределение вероятностей состояний может быть представлено в

мультипликативном виде (product-form solution) $\pi(x_1, \dots, x_n) = \text{const} \cdot \prod_{i=1}^n \pi(x_i)$.

Часто это справедливо для полноступных (complete sharing, CS) моделей с однородным трафиком – либо только потоковым, либо только эластичным. Однако существуют и исключения из этого правила.

Модель с одноадресным трафиком – это мультисервисная модель Эрланга со стационарным распределением в мультипликативном виде

$$\pi(n_1, \dots, n_K) = G^{-1}(\mathcal{X}) \cdot \prod_{k=1}^K \left(\frac{\lambda_k}{\mu_k} \right)^{n_k} \frac{1}{n_k!}, \quad (n_1, \dots, n_K) \in \mathcal{X}, \quad (1.1)$$

где K число классов трафика, λ_k интенсивность пуассоновского потока запросов на передачу трафика, μ_k^{-1} среднее экспоненциальное время передачи трафика, n_k число обслуживаемых сессий или число пользователей, $\mathcal{X}$ пространство состояний случайного процесса (СП), а $G^{-1}(\mathcal{X})$ нормировочная константа.

Модели с многоадресным трафиком рассматриваются с разными дисциплинами обслуживания «прозрачных» заявок [155–158]. Дисциплина П1 (Tr1) освобождает ресурс после завершения обслуживания пользователя, инициировавшего сессию, например, это видеоконференция. Дисциплина П2 (Tr2) освобождает ресурс после завершения обслуживания последнего из присутствующих на сессии пользователей, например, телевидение. Стационарное распределение имеет мультипликативный вид

$$\pi(n_1, \dots, n_K) = G^{-1}(\mathcal{X}) \cdot \prod_{k=1}^K \left(\frac{\lambda_k}{\alpha_k} \right)^{n_k}, \quad (n_1, \dots, n_K) \in \mathcal{X}, \quad (1.2)$$

$$\alpha_k = \mu_k, \quad \text{дисциплина П1,}$$

$$\alpha_k = \frac{\lambda_k}{\exp(\lambda_k / \mu_k) - 1}, \quad \text{дисциплина П2,}$$

где n_k также число обслуживаемых сессий, но фактически $n_k = 1$, когда есть хотя бы один пользователь, либо $n_k = 0$, когда пользователей нет.

В случае совместного обслуживания одноадресного и многоадресного трафика стационарное распределение также будет мультипликативного вида [159,160]. Это происходит из-за обратимости описывающей модель марковского СП

$$\pi(n_1^u, \dots, n_{K_u}^u, n_1^m, \dots, n_{K_m}^m) = G^{-1}(\mathcal{X}) \cdot \prod_{k=1}^{K_u} \left(\frac{\lambda_k^u}{\mu_k^u} \right)^{n_k^u} \frac{1}{n_k^u!} \cdot \prod_{k=1}^{K_m} (\alpha_k^m)^{n_k^m}, \quad (1.3)$$

$$(n_1^u, \dots, n_{K_u}^u, n_1^m, \dots, n_{K_m}^m) \in \mathcal{X},$$

где индексы u и m обозначают одноадресный и многоадресный трафик соответственно. Автор участвовал в исследовании этой модели [161].

Вид стационарного распределения для модели с эластичным трафиком зависит от того, как ресурс делится между сессиями [162–177]. Для эгалитарного разделения процессора (egalitarian PS, EPS), если ресурс делится поровну между сессиями, то СП является обратимым. В кандидатской диссертации [178] автор исследовал такую систему, где требуется минимальная скорость передачи трафика, стационарное распределение имеет мультипликативный вид

$$\pi(n_1, \dots, n_K) = G^{-1}(\mathcal{X}) \cdot (n_1 + \dots + n_K)! \prod_{k=1}^K \left(\frac{\lambda_k}{V \mu_k} \right)^{n_k} \frac{1}{n_k!}, \quad (n_1, \dots, n_K) \in \mathcal{X}, \quad (1.4)$$

где $1/\mu_k$ – средний экспоненциальный объем трафика.

Для дискриминаторного разделения процессора (discriminatory PS, DPS), если ресурс делится между сессиями разных классов пропорционально $\omega_i / \sum_i \omega_i n_i$ весовым коэффициентам ω_i , то СП уже не будет обратимым. Автор совместно с коллегами исследовал такую систему с двумя классами эластичного трафика, где требуется минимальная и есть пороги на максимальную скорость передачи данных [178–181]. Эти пороги выступают в роли весовых коэффициентов. Были предложены приближенные формулы для расчета стационарного распределения.

1.2.2 Модели с потоковым и эластичным трафиком

В отличие от моделей с однородным трафиком, где для полноступных базовых схем обслуживания СП является обратимым, полноступные модели с разнородным трафиком – потоковым и эластичным – уже не имеют стационарного распределения в мультипликативном виде [182,183]. Для расчета стационарного распределения в таких моделях, помимо матричного представления, в литературе предлагаются различные приближенные формулы. Аналогичный подход был использован автором в кандидатской диссертации при исследовании модели с одноадресным и эластичным трафиком [178]. Один вариант приближенных формул основан на мультипликативном виде выражения $\pi(n_u, n_e) \approx \text{const} \cdot \pi(n_u) \cdot \pi(n_u)$, а второй – на использовании условного стационарного распределения $\pi(n_u, n_e) \approx \pi(n_e | n_u) \cdot \pi(n_u)$ и $\pi(n_u, n_e) \approx \pi(n_u | n_e) \cdot \pi(n_e)$.

Однако для некоторых неполноступных моделей СП может быть обратимым. Автор участвовал в исследовании такой модели, где на отдельных интервалах времени схема обслуживания фактически является схемой с полным разделением (complete partitioning, CP) [184–186]. Изменение деления ресурса между одноадресным и эластичным трафиком происходит в зависимости от числа сессий эластичного трафика, но не непрерывно, а по пороговым значениям. Такая система применяется для анализа совместного обслуживания одноадресного трафика, генерируемого пользователями (human-to-human, H2H), и эластичного трафика, создаваемого устройствами межмашинного взаимодействия M2M.

Для одноадресного трафика задана постоянная скорость b_u передачи, а для эластичного – требование b_e к минимальной скорости. В общем объеме V ресурса зарезервирован объем V_u для одноадресного трафика. Эластичный трафик занимает ресурс, кратный объему V_e , на котором может одновременно обслуживаться максимум $N_e = \lfloor V_e / b_e \rfloor$ сессий. Таким образом, пространство состояний СП имеет вид

$$\mathcal{X} = \left\{ (n_u, n_e) : n_u \geq 0, n_e \geq 0, n_u b_u \leq V - \left\lceil \frac{n_e}{N_e} \right\rceil V_e, n_e b_e \leq V - V_u \right\}, \quad (1.5)$$

а стационарное распределение – мультипликативное

$$\pi(n_u, n_e) = G^{-1}(\mathcal{X}) \cdot \left(\frac{\lambda_u}{\mu_u} \right)^{n_u} \frac{1}{n_u!} \cdot \left(\frac{\lambda_e}{\mu_e} \frac{1}{N_e b_e} \right)^{n_e} \left(\prod_{i=1}^{n_e} \left\lceil \frac{i}{N_e} \right\rceil \right)^{-1}, \quad (n_u, n_e) \in \mathcal{X}. \quad (1.6)$$

1.2.3 Полнодоступная модель «тройной услуги»

Обслуживание трех типов трафика – одноадресного, многоадресного и эластичного – при полнодоступной схеме принципиально не меняет картину: СП также необратим, а стационарное распределение не представимо в мультипликативном виде. Однако такая модель охватывает услуги разных типов и разные режимы передачи данных, и тем самым хорошо описывает доступ к ресурсу пользователей, которым предоставляются различные услуги в отсутствии приоритетов. Автор исследовал такую модель «тройной услуги» в кандидатской диссертации [178], схема которой показана на рисунке 1.1 [187–189].

Помимо уже упомянутых для одноадресного и эластичного трафика приближенных формул, рекуррентный алгоритм сокращает размерность системы уравнений равновесия (СУР) с

$$|\mathcal{X}| = \left| \left\{ (n_u, n_m, n_e) : n_u \geq 0, n_m = 0, 1, n_e \geq 0, n_u b_u + n_m b_m + n_e b_e \leq V \right\} \right|$$

до

$$\left| \left\{ (n_u, 0, 0) : n_u = 0, \dots, \lfloor V/b_u \rfloor \right\} \cup \left\{ (n_u, 1, 0) : n_u = 0, \dots, \lfloor (V - b_m)/b_u \rfloor \right\} \right|$$

за счет линейных преобразований СУР.

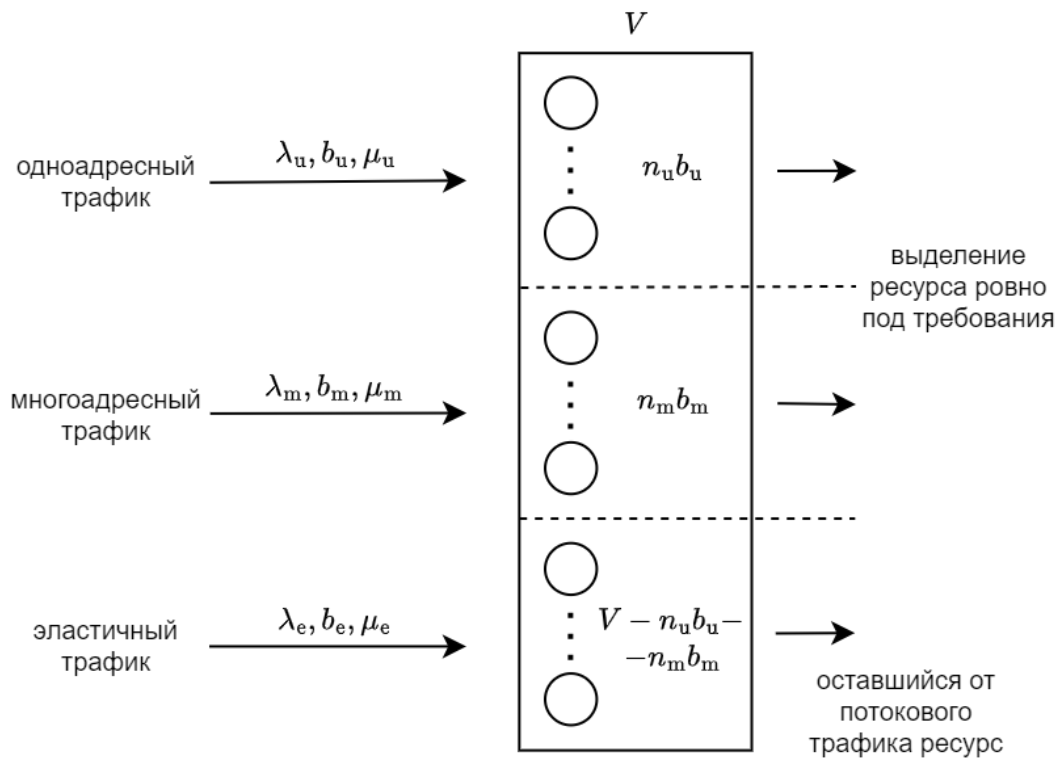


Рисунок 1.1 – Схема модели «тройной услуги»

Завершая обсуждение базовых моделей обслуживания трафика в беспроводных сетях, отметим, что пока не рассматривали вопросы приоритизации услуг, генерирующих этот трафик. Не обсуждали также, что подразумевается под ресурсом, а имели в виду пиковую пропускную способность соты. При исследовании доступа и предоставления различных услуг пользователям в современных сетях 5G необходимо учитывать приоритеты услуг друг относительно друга. Важно также соответствующим образом моделировать ресурс с учетом особенностей беспроводной среды передачи данных. Например, нужно учитывать, что пользователям, находящимся на разном расстоянии от базовой станции, будет доступна различная максимальная скорость передачи данных. Чем ближе устройство к базовой станции, тем скорость выше.

Технология нарезки сети и, как частный случай, нарезка радиоресурсов вносит свои коррективы в организацию приоритетного доступа трафика к ресурсу. Трафик различных классов передается внутри отдельных сегментов изолированно, но во времени такое распределение ресурса может меняться. Контроллер, осуществляющий мониторинг за этим процессом, принимает решение о

необходимости перераспределения ресурса. Таким образом, контроллер вносит корректировки в схему обслуживания трафика в каждом из сегментов – его действия будут являться более приоритетными относительно событий, связанных с сегментами сети.

В условиях ограниченности диапазона радиочастот используются системы совместного доступа к радиочастотам – как лицензируемым, так и нелицензируемым. В первом случае могут быть выделены диапазоны для совместного использования двумя мобильными операторами, а во втором – это обслуживание трафика в сети Wi-Fi. В таких условиях трафик владельца лицензированных радиочастот или трафик сети Wi-Fi будет являться приоритетным, что также необходимо учитывать при моделировании.

В последующих разделах этой главы обсудим особенности моделирования ресурса в беспроводной сети (раздел 1.3), системы нарезки радиоресурсов (раздел 1.4) и системы совместного доступа к радиочастотам (раздел 1.5).

1.3 ОСОБЕННОСТИ МОДЕЛИРОВАНИЯ РАДИОРЕСУРСА

1.3.1 Модель требования к ресурсу

В сетях 4G и 5G используется технология OFDMA (orthogonal frequency-division multiple access, многопользовательское мультиплексирование с ортогональным частотным разделением каналов), которая позволяет разделить диапазон радиочастот F , выделенный для передающей антенны базовой станции, между пользователями по частоте и времени одновременно (см. спецификацию 3GPP TS 38.211).

В сетях 5G диапазон частот разделен на поднесущие шириной $2^i \cdot 15$ кГц, $i = 0, 1, \dots, 6$. Время разделено на кадры длительностью 10 мс, каждый из которых состоит из 10 подкадров по 1 мс. Половина подкадров предназначена для нисходящей передачи данных от базовой станции к устройству (downlink, DL), а

другая половина – для восходящей передачи от устройства к базовой станции (uplink, UL). Подкадр делится на 2^i слотов длительностью $1/2^i$ мс, $i = 0, 1, \dots, 6$, а слот может быть разделен на более короткие мини-слоты.

В сетях 4G минимальной единицей частотно-временного ресурса является ресурсный блок (resource block, RB) размерности 12 поднесущих шириной 15 кГц на 1 слот длительностью 0,5 мс. В сетях 5G ресурсный блок определен только по частоте и объединяет 12 поднесущих шириной $2^i \cdot 15$ кГц, $i = 0, 1, \dots, 6$. По времени ресурс может выделяться на уровне мини-слотов, а планирование происходит на $T = 1$ кадр вперед.

Помимо выделения для каждого пользователя определенного числа ресурсных блоков, также можно настроить мощность передающей антенны как на базовой станции, так и на устройстве пользователя. Таким образом, в беспроводной сети требование к ресурсу определяется тройкой параметров

$$\mathbf{y} = (y_f, y_t, y_p), \quad y_f \in [0, F], \quad y_t \in [0, T = 1], \quad y_p \in [0, P], \quad (1.7)$$

где P максимальная мощность передающей антенны. Для упрощения дальнейшего анализа будем считать параметры (y_f, y_t, y_p) непрерывными, а y_t нормированным, т.е. представляющим собой долю от длительности кадра, выделенную пользователю.

Множество вариантов занятия ресурса для $n > 0$ сессий в случае разделения ресурса по частоте и времени можно записать как

$$\mathcal{Y}(n) = \left\{ (\mathbf{y}_1, \dots, \mathbf{y}_n) : \mathbf{y}_i = (y_i^f, y_i^t, y_i^p), \quad y_i^f \in (0, F], \quad y_i^t \in (0, 1], \quad y_i^p \in (0, P], \right. \\ \left. \sum_{i=1}^n y_i^f \leq F, \quad \sum_{i=1}^n y_i^t \leq 1, \quad y_i^p \leq P, \quad i = 1, \dots, n \right\}, \quad n > 0, \quad (1.8)$$

а множество всех вариантов занятия ресурса имеет вид

$$\mathcal{Y} = \{(0, 0, 0)\} \cup \bigcup_{n=1}^{\infty} \mathcal{Y}(n). \quad (1.9)$$

Пиковая пропускная способность C , т.е. максимально возможная скорость передачи данных для единственного пользователя, рассчитывается по формуле Шеннона-Хартли и формуле Фрииса

$$\begin{aligned} C &= F \log_2(1 + \text{SINR}), \\ \text{SINR} &= \frac{P_r}{N_0 + I_0}, \\ P_r &= \frac{P_t G_t G_r}{\text{PL}}, \\ P_r(\text{дБ}) &= P_t(\text{дБ}) + G_t(\text{дБи}) + G_r(\text{дБи}) - \text{PL}(\text{дБ}), \end{aligned} \tag{1.10}$$

где SINR отношение уровня сигнала к уровню шума и интерференции (signal-to-interference-plus-noise ratio, SINR), P_r мощность, принятая на приемной антенне, N_0 мощность шума, I_0 мощность интерференции, P_t мощность передающей антенны, G_t коэффициент усиления передающей антенны, G_r коэффициент усиления приемной антенны, PL мощность затухания сигнала (path loss, PL). Отметим, что в диссертационной работе рассматриваются только постоянные шум, интерференция и коэффициенты усиления антенн.

Если пользователей несколько, то каждый из них будет занимать часть y_t кадра, и максимальная скорость станет меньше пиковой – $c = y_t C$.

Моделью требования к ресурсу назовем систему трансцендентных уравнений, которая задает тройку параметров $(y_f, y_t, y_p) \in \mathcal{Y}$

$$\begin{aligned} b &= y_t y_f \log_2 \left(1 + y_p \frac{G_t G_r}{\text{PL}} \frac{1}{N_0 + I_0} \right), \\ A_j(y_f, y_t, y_p) &= 0, \quad j = 1, \dots, J, \end{aligned} \tag{1.11}$$

где $A_j(y_f, y_t, y_p)$ некоторая функция от (y_f, y_t, y_p) , а b необходимая скорость передачи данных.

Рассмотрим две модели требования к ресурсу, которые автор использовал в диссертационной работе. Обе модели не предусматривают разделение ресурса по частоте, т.е. $y_f = F$. Первая – политика постоянной мощности сигнала (full power, FP). При ее использовании мощность $y_p = P$ передающей антенны устанавливается на максимальном уровне, а доля кадра, выделяемая каждому пользователю, определяется таким образом, чтобы обеспечить скорость b . Схема занятия ресурса в этом случае показана на рисунке 1.2, а требование к ресурсу имеет вид

$$y(n) = \left(y_f = F, y_t = \frac{b}{F} \left[\log_2 \left(1 + \frac{PG_t G_r}{N_0 + I_0} \cdot \frac{1}{PL} \right) \right]^{-1}, y_p = P \right), \quad n > 0. \quad (1.12)$$

Вторая модель – политика снижения мощности сигнала (round robin, RR). В этом случае кадр делится поровну $y_t = 1/n$ между всеми пользователями, а мощность сигнала устанавливается так, чтобы обеспечить необходимую скорость b передачи данных для каждого пользователя. Схема занятия ресурса при использовании этой политики представлена на рисунке 1.3, а требование к ресурсу имеет вид

$$y(n) = \left(y_f = F, y_t = \frac{1}{n}, y_p = \frac{PL(N_0 + I_0)}{G_t G_r} (2^{bn/F} - 1) \right), \quad n > 0. \quad (1.13)$$

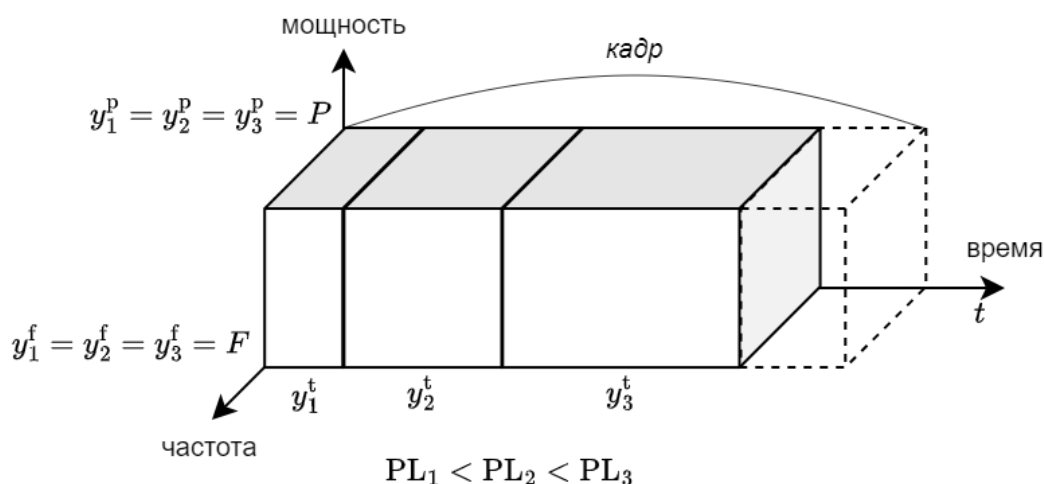


Рисунок 1.2 – Занятие ресурса для модели требования с постоянной мощностью сигнала

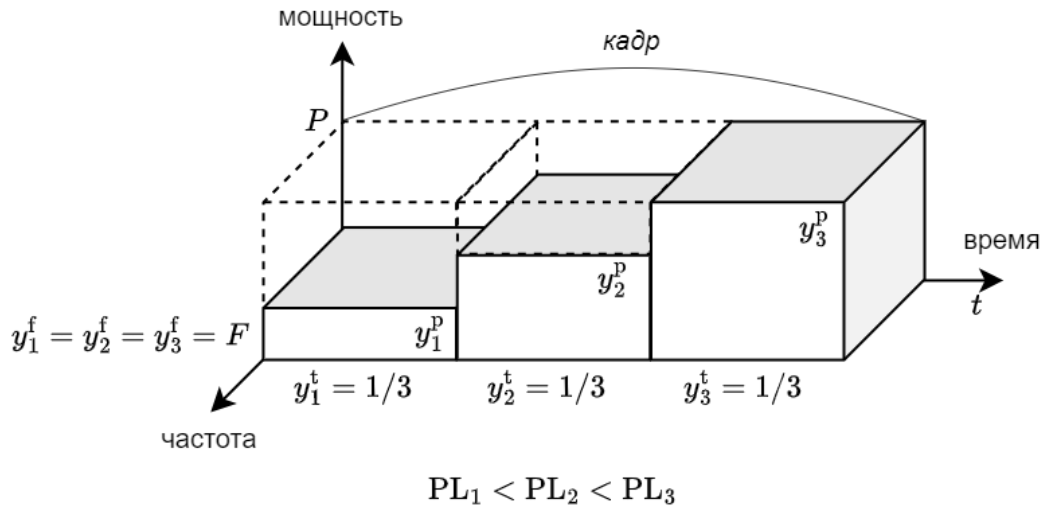


Рисунок 1.3 – Занятие ресурса для модели требования со снижением мощности сигнала

1.3.2 Модель затухания сигнала для анализа случайного требования к ресурсу

В диссертационной работе рассматриваются модели с требованием к ресурсу, где мощность затухания сигнала PL – переменная величина. Базовая модель затухания сигнала в свободном пространстве (free-space PL , FSPL) предполагает распространение сигнала в среде от передающей до приемной антенны в условиях прямой видимости без препятствий в идеальных условиях

$$PL = FSPL = \left(\frac{4\pi d}{\lambda} \right)^2 = \left(\frac{4\pi f}{c} \right)^2, \quad (1.14)$$

$$PL(\text{дБи}) = FSPL(\text{дБи}) = 20 \log_{10}(d) + 20 \log_{10}(f) - 147,55,$$

где d расстояние между передающей и приемной антеннами, λ длина волны, f несущая частота, c скорость света.

Однако в реальности на мощность затухания сигнала влияют различные факторы, такие как рельеф местности, постройки и преграды на пути распространения сигнала [190–195].

Автор участвовал в натурном эксперименте, проведенном в Университете Брно (Чехия), где исследовалось качество сигнала с использованием технологии LoRaWAN (long range low power wide area network) в окрестностях университета [196]. Вид функции мощности затухания сигнала от передающей антенны в зависимости от расстояния до принимающего устройства представлен на рисунке 1.4.

Существуют различные модели затухания сигнала – для городских районов, сельской местности, в помещении, с учетом особенностей распространения сигнала в условиях прямой видимости (line-of-sight, LOS) с устойчивым уровнем сигнала или вне зоны прямой видимости (non-LOS, NLOS), для различных длин волн и частот. В спецификации 3GPP TR 38.901 приведены эмпирические формулы для сетей 5G и частот от 0,5 до 100 ГГц для различных сценариев развертывания сети. Автор с коллегами в работе [197] описал соответствующую модель затухания сигнала для случайного расположения устройств пользователей в зоне покрытия базовой станции.

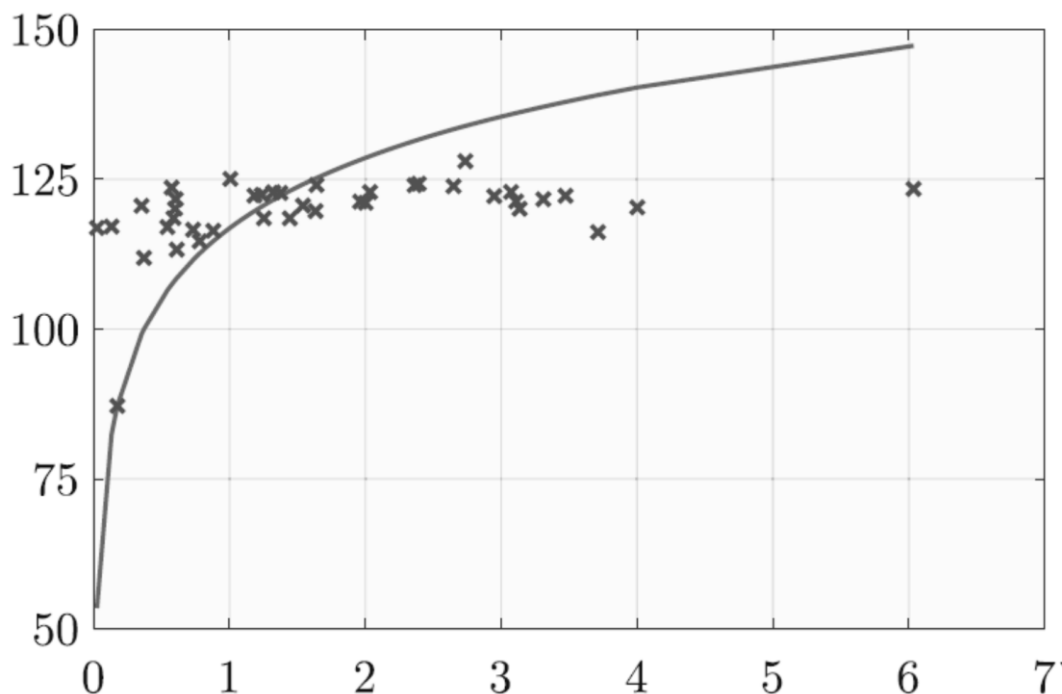


Рисунок 1.4 – Вид функции мощности затухания сигнала (дБ) в зависимости от расстояния (км)

Пусть передающая антенна базовой станции расположена на высоте h_{BS} , использует несущую частоту f_c и создает покрытие радиуса R . Приемная антенна устройства пользователя располагается на высоте h_{UT} и находится на расстоянии d от базовой станции на плоскости и с вероятностью $\text{Pr}_{\text{LOS}}(d)$ в зоне прямой видимости LOS. При возникновении препятствия на пути распространения сигнала, так называемого блокатора, устройство попадает в зону не прямой видимости NLOS. В таком случае мощность затухания сигнала имеет вид

$$\text{PL}(d) = \text{PL}_{\text{LOS}}(d) \cdot \text{Pr}_{\text{LOS}}(d) + \text{PL}_{\text{NLOS}}(d) \cdot [1 - \text{Pr}_{\text{LOS}}(d)], \quad (1.15)$$

$$\text{PL}_{\text{LOS}}(d) = \begin{cases} \text{PL}_1^{\text{LOS}}(d), & 0 = d_0 \leq d < d_1, \\ \text{PL}_2^{\text{LOS}}(d), & d_1 \leq d < d_2, \\ \dots & \dots \\ \text{PL}_J^{\text{LOS}}(d), & d_{J-1} \leq d \leq d_J = R, \end{cases} \quad (1.16)$$

$$\text{PL}_{\text{NLOS}}(d) = \max(\text{PL}_{\text{LOS}}(d), \text{PL}_1^{\text{NLOS}}(d), \dots, \text{PL}_K^{\text{NLOS}}(d)), \quad (1.17)$$

$$\begin{aligned} \text{PL}_m^l(d) &= \alpha_m^l \cdot D^{\beta_m^l}(d), \quad \text{PL}_m^l(d) (\text{дБ}) = \alpha_m^l (\text{дБ}) + \beta_m^l (\text{дБ}) \cdot \log_{10} D(d), \\ l &= \text{"LOS"}, \quad m = j = 1, \dots, J, \quad l = \text{"NLOS"}, \quad m = k = 1, \dots, K, \end{aligned} \quad (1.18)$$

$$\text{Pr}_{\text{LOS}}(d) = \begin{cases} \text{Pr}_1^{\text{LOS}}(d), & 0 = r_0 \leq d < r_1, \\ \text{Pr}_2^{\text{LOS}}(d), & r_1 \leq d < r_2, \\ \dots & \dots \\ \text{Pr}_I^{\text{LOS}}(d), & r_{I-1} \leq d \leq r_I = R, \end{cases} \quad (1.19)$$

$$D(d) = \sqrt{d^2 - (h_{\text{BS}} - h_{\text{UT}})^2}, \quad (1.20)$$

где $\text{PL}_{\text{LOS}}(d)$ и $\text{PL}_{\text{NLOS}}(d)$ мощности затухания сигнала при прямой LOS и не прямой видимости NLOS, $\text{PL}_m^l(d)$ компоненты этих функций зависят от конкретных констант α_m^l и β_m^l , которые определяются сценарием развертывания сети согласно спецификации 3GPP TR 38.901, D расстояние от базовой станции до пользователя в пространстве.

Если расстояние от базовой станции до пользователя на плоскости является случайной величиной (СВ) ξ_d с функцией распределения (ФР) $F_{\xi_d}(x)$, то мощность затухания сигнала будет функцией от ξ_d с ФР $F_{\xi_{PL}}(x)$

$$F_{\xi_{PL}}(x) = \Pr(\xi_{PL} \leq x) = \Pr(\xi_{LOS} \cdot \xi_{P_{TLOS}} + \xi_{NLOS} \cdot [1 - \xi_{P_{TLOS}}] \leq x) \approx \int_0^\infty F_{\xi_{LOS}}\left(\frac{x}{E[\xi_{P_{TLOS}}]} - y\right) dF_{\xi_{NLOS}}\left(\frac{y}{1 - E[\xi_{P_{TLOS}}]}\right), \quad (1.21)$$

$$F_{\xi_{LOS}}(x) = \Pr(\xi_{LOS} \leq x) = \sum_{j=1}^J F_{\xi_j^{LOS}}(x) \cdot [F_{\xi_d}(d_j) - F_{\xi_d}(d_{j-1})], \quad (1.22)$$

$$F_{\xi_{NLOS}}(x) = \Pr(\xi_{NLOS} \leq x) = F_{\xi_{LOS}}(x) \cdot \prod_{k=1}^K F_{\xi_k^{NLOS}}(x), \quad (1.23)$$

$$F_{\xi_m^l}(x) = \Pr(\xi_m^l \leq x) = F_{\xi_D}\left(\left[x/\alpha_m^l\right]^{1/\beta_m^l}\right), \quad (1.24)$$

$$l = \text{"LOS"}, m = j = 1, \dots, J, \quad l = \text{"NLOS"}, m = k = 1, \dots, K,$$

$$F_{\xi_D}(x) = \Pr(\xi_D \leq x) = F_{\xi_d}\left(\sqrt{x^2 - (h_{BS} - h_{UT})^2}\right), \quad (1.25)$$

$$E[\xi_{P_{TLOS}}] = \int_0^R \Pr_{LOS}(x) dF_{\xi_d}(x). \quad (1.26)$$

1.3.3 Ресурсная система массового обслуживания

В условиях, когда пользователи случайным образом расположены в зоне покрытия базовой станции и мощность затухания сигнала также случайна, требование к ресурсу становится СВ $\xi = (\xi_f, \xi_t, \xi_p)$ с соотношением

$$b = \xi_t \xi_f \log_2 \left(1 + \xi_p \cdot G_t G_r [\xi_{PL}(N_0 + I_0)]^{-1} \right) \quad (1.27)$$

между компонентами вектора.

Вероятность $\Pr((\xi_1, \dots, \xi_n) \in \mathcal{Y}(n))$ занятия ресурса $n > 0$ сессиями определяет возможность одновременного обслуживания такого числа пользователей. Для модели FR требования к ресурсу с постоянной мощностью сигнала эта вероятность рассчитывается по приближенной формуле

$$\Pr((\xi_1, \dots, \xi_n) \in \mathcal{Y}(n)) = \Pr\left(\sum_{i=1}^n \xi_i^t \leq 1\right) \approx \Phi\left(\frac{1 - \mu n}{\sigma \sqrt{n}}\right), \quad n > 0,$$

$$\mu = \frac{b}{F} \int_0^\infty \left[\log_2 \left(1 + \frac{PG_t G_r}{N_0 + I_0} \cdot \frac{1}{x} \right) \right]^{-1} dF_{\xi_{PL}}(x),$$

$$\sigma^2 = \left(\frac{b}{F} \right)^2 \int_0^\infty \left[\log_2 \left(1 + \frac{PG_t G_r}{N_0 + I_0} \cdot \frac{1}{x} \right) \right]^{-2} dF_{\xi_{PL}}(x) - \mu^2,$$

где $\Phi(x)$ ФР для стандартного нормального распределения.

Для модели RR со снижением мощности сигнала вероятность занятия ресурса $n > 0$ пользователями имеет вид

$$\Pr((\xi_1, \dots, \xi_n) \in \mathcal{Y}(n)) = \Pr(\xi_i^p \leq P, \quad i = 1, \dots, n) = \left[F_{\xi_{PL}} \left(\frac{PG_t G_r}{N_0 + I_0} \frac{1}{2^{bn/F} - 1} \right) \right]^n,$$

$$n > 0.$$

Определим также вероятность того, что еще одному пользователю может быть выделен ресурс, при условии обслуживания $n \geq 0$ сессий

$$p(0) = \Pr(\xi_1 \in \mathcal{Y}(1)),$$

$$p(n) = \frac{\Pr((\xi_1, \dots, \xi_n, \xi_{n+1}) \in \mathcal{Y}(n+1))}{\Pr((\xi_1, \dots, \xi_n) \in \mathcal{Y}(n))}, \quad n > 0.$$

Эти условные вероятности используются автором в СМО для учета в структурных ограничениях моделей не только обслуживающих приборов, но и трехкомпонентного ресурса.

Ресурсные СМО (РеСМО) [198–214] и модели стохастической геометрии широко применяются для анализа беспроводных сетей. В зависимости от задачи ресурс в РеСМО может моделировать разные сущности сети, такие как длительность кадра по времени, мощность сигнала в ваттах, ресурсные блоки и др. В диссертационной работе автор использовал РеСМО для учета в исследуемых системах двух моделей требования к ресурсу: с постоянной мощностью FR и снижением мощности RR сигнала. Фактически, это РеСМО с двумя типами ресурсов, которые зависят друг от друга.

Проиллюстрируем вид стационарного распределения для базовой РеСМО – фактически первой модели Эрланга со случайным требованием к ресурсу. В работах автора с коллегами [215,216] приведен анализ описывающего функционирование модели СП с детальными состояниями – не только общее число пользователей, но и их распределение по уровням качества сигнала в соте (channel quality indicator, CQI). Стационарное распределение для СП с агрегированными по общему числу пользователей состояниями имеет вид

$$\pi(n) = \left(\frac{\lambda \cdot p(n)}{\mu} \right)^n \frac{1}{n!} \cdot \left[\sum_{i=0}^V \left(\frac{\lambda \cdot p(i)}{\mu} \right)^i \frac{1}{i!} \right]^{-1}, \quad n = 0, \dots, V, \quad (1.31)$$

а вероятность блокировки вычисляется по формуле

$$B = \sum_{n=0}^V \pi(n) [1 - p(n)]. \quad (1.32)$$

Автор с коллегами применила такую РеСМО для анализа совместного обслуживания трафика Н2Н и М2М, при этом в качестве модели требования к ресурсу для Н2Н применялась модель FR с постоянной мощностью, а для М2М – RR со снижением мощности [217].

1.4 СИСТЕМА НАРЕЗКИ РАДИОРЕСУРСОВ

1.4.1 Особенности нарезки радиоресурсов

Технология нарезки сети и радиоресурсов позволяет организовать приоритетный доступ к ресурсам беспроводной сети [218–238]. В этом процессе участвуют три стороны, так называемые бизнес-роли (см. спецификации 3GPP TS 28.552 и 3GPP 28.554). Базовый оператор (network slice provider, network operator) владеет ресурсами сети и предоставляет логические сегменты сети (network slice) поставщикам услуг. Поставщик услуг (network slice service provider, communication service provider) предоставляет услуги конечным пользователям. Конечный пользователь (network slice service user, communication service customer) получает доступ к услугам. При этом поставщик услуг может сам создавать сегменты сети под конкретные свои услуги или типы услуг.

Взаимоотношения между этими сторонами регулируются соглашениями о качестве обслуживания (service level agreement, SLA). Одно из важных условий – логическая изоляция сегментов сети (slice isolation), чтобы обеспечение качества обслуживания и производительности в одном сегменте не влияло на другие. Разделение ресурса между сегментами сети может быть фиксированным во времени или динамическим. Контроллер управляет перераспределением ресурса и осуществляет мониторинг.

Решение, принимаемое контроллером, зависит от типа ресурса и показателей, относящихся к бизнес-ролям. Например, это повышение пропускной способности и спектральной эффективности, уменьшение задержек передачи данных при управлении занятием диапазона радиочастот, снижение уровня интерференции и повышение энергоэффективности при управлении мощностью. С математической точки зрения, это многокритериальная задача оптимизации. В литературе для ее решения часто применяется аппарат МППР и алгоритмы машинного обучения с подкреплением (reinforcement learning).

Если необходимо обеспечить фиксированное распределение ресурса $V = V_1 + \dots + V_K$ между сегментами, эту задачу можно решить один раз в период планирования сети. В этом случае стационарное распределение числа пользователей в каждом сегменте, например, для нетерпеливого эластичного трафика с задержками, будет иметь мультипликативный вид

$$\pi(n_1, \dots, n_K) = G^{-1}(\mathcal{X}) \cdot \prod_{k=1}^K \frac{\lambda_k^{n_k} \left(\frac{1}{V_k \mu_k} \right)^{\min(n_k, \lfloor V_k/b_k \rfloor)}}{n_k - \lfloor V_k/b_k \rfloor \prod_{i=1}^K (V_i \mu_i + i \varepsilon_i)}, \quad (n_1, \dots, n_K) \in \mathcal{X}, \quad (1.33)$$

где b_k требование к минимальной скорости передачи трафика, $1/\varepsilon_k$ среднее экспоненциальное время на ожидание начала обслуживания нетерпеливого трафика, $\prod_{i \in \mathcal{O}} x_i = 1$ [239–241].

1.4.2 Нестационарные требования к объему сегментов

Динамическое перераспределение ресурса необходимо, когда потребности в услугах конечных пользователей меняются во времени, имеют нестационарный характер. Например, на рисунке 1.5 представлен средний суточный профиль трафика в регионе Центральной и Восточной Европы в 2020 году для различных классов услуг (см. спецификацию ITU-R M.2370). Вернемся к этому рисунку в разделе 1.5.

На рисунке 1.6 показан профиль трафика оператора Vodafone Portugal за 25 дней в 2018 году. В разделе 1.1 уже упоминали этот профиль. Автор с коллегами решал задачу прогнозирования объема трафика в восходящем и нисходящем направлении, а также суммарного объема и объема по разным услугам в отдельности [80]. В исследовании использовались методы математической статистики, а именно модели прогнозирования временных рядов – сезонная интегрированная модель авторегрессии – скользящего среднего SARIMA (seasonal

autoregressive integrated moving averag) и модель тройного экспоненциального сглаживания Хольта-Винтерса (Holt Winters exponential smoothing).

Еще одной моделью описания нестационарного профиля трафика является его аппроксимация тригонометрическим рядом Фурье. Например, для набора данных оператора Vodafone Portugal для услуги передачи файлов в восходящем направлении модель трафика в зависимости от времени получилась следующей

$$a(t) = 0,6521 - 0,08888 \cdot \cos(wt) - 0,3963 \cdot \sin(wt) + \\ + 0,1384 \cdot \cos(2wt) - 0,1411 \cdot \sin(2wt) - \\ - 0,01446 \cdot \cos(3wt) + 0,05875 \cdot \sin(3wt),$$

где $w = 0,2603$ [242]. Данные были нормированы, а вид полученной функции проиллюстрирован на рисунке 1.7.

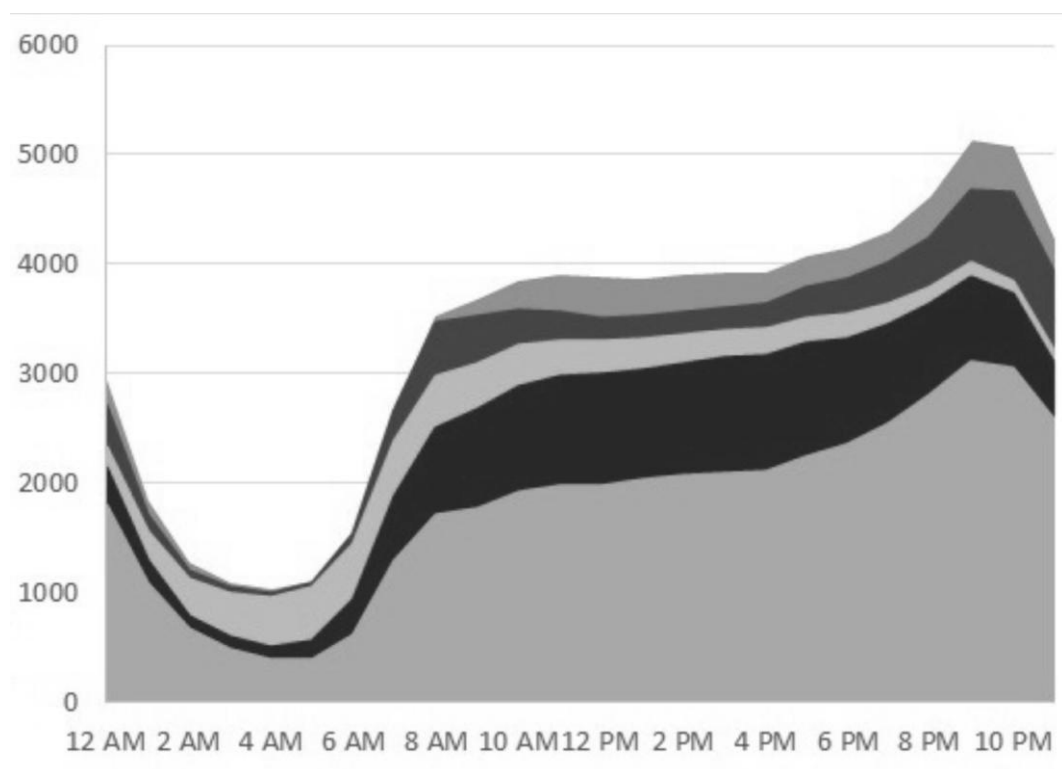


Рисунок 1.5 – Вид суточного профиля трафика

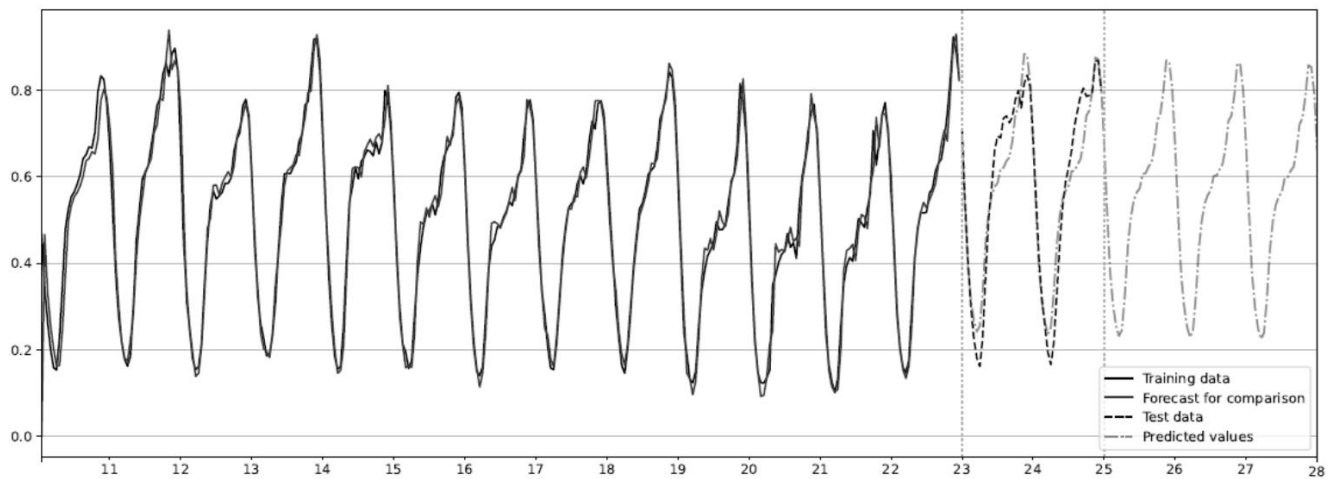


Рисунок 1.6 – Прогнозирование профиля трафика

Такая аппроксимация тригонометрическим рядом Фурье удобна для последующего использования в анализе СМО с нестационарными интенсивностями [243–252]. Автор участвовал в исследовании СМО с допускающим задержки нетерпеливым эластичным трафиком. Входящий поток запросов на передачу трафика в системе нестационарный, объем и время на ожидание начала обслуживания нетерпеливого трафика также нестационарные [242]. На рисунке 1.8 показана зависимость среднего числа ожидающих начала обслуживания сессий трафика от времени. После стабилизации режима около $t \approx 600$ характеристика периодически колеблется от 65 до 75 сессий.

Этот пример иллюстрирует влияние нестационарного профиля трафика на показатели эффективности системы. Там, где такие периодические колебания существенны, рекомендуется применять динамическое перераспределение ресурса в нарезке сети.

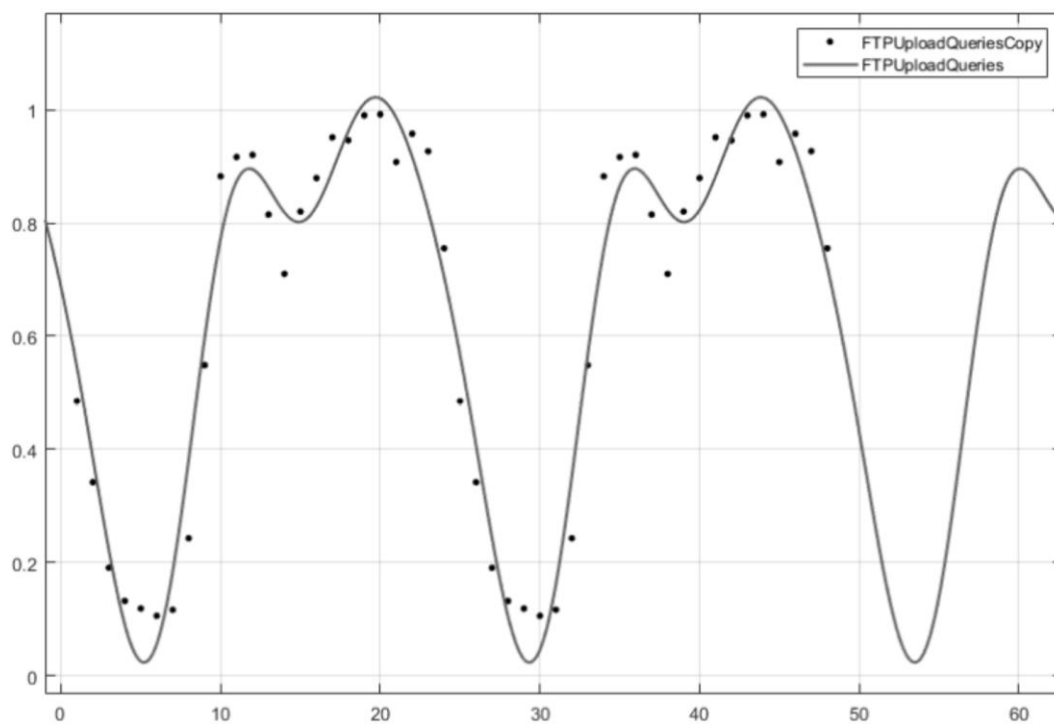


Рисунок 1.7 – Аппроксимация профиля трафика рядом Фурье

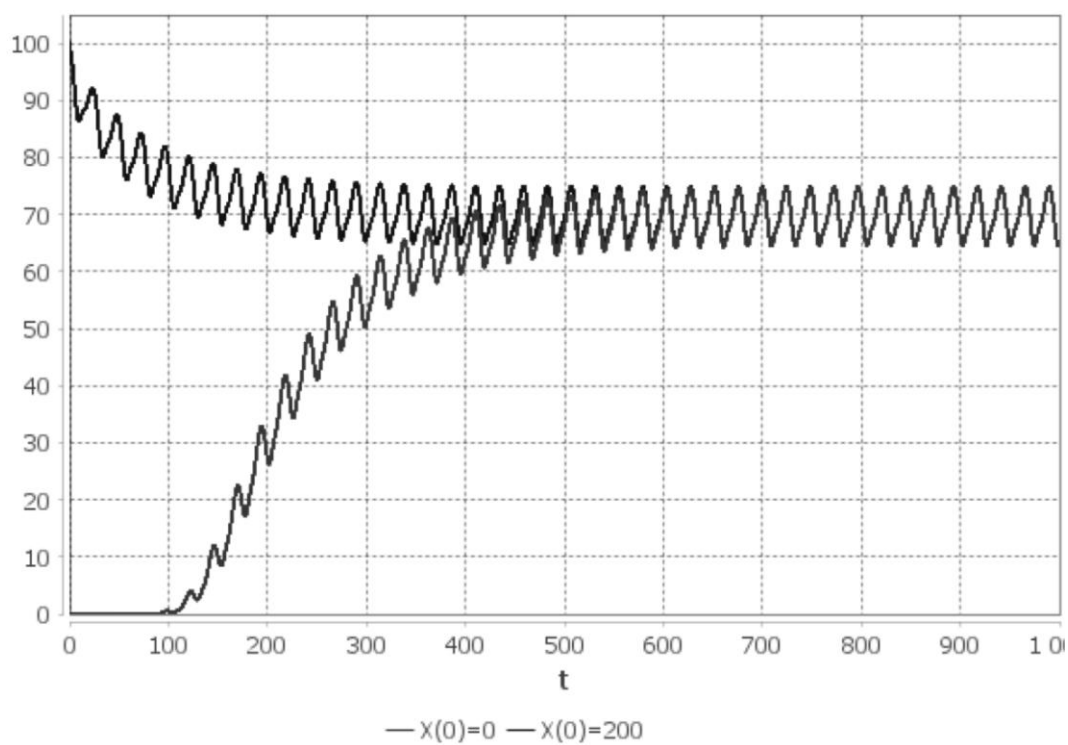


Рисунок 1.8 – Среднее число ожидающих начала обслуживания пользователей для системы с нестационарными интенсивностями

1.4.3 Принципы перераспределения ресурса

Как уже было отмечено, за динамическое перераспределение ресурса между сегментами сети отвечает контроллер. Он отправляет сообщения с заданной периодичностью, чтобы принять решение, нужно ли изменить разделение ресурса. Поток служебных сообщений контроллера создает дополнительную сигнальную нагрузку на сеть, и перераспределение ресурса происходит с некоторой задержкой. В этом случае фиксированное разделение ресурса кажется более привлекательным, чем крайний случай динамического с постоянным мониторингом состояния сети и возможностью перераспределения при любом изменении в системе.

Однако при выборе способа разделения ресурса необходимо учитывать и другие показатели эффективности. Например, насколько текущее разделение ресурса отличается от изначально рассчитанного, указанного в соглашениях SLA между базовым оператором и поставщиками услуг. Важно также эффективно использовать ресурсы, чтобы минимизировать их простой, когда есть пользователи других поставщиков услуг, ожидающие начала обслуживания. Схема с иллюстрацией перераспределения ресурса по перечисленным принципам представлена на рисунке 1.9.

Выбор периодичности отправки сообщений контроллера зависит от других показателей эффективности и неоднозначен. Например, в случае фиксированного разделения ресурса, строго соответствующему базовому по SLA, использование ресурса будет неэффективным. С другой стороны, при непрерывном мониторинге сети простой ресурса будет исключен, но отклонение от базового разделения может быть значительным.

Таким образом, решение будет зависеть от важности того или иного принципа для базового оператора. При переходе от многокритериальной задачи оптимизации показателей эффективности к оптимизации взвешенной суммы функций по каждому из них, значимость каждого показателя можно учесть в соответствующих весах. Данная взвешенная сумма будет являться функцией

вознаграждения МППР, а для анализа системы могут применяться управляемые СМО (УСМО), описываемые с помощью МППР [253–265].

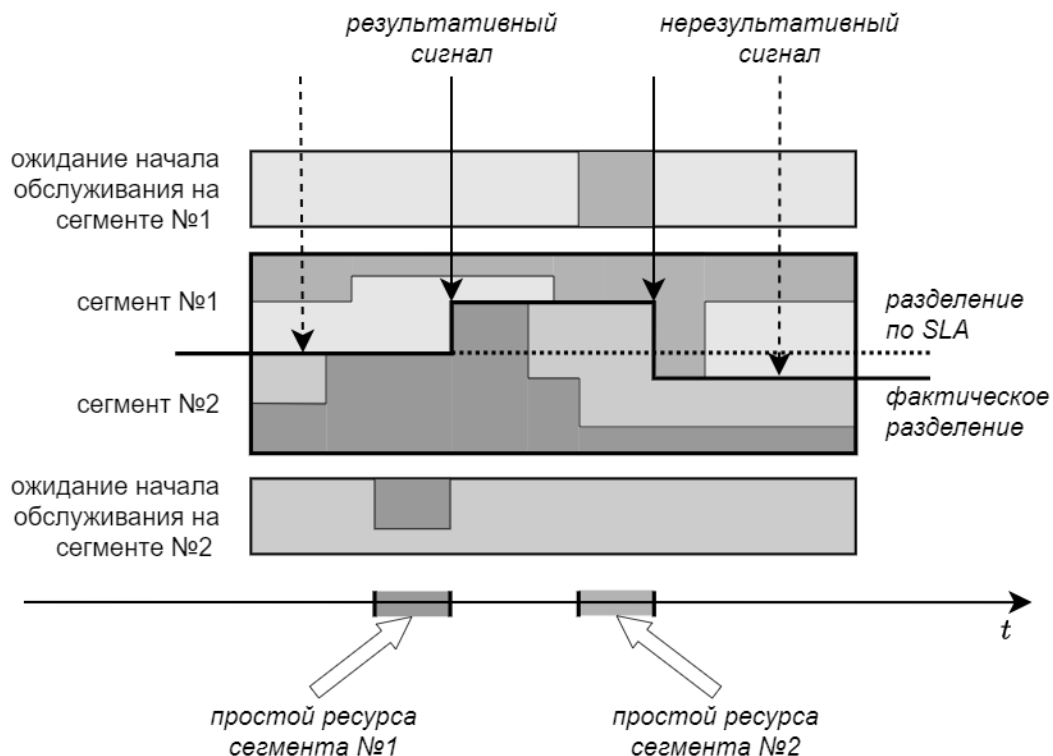


Рисунок 1.9 – Схема перераспределения ресурса при нарезке сети

1.5 СИСТЕМА СОВМЕСТНОГО ДОСТУПА К РАДИОЧАСТОТАМ

1.5.1 Особенности использования нелицензируемых и лицензируемых радиочастот

Система совместного доступа к нелицензируемым и лицензируемым радиочастотам позволяет операторам расширить доступный им диапазон частот [266–268]. Это достигается за счет совместного использования частот с другими сторонами: пользователями сети Wi-Fi и владельцами лицензированных радиочастот соответственно. Однако, несмотря на преимущества такой системы, у нее есть особенности – оператору необходимо учитывать, что он не имеет приоритетного доступа к радиоресурсам. Это означает, что при планировании доступа к частотам он не может обеспечить качество обслуживания своих

пользователей на том же уровне, как если бы он это делал на базе своего индивидуального диапазона частот.

Для нелицензированных радиочастот используется система совместного доступа long-term evolution unlicensed (LTE-U), license assisted access (LAA), enhanced LAA (eLAA), further enhanced LAA (feLAA), 5G new radio unlicensed (5G NR-U) (см. спецификации 3GPP TR 38.849, 3GPP TR 37.890) [269–275]. При обслуживании пользователей оператора 4G/5G в сети Wi-Fi необходимо учитывать принципы управления доступом в этой сети. Это означает, что все пользователи не различимы с точки зрения доступа к радиоресурсам, и обеспечить гарантированное качество обслуживания для пользователей оператора становится проблематичным. Схема принципа такой выгрузки трафика в сеть Wi-Fi и сосуществования двух типов трафика показана на рисунке 1.10.

Совместный доступ к лицензированным радиочастотам реализует система licensed shared access (LSA), evolved LSA (eLSA), 5G shared access [276–280]. Особенностью использования оператором 4G/5G лицензированных радиочастот является то, что он находится в неприоритетном положении – абсолютный приоритет у владельца диапазона радиочастот. Для системы LSA это частоты 2300–2400 МГц, которые в зависимости от страны используются для различных целей, таких как радиотелеметрия с самолетами, экстренная связь для обеспечения общественной безопасности и помощи при бедствиях (public protection and disaster relief, PPDR), запись аудио и видео контента журналистами (electronic news gathering, ENG), любительская радиосвязь. Соответственно, в те интервалы времени, когда радиочастоты нужны владельцу в необходимом объеме, они становятся недоступны для оператора, что схематично изображено на рисунке 1.11.

При моделировании такой схемы доступа к ресурсу переменного объема удобно использовать СМО с ненадежными приборами и СМО в случайной среде [281–298]. В этой модели время, в течение которого владельцу нужны радиочастоты, соответствует времени отказа ресурса.

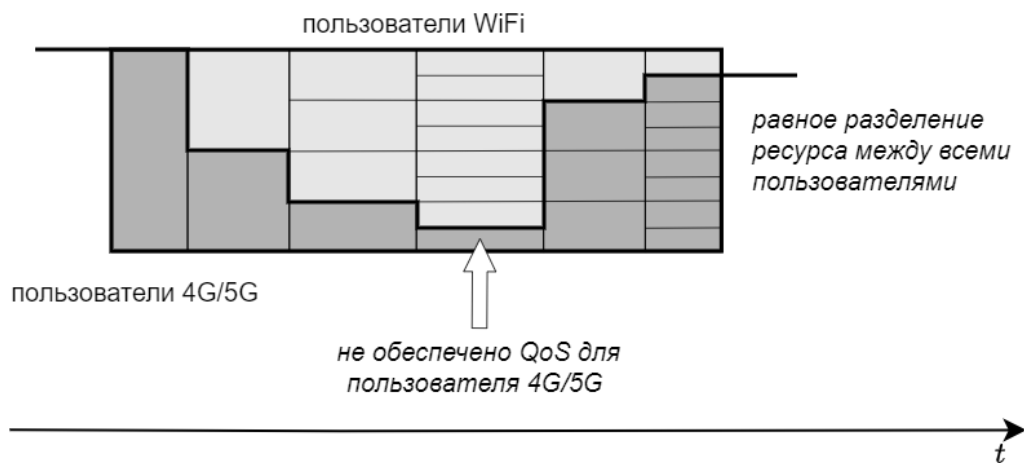


Рисунок 1.10 – Схема обслуживания трафика в нелицензируемым радиочастотах

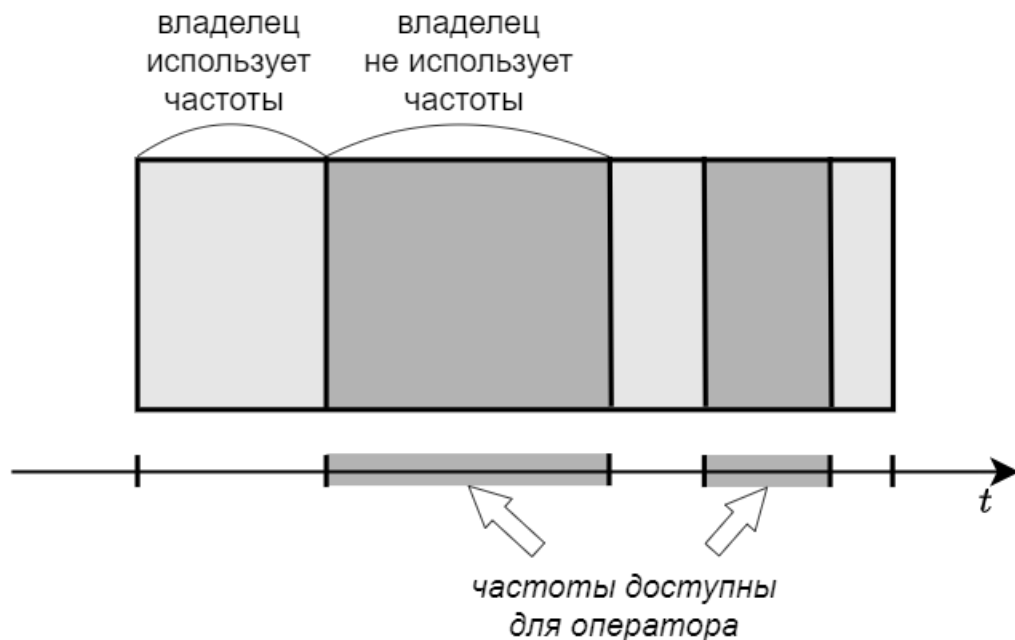


Рисунок 1.11 – Схема совместного использования лицензируемых радиочастот

1.5.2 Механизм снижения мощности сигнала

При совместном доступе к лицензируемым радиочастотам, когда радиоресурсы понадобятся владельцу диапазона, оператору необходимо будет принять решение по изменению схемы обслуживания своих пользователей. Это решение будет зависеть от потребностей владельца и требований к объему ресурса, необходимого пользователям оператора. Учитывая структуру ресурса, о которой

говорилось в разделе 1.3, владелец может занимать его по-разному. Один из вариантов – уменьшить мощность передающей антенны, чтобы пользователи оператора не создавали интерференцию владельцу.

Автор принимал участие в натурном эксперименте, который проводился в Университете Брно (Чехия) [299]. Целью эксперимента было исследование механизма снижения мощности сигнала, который доступен для пользователей оператора. Эксперимент, схема которого показана на рисунке 1.12, проводился внутри здания университета на одном из этажей. Измерительная установка, которая является владельцем совместного диапазона радиочастот, перемещается по коридорам этажа. Второй диапазон предназначен только для четырех неподвижных пользовательских устройств. На рисунке 1.13 показана мощность сигнала, доступная для пользователей, в зависимости от положения измерительной установки.

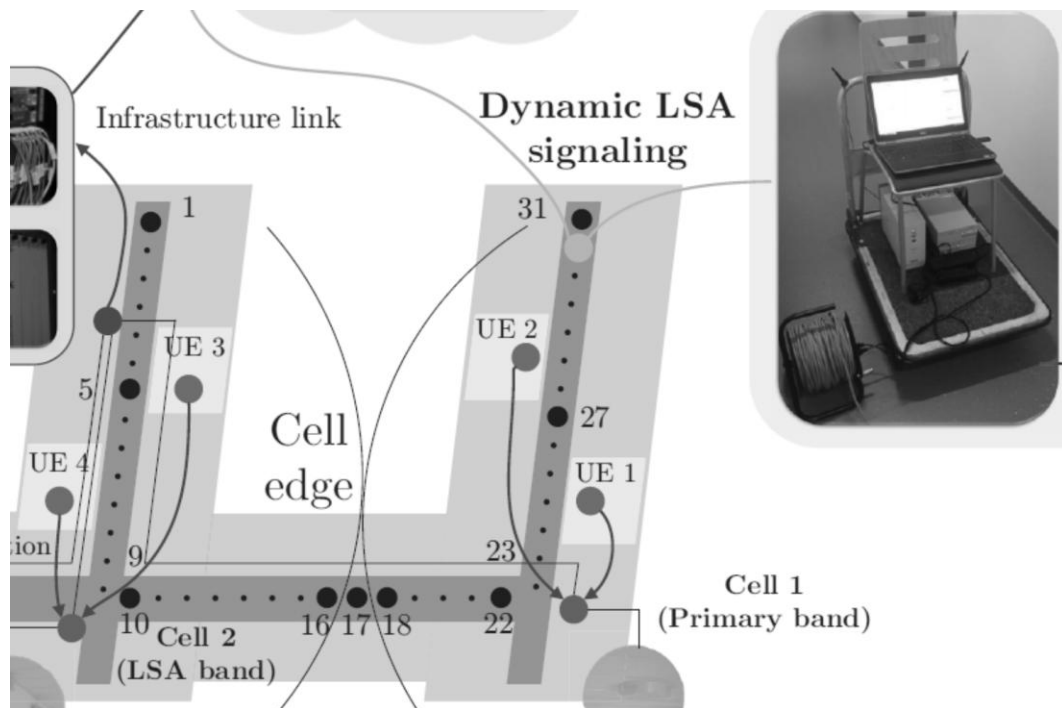


Рисунок 1.12 – Схема натурного эксперимента по совместному доступу к лицензируемым радиочастотам

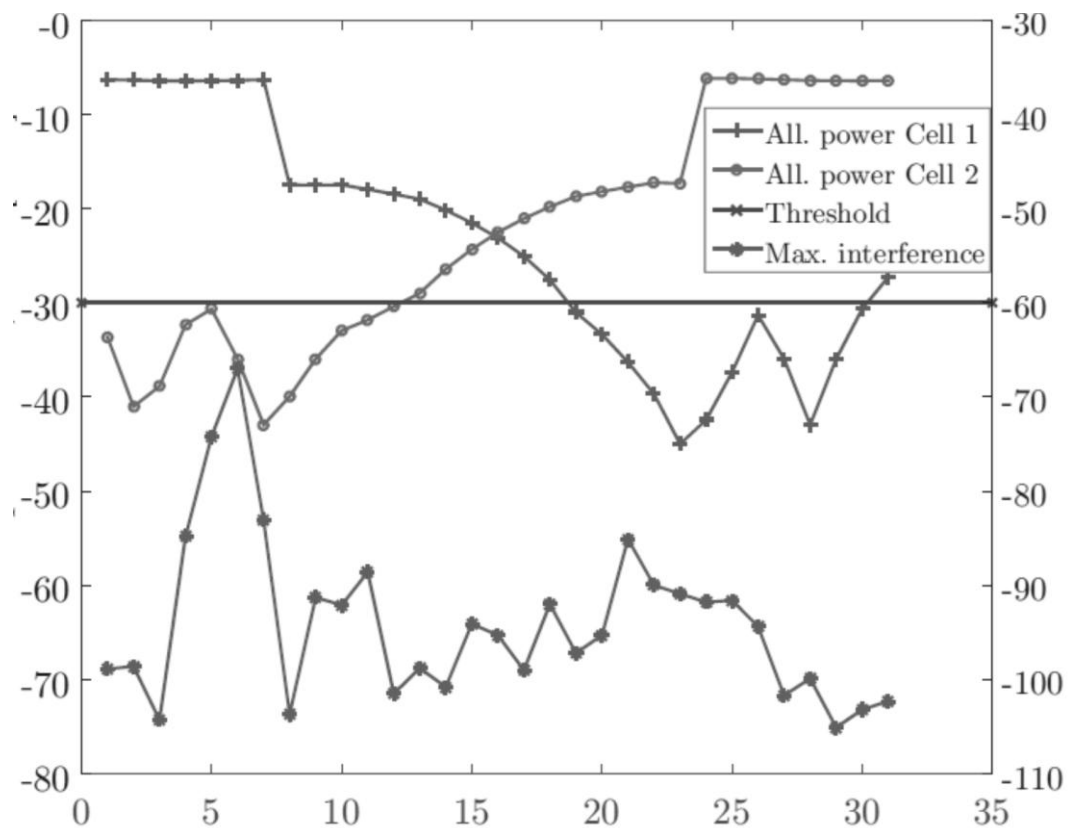


Рисунок 1.13 – Доступная мощность сигнала (дБ) в натурном эксперименте в зависимости от местоположения владельца

Автор совместно с коллегами исследовал механизм снижения мощности сигнала и в виде математической модели [300,301]. Модель построена для случая, когда лицензированный диапазон радиочастот используется для радиотелеметрии с самолетами. На рисунке 1.14 показана схема сценария, в котором самолет взлетает и по мере его удаления от земли «шар» вокруг него, где не должно быть интерференции, становится меньше. Для оператора во времени регулируется доступная ему мощность. На рисунке 1.15 показаны пороги мощности в зависимости от времени после взлета самолета. Они различаются местоположением пользователей относительно траектории полета самолета. Худший случай – это когда пользователи находятся на проекции траектории полета.

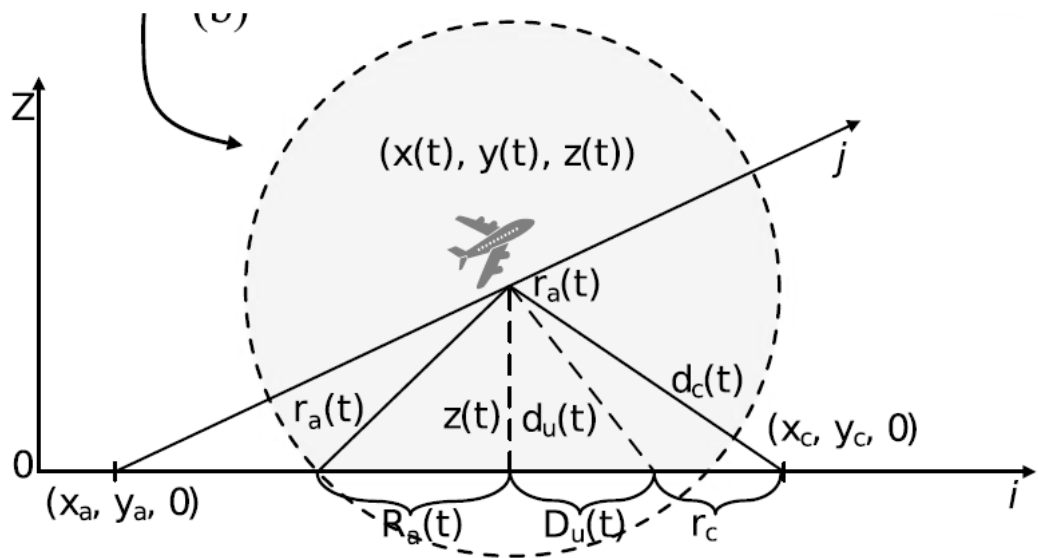


Рисунок 1.14 – Схема сценария радиотелеметрии с самолетами

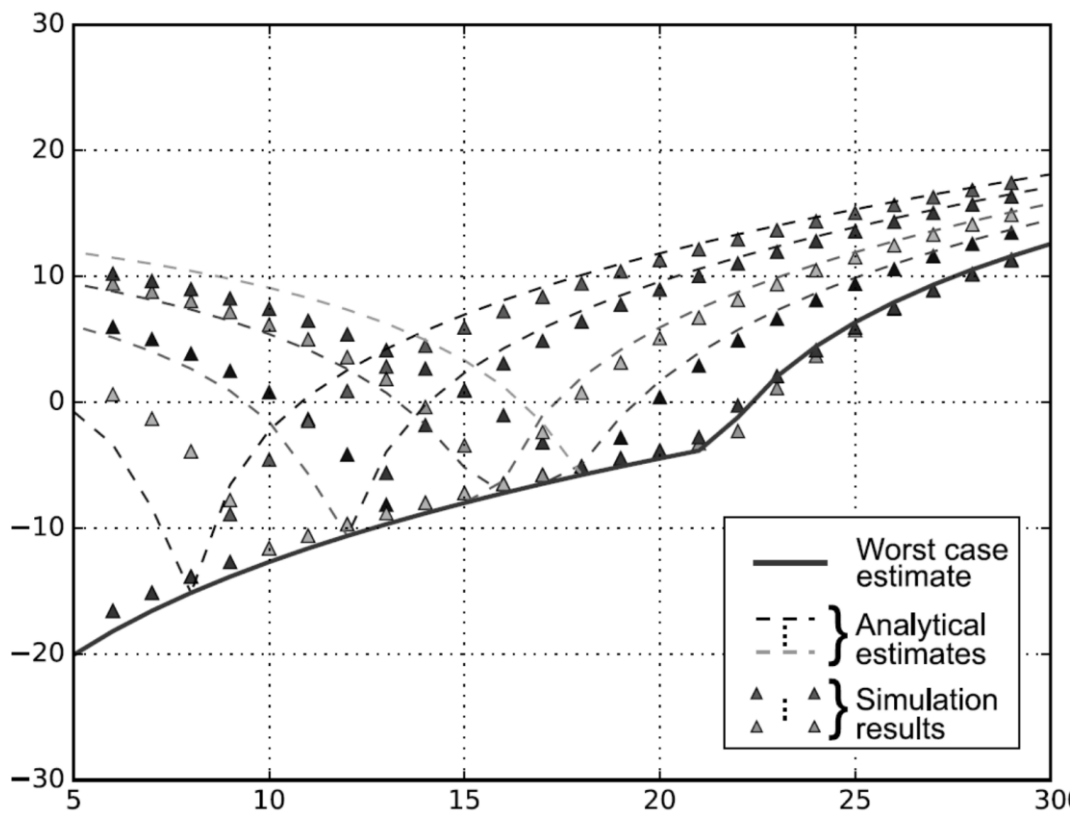


Рисунок 1.15 – Доступная мощность сигнала (дБ) для сценария радиотелеметрии с самолетами в зависимости от времени

1.5.3 Замечание к нестационарному режиму

Потребности владельца радиочастот, к которым имеет доступ и оператор, могут меняться во времени. Например, для сценария радиотелеметрии с самолетом частота вылетов самолетов может различаться в разное время суток. Как уже упоминалось в разделе 1.4, трафик пользователей также является нестационарным. В таких случаях полезным инструментом может быть аппарат СМО с нестационарными интенсивностями.

Автор участвовал в исследовании двух моделей доступа к радиоресурсам, которые используются совместно с аэропортом для радиотелеметрии с самолетами. Модели построены в виде СМО с нестационарными интенсивностями для двух сценариев: трафика, не допускающего задержек [302], и трафика, допускающего задержки [303]. В качестве модели трафика использовались данные с рисунка 1.5. Аппроксимация выполнена с помощью тригонометрического ряда Фурье. Проведен анализ поведения показателей эффективности моделей, включая вероятность блокировки запроса пользователя на передачу трафика. На рисунках 1.16 и 1.17 показаны вероятности блокировки для модели с очередью. Для модели без очереди получены оценки скорости сходимости для квазистационарного распределения вероятностей состояний системы и среднего числа обслуживаемых сессий.

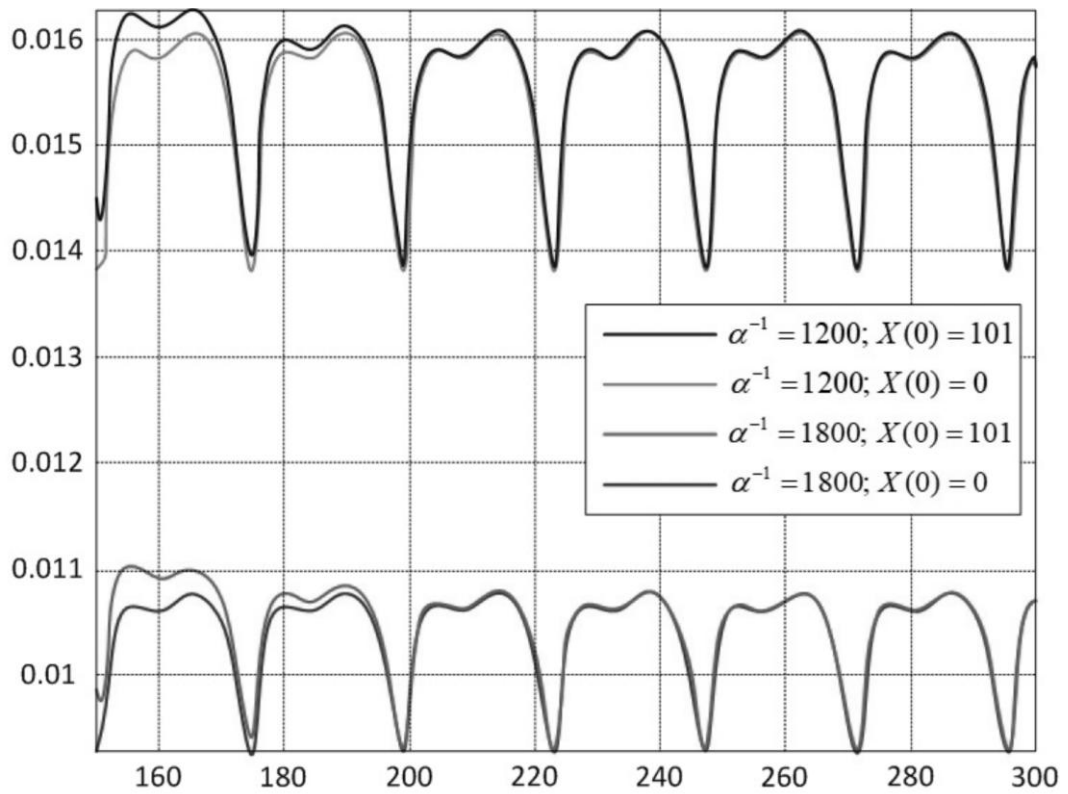


Рисунок 1.16 – Вероятность блокировки для системы с нестационарными интенсивностями для разного времени доступности радиочастот в зависимости от времени (с)

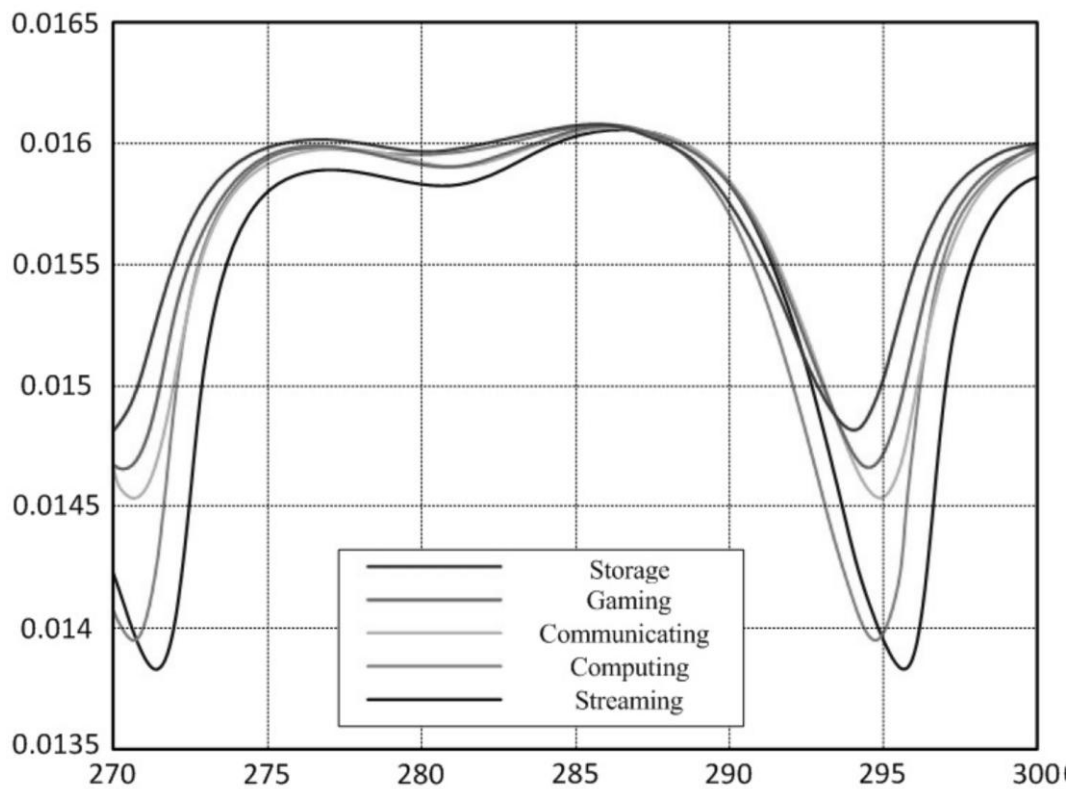


Рисунок 1.17 – Вероятность блокировки для системы с нестационарными интенсивностями для разных услуг в зависимости от времени (с)

1.6 ОПИСАНИЕ КОМПЛЕКСА МОДЕЛЕЙ

1.6.1 Нотация приоритетов в символике Башарина-Кендала

Вернемся к вопросу приоритизации в беспроводных сетях. Как отмечено в предыдущих разделах 1.2-1.5, услуги имеют различный приоритет доступа к радиоресурсам. Кроме того, есть события во внешней среде, которые также могут быть приоритетными относительно процесса обслуживания пользователей оператора. Это, например, возникновение препятствий на пути распространения сигнала – так называемые блокировки пути прямой видимости. Или трафик сети нелицензируемых и лицензируемых радиочастот. Или служебные сообщения контроллера, которые сигнализируют о необходимости мониторинга перераспределения ресурса при нарезке сети.

Приоритетным назовем событие, связанное с услугами различных классов или внешней средой, которое влияет на объем доступного ресурса для сессий различных классов или на объем занимаемого сессиями ресурса. В качестве примеров таких событий и их последствий для системы можно привести:

- прерывание обслуживания при поступлении запроса на предоставление услуги;
- увеличение скорости предоставления услуги при завершении обслуживания;
- снижение пропускной способности соты из-за блокировки пути прямой видимости;
- возобновление обслуживания после недоступности ресурса;
- перераспределение ресурса между сегментами сети при поступлении сообщения контроллера;
- снижение мощности базовой станции из-за доступа к ресурсу владельца лицензированных радиочастот.

Для моделирования и анализа беспроводных сетей используется аппарат теории массового обслуживания и математической теории телетрафика [304–310]. Этот аппарат развивается и адаптируется под новые особенности управления приоритетным доступом и обслуживанием пользователей. В том числе, за счет использования моделей стохастической геометрии. Приоритетные СМО учитывают различные варианты приоритетов [311–327]. Они позволяют исследовать мультисервисные модели с различными политиками доступа к обслуживающим приборам и очереди. В том числе, за счет вытеснения более приоритетными заявками менее приоритетных.

Для более удобного описания моделей, помимо текстовых формулировок и схем, используется кодирование СМО обозначениями Башарина – Кендалла в виде $A | B | \bullet | \bullet | f_i^j$. В этой системе используются следующие элементы: A входящий поток заявок, B обслуживание, третий и четвертый элементы – системы обслуживающих приборов и очередей, f_i^j дисциплина приоритетного обслуживания. Символика, используемая для обозначения приоритетов, введена профессором Башариным Г.П. в работах [328–330].

В рамках классификации Башарина – Кендалла индекс i обозначает дисциплину выбора заявки из очереди, т.е. фактически доступ к приборам: $i = 0$ без приоритета, $i = 1$ относительный приоритет, $i = 2$ абсолютный приоритет – вытеснение. Индекс j определяет дисциплину постановки заявки в очередь, т.е. доступ к очереди: $j = 0$ без приоритета, $j = 2$ абсолютный приоритет.

В некоторых работах Башарина Г.П. используются дополнительные индексы. Подробнее см. в приложении «Обзор приоритетных систем» монографии Геденко Б.В. [312]. В обозначении f^{0v} индекс v описывает дисциплину выбора заявки, которая будет потеряна при постановке поступившей заявки в очередь: $v = 0$ поступившая заявка, $v = 1$ заявка того же класса из очереди, которая ждет больше всех, $v = 2$ заявка того же класса из очереди, которая ждет меньше всех. Индекс w в f^{2w} описывает дисциплину выбора заявки, которая будет вытеснена из очереди: первая, последняя, любая. Индекс m в $_m f_2$ и $_m f^2$ описывает дисциплину,

определяющую, что делать с вытесненной заявкой из очереди или прибора. Оба индекса w и t применимы для абсолютного приоритета. Системы с возобновлением обслуживания – это например, системы с повторными вызовами или орбитами [331–350].

Приоритет заявок в системе определяется их классом, а классы заявок расположены в порядке приоритетов, т.е. приоритет заявки класса k_1 выше приоритета заявки класса k_2 , если $k_1 < k_2$. В работах Гнеденко Б.В. [312] под дополнительным приоритетным потоком понимается также поток отказов в СМО с ненадежными приборами. Фактически, помимо внутренних приоритетных событий, связанных с классами заявок, добавляются внешние события.

Для кодирования СМО с более гибким приоритетным доступом к приборам – не линейно упорядоченными приоритетами для классов заявок, а заданием приоритетов для каждой пары классов, Алиев Т.И. [351] предложил перейти от индекса i в f_i^j Башарина Г.П. к матрице приоритетов Q . Элементы этой матрицы q_{kl} соответствуют идентификатору занятия ресурса (pre-emption capability), о котором говорилось в разделе 1.1. Они описывают приоритет доступа заявок класса k к приборам относительно обслуживающихся заявок класса l : $q_{kl} = 0$ без приоритета, $q_{kl} = 1$ относительный приоритет, $q_{kl} = 2$ абсолютный приоритет.

Возможность учесть в кодировании СМО не только абсолютный приоритет, т.е. вытеснение заявки с приборов или прерывание обслуживания, но и снижение качества обслуживания, например, скорости, можно также отразить в матрице приоритетов. Автор в ранних работах [352–355] предложил использовать матрицу приоритетов P с элементами $p_{kl} = (d_{kl}, i_{kl})$, где d_{kl} и i_{kl} индикаторы возможности снижения скорости и прерывания обслуживания заявки класса l при поступлении заявки класса k . Предполагается, что сначала проверяется возможность снизить скорость, и только потом прерывается обслуживание.

Таким образом, можно выделить пять основных составляющих – компонент приоритетного доступа:

- (1) События приоритетного доступа – доступ заявок к приборам или очереди, отказ приборов.
- (2) Механизм реализации приоритета – без приоритета, относительный приоритет, абсолютный приоритет, снижение скорости обслуживания.
- (3) Траектория заявки после вытеснения.
- (4) Порядок применения механизмов – сначала снижение скорости, потом прерывание обслуживания.
- (5) Принцип выбора заявок для применения механизма – равновероятно, упорядоченно.

1.6.2 Матричная нотация приоритетов

В рамках диссертационной работы предлагается использовать модифицированную матрицу приоритетов $\mathbf{F}$ для учета особенностей приоритизации услуг в беспроводных сетях, которые были рассмотрены в разделах 1.1–1.5 [356,357]. Это позволит предоставить читателям дополнительную информацию о структуре схемы приоритетного доступа в каждой из исследуемых моделей. СМО будем кодировать в нотации Башарина – Кендалла как $A | B | \bullet | \bullet | \mathbf{F}$. В матрице $\mathbf{F}$ сохраним пять компонент, описанных выше, а также отразим различные виды приоритетных внешних событий, механизмы реализации приоритетов для разных типов ресурсов и приборов, включая группы ресурсов при нарезке сети. Кроме того, учтем возможность последовательного применения различных механизмов.

В строках матрицы $\mathbf{F}$ расположим события приоритетного доступа. Пусть $\mathcal{E} = \mathcal{E}_0 \cup \mathcal{E}_1$ множество случайных событий приоритетного доступа, где $\mathcal{E}_0$ события без приоритета, $\mathcal{E}_1$ приоритетные события. Примеры таких событий:

- $e = 1$ возникновение блокировки пути прямой видимости;
- $e = 2$ поступление сигнала на перераспределение ресурса;
- $e = 3$ получение доступа к ресурсу владельца лицензированных радиочастот;
- $e = 4$ поступление приоритетной заявки;
- $e = 5$ поступление неприоритетной заявки;
- $e = 6$ попытка начать обслуживание неприоритетной заявки;
- $e = 7$ попытка возобновить обслуживание неприоритетной заявки.

Отметим, что конкретный перечень событий приоритетного доступа фактически соотносится со списком событий, которые используются при дискретно-событийном стохастическом моделировании (discrete-event simulation, DES) системы. Напомним, что это внешние события, которые влияют на объем доступного ресурса для заявок разных классов, на объем занимаемого заявками ресурса, на доступ к ресурсу и т.д. Это также события, связанные с заявками разных классов, включая их состояние во время пребывания в системе, например заявки, которые возобновляют обслуживание.

В столбцах матрицы F расположим механизмы реализации приоритета с указанием, в случае абсолютного приоритета, траектории заявки после вытеснения. Пусть $\mathcal{P} = \bigcup_{e \in \mathcal{E}} \mathcal{P}_e$ множество приоритетов, где множества $\mathcal{P}_e$ могут пересекаться и определяют приоритеты для события $e \in \mathcal{E}$. Механизм $p \neq 0$ для приоритетного события определяет, каким образом осуществляется доступ к ресурсам и другим компонентам системы: относительный или абсолютный приоритеты. Примеры механизмов:

- $p = 0$ без приоритета;
- $p = 1$ изменение пропускной способности первого сегмента сети;
- $p = 2$ изменение пропускной способности второго сегмента сети;
- $p = 3$ уменьшение числа ресурсных блоков;
- $p = 4$ уменьшение мощности сигнала базовой станции;
- $p = 5$ снижение мощности сигнала для неприоритетного трафика;
- $p = 6$ снижение скорости передачи неприоритетного трафика;

$p = 7$ прерывание обслуживания неприоритетного трафика.

Под структурой системы будем понимать компоненты, к которым организован приоритетный доступ. Например, это могут быть места для обслуживания заявок различных классов (ресурсы), места для ожидания начала обслуживания, места для ожидания возобновления обслуживания, включая объем этих мест. Тип ресурса зависит от модели и может включать в себя ресурсные блоки, пропускную способность, мощность сигнала, диапазон радиочастот или длительность кадра. Абсолютный приоритет может быть реализован путем снижения или освобождения объема занимаемого ресурса.

В элементах матрицы приоритетов

$$\mathbf{F}: f_{ep} = (n_{ep}, w_{ep}), \quad e \in \mathcal{E}, \quad p \in \mathcal{P}, \quad (1.34)$$

представим остальные две компоненты приоритетного доступа. Индекс n_{ep} обозначает порядковый номер применения механизма:

$n_{ep} = 0$ нет приоритета;

$n_{ep_1} < n_{ep_2}$ сначала применяется механизм p_1 , а затем механизм p_2 ;

$n_{ep_1} = n_{ep_2}$ механизмы p_1 и p_2 применяются одновременно.

Элемент w_{ep} описывает принцип выбора заявок для применения механизма, например:

$w_{ep} = 1$ одна сессия равновероятно;

$w_{ep} = 2$ все сессии;

$w_{ep} = 3$ одна сессия с максимальной скоростью;

$w_{ep} = 4$ одна сессия с минимальной скоростью.

Таким образом, в диссертационной работе матрицы приоритетов $\mathbf{F}$ дополняют описания исследуемых моделей беспроводных сетей. Матрица $\mathbf{F}$ даст читателю представление обо всех компонентах приоритизации в модели.

Проиллюстрируем вид матрицы приоритетов **F** на примере. Рассмотрим систему $M_2 | M_2 | C | 0 | f_2$ с абсолютным приоритетом доступа к ресурсным блокам. Если абсолютный приоритет достигается за счет прерывания обслуживания неприоритетного трафика, то матрица приоритетов не несет дополнительной информации по сравнению с обозначением f_2 :

F			неприоритетный трафик обслуживается прерывание
			$p = 1$
приоритетный трафик	поступление	$e = 1$	$(n_{11} = 1, w_{11} = 1)$
	завершение обслуживания	$e = 2$	0
неприоритет. трафик	поступление	$e = 3$	0
	завершение обслуживания	$e = 4$	0

Однако если схема доступа усложняется, матрица **F** может расширяться, отражая особенности приоритетного доступа. Например, при нехватке ресурса для обслуживания приоритетного трафика сначала можно попытаться снизить скорость неприоритетного трафика до заданного порога. Если этого недостаточно, то обслуживание сессии прерывается. После завершения обслуживания как приоритетного, так и неприоритетного трафика скорость передачи трафика может быть увеличена:

F			неприоритетный трафик обслуживается		
			сниж. скор.	прерывание	повыш. скор.
			$p = 1$	$p = 2$	$p = 3$
приор. трафик	поступ.	$e = 1$	$n_{11} = 1$	$(n_{12} = 2, w_{12} = 1)$	0
	заверш. обслуж.	$e = 2$	0	0	$n_{23} = 1$
непр. трафик	поступ.	$e = 3$	0	0	0
	заверш. обслуж.	$e = 4$	0	0	$n_{43} = 1$

Отметим, что анализ моделей с данными матрицами приоритетов изложен в главе 2.

1.6.3 Комплекс моделей приоритетного доступа

Исходя из изложенного в первой главе, объектом исследования является беспроводная сеть 4/5G с приоритизацией услуг, механизмами доступа к ресурсам и блокировками пути прямой видимости. Предметом исследования являются модели приоритетного доступа к радиоресурсам сетей 4/5G в условиях обслуживания эластичного и потокового трафика.

1. В качестве услуг рассматриваются сверхнадежная передача данных с малой задержкой URLLC и широкополосная связь (enhanced mobile broadband, eMBB), видео по требованию и видеоконференция (глава 2).
2. Механизмы доступа к ресурсам исследуются в условиях нарезки радиоресурсов и совместного доступа к нелицензированным (глава 3) и к лицензированным радиочастотам (глава 5).
3. Блокировки пути прямой видимости рассматриваются для сети миллиметрового диапазона (mmWave) (глава 4).

Безусловно, существует множество механизмов доступа и обслуживания эластичного и потокового трафика в беспроводных сетях, специфичных для различных классов услуг. Диссертационная работа не нацелена исследовать все возможные варианты приоритизации, охватив при этом одновременно все перечисленные особенности сетей. Задача состоит в том, чтобы отразить в моделях математической теории телетрафика ключевые особенности приоритетного доступа для каждого сценария отдельно. Это позволит проанализировать влияние этих особенностей на обслуживание эластичного и потокового трафика с помощью различных показателей эффективности приоритетного доступа.

Схемы приоритетного доступа для моделей с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса (глава 2, разделы 2.1–2.4):

- (а) прерывание обслуживания широкополосного трафика [358];
- (б) приостановка обслуживания широкополосного трафика [359,360];
- (в) снижение мощности и прерывание обслуживания широкополосного трафика [361];
- (г) снижение мощности и скорости передачи широкополосного трафика [362,363].

Автор исследовал и другие схемы, изложенные в работах [364,365,366,367].

Схемы приоритетного доступа для моделей с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса (глава 2, разделы 2.5–2.6):

- (а) прерывание обслуживания одноадресного трафика и одна дисциплина обслуживания многоадресного трафика [368,369];
- (б) прерывание обслуживания одноадресного трафика и две дисциплины обслуживания многоадресного трафика [370];
- (в) прерывание обслуживания одноадресного трафика и снижение скорости передачи многоадресного трафика [371].

Схемы приоритетного доступа для моделей с эластичным трафиком и полнодоступным ресурсом постоянного объема (глава 3, разделы 3.1–3.2):

- (а) снижение скорости передачи трафика [372];
- (б) прерывание обслуживания трафика [372,373].

Схемы приоритетного доступа для моделей с эластичным трафиком и перераспределением ресурса постоянного объема (глава 3, разделы 3.3–3.6):

- (а) фиксированная политика перераспределения ресурса между классами трафика [374,375];
- (б) политика управления перераспределением ресурса между классами трафика [376,377,378,379,380,381].

Автор исследовал и другие схемы, изложенные в работах [382, 383].

Схемы приоритетного доступа для моделей с эластичным трафиком и полным отказом изменяемого объема ресурса (глава 4, разделы 4.1–4.5):

- (а) приостановка обслуживания трафика и пуассоновский входящий поток [384];
- (б) приостановка обслуживания трафика и марковский входящий поток [385,386].

Схемы приоритетного доступа для моделей с эластичным трафиком и частичным отказом изменяемого объема ресурса (глава 4, разделы 4.1, 4.6):

- (в) снижение скорости передачи трафика и марковский входящий поток [387].

Схемы приоритетного доступа для моделей с потоковым трафиком и полным отказом изменяемого объема ресурса (глава 5, разделы 5.1–5.3):

- (а) приостановка обслуживания трафика без отказов ресурса на периоде простоя системы [388];
- (б) приостановка обслуживания трафика с отказами ресурса на периоде простоя системы [389].

Схемы приоритетного доступа для моделей с потоковым трафиком и частичным отказом изменяемого объема ресурса (глава 5, разделы 5.1, 5.4–5.6):

- (в) приостановка обслуживания трафика [390];
- (г) приостановка обслуживания трафика и гистерезисный доступ [391];
- (в) снижение мощности и скорости передачи трафика [392,393,394];
- (г) постоянная мощность и прерывание обслуживания трафика [395,396].

Автор исследовал и другие схемы, изложенные в работах [397,398,399,400].

Структура работы представлена в таблице Таблица 1.3. В работе используется сквозная система основных обозначений, перечисленных в таблице 1.4., механизмов реализации приоритета (см. таблицу 1.5) и принципов выбора заявок для применения механизмов (см. таблицу 1.6). Отметим также, что матрицы **F** приоритетов перенумерованы в порядке встречаемости.

Дополнительные списки обозначений можно найти в каждой главе. Основные результаты работы сформулированы в виде теорем со вспомогательными леммами.

Таблица 1.3 – Структура разделов работы

	Потоковый трафик	Эластичный трафик
Постоянный объем ресурса	Узкополосный и широкополосный трафик (глава 2, разделы 2.1–2.4) Одноадресный и многоадресный трафик (глава 2, разделы 2.5–2.6)	Полнодоступный ресурс (глава 3, разделы 3.1–3.2) Перераспределение ресурса (глава 3, разделы 3.3–3.6)
Изменяемый объем ресурса	Полный отказ ресурса (глава 5, разделы 5.1–5.3) Частичный отказ ресурса (глава 5, разделы 5.1, 5.4-5.6)	Полный отказ ресурса (глава 4, разделы 4.1–4.5) Частичный отказ ресурса (глава 4, разделы 4.1, 4.6)

Таблица 1.4 – Обозначения, используемые в работе

Обозначение	Параметр
λ	Интенсивность пуассоновского / марковского потока запросов на передачу трафика
$1/\mu$	Среднее экспоненциальное время передачи потокового трафика / средний экспоненциальный объем эластичного трафика
b	Скорость передачи трафика
V	Объем надежного ресурса – число надежных приборов
C	Объем ненадежного ресурса – число ненадежных приборов
α	Интенсивность пуассоновского потока отказов ненадежного ресурса
β	Интенсивность пуассоновского потока восстановлений ненадежного ресурса
R	Число мест в очереди / на орбите
F	Диапазон радиочастот
$T = 1$	Длительность кадра
P	Максимальная мощность передающей антенны
G_t	Коэффициент усиления передающей антенны
G_r	Коэффициент усиления приемной антенны
PL	Мощность затухания сигнала
N_0	Мощность шума
I_0	Мощность интерференции

Обозначение	Параметр
y_f	Требование к ширине радиочастот
y_t	Требование к доле длительности кадра
y_p	Требование к мощности передающей антенны
$y = (y_f, y_t, y_p)$	Требование к ресурсу
$X(t)$	Случайный процесс
x	Состояние системы
$\mathcal{X}$	Пространство состояний системы
$\pi(x)$	Стационарная вероятность состояния системы
Q	Матрица интенсивностей переходов

Таблица 1.5 – Механизмы реализации приоритета, используемые в работе

Обозначение	Механизм
$p = 1$	Прерывание обслуживания
$p = 2$	Приостановка обслуживания
$p = 3$	Снижение скорости передачи
$p = 4$	Повышение скорости передачи
$p = 5$	Уменьшение доли продолжительности кадра
$p = 6$	Увеличение доли продолжительности кадра
$p = 7$	Снижение мощности сигнала
$p = 8$	Повышение мощности сигнала
$p = 9$	Относительный приоритет
$p = 10$	Уменьшение пропускной способности
$p = 11$	Увеличение пропускной способности
$p = 12$	Пропускная способность недоступна
$p = 13$	Пропускная способность доступна
$p = 14$	Перемещение (стягивание) на надежный ресурс

Таблица 1.6 – Принципы выбора заявок для применения механизмов, используемые в работе

Обозначение	Принцип
$w = 1$	Одна сессия равновероятно
$w = 2$	Все сессии
$w = 3$	Одна сессия с максимальной скоростью
$w = 4$	Одна сессия с минимальной скоростью
$w = 5$	Минимально необходимое число сессий равновероятно
$w = 5$	Максимальное число сессий равновероятно
$w = 6$	С заданной вероятностью
$w = 7$	С заданной вероятностью все сессии

Глава 2 МОДЕЛИ С ПОТОКОВЫМ ТРАФИКОМ И ПОСТОЯННЫМ ОБЪЕМОМ РЕСУРСА

2.1 МОДЕЛЬ СВЕРХНАДЕЖНОЙ ПЕРЕДАЧИ С МАЛОЙ ЗАДЕРЖКОЙ И ШИРОКОПОЛОСНОЙ СВЯЗИ

2.1.1 Системная модель

В разделах 2.1–2.4 главы 2 исследованы модели приоритетного доступа двух классов одноадресного потокового трафика, генерируемого при сверхнадежной передаче данных с малой задержкой URLLC и широкополосной связи eMBB. Приоритетным является трафик URLLC, объем ресурса постоянный. Рассмотрим сценарии, где передача трафика URLLC занимает существенно меньше времени, чем передача трафика eMBB. Поскольку объем данных URLLC невелик, для обеспечения малой задержки передачи достаточно низкой скорости. В то же время трафик eMBB характеризуется высокой скоростью. Поэтому трафик URLLC будем называть узкополосным, а трафик eMBB – широкополосным.

Рассмотрим несколько сценариев обслуживания широкополосного трафика. В первом случае передача трафика не допускает задержек, а во втором – допускает. Скорость передачи широкополосного трафика может быть постоянной или изменяться по уровням скорости. Для обеспечения приоритета узкополосного трафика применяются различные механизмы доступа широкополосного трафика: прерывание и приостановка обслуживания, снижение скорости передачи и мощности сигнала.

В таблице 2.1 приведены основные обозначения, параметры и компоненты состояния системы. Для моделирования механизма приостановки обслуживания используется СМО с двумя орбитами, а для механизма снижения мощности сигнала – РеСМО с двумя типами ресурсов.

Таблица 2.1 – Обозначения, используемые для моделей с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса (глава 2, разделы 2.1–2.4)

Обозначение	Параметр
λ_1	Интенсивность пуассоновского потока запросов на передачу узкополосного трафика
λ_2	Интенсивность пуассоновского потока запросов на передачу широкополосного трафика
$d = 1$	Скорость передачи узкополосного трафика
K	Число уровней скорости, на которых может передаваться широкополосный трафик
$b = b_1 > \dots > b_k > \dots > b_K,$ $b_k = b_{k-1} - 1, k = 2, \dots, K$	Скорости передачи широкополосного трафика
$1/\mu_1$	Среднее экспоненциальное время передачи узкополосного трафика
$1/\mu_2$	Среднее экспоненциальное время передачи широкополосного трафика
$M = \lfloor V/b \rfloor = \lfloor V/b_1 \rfloor$	Максимальное число обслуживаемых сессий широкополосного трафика
$R_1 = M$	Число мест на орбите для приостановленных сессий широкополосного трафика
$R_2 \geq M$	Число мест на орбите для ожидающих начала обслуживания сессий широкополосного трафика
$1/\gamma_1$	Среднее экспоненциальное время между повторными попытками возобновить обслуживание сессии широкополосного трафика
$1/\gamma_2$	Среднее экспоненциальное время между повторными попытками начать обслуживание сессии широкополосного трафика
n	Число обслуживаемых сессий узкополосного трафика
$m = (m_1, \dots, m_K)$	Число обслуживаемых сессий широкополосного трафика
q	Число приостановленных сессий широкополосного трафика
r	Число ожидающих начала обслуживания сессий широкополосного трафика

2.1.2 Схемы приоритетного доступа

Рассмотрим следующие схемы приоритетного доступа для моделей с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса:

- (а) прерывание обслуживания широкополосного трафика;
- (б) приостановка обслуживания широкополосного трафика;
- (в) снижение мощности и прерывание обслуживания широкополосного трафика;
- (г) снижение мощности и скорости передачи широкополосного трафика.

В этих схемах используются механизмы $p \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \subset \mathcal{P}$ реализации приоритета и принципы $w \in \{1, 2, 3, 4\}$ выбора сессии для применения механизма из параграфа 1.6.3.

Опишем особенности исследуемых моделей с перечисленными схемами приоритетного доступа через матрицы приоритетов.

Схема (а) – базовая система с одним механизмом приоритетного доступа – прерыванием обслуживания широкополосного трафика. Это СМО $M_2 \mid M_2 \mid V \mid 0 \mid F_1$ с матрицей приоритетов

F_1				широкополос. трафик обслуживается
		$e \backslash p$		1
узкополос. трафик	поступление	1		(1,1)
	завершение обслуживания	2		0
широкополос. трафик	поступление	3		0
	завершение обслуживания	4		0

Схема (б). В этой системе широкополосный трафик может находиться в трех состояниях: обслуживания, ожидания возобновления обслуживания и ожидания начала обслуживания. При этом попытки возобновить обслуживание имеют относительный приоритет перед попытками начать обслуживание. Это СМО $M_2 \mid M_2 \mid V \mid R_1, R_2 \mid F_2$ с матрицей приоритетов

F_2			широкополос. трафик	
			обслуживается	ожидает начала обслуживания
		$\begin{matrix} p \\ e \end{matrix}$	2	9
узкополосный трафик	поступление	1	(1,1)	0
	завершение обслуживания	2	0	0
широкополосный трафик	поступление	3	0	0
	попытка возобновить обслуживание	4	0	1
	попытка начать обслуживание	5	0	0
	завершение обслуживания	6	0	0

Схема (в). Для третьей системы со снижением мощности и прерыванием обслуживания широкополосного трафика в качестве модели требования к ресурсу выберем политику снижения мощности сигнала RR (см. раздел 1.3). Эта политика предполагает равное разделение кадра между всеми сессиями и без разделения ресурса по частоте. Равное или эгалитарное разделение кадра RR напоминает дисциплину EPS для эластичного трафика. Мощность сигнала устанавливается таким образом, чтобы обеспечить необходимую скорость $d=1$ передачи узкополосного трафика и скорость b передачи широкополосного трафика для каждой сессии. Это СМО $M_2 \mid M_2 \mid V, \langle T, P \rangle \mid 0 \mid F_3$ с матрицей приоритетов

F_3			узкополосный трафик		широкополосный трафик		
			обслуживается		обслуживается		
		$\begin{matrix} p \\ e \end{matrix}$	5 ↓ 8	6 ↓ 7	5 ↓ 8	1	6 ↓ 7
узкополосный трафик	поступление	1	(1,2)	0	(1,2)	(2,1)	0
	завершение обслуживания	2	0	(1,2)	0	0	(1,2)
широкополосный трафик	поступление	3	(1,2)	0	(1,2)	0	0
	завершение обслуживания	4	0	(1,2)	0	0	(1,2)

Отметим, что в модели применение одного механизма приоритетного доступа может автоматически повлечь за собой применение второго механизма. В таких случаях для наглядности в описании столбцов с механизмами p используется символ $\Downarrow$. Например, уменьшение доли длительности кадра $p = 5$ влечет за собой повышение мощности сигнала $p = 8$.

Схема (г). Последняя система – со снижением мощности и скорости передачи широкополосного трафика. Применена политика снижения мощности сигнала WRR (weighted RR) с дискриминаторным разделением кадра между всеми сессиями и без разделения ресурса по частоте. Дискриминаторное разделение кадра WRR напоминает дисциплину DPS для эластичного трафика с весовыми коэффициентами, соответствующим скоростям b_k широкополосного трафика. Мощность сигнала для узкополосного трафика постоянна и равна P , а для широкополосного трафика устанавливается так, чтобы обеспечить необходимую скорость b передачи для каждой сессии. Это СМО $M_2 \mid M_2 \mid V, \langle T, P \rangle \mid 0 \mid F_4$ с матрицей приоритетов

F_4		$e \backslash p$	узкополосный трафик обслуживается		широкополосный трафик обслуживается				
			5 $\Downarrow$ 3	6 $\Downarrow$ 4	5 $\Downarrow$ 8	3 $\Downarrow$ 5	1	6 $\Downarrow$ 7	4 $\Downarrow$ 6
узкополосный трафик	поступление	1	(1,2)	(3,2)	(1,2)	(2,3)	(3,1)	(3,2)	0
	завершение обслуживания	2	0	(2,2)	0	0	0	(2,2)	(1,4)
широкополосный трафик	поступление	3	(1,2)	0	(1,2)	0	0	0	0
	завершение обслуживания	4	0	(1,2)	0	0	0	(1,2)	0

Автор исследовал и другие схемы приоритетного доступа для моделей с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса. Однако чтобы сохранить относительную соразмерность глав диссертационной работы, эти схемы не включены в текст работы:

- (д) снижение скорости передачи всех сессий и прерывание обслуживания широкополосного трафика;
- (е) снижение скорости передачи одной из сессий и прерывание обслуживания широкополосного трафика.

2.1.3 Модели математической теории телетрафика

Функционирование систем опишем при помощи СП $X(t)$ над пространствами состояний:

модели (а) и (в)

$$\mathcal{X} = \{(n, m): n \geq 0, m \geq 0, n + bm \leq V\}; \quad (2.1)$$

модель (б)

$$\mathcal{X} = \{(n, m, q, r): n \geq 0, m \geq 0, 0 \leq q \leq R_1, 0 \leq r \leq R_2, m + q \leq M, n + bm \leq V\}; \quad (2.2)$$

модель (г)

$$\mathcal{X} = \{(n, m_1, \dots, m_K): n \geq 0, m_k \geq 0, k = 1, \dots, K, \sum_{k=1}^K m_k \leq M, v(n, m_1, \dots, m_K) \leq V\}, \quad (2.3)$$

где $v(n, m_1, \dots, m_K) = v(\mathbf{x})$ суммарное требование к скорости передачи узкополосного и широкополосного трафика в состоянии $\mathbf{x} = (n, m_1, \dots, m_K)$

$$v(\mathbf{x}) = v(n, m_1, \dots, m_K) = n + \sum_{k=1}^K m_k b_k, \quad \mathbf{x} = (n, m_1, \dots, m_K) \in \mathcal{X}. \quad (2.4)$$

Ненулевые элементы матриц интенсивностей переходов имеют вид

модель (а)

$$\begin{aligned} Q[(n, m), (n+1, m)] &= \lambda_1, \quad n+1+bm \leq V, \\ Q[(n, m), (n+1, m-1)] &= \lambda_1, \quad n+1+bm > V, \quad m > 0, \\ Q[(n, m), (n, m+1)] &= \lambda_2, \quad n+b(m+1) \leq V, \\ Q[(n, m), (n-1, m)] &= n\mu_1, \quad n > 0, \\ Q[(n, m), (n, m-1)] &= m\mu_2, \quad m > 0; \end{aligned} \quad (2.5)$$

модель (б)

$$\begin{aligned}
 Q[(n, m, q, r), (n+1, m, q, r)] &= \lambda_1, \quad n+1+bm \leq V, \\
 Q[(n, m, q, r), (n+1, m-1, q+1, r)] &= \lambda_1, \quad n+1+bm > V, \quad m > 0, \\
 Q[(n, m, 0, 0), (n, m+1, 0, 0)] &= \lambda_2, \quad n+b(m+1) \leq V, \\
 Q[(n, m, q, r), (n, m, q, r+1)] &= \lambda_2, \quad n+b(m+1) \leq V, \quad (q > 0 \vee r > 0), \quad r < R_2, \\
 Q[(n, m, q, r), (n, m, q, r+1)] &= \lambda_2, \quad n+b(m+1) > V, \quad r < R_2, \\
 Q[(n, m, q, r), (n-1, m, q, r)] &= n\mu_1, \quad n > 0, \\
 Q[(n, m, q, r), (n, m-1, q, r)] &= m\mu_2, \quad m > 0, \\
 Q[(n, m, q, r), (n, m+1, q-1, r)] &= \gamma_1, \quad q > 0, \quad n+b(m+1) \leq V, \\
 Q[(n, m, 0, r), (n, m+1, 0, r-1)] &= \gamma_2, \quad r > 0, \quad n+b(m+1) \leq V;
 \end{aligned} \tag{2.6}$$

модель (в)

$$\begin{aligned}
 Q[(n, m), (n+1, m)] &= \lambda_1 \cdot p(n, m), \quad n+1+bm \leq V, \\
 Q[(n, m), (n+1, m-1)] &= \lambda_1 \cdot p(n, m), \quad n+1+bm > V, \quad m > 0, \\
 Q[(n, m), (n, m+1)] &= \lambda_2 \cdot p(n, m), \quad n+b(m+1) \leq V, \\
 Q[(n, m), (n-1, m)] &= n\mu_1, \quad n > 0, \\
 Q[(n, m), (n, m-1)] &= m\mu_2, \quad m > 0;
 \end{aligned} \tag{2.7}$$

модель (г)

$$\begin{aligned}
 Q[(n, \mathbf{m}), (n+1, \mathbf{m})] &= \lambda_1 \cdot p_1(n, \mathbf{m}), \quad v(\mathbf{x})+1 \leq V, \\
 Q[(n, \mathbf{m}), (n+1, \mathbf{m} - \mathbf{e}_k + \mathbf{e}_{k+1})] &= \lambda_1 \cdot p_1(n, \mathbf{m}), \\
 v(\mathbf{x})+1 > V, \quad \sum_{i=1}^{k-1} m_i &= 0, \quad m_k > 0, \quad k = 1, \dots, K-1, \\
 Q[(n, \mathbf{m}), (n+1, \mathbf{m} - \mathbf{e}_K)] &= \lambda_1 \cdot p_1(n, \mathbf{m}), \quad v(\mathbf{x})+1 > V, \quad \sum_{k=1}^{K-1} m_k = 0, \quad m_K > 0, \\
 Q[(n, \mathbf{m}), (n, \mathbf{m} + \mathbf{e}_1)] &= \lambda_2 \cdot p_2(n, \mathbf{m}), \quad v(\mathbf{x})+b_1 \leq V, \\
 Q[(n, \mathbf{m}), (n-1, \mathbf{m} - \mathbf{e}_k + \mathbf{e}_{k-1})] &= n\mu_1, \quad n > 0, \quad m_k > 0, \quad \sum_{i=k+1}^K m_i = 0, \quad k = 2, \dots, K, \\
 Q[(n, \mathbf{m}), (n-1, \mathbf{m})] &= n\mu_1, \quad n > 0, \quad \sum_{k=2}^K m_k = 0, \\
 Q[(n, \mathbf{m}), (n, \mathbf{m} - \mathbf{e}_k)] &= m_k \mu_2, \quad m_k > 0.
 \end{aligned} \tag{2.8}$$

Напомним, что в параграфе 1.3.3 была определена вероятность $p(\bullet)$ того, что еще одной сессии может быть выделен ресурс, при условии обслуживания текущего числа сессий – см. формулу (1.30). В рассматриваемых ресурсных системах аналогичные вероятности также будут участвовать в интенсивностях переходов СП – там, где при переходе в новое состояние возникает увеличенная потребность в ресурсе. Это происходит тогда, когда поступает новый запрос на передачу трафика. Определение и формулы для расчета вероятностей $p(\bullet)$ будут приведены ниже в разделе 2.3.

В следующих разделах 2.2–2.4 описаны результаты исследования моделей схем приоритетного доступа с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса. Матрицы интенсивностей переходов представлены в блочно-трехдиагональном виде. Это позволяет рассчитать стационарное распределение с помощью матричного алгоритма для моделей (а) и (б) (раздел 2.2). В качестве альтернативы матричному расчету стационарного распределения предложен алгоритм снижения размерности СУР в скалярном виде для модели (а) (раздел 2.2).

Для ресурсных систем (в) и (г) найдены вероятности занятия ресурса и условные вероятности выделения ресурса (раздел 2.3). Приведены показатели эффективности приоритетного доступа: вероятности прерывания и приостановки обслуживания, средней скорости передачи трафика (раздел 2.4). Проиллюстрировано влияние на показатели параметра потока попыток возобновления обслуживания широкополосного трафика для модели (б), уровней скорости широкополосного трафика для модели (г), моделей затухания сигнала для моделей (в) и (г).

2.2 СИСТЕМА С УЗКОПОЛОСНЫМ И ШИРОКОПОЛОСНЫМ ТРАФИКОМ

2.2.1 Матрица интенсивностей переходов в блочно-трехдиагональном виде

Лемма 2.1. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания широкополосного трафика, если на пространстве $\mathcal{X}$ состояний СП $X(t)$ введен лексикографический порядок

$$\begin{aligned} x = (n, m) < (n', m') = x' &\Leftrightarrow \\ (m < m') \vee \\ \vee (m = m', n < n'), \quad x, x' \in \mathcal{X}, \end{aligned} \quad (2.9)$$

тогда матрица интенсивностей переходов СП $X(t)$ представима в блочно-трехдиагональном виде

$$\begin{aligned} Q = \text{diag}(Q_{00}, Q_{01}, \dots, Q_{0M}) + \\ + \text{diag}^+(Q_{11}, Q_{12}, \dots, Q_{1M}) + \text{diag}^-(Q_{21}, Q_{22}, \dots, Q_{2M}), \end{aligned} \quad (2.10)$$

где ненулевые недиагональные элементы блоков определяются по формулам

$$\begin{aligned} Q_{0m}[(n, m), (n', m')] = \begin{cases} \lambda_1, & n' = n + 1, m' = m, \\ n\mu_1, & n' = n - 1, m' = m, \end{cases} \\ \dim Q_{0m} = (V - bm + 1) \times (V - bm + 1), \quad m = 0, \dots, M, \end{aligned} \quad (2.11)$$

$$\begin{aligned} Q_{1m}[(n, m), (n', m')] = \lambda_2, \quad n' = n, m' = m + 1, \\ \dim Q_{1m} = (V - b(m - 1) + 1) \times (V - bm + 1), \quad m = 1, \dots, M, \end{aligned} \quad (2.12)$$

$$\begin{aligned} Q_{2m}[(n, m), (n', m')] = \begin{cases} \lambda_1, & n' = n + 1, m' = m - 1, n + bm = V, \\ m\mu_2, & n' = n, m' = m - 1, \end{cases} \\ \dim Q_{2m} = (V - bm + 1) \times (V - b(m - 1) + 1), \quad m = 1, \dots, M. \end{aligned} \quad (2.13)$$

Доказательство.

а) Введение лексикографического порядка

Разобьем пространство $\mathcal{X}$ состояний (2.1) СП $X(t)$ на подмножества по числу m обслуживаемых сессий широкополосного трафика

$$\mathcal{X} = \bigcup_{m=0}^M \mathcal{X}(m),$$

$$\mathcal{X}(m) = \{(n, i) \in \mathcal{X} : i = m\} = \{(n, m) \in \mathcal{X}\} = \{(0, m), (1, m), \dots, (n, m), \dots, (V - bm, m)\}.$$

Введем лексикографический порядок на множестве $\mathcal{X}$ в порядке объединения $\mathcal{X}(m)$

$$\begin{aligned} \mathcal{X} = & \{(0, 0), (1, 0), \dots, (n, 0), \dots, (V, 0), \\ & (0, 1), (1, 1), \dots, (n, 1), \dots, (V - b, 1), \\ & \dots, \\ & (0, m), (1, m), \dots, (n, m), \dots, (V - bm, m), \\ & \dots, \\ & (0, M - 1), (1, M - 1), \dots, (n, M - 1), \dots, (b, M - 1), \\ & (0, M)\}. \end{aligned}$$

б) Описание блоков матрицы

С учетом введенного лексикографического порядка и ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (2.5), матрица $\mathbf{Q}$ имеет блочно-трехдиагональный вид

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{00} & \mathbf{Q}_{11} & & & & & \\ \mathbf{Q}_{21} & \mathbf{Q}_{01} & \mathbf{Q}_{12} & & & & \\ & \mathbf{Q}_{22} & \ddots & \ddots & & & \\ & & \ddots & \ddots & \mathbf{Q}_{1m} & & \\ & & & \mathbf{Q}_{2m} & \mathbf{Q}_{0m} & \mathbf{Q}_{1,m+1} & \\ & & & & \mathbf{Q}_{2,m+1} & \ddots & \ddots \\ & & & & & \ddots & \ddots & \mathbf{Q}_{1M} \\ & & & & & & \mathbf{Q}_{2M} & \mathbf{Q}_{0M} \end{bmatrix}.$$

Блоки матрицы $\mathbf{Q}$ имеют следующую структуру

$$\mathbf{Q}_{1m} = \begin{bmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_2 \\ \hline 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}, \quad m=1, \dots, M,$$

$$\mathbf{Q}_{2m} = \begin{bmatrix} m\mu_2 & 0 & & & 0 & \cdots & 0 \\ & \ddots & \ddots & & \vdots & & \vdots \\ & & m\mu_2 & 0 & \vdots & & \vdots \\ & & & m\mu_2 & \lambda_2 & 0 & \cdots & 0 \\ & & & & & 0 & \cdots & 0 \end{bmatrix}, \quad m=1, \dots, M,$$

$$\mathbf{Q}_{00} = \begin{bmatrix} -(\lambda_1 + \lambda_2) & \lambda_1 & & & \\ \mu_1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \lambda_1 \\ & & & V\mu_1 & -V\mu_1 \end{bmatrix},$$

$$\mathbf{Q}_{0m} = \begin{bmatrix} -(\lambda_1 + \lambda_2 + m\mu_2) & \lambda_1 & & & \\ \mu_1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \lambda_1 \\ & & & V\mu_1 & -(\lambda_2 + V\mu_1 + m\mu_2) \end{bmatrix}, \quad m=1, \dots, M-1,$$

$$\mathbf{Q}_{0M} = \begin{bmatrix} -(\lambda_1 + M\mu_2) & \lambda_1 & & & \\ \mu_1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \lambda_1 \\ & & & V\mu_1 & -(\lambda_2 + V\mu_1 + M\mu_2) \end{bmatrix}.$$

■

Лемма 2.2. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой с приостановкой обслуживания широкополосного трафика, если на пространстве $\mathcal{X}$ состояний СП $X(t)$ введен лексикографический порядок

$$\begin{aligned} \mathbf{x} = (n, m, q, r) < (n', m', q', r') = \mathbf{x}' \Leftrightarrow \\ (r < r') \vee \\ \vee (r = r', q < q') \vee \\ \vee (r = r', q = q', m < m') \vee \\ \vee (r = r', q = q', m = m', n < n'), \quad \mathbf{x}, \mathbf{x}' \in \mathcal{X}, \end{aligned} \quad (2.14)$$

тогда матрица интенсивностей переходов СП $X(t)$ представима в блочно-трехдиагональном виде

$$\begin{aligned} \mathbf{Q} = \text{diag}(\mathbf{Q}_{00}, \mathbf{Q}_{01}, \dots, \mathbf{Q}_{0R_2}) + \\ + \text{diag}^+(\mathbf{Q}_{11}, \mathbf{Q}_{12}, \dots, \mathbf{Q}_{1R_2}) + \text{diag}^-(\mathbf{Q}_{21}, \mathbf{Q}_{22}, \dots, \mathbf{Q}_{2R_2}), \end{aligned} \quad (2.15)$$

где ненулевые недиагональные элементы блоков определяются по формулам

$$\begin{aligned} \mathbf{Q}_{0r}[(n, m, q, r), (n', m', q', r')] = \\ = \begin{cases} \lambda_1, & (n' = n + 1, m' = m, q' = q, r' = r) \vee \\ & \vee (n' = n + 1, m' = m - 1, q' = q + 1, r' = r, n + 1 + bm > V), \\ \lambda_2, & n' = n, m' = m + 1, q' = q = 0, r' = r = 0, \\ n\mu_1, & n' = n - 1, m' = m, q' = q, r' = r, \\ m\mu_2, & n' = n, m' = m - 1, q' = q, r' = r, \\ \gamma_1, & n' = n, m' = m + 1, q' = q - 1, r' = r, \end{cases} \end{aligned} \quad (2.16)$$

$$\begin{aligned} \mathbf{Q}_{1r}[(n, m, q, r), (n', m', q', r')] = \lambda_2, \\ (n' = n, m' = m, q' = q, r' = r + 1) \wedge \\ \wedge ([n + b(m + 1) \leq V, (q > 0 \vee r > 0)] \vee [n + b(m + 1) > V]), \end{aligned} \quad (2.17)$$

$$\begin{aligned} \mathbf{Q}_{2r}[(n, m, q, r), (n', m', q', r')] = \gamma_2, \quad n' = n, m' = m, q' = q = 0, r' = r - 1, \\ \dim \mathbf{Q}_{ir} = \sum_{q=0}^{R_1} \sum_{m=0}^{M-q} (V - bm + 1) \times \sum_{q=0}^{R_1} \sum_{m=0}^{M-q} (V - bm + 1), \end{aligned} \quad (2.18)$$

$$i = 0, r = 0, \dots, R_2, \quad i = 1, 2, r = 1, \dots, R_2.$$

Доказательство.

Разобьем пространство $\mathcal{X}$ состояний (2.2) СП $X(t)$ на подмножества по числу r ожидающих начала обслуживания сессий широкополосного трафика

$$\mathcal{X} = \bigcup_{r=0}^{R_2} \mathcal{X}(r), \quad \mathcal{X}(r) = \{(n, m, q, i) \in \mathcal{X} : i = r\} = \{(n, m, q, r) \in \mathcal{X}\}.$$

Введем лексикографический порядок на множестве $\mathcal{X}$ в порядке объединения $\mathcal{X}(r)$, тогда с учетом ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (2.6), матрица $\mathbf{Q}$ имеет блочно-трехдиагональный вид (2.15).

■

2.2.2 Матричный алгоритм расчета стационарного распределения

Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания широкополосного трафика, в соответствии с лексикографическим порядком (2.9), представим стационарное распределение СП $X(t)$ в виде

$$\begin{aligned} \pi &= (\pi_0, \pi_1, \dots, \pi_m, \dots, \pi_M), \\ \pi_m &= (\pi(0, m), \pi(1, m), \dots, \pi(n, m), \dots, \pi(V - bm, m)). \end{aligned} \tag{2.19}$$

Тогда с учетом леммы 2.1, справедлива следующая теорема.

Теорема 2.1. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания широкополосного трафика, стационарное распределение СП $X(t)$ рассчитывается в матричном виде по формуле

$$\pi_m = \pi_M \prod_{i=1}^{M-m} \mathbf{A}_{M-i}, \quad m = 0, \dots, M-1, \tag{2.20}$$

где π_M является единственным решением системы уравнений

$$\begin{aligned}\pi_M (\mathbf{A}_{M-1} \mathbf{Q}_{1M} + \mathbf{Q}_{0M}) &= \mathbf{0}, \\ \pi_M \left(\sum_{m=0}^M \prod_{i=1}^{M-m} \mathbf{A}_{M-i} \mathbf{1}^T \right) &= 1,\end{aligned}\tag{2.21}$$

а матрицы $\mathbf{A}_m$ вычисляются по рекуррентным соотношениям

$$\begin{aligned}\mathbf{A}_0 &= -\mathbf{Q}_{21} \mathbf{Q}_{00}^{-1}, \\ \mathbf{A}_m &= -\mathbf{Q}_{2,m+1} (\mathbf{A}_{m-1} \mathbf{Q}_{1m} + \mathbf{Q}_{0m})^{-1}, \quad m=1, \dots, M-1.\end{aligned}\tag{2.22}$$

Доказательство.

а) *Запись СУР в матричном виде*

Запишем СУР $\pi \mathbf{Q} = \mathbf{0}$ с учетом блочно-трехдиагонального вида матрицы $\mathbf{Q}$ интенсивностей переходов

$$\begin{aligned}\pi_0 \mathbf{Q}_{00} + \pi_1 \mathbf{Q}_{21} &= \mathbf{0}, \\ \pi_{m-1} \mathbf{Q}_{1m} + \pi_m \mathbf{Q}_{0m} + \pi_{m+1} \mathbf{Q}_{2,m+1} &= \mathbf{0}, \quad m=1, \dots, M-1, \\ \pi_{M-1} \mathbf{Q}_{1M} + \pi_M \mathbf{Q}_{0M} &= \mathbf{0}.\end{aligned}$$

б) *Получение зависимости π_m от π_{m+1}*

Выразим π_0 через π_1 из матричного уравнения СУР при $m=0$

$$\pi_0 = -\pi_1 \mathbf{Q}_{21} \mathbf{Q}_{00}^{-1}$$

и обозначим $\mathbf{A}_0 = -\mathbf{Q}_{21} \mathbf{Q}_{00}^{-1}$, тогда

$$\pi_0 = \pi_1 \mathbf{A}_0.$$

Получим зависимость π_{m-1} от π_m . Матричное уравнение СУР при $m=1$ имеет вид

$$\pi_0 \mathbf{Q}_{11} + \pi_1 \mathbf{Q}_{01} + \pi_2 \mathbf{Q}_{22} = \mathbf{0}.$$

Подставим полученное соотношение для π_0

$$\pi_1 = -\pi_2 \mathbf{Q}_{22} (\mathbf{A}_0 \mathbf{Q}_{11} + \mathbf{Q}_{01})^{-1}$$

и обозначим $\mathbf{A}_1 = -\mathbf{Q}_{22} (\mathbf{A}_0 \mathbf{Q}_{11} + \mathbf{Q}_{01})^{-1}$, тогда

$$\pi_1 = \pi_2 \mathbf{A}_1.$$

Предположим, что соотношение верно для $m-1$

$$\pi_{m-1} = \pi_m \mathbf{A}_{m-1},$$

$$\mathbf{A}_{m-1} = -\mathbf{Q}_{2,m} \left(\mathbf{A}_{m-2} \mathbf{Q}_{1,m-1} + \mathbf{Q}_{0,m-1} \right)^{-1}.$$

Подставим его в матричное уравнение СУР при m

$$\pi_m = -\pi_{m+1} \mathbf{Q}_{2,m+1} \left(\mathbf{A}_{m-1} \mathbf{Q}_{1m} + \mathbf{Q}_{0m} \right)^{-1}$$

и обозначим $\mathbf{A}_m = -\mathbf{Q}_{2,m+1} \left(\mathbf{A}_{m-1} \mathbf{Q}_{1m} + \mathbf{Q}_{0m} \right)^{-1}$, тогда

$$\pi_m = \pi_{m+1} \mathbf{A}_m.$$

Следовательно, сделанное предположение о зависимости π_{m-1} от π_m верно, таким образом

$$\pi_m = \pi_{m+1} \mathbf{A}_m, \quad m = 0, \dots, M-1,$$

$$\mathbf{A}_0 = -\mathbf{Q}_{21} \mathbf{Q}_{00}^{-1},$$

$$\mathbf{A}_m = -\mathbf{Q}_{2,m+1} \left(\mathbf{A}_{m-1} \mathbf{Q}_{1m} + \mathbf{Q}_{0m} \right)^{-1}, \quad m = 1, \dots, M-1.$$

в) Выражение π_m через π_M

Выразим π_m через π_M , рекуррентно подставляя полученное соотношение

$$\pi_m = \pi_{m+1} \mathbf{A}_m = \pi_{m+2} \mathbf{A}_{m+1} \mathbf{A}_m = \dots = \pi_M \mathbf{A}_{M-1} \dots \mathbf{A}_{m+1} \mathbf{A}_m = \pi_M \prod_{i=1}^{M-m} \mathbf{A}_{M-i},$$

$$m = 0, \dots, M-1.$$

г) Получение системы уравнений относительно π_M

Подставим данное соотношение в матричное уравнение СУР при $m = M$ и получим уравнение относительно π_M

$$\pi_M \left(\mathbf{A}_{M-1} \mathbf{Q}_{1M} + \mathbf{Q}_{0M} \right) = \mathbf{0}.$$

Запишем условие нормировки $\pi \mathbf{1}^T = 1$

$$\sum_{m=0}^M \pi_m \mathbf{1}^T = 1$$

и подставим полученное выражение π_m через π_M

$$\pi_M \left(\sum_{m=0}^{M-1} \prod_{i=1}^{M-m} \mathbf{A}_{M-i} \mathbf{1}^T + \mathbf{1}^T \right) = 1.$$

Таким образом, получаем систему уравнений относительно π_M

$$\pi_M (\mathbf{A}_{M-1} \mathbf{Q}_{1M} + \mathbf{Q}_{0M}) = \mathbf{0},$$

$$\pi_M \left(\sum_{m=0}^M \prod_{i=1}^{M-m} \mathbf{A}_{M-i} \mathbf{1}^T \right) = 1.$$

■

Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой с приостановкой обслуживания широкополосного трафика, в соответствии с лексикографическим порядком (2.14), представим стационарное распределение СП $\mathbf{X}(t)$ в виде $\pi = (\pi_0, \dots, \pi_{R_2})$. Тогда с учетом леммы 2.2, справедлива следующая теорема.

Теорема 2.2. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой с приостановкой обслуживания широкополосного трафика, стационарное распределение СП $\mathbf{X}(t)$ рассчитывается в матричном виде по формуле

$$\pi_r = \pi_{R_2} \prod_{i=1}^{R_2-r} \mathbf{A}_{R_2-i}, \quad r = 0, \dots, R_2 - 1, \quad (2.23)$$

где π_{R_2} является единственным решением системы уравнений

$$\begin{aligned} \pi_{R_2} (\mathbf{A}_{R_2-1} \mathbf{Q}_{1R_2} + \mathbf{Q}_{0R_2}) &= \mathbf{0}, \\ \pi_{R_2} \sum_{r=0}^{R_2} \prod_{i=1}^{R_2-r} \mathbf{A}_{R_2-i} \mathbf{1}^T &= 1, \end{aligned} \quad (2.24)$$

а матрицы $\mathbf{A}_r$ вычисляются по рекуррентным соотношениям

$$\begin{aligned} \mathbf{A}_0 &= -\mathbf{Q}_{21} \mathbf{Q}_{00}^{-1}, \\ \mathbf{A}_r &= -\mathbf{Q}_{2,r+1} (\mathbf{A}_{r-1} \mathbf{Q}_{1r} + \mathbf{Q}_{0r})^{-1}, \quad r = 1, \dots, R_2 - 1. \end{aligned} \quad (2.25)$$

Доказательство.

Доказательство можно провести аналогично теореме 2.1, заменив $M \rightarrow R_2$.

■

2.2.3 Алгоритм снижения размерности системы уравнений равновесия

Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания широкополосного трафика, также справедлива следующая теорема.

Теорема 2.3. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания широкополосного трафика, стационарное распределение СП $X(t)$ рассчитывается по формуле

$$\pi(n, m) = \frac{p(n, m)}{\sum_{(i, j) \in \mathcal{X}} p(i, j)}, \quad (n, m) \in \mathcal{X}, \quad (2.26)$$

$$p(n, m) = \sum_{i=0}^{M-1} A_i(n, m) \cdot X_i + A_M(n, m), \quad (n, m) \in \mathcal{X}, \quad (2.27)$$

где вектор $(X_1, X_2, \dots, X_{M-1})$ является единственным решением системы уравнений

$$\sum_{i=0}^{M-1} B_i(m) \cdot X_i + B_M(m) = 0, \quad m = 1, \dots, M-1, \quad (2.28)$$

$$\begin{aligned} B_i(m) = & A_i(V - bm, m) \cdot [\lambda_1 + (V - bm)\mu_1 + m\mu_2] - \\ & - A_i(V - bm, m-1) \cdot \lambda_2 - A_i(V - bm - 1, m) \cdot \lambda_1, \\ & m = 1, \dots, M-1, \quad i = 0, \dots, M, \end{aligned} \quad (2.29)$$

а коэффициенты $A_i(n, m)$ вычисляются по рекуррентным соотношениям

$$\begin{aligned} A_m(0, m) &= 1, \quad m = 0, \dots, M, \\ A_i(0, m) &= 0, \quad m, i = 0, \dots, M, \quad i \neq m, \end{aligned} \quad (2.30)$$

$$\begin{aligned}
 A_i(n, m) \cdot n\mu_1 = & \\
 = A_i(n-1, m) \cdot & \left[\lambda_1 + \lambda_2 \cdot 1\{n-1+b(m+1) \leq V\} + (n-1)\mu_1 + m\mu_2 \right] - \\
 & - A_i(n-1, m-1) \cdot \lambda_2 \cdot 1\{m > 0\} - \\
 & - A_i(n-1, m+1) \cdot (m+1)\mu_2 \cdot 1\{n-1+b(m+1) \leq V\} - \\
 & - A_i(n-2, m) \cdot \lambda_1 \cdot 1\{n > 1\} - \\
 & - A_i(n-2, m+1) \cdot \lambda_1 \cdot 1\{n > 1, n-2+b(m+1) = V\}, \\
 (n, m) \in \mathcal{X}: & n > 0, i = 0, \dots, M.
 \end{aligned}$$

Доказательство.

а) Получение зависимости $\pi(n, m)$ от $\pi(n-1, i)$ и $\pi(n-2, j)$

Разобьем пространство $\mathcal{X}$ состояний (2.1) СП $X(t)$ на подмножества

$$\begin{aligned}
 \mathcal{X} &= \mathcal{X}_0 \cup \mathcal{X}_1 \cup \mathcal{X}_2, \\
 \mathcal{X}_0 &= \{(n, m) \in \mathcal{X}: n = 0\} = \{(0, m) \in \mathcal{X}\} = \{(0, 0), (0, 1), \dots, (0, m), \dots, (0, M)\}, \\
 \mathcal{X}_1 &= \{(n, m) \in \mathcal{X}: 0 < n < V - bm\}, \\
 \mathcal{X}_2 &= \{(n, m) \in \mathcal{X}: n = V - bm, m < M\} = \{(V - bm, m) \in \mathcal{X}, m < M\} = \\
 &= \{(V, 0), (V, 1), \dots, (V - bm, m), \dots, (b, M - 1)\}.
 \end{aligned}$$

Запишем СУР $\pi\mathbf{Q} = \mathbf{0}$ в виде СУГБ с учетом ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (2.5)

$$\begin{aligned}
 \pi(n, m) \cdot & \left[\lambda_1 \cdot 1\{n+1+bm \leq V\} + \lambda_1 \cdot 1\{n+1+bm > V, m > 0\} + \right. \\
 & \left. + \lambda_2 \cdot 1\{n+b(m+1) \leq V\} + n\mu_1 + m\mu_2 \right] = \\
 = \pi(n-1, m) \cdot & \lambda_1 \cdot 1\{n > 0\} + \\
 & + \pi(n-1, m+1) \cdot \lambda_1 \cdot 1\{n > 0, n-1+b(m+1) = V\} + \\
 & + \pi(n, m-1) \cdot \lambda_2 \cdot 1\{m > 0\} + \\
 & + \pi(n+1, m) \cdot (n+1)\mu_1 \cdot 1\{n+1+bm \leq V\} + \\
 & + \pi(n, m+1) \cdot (m+1)\mu_2 \cdot 1\{n+b(m+1) \leq V\}, \quad (n, m) \in \mathcal{X}.
 \end{aligned}$$

Выразим $\pi(n+1, m)$ из уравнений СУГБ при $(n, m) \in \mathcal{X}_0 \cup \mathcal{X}_1$

$$\begin{aligned} \pi(n+1, m) \cdot (n+1) \mu_1 = & \\ = \pi(n, m) \cdot \left[\lambda_1 + \lambda_2 \cdot 1\{n+b(m+1) \leq V\} + n\mu_1 + m\mu_2 \right] - & \\ - \pi(n-1, m) \cdot \lambda_1 \cdot 1\{n > 0\} - & \\ - \pi(n-1, m+1) \cdot \lambda_1 \cdot 1\{n > 0, n-1+b(m+1) = V\} - & \\ - \pi(n, m-1) \cdot \lambda_2 \cdot 1\{m > 0\} - & \\ - \pi(n, m+1) \cdot (m+1) \mu_2 \cdot 1\{n+b(m+1) \leq V\}, & \\ (n, m) \in \mathcal{X}_0 \cup \mathcal{X}_1. & \end{aligned}$$

Заменим $n+1$ на n , тогда уравнения будут справедливы для $(n-1, m) \in \mathcal{X}_0 \cup \mathcal{X}_1$, то

есть для $(n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2$, и $\pi(n, m)$ представляет собой линейную комбинацию

$$\pi(n-1, i) \text{ и } \pi(n-2, j)$$

$$\begin{aligned} \pi(n, m) \cdot n \mu_1 = & \\ = \pi(n-1, m) \cdot \left[\lambda_1 + \lambda_2 \cdot 1\{n-1+b(m+1) \leq V\} + (n-1)\mu_1 + m\mu_2 \right] - & \\ - \pi(n-1, m-1) \cdot \lambda_2 \cdot 1\{m > 0\} - & \\ - \pi(n-1, m+1) \cdot (m+1) \mu_2 \cdot 1\{n-1+b(m+1) \leq V\} - & \\ - \pi(n-2, m) \cdot \lambda_1 \cdot 1\{n-1 > 0\} - & \\ - \pi(n-2, m+1) \cdot \lambda_1 \cdot 1\{n-1 > 0, n-2+b(m+1) = V\}, & \\ (n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2. & \end{aligned}$$

б) Выражение $\pi(n, m)$, $n > 0$ через $\pi(0, i)$

Воспользовавшись рекуррентно полученным соотношением, можно представить $\pi(n, m)$ в виде линейной комбинации $\pi(0, i)$

$$\pi(n, m) = \sum_{i=0}^M A_i(n, m) \cdot \pi(0, i), \quad (n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2.$$

Обозначим

$$X_m = \pi(0, m), \quad m = 0, \dots, M,$$

тогда

$$\pi(n, m) = \sum_{i=0}^M A_i(n, m) \cdot X_i, \quad (n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2.$$

Доопределим $A_i(0, m)$ для $(n, m) \in \mathcal{X}_0$

$$A_m(0, m) = 1, \quad m = 0, \dots, M,$$

$$A_i(0, m) = 0, \quad m, i = 0, \dots, M, \quad i \neq m,$$

тогда

$$\pi(n, m) = \sum_{i=0}^M A_i(n, m) \cdot X_i, \quad (n, m) \in \mathcal{X}.$$

Подставим полученный вид $\pi(n, m)$ в СУГБ, и соотнеся коэффициенты при X_i в левой и правой частях равенств, получим рекуррентные соотношения для расчета $A_i(n, m)$

$$\begin{aligned} A_i(n, m) \cdot n\mu_1 = & \\ = A_i(n-1, m) \cdot & \left[\lambda_1 + \lambda_2 \cdot 1\{n-1+b(m+1) \leq V\} + (n-1)\mu_1 + m\mu_2 \right] - \\ & - A_i(n-1, m-1) \cdot \lambda_2 \cdot 1\{m > 0\} - \\ & - A_i(n-1, m+1) \cdot (m+1)\mu_2 \cdot 1\{n-1+b(m+1) \leq V\} - \\ & - A_i(n-2, m) \cdot \lambda_1 \cdot 1\{n > 1\} - \\ & - A_i(n-2, m+1) \cdot \lambda_1 \cdot 1\{n > 1, n-2+b(m+1) = V\}, \\ & (n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2, \quad i = 0, \dots, M. \end{aligned}$$

в) Получение системы уравнений относительно $\pi(0, m)$

Уравнения СУГБ при $(n, m) \in \mathcal{X}_2$ имеют вид

$$\begin{aligned} \pi(V-bm, m) \cdot & \left[\lambda_1 + (V-bm)\mu_1 + m\mu_2 \right] - \\ & - \pi(V-bm, m-1) \cdot \lambda_2 - \\ & - \pi(V-bm-1, m) \cdot \lambda_1 = 0, \quad m = 1, \dots, M-1. \end{aligned}$$

Подставим $\pi(n, m)$ в виде линейной комбинации $\pi(0, i)$ в полученную систему уравнений

$$\begin{aligned} & \sum_{i=0}^M A_i(V - bm, m) \cdot X_i \cdot [\lambda_1 + (V - bm)\mu_1 + m\mu_2] - \\ & - \sum_{i=0}^M A_i(V - bm, m-1) \cdot X_i \cdot \lambda_2 - \\ & - \sum_{i=0}^M A_i(V - bm - 1, m) \cdot X_i \cdot \lambda_1 = 0, \quad m = 1, \dots, M-1. \end{aligned}$$

Обозначим выражения при X_i через $B_i(m)$

$$\begin{aligned} B_i(m) = & A_i(V - bm, m) \cdot [\lambda_1 + (V - bm)\mu_1 + m\mu_2] - \\ & - A_i(V - bm, m-1) \cdot \lambda_2 - \\ & - A_i(V - bm - 1, m) \cdot \lambda_1, \quad m = 1, \dots, M-1, \quad i = 0, \dots, M, \end{aligned}$$

тогда

$$\sum_{i=0}^M B_i(m) \cdot X_i = 0, \quad m = 1, \dots, M-1.$$

Запишем условие нормировки $\pi \mathbf{1}^T = 1$

$$\sum_{(n,m) \in \mathcal{X}} \pi(n, m) = 1,$$

примем $X_M = 1$ и получаем систему уравнений относительно $(X_1, X_2, \dots, X_{M-1})$

$$\sum_{i=0}^{M-1} B_i(m) \cdot X_i + B_M(m) = 0, \quad m = 1, \dots, M-1.$$

При этом по окончании вычислений, вероятности $\pi(n, m)$ надо будет отнормировать

$$\pi(n, m) \leftarrow \frac{\pi(n, m)}{\sum_{(j,i) \in \mathcal{X}} \pi(j, i)}, \quad (n, m) \in \mathcal{X}.$$

■

2.3 РЕСУРСНАЯ СИСТЕМА СО СНИЖЕНИЕМ МОЩНОСТИ

2.3.1 Модель требования к ресурсу

Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и прерыванием обслуживания широкополосного трафика, требование к ресурсу для сессии узкополосного и широкополосного трафика в состоянии (n, m) имеет вид соответственно

$$\begin{aligned} y_1(n, m) &= \left(y_1^f = F, y_1^t = \frac{1}{n+m}, y_1^p = \frac{PL(N_0 + I_0)}{G_t G_r} \left(2^{(n+m)/F} - 1 \right) \right), \\ (n, m) &\in \mathcal{X} \setminus \{\mathbf{0}\}, \end{aligned} \quad (2.31)$$

$$\begin{aligned} y_2(n, m) &= \left(y_2^f = F, y_2^t = \frac{1}{n+m}, y_2^p = \frac{PL(N_0 + I_0)}{G_t G_r} \left(2^{b(n+m)/F} - 1 \right) \right), \\ (n, m) &\in \mathcal{X} \setminus \{\mathbf{0}\}, \end{aligned} \quad (2.32)$$

$$y_1(\mathbf{0}) = y_2(\mathbf{0}) = (0, 0, 0). \quad (2.33)$$

Тогда множество вариантов занятия ресурса в состоянии (n, m) представимо в виде

$$\begin{aligned} \mathcal{Y}(n, m) &= \left\{ (y_{11}, \dots, y_{1n}, y_{21}, \dots, y_{2m}) = (y_1[n], y_2[m]) : \right. \\ &\quad y_{1i} = \left(y_{1i}^f, y_{1i}^t, y_{1i}^p \right), \quad y_{1i}^f = F, \quad y_{1i}^t = \frac{1}{n+m}, \\ &\quad y_{1i}^p = \frac{PL_{1i}(N_0 + I_0)}{G_t G_r} \left(2^{(n+m)/F} - 1 \right), \quad PL_{1i} \geq 1, \quad y_{1i}^p \leq P, \quad i = 1, \dots, n, \\ &\quad y_{2i} = \left(y_{2i}^f, y_{2i}^t, y_{2i}^p \right), \quad y_{2i}^f = F, \quad y_{2i}^t = \frac{1}{n+m}, \\ &\quad y_{2i}^p = \frac{PL_{2i}(N_0 + I_0)}{G_t G_r} \left(2^{b(n+m)/F} - 1 \right), \quad PL_{2i} \geq 1, \quad y_{2i}^p \leq P, \quad i = 1, \dots, m \left. \right\}, \\ &\quad (n, m) \in \mathcal{X} \setminus \{\mathbf{0}\}, \\ &\quad \mathcal{Y}(\mathbf{0}) = \{(0, 0, 0)\}, \\ &\quad \mathcal{Y}(n, m) = \emptyset, \quad (n, m) \notin \mathcal{X}. \end{aligned} \quad (2.34)$$

Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и скорости передачи широкополосного трафика, требование к ресурсу для сессии узкополосного и широкополосного трафика в состоянии $(n, m_1, \dots, m_K)$ имеет вид соответственно

$$\mathbf{y}_1(n, m_1, \dots, m_K) = \left(y_1^f = F, y_1^t = \frac{1}{v(n, m_1, \dots, m_K)}, y_1^p = P \right), \quad (2.35)$$

$$(n, m_1, \dots, m_K) \in \mathcal{X} \setminus \{\mathbf{0}\},$$

$$\begin{aligned} \mathbf{y}_{2k}(n, m_1, \dots, m_K) = \\ = \left(y_{2k}^f = F, y_{2k}^t = \frac{b_k}{v(n, m_1, \dots, m_K)}, y_{2k}^p = \frac{PL(N_0 + I_0)}{G_t G_r} \left(2^{v(n, m_1, \dots, m_K)/F} - 1 \right) \right), \end{aligned} \quad (2.36)$$

$$(n, m_1, \dots, m_K) \in \mathcal{X} \setminus \{\mathbf{0}\},$$

$$\mathbf{y}_1(\mathbf{0}) = \mathbf{y}_{21}(\mathbf{0}) = \dots = \mathbf{y}_{2K}(\mathbf{0}) = (0, 0, 0). \quad (2.37)$$

Тогда множество вариантов занятия ресурса в состоянии $(n, m_1, \dots, m_K)$ представимо в виде

$$\begin{aligned} \mathcal{Y}(n, m_1, \dots, m_K) = \\ = \left\{ (\mathbf{y}_{11}, \dots, \mathbf{y}_{1n}, \mathbf{y}_{211}, \dots, \mathbf{y}_{21m_1}, \dots, \mathbf{y}_{2k1}, \dots, \mathbf{y}_{2km_k}, \dots, \mathbf{y}_{2K1}, \dots, \mathbf{y}_{2Km_K}) \right\} = \\ = (\mathbf{y}_1[n], \mathbf{y}_{21}[m_1], \dots, \mathbf{y}_{2k}[m_k], \dots, \mathbf{y}_{2K}[m_K]): \\ \mathbf{y}_{1i} = (y_{1i}^f, y_{1i}^t, y_{1i}^p), \quad y_{1i}^f = F, \quad y_{1i}^t = \frac{1}{v(n, m_1, \dots, m_K)}, \quad y_{1i}^p = P, \quad i = 1, \dots, n, \\ \mathbf{y}_{2ki} = (y_{2ki}^f, y_{2ki}^t, y_{2ki}^p), \\ y_{2ki}^f = F, \quad y_{2ki}^t = \frac{b_k}{v(n, m_1, \dots, m_K)}, \quad y_{2ki}^p = \frac{PL_{2ki}(N_0 + I_0)}{G_t G_r} \left(2^{v(n, m_1, \dots, m_K)/F} - 1 \right), \\ PL_{2ki} \geq 1, \quad y_{2ki}^p \leq P, \quad i = 1, \dots, m_k, \quad k = 1, \dots, K, \quad (n, m_1, \dots, m_K) \in \mathcal{X} \setminus \{\mathbf{0}\}, \\ \mathcal{Y}(\mathbf{0}) = \{(0, 0, 0)\}, \\ \mathcal{Y}(n, m_1, \dots, m_K) = \emptyset, \quad (n, m_1, \dots, m_K) \notin \mathcal{X}. \end{aligned} \quad (2.38)$$

2.3.2 Вероятности занятия ресурса

Пользователи случайным образом располагаются в зоне покрытия базовой станции, поэтому мощность ξ_{PL} затухания сигнала с ФР $F_{\xi_{\text{PL}}}(x)$ и требование к ресурсу $\xi = (\xi_f, \xi_t, \xi_p)$ также являются СВ. Модель затухания сигнала и вид $F_{\xi_{\text{PL}}}(x)$ описаны в параграфе 1.3.2.

Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и прерыванием обслуживания широкополосного трафика, тогда справедлива следующая лемма.

Лемма 2.3. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и прерыванием обслуживания широкополосного трафика, вероятности занятия ресурса рассчитываются по формуле

$$\Pr((\xi_1[n], \xi_2[m]) \in \mathcal{Y}(n, m)) = \left[F_{\xi_{\text{PL}}} \left(\frac{PG_t G_r}{N_0 + I_0} \frac{1}{2^{(n+m)/F} - 1} \right) \right]^n \cdot \left[F_{\xi_{\text{PL}}} \left(\frac{PG_t G_r}{N_0 + I_0} \frac{1}{2^{b(n+m)/F} - 1} \right) \right]^m, \quad (n, m) \in \mathcal{X}. \quad (2.39)$$

Доказательство.

Из вида (2.34) множества $\mathcal{Y}(n, m)$ вариантов занятия ресурса, получим

$$\Pr((\xi_1[n], \xi_2[m]) \in \mathcal{Y}(n, m)) = \Pr(\xi_{1i}^p \leq P, i = 1, \dots, n, \xi_{2i}^p \leq P, i = 1, \dots, m), \quad (n, m) \in \mathcal{X}.$$

Расположение пользователей в зоне покрытия базовой станции случайное, и используемая модель затухания сигнала не предполагает дополнительные зависимости компонент требования к ресурсу одной сессии от параметров других сессий, в частности интерференция постоянная. Поэтому СВ ξ_i^p независимы и события $\xi_i^p \leq P$ также независимы в совокупности

$$\Pr((\xi_1[n], \xi_2[m]) \in \mathcal{Y}(n, m)) = \prod_{i=1}^n \Pr(\xi_{1i}^p \leq P) \cdot \prod_{i=1}^m \Pr(\xi_{2i}^p \leq P), \quad (n, m) \in \mathcal{X}.$$

Таким образом, из вида (2.31) требования к ресурсу следует

$$\begin{aligned}
 \Pr\left(\left(\xi_1[n], \xi_2[m]\right) \in \mathcal{Y}(n, m)\right) &= \\
 &= \prod_{i=1}^n \Pr\left(\frac{\xi_{1i}^{\text{PL}}(N_0 + I_0)}{G_t G_r} \left(2^{(n+m)/F} - 1\right) \leq P\right) \cdot \prod_{i=1}^m \Pr\left(\frac{\xi_{2i}^{\text{PL}}(N_0 + I_0)}{G_t G_r} \left(2^{b(n+m)/F} - 1\right) \leq P\right) = \\
 &= \prod_{i=1}^n \Pr\left(\xi_{1i}^{\text{PL}} \leq \frac{P G_t G_r}{N_0 + I_0} \frac{1}{2^{(n+m)/F} - 1}\right) \cdot \prod_{i=1}^m \Pr\left(\xi_{2i}^{\text{PL}} \leq \frac{P G_t G_r}{N_0 + I_0} \frac{1}{2^{b(n+m)/F} - 1}\right), \\
 &(n, m) \in \mathcal{X}.
 \end{aligned}$$

Полученные множители – вероятности представляют собой ФР СВ $\xi_{\bullet}^{\text{PL}}$, которые являются одинаково распределенными СВ с ФР $F_{\xi_{\text{PL}}}(x)$, поэтому

$$\begin{aligned}
 \Pr\left(\left(\xi_1[n], \xi_2[m]\right) \in \mathcal{Y}(n, m)\right) &= \\
 &= \prod_{i=1}^n F_{\xi_{1i}^{\text{PL}}}\left(\frac{P G_t G_r}{N_0 + I_0} \frac{1}{2^{(n+m)/F} - 1}\right) \cdot \prod_{i=1}^m F_{\xi_{2i}^{\text{PL}}}\left(\frac{P G_t G_r}{N_0 + I_0} \frac{1}{2^{b(n+m)/F} - 1}\right) = \\
 &= \left[F_{\xi_{\text{PL}}}\left(\frac{P G_t G_r}{N_0 + I_0} \frac{1}{2^{(n+m)/F} - 1}\right)\right]^n \cdot \left[F_{\xi_{\text{PL}}}\left(\frac{P G_t G_r}{N_0 + I_0} \frac{1}{2^{b(n+m)/F} - 1}\right)\right]^m, \quad (n, m) \in \mathcal{X}.
 \end{aligned}$$

■

Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и скорости передачи широкополосного трафика, также справедлива следующая лемма.

Лемма 2.4. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и скорости передачи широкополосного трафика, вероятности занятия ресурса рассчитываются по формуле

$$\begin{aligned}
 \Pr\left(\left(\xi_1[n], \xi_{21}[m_1], \dots, \xi_{2K}[m_K]\right) \in \mathcal{Y}(n, m_1, \dots, m_K)\right) &= \\
 &= \left[F_{\xi_{\text{PL}}}\left(\frac{P G_t G_r}{N_0 + I_0} \frac{1}{2^{v(n, m_1, \dots, m_K)/F} - 1}\right)\right]^{\sum_{k=1}^K m_k}, \quad (n, m_1, \dots, m_K) \in \mathcal{X}.
 \end{aligned} \tag{2.40}$$

Доказательство.

Доказательство можно провести аналогично лемме 2.3. Отметим, что множитель, соответствующий узкополосному трафику, отсутствует, поскольку мощность для него не меняется.

■

2.3.3 Условные вероятности выделения ресурса

Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и прерыванием обслуживания широкополосного трафика, обозначим вероятности того, что еще одной сессии узкополосного и широкополосного трафика соответственно может быть выделен ресурс, при условии нахождения в состоянии (n, m)

$$\begin{aligned} p_1(n, m) &= \Pr\left(\left(\xi_1[n+1], \xi_2[m]\right) \in \mathcal{Y}(n+1, m) \mid \left(\xi_1[n], \xi_2[m]\right) \in \mathcal{Y}(n, m)\right), \\ &\quad (n, m) \in \mathcal{X}, \\ p_2(n, m) &= \Pr\left(\left(\xi_1[n], \xi_2[m+1]\right) \in \mathcal{Y}(n, m+1) \mid \left(\xi_1[n], \xi_2[m]\right) \in \mathcal{Y}(n, m)\right), \\ &\quad (n, m) \in \mathcal{X}. \end{aligned} \tag{2.41}$$

Тогда с учетом леммы 2.3, справедлива следующая теорема.

Теорема 2.4. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и прерыванием обслуживания широкополосного трафика, условные вероятности того, что еще одной сессии узкополосного и широкополосного трафика соответственно может быть выделен ресурс, рассчитываются по формулам

$$\begin{aligned} p_1(\mathbf{0}) &= F_{\xi_{\text{PL}}}\left(\frac{X}{2^{1/F} - 1}\right), \\ p_1(n, m) &= \left[F_{\xi_{\text{PL}}}\left(\frac{X}{2^{(n+m+1)/F} - 1}\right)\right]^{n+1} \cdot \left[F_{\xi_{\text{PL}}}\left(\frac{X}{2^{b(n+m+1)/F} - 1}\right)\right]^m \times \\ &\quad \times \left[F_{\xi_{\text{PL}}}\left(\frac{X}{2^{(n+m)/F} - 1}\right)\right]^{-n} \cdot \left[F_{\xi_{\text{PL}}}\left(\frac{X}{2^{b(n+m)/F} - 1}\right)\right]^{-m}, \quad (n, m) \in \mathcal{X} \setminus \{\mathbf{0}\}, \end{aligned} \tag{2.42}$$

$$\begin{aligned}
 p_2(\mathbf{0}) &= F_{\xi_{\text{PL}}} \left(\frac{X}{2^{b/F} - 1} \right), \\
 p_2(n, m) &= \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{(n+m+1)/F} - 1} \right) \right]^n \cdot \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{b(n+m+1)/F} - 1} \right) \right]^{m+1} \times \\
 &\quad \times \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{(n+m)/F} - 1} \right) \right]^{-n} \cdot \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{b(n+m)/F} - 1} \right) \right]^{-m}, \quad (n, m) \in \mathcal{X} \setminus \{\mathbf{0}\},
 \end{aligned} \tag{2.43}$$

где $X = PG_t G_r (N_0 + I_0)^{-1}$.

Замечание. В формулах $p_1(n, m)$ и $p_2(n, m)$ при $(n, m) \in \mathcal{X} \setminus \{\mathbf{0}\}$, для краткости записи, будем иметь в виду, что при $n = 0$ или $m = 0$, $-n = 0$ или $-m = 0$ соответственно, т.е. не происходит деления на ноль.

Доказательство.

По определению условной вероятности, запишем

$$\begin{aligned}
 p_1(n, m) &= \frac{\Pr((\xi_1[n+1], \xi_2[m]) \in \mathcal{Y}(n+1, m), (\xi_1[n], \xi_2[m]) \in \mathcal{Y}(n, m))}{\Pr((\xi_1[n], \xi_2[m]) \in \mathcal{Y}(n, m))} = \\
 &= \frac{\Pr((\xi_1[n+1], \xi_2[m]) \in \mathcal{Y}(n+1, m))}{\Pr((\xi_1[n], \xi_2[m]) \in \mathcal{Y}(n, m))}, \quad (n, m) \in \mathcal{X}, \\
 p_2(n, m) &= \frac{\Pr((\xi_1[n], \xi_2[m+1]) \in \mathcal{Y}(n, m+1), (\xi_1[n], \xi_2[m]) \in \mathcal{Y}(n, m))}{\Pr((\xi_1[n], \xi_2[m]) \in \mathcal{Y}(n, m))} = \\
 &= \frac{\Pr((\xi_1[n], \xi_2[m+1]) \in \mathcal{Y}(n, m+1))}{\Pr((\xi_1[n], \xi_2[m]) \in \mathcal{Y}(n, m))}, \quad (n, m) \in \mathcal{X}.
 \end{aligned}$$

Для состояния $(n, m) = \mathbf{0}$ ввиду того, что $\Pr((\xi_1[0], \xi_2[0]) \in \mathcal{Y}(0, 0)) = 1$, получим

$$p_1(\mathbf{0}) = \Pr(\xi_1 \in \mathcal{Y}(1, 0)),$$

$$p_2(\mathbf{0}) = \Pr(\xi_2 \in \mathcal{Y}(0, 1)).$$

Таким образом, применение леммы 2.3 завершает доказательство. ■

Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и прерыванием обслуживания широкополосного трафика, обозначим вероятности того, что еще одной сессии узкополосного и широкополосного трафика соответственно может быть выделен ресурс, при условии нахождения в состоянии $(n, m_1, \dots, m_K)$

$$\begin{aligned} p_1(n, \mathbf{m}) &= \Pr\left(\left(\xi_1[n+1], \xi_{21}[m_1], \dots, \xi_{2K}[m_K]\right) \in \mathcal{Y}(n+1, m_1, \dots, m_K) \mid \right. \\ &\quad \left. \mid \left(\xi_1[n], \xi_{21}[m_1], \dots, \xi_{2K}[m_K]\right) \in \mathcal{Y}(n, m_1, \dots, m_K)\right), \\ &\quad (n, m_1, \dots, m_K) \in \mathcal{X}, \\ p_2(n, \mathbf{m}) &= \Pr\left(\left(\xi_1[n], \xi_{21}[m_1+1], \dots, \xi_{2K}[m_K]\right) \in \mathcal{Y}(n, m_1+1, \dots, m_K) \mid \right. \\ &\quad \left. \mid \left(\xi_1[n], \xi_{21}[m_1], \dots, \xi_{2K}[m_K]\right) \in \mathcal{Y}(n, m_1, \dots, m_K)\right), \\ &\quad (n, m_1, \dots, m_K) \in \mathcal{X}. \end{aligned} \quad (2.44)$$

Тогда с учетом леммы 2.4, справедлива следующая теорема.

Теорема 2.5. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и скорости передачи широкополосного трафика, условные вероятности того, что еще одной сессии узкополосного и широкополосного трафика соответственно может быть выделен ресурс, рассчитываются по формулам

$$\begin{aligned} p_1(0, \mathbf{0}) &= 1, \\ p_1(n, \mathbf{m}) &= \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{(v(n, \mathbf{m})+1)/F} - 1} \right) \right]^{m_\bullet} \cdot \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{v(n, \mathbf{m})/F} - 1} \right) \right]^{-m_\bullet}, \\ &\quad (n, \mathbf{m}) \in \mathcal{X} \setminus \{(0, \mathbf{0})\}, \end{aligned} \quad (2.45)$$

$$\begin{aligned} p_2(0, \mathbf{0}) &= F_{\xi_{\text{PL}}} \left(\frac{X}{2^{b_1/F} - 1} \right), \\ p_2(n, \mathbf{m}) &= \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{(v(n, \mathbf{m})+b_1)/F} - 1} \right) \right]^{m_\bullet+1} \cdot \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{v(n, \mathbf{m})/F} - 1} \right) \right]^{-m_\bullet}, \\ &\quad (n, \mathbf{m}) \in \mathcal{X} \setminus \{(0, \mathbf{0})\}, \end{aligned} \quad (2.46)$$

где $X = PG_t G_r (N_0 + I_0)^{-1}$ и $m_\bullet = \sum_{k=1}^K m_k$.

Доказательство.

Доказательство можно провести аналогично теореме 2.4 с применением леммы 2.4

$$p_1(\mathbf{0}) = \Pr(\xi_{11} \in \mathcal{Y}(1, 0, \dots, 0)),$$

$$p_1(n, m_1, \dots, m_K) = \frac{\Pr((\xi_1[n+1], \xi_{21}[m_1], \dots, \xi_{2K}[m_K]) \in \mathcal{Y}(n+1, m_1, \dots, m_K))}{\Pr((\xi_1[n], \xi_{21}[m_1], \dots, \xi_{2K}[m_K]) \in \mathcal{Y}(n, m_1, \dots, m_K))},$$

$$(n, m_1, \dots, m_K) \in \mathcal{X} \setminus \{\mathbf{0}\},$$

$$p_2(\mathbf{0}) = \Pr(\xi_{21} \in \mathcal{Y}(0, 1, 0, \dots, 0)),$$

$$p_2(n, m_1, \dots, m_K) = \frac{\Pr((\xi_1[n], \xi_{21}[m_1+1], \xi_{22}[m_2], \dots, \xi_{2K}[m_K]) \in \mathcal{Y}(n, m_1+1, \dots, m_K))}{\Pr((\xi_1[n], \xi_{21}[m_1], \dots, \xi_{2K}[m_K]) \in \mathcal{Y}(n, m_1, \dots, m_K))},$$

$$(n, m_1, \dots, m_K) \in \mathcal{X} \setminus \{\mathbf{0}\}.$$

■

2.4 АНАЛИЗ ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ ПЕРЕДАЧИ ШИРОКОПОЛОСНОГО ТРАФИКА

2.4.1 Вероятность приостановки обслуживания

Рассмотрим одну из характеристик систем с приоритетным обслуживанием, которая отражает влияние более приоритетного трафика на менее приоритетный. В системе с приостановкой обслуживания этой характеристикой является вероятность такого события. Под вероятностью приостановки обслуживания сессии широкополосного трафика будем понимать вероятность того, что обслуживание конкретной сессии широкополосного трафика будет приостановлено.

Следствие 2.1. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой с приостановкой обслуживания широкополосного трафика, вероятность приостановки обслуживания сессии широкополосного трафика рассчитывается по формуле

$$P_{dl} = \sum_{m=1}^M \sum_{q=0}^{R_1-m} \sum_{r=0}^{R_2-1} \frac{\lambda_1}{\lambda_1 + \lambda_2 + (V - bm)\mu_1 + m\mu_2} \cdot \frac{1}{m} \cdot \pi(V - bm, m, q, r) + \\ + \sum_{m=1}^M \sum_{q=0}^{R_1-m} \frac{\lambda_1}{\lambda_1 + (V - bm)\mu_1 + m\mu_2} \cdot \frac{1}{m} \cdot \pi(V - bm, m, q, R_2). \quad (2.47)$$

Доказательство.

Обозначим следующие события:

искомое событие $\mathcal{A}$: «приостановлено обслуживание конкретной сессии широкополосного трафика»;

событие $\mathcal{B}(\mathbf{x})$: «система находится в состоянии $\mathbf{x}$ »;

событие $\mathcal{C}$: «поступил запрос на передачу узкополосного трафика»;

событие $\mathcal{D}$: «выбрана конкретная сессия широкополосного трафика из всех обслуживаемых».

По формуле полной вероятности, вероятность искомого события $\mathcal{A}$ имеет вид

$$\Pr(\mathcal{A}) = \sum_{\mathbf{x} \in \mathcal{X}} \Pr(\mathcal{A} \mid \mathcal{B}(\mathbf{x})) \cdot \Pr(\mathcal{B}(\mathbf{x})).$$

Условные вероятности $\Pr(\mathcal{A} \mid \mathcal{B}(\mathbf{x}))$ равны 0, если система находится в состоянии $\mathbf{x}$, когда приостановка обслуживания невозможна. Поэтому обозначим $\mathcal{X}_1 \subseteq \mathcal{X}$ подмножество пространства состояний, где приостановка обслуживания возможна $\mathcal{X}_1 = \{\mathbf{x} = (n, m, q, r) \in \mathcal{X} : n + 1 + bm > V, m > 0\}$,

тогда

$$\Pr(\mathcal{A}) = \sum_{\mathbf{x} \in \mathcal{X}_1} \Pr(\mathcal{A} \mid \mathcal{B}(\mathbf{x})) \cdot \Pr(\mathcal{B}(\mathbf{x})).$$

В состоянии $\mathbf{x} \in \mathcal{X}_1$ событие «приостановлено обслуживание конкретной сессии широкополосного трафика» эквивалентно совместному выполнению двух событий «поступил запрос на передачу узкополосного трафика» и «выбрана конкретная сессия широкополосного трафика из всех обслуживаемых»

$$\Pr(\mathcal{A} \mid \mathcal{B}(\mathbf{x})) = \Pr(\mathcal{C} \cap \mathcal{D} \mid \mathcal{B}(\mathbf{x})), \quad \mathbf{x} \in \mathcal{X}_1.$$

Таким образом,

$$\begin{aligned} \Pr(\mathcal{A}) &= \sum_{\mathbf{x} \in \mathcal{X}_1} \Pr(\mathcal{C} \cap \mathcal{D} \mid \mathcal{B}(\mathbf{x})) \cdot \Pr(\mathcal{B}(\mathbf{x})) = \\ &= \sum_{\mathbf{x} \in \mathcal{X}_1} \Pr(\mathcal{C} \mid \mathcal{B}(\mathbf{x})) \cdot \Pr(\mathcal{D} \mid \mathcal{B}(\mathbf{x})) \cdot \Pr(\mathcal{B}(\mathbf{x})) = \\ &= \sum_{\mathbf{x} \in \mathcal{X}_1} \Pr(\mathcal{C} \mid \mathcal{B}(\mathbf{x})) \cdot \Pr(\mathcal{D} \mid \mathcal{B}(\mathbf{x})) \cdot \pi(\mathbf{x}). \end{aligned}$$

Найдем условную вероятность $\Pr(\mathcal{C} \mid \mathcal{B}(\mathbf{x}))$. Система построена в экспоненциальных предположениях – входящие потоки, потоки попыток возобновить и начать обслуживание пуассоновские, обслуживание экспоненциальное, поэтому $\Pr(\mathcal{C} \mid \mathcal{B}(\mathbf{x}))$ можно интерпретировать как вероятность того, что минимум независимых экспоненциально распределенных СВ – времен до наступления возможных случайных событий в состоянии $\mathbf{x} \in \mathcal{X}_1$ – будет время до поступления запроса на передачу узкополосного трафика. Согласно соответствующему свойству экспоненциального распределения, получим

$$\begin{aligned} \Pr(\mathcal{C} \mid \mathcal{B}(\mathbf{x})) &= \\ &= \frac{\lambda_1}{\lambda_1 + \lambda_2 \cdot 1\{r < R_2\} + (V - bm)\mu_1 + m\mu_2}, \quad \mathbf{x} \in \mathcal{X}_1. \end{aligned}$$

Условную вероятность $\Pr(\mathcal{D} \mid \mathcal{B}(\mathbf{x}))$, поскольку сессия выбирается равновероятно ($w=1$), запишем с использованием числа сочетаний из всех m обслуживаемых сессий по 1 выбранной конкретной сессии

$$\Pr(\mathcal{D} \mid \mathcal{B}(\mathbf{x})) = 1 / \binom{m}{1} = \frac{1}{m}, \quad \mathbf{x} \in \mathcal{X}_1.$$

■

Рассмотрим вероятность P_{dl} приостановки обслуживания сессии широкополосного трафика. На рисунках 2.1 и 2.2 представлены графики, которые иллюстрируют эту вероятность в зависимости различных факторов. На рисунке 2.1 – от доли интенсивности потока запросов на передачу узкополосного трафика, а на рисунке 2.2 – от интенсивности повторных попыток начать обслуживание сессии широкополосного трафика.

Первый сценарий актуален в условиях роста числа устройств, требующих сверхнадежной передачи данных с малой задержкой. Схема обслуживания широкополосного трафика с двумя орбитами демонстрирует более низкую вероятность приостановки обслуживания по сравнению со схемой без орбит или только с одной орбитой для приостановленных сессий. В зависимости от соотношения между входящими потоками запросов, вероятность приостановки обслуживания сначала возрастает, потом убывает, а затем снова возрастает.

Во втором сценарии видно, как по-разному ведет себя характеристика на этих трех интервалах с точки зрения выбора параметра γ_1 попыток возобновления обслуживания широкополосного трафика. При этом важно учитывать вероятность блокировки запроса на передачу широкополосного трафика

$$P_{bl} = \sum_{m=0}^M \sum_{n=V-bm-b+1}^{V-bm} \sum_{q=0}^{M-m} \pi(n, m, q, R_2). \quad (2.48)$$

Чем чаще предпринимаются попытки, тем ниже вероятность блокировки, но тем выше вероятность приостановки обслуживания.

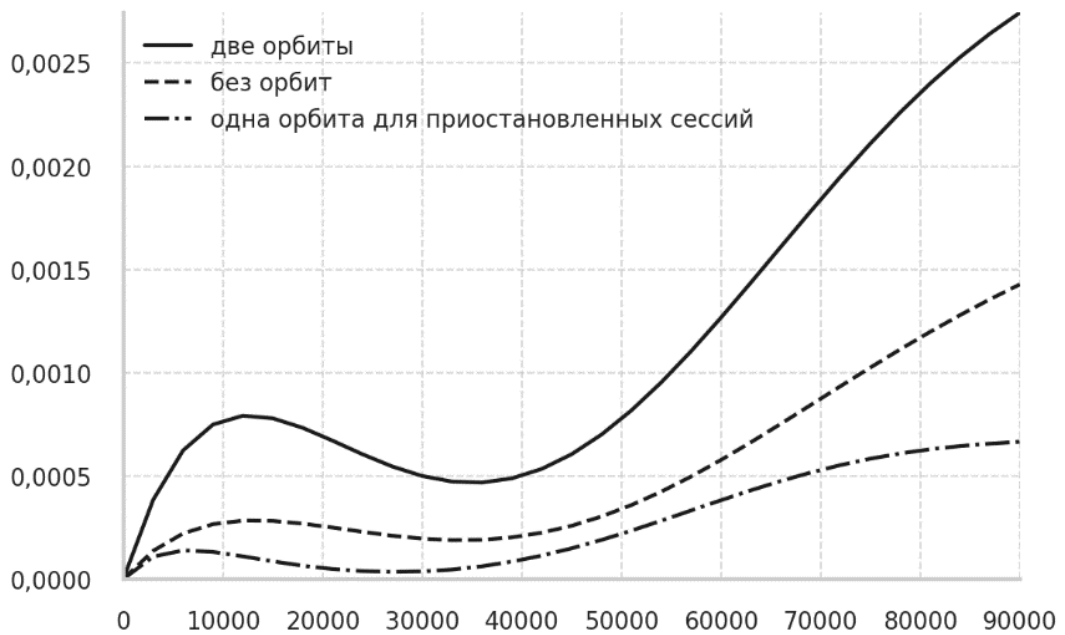


Рисунок 2.1 – Вероятность приостановки обслуживания сессии широкополосного трафика в зависимости от доли интенсивности потока запросов на передачу узкополосного трафика для разных моделей

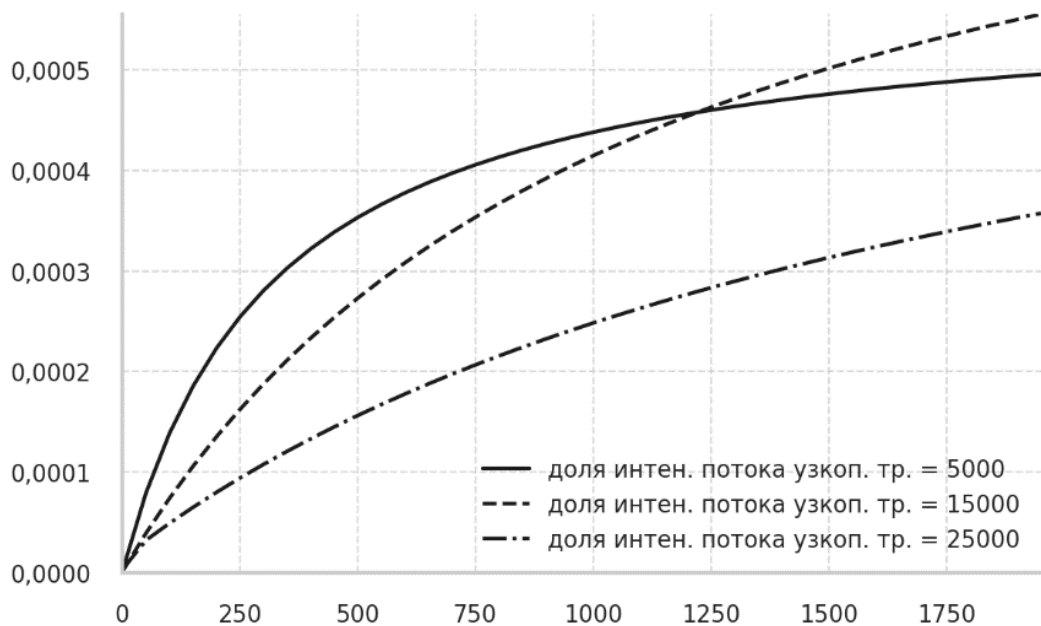


Рисунок 2.2 – Вероятность приостановки обслуживания сессии широкополосного трафика в зависимости от интенсивности повторных попыток начать обслуживание сессии широкополосного трафика (1/с) для разной доли интенсивности потока запросов на передачу узкополосного трафика

2.4.2 Средняя скорость передачи трафика

Следствие 2.2. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и скорости передачи широкополосного трафика, следующие показатели эффективности рассчитываются по формулам:

средняя скорость передачи широкополосного трафика

$$\bar{b} = \left[\sum_{(n,m) \in \mathcal{X}: m \neq \mathbf{0}} \left(\sum_{k=1}^K m_k b_k \cdot \pi(n, \mathbf{m}) \right) \cdot \left(\sum_{k=1}^K m_k \right)^{-1} \right] \cdot \left[1 - \sum_{n=0}^V \pi(n, \mathbf{0}) \right]^{-1}; \quad (2.49)$$

вероятность прерывания обслуживания сессии широкополосного трафика

$$P_{\text{in}} = \sum_{m_K=1}^M \frac{\lambda_1}{\lambda_1 + (V - b_K m_K) \mu_1 + m_K \mu_2} \cdot \frac{1}{m_K} \cdot \pi(V - b_K m_K, \mathbf{0}, m_K); \quad (2.50)$$

вероятность блокировки сессии широкополосного трафика

$$P_{\text{bl}} = \sum_{(n,m) \in \mathcal{B}} \pi(n, \mathbf{m}), \quad \mathcal{B} = \{(n, \mathbf{m}) \in \mathcal{X}: v(n, \mathbf{m}) + b_1 > V\}. \quad (2.51)$$

Задачу максимизации средней скорости широкополосного трафика по уровням скорости и при ограничениях на скорость, вероятность прерывания и вероятность блокировки сессии широкополосного трафика запишем следующим образом

$$\begin{aligned} (K^*, \mathbf{b}^*) &= \arg \max_{K, \mathbf{b}} \bar{b}(K, \mathbf{b}), \\ 1 \leq K \leq \hat{K}, \quad b_k &\in \{\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_{\hat{K}}\}, \quad k = 1, \dots, K, \\ b_1 &> \dots > b_k > \dots > b_K \geq \hat{b} > 0, \quad b_k = b_{k-1} - 1, \quad k = 2, \dots, K, \\ P_{\text{in}}(K, \mathbf{b}) &\leq \hat{P}_{\text{in}}, \\ P_{\text{bl}}(K, \mathbf{b}) &\leq \hat{P}_{\text{bl}}. \end{aligned} \quad (2.52)$$

На рисунках 2.3 и 2.4 проиллюстрированы средняя скорость и вероятность прерывания обслуживания для разных наборов уровней скорости.

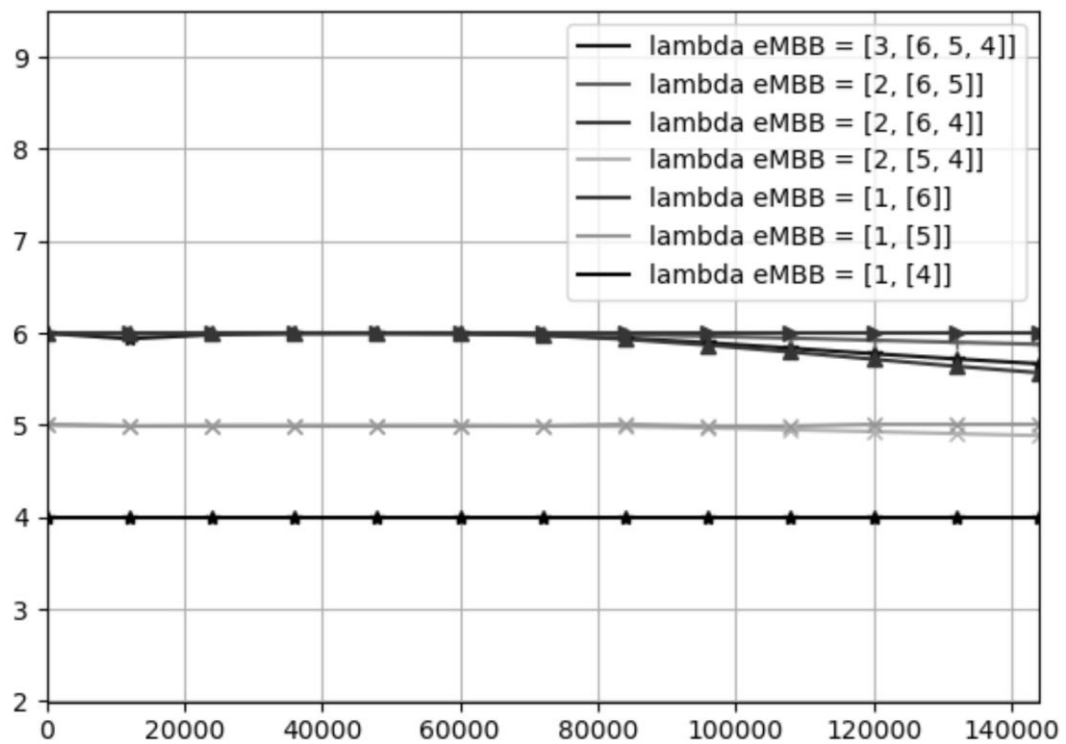


Рисунок 2.3 – Средняя скорость передачи широкополосного трафика в зависимости от доли интенсивности потока запросов на передачу узкополосного трафика для разных наборов уровней скорости

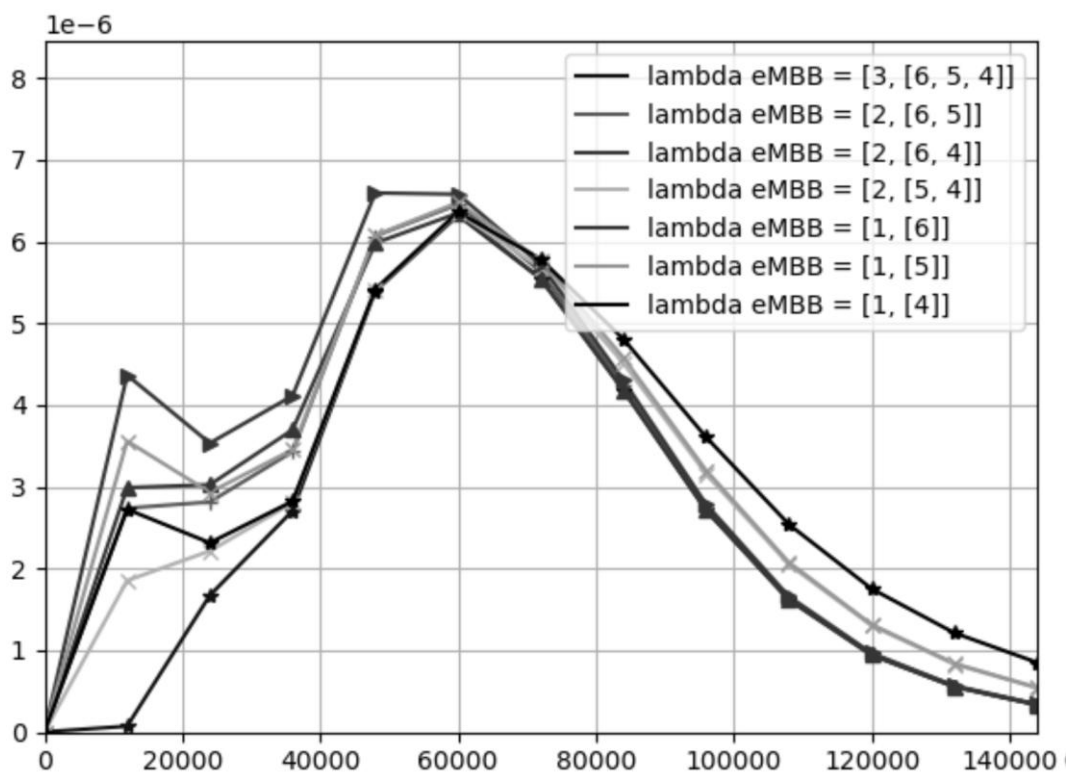


Рисунок 2.4 – Вероятность прерывания обслуживания сессии широкополосного трафика в зависимости от доли интенсивности потока запросов на передачу узкополосного трафика для разных наборов уровней скорости

2.4.3 Вероятность прерывания обслуживания

На поведение показателей эффективности приоритетного доступа безусловно влияет затухание мощности сигнала. Формулы для их расчета выписываются с учетом условного распределения вероятностей того, что еще одной сессии может быть выделен ресурс (см. раздел 2.3).

Завершим раздел численным примером, иллюстрирующим поведение показателей эффективности обслуживания широкополосного трафика с учетом затухания сигнала. Формулы для их расчета выписываются с учетом условного распределения вероятностей того, что еще одной сессии может быть выделен ресурс.

Следствие 2.3. Для модели с узкополосным и широкополосным потоковым трафиком и постоянным объемом ресурса, схемой со снижением мощности и скорости передачи широкополосного трафика, вероятность прерывания обслуживания сессии широкополосного трафика рассчитывается по формуле

$$P_{in} = \sum_{m_K=1}^M \left[\frac{\lambda_1 \cdot p_1(V - b_K m_K, \mathbf{0}, m_K)}{\lambda_1 \cdot p_1(V - b_K m_K, \mathbf{0}, m_K) + (V - b_K m_K) \mu_1 + m_K \mu_2} \times \right. \\ \left. \times \frac{1}{m_K} \cdot \pi(V - b_K m_K, \mathbf{0}, m_K) \right]. \quad (2.53)$$

Следствие 2.4. Если все устройства пользователей при любом своем местоположении в зоне покрытия базовой станции находятся в прямой видимости, тогда ФР в теореме 2.5 рассчитывается по формуле

$$F_{\xi_{PL}}(x) = F_{\xi_d} \left(\sqrt{[x/\alpha]^{2/\beta} - (h_{BS} - h_{UT})^2} \right). \quad (2.54)$$

Если при этом пользователи расположены равномерно в зоне покрытия базовой станции, тогда

$$F_{\xi_{PL}}(x) = \frac{[x/\alpha]^{2/\beta} - (h_{BS} - h_{UT})^2}{d_{\max}^2}, \quad (2.55)$$

где $d_{\max}$ радиус зоны покрытия базовой станции.

На рисунке 2.5 проиллюстрирован вид вероятности P_{in} прерывания обслуживания сессии широкополосного трафика соответственно для двух сценариев развертывания сети согласно спецификации 3GPP TR 38.901 – макросота (urban macro, UMa) и микросота (urban micro, UMi) в городе.

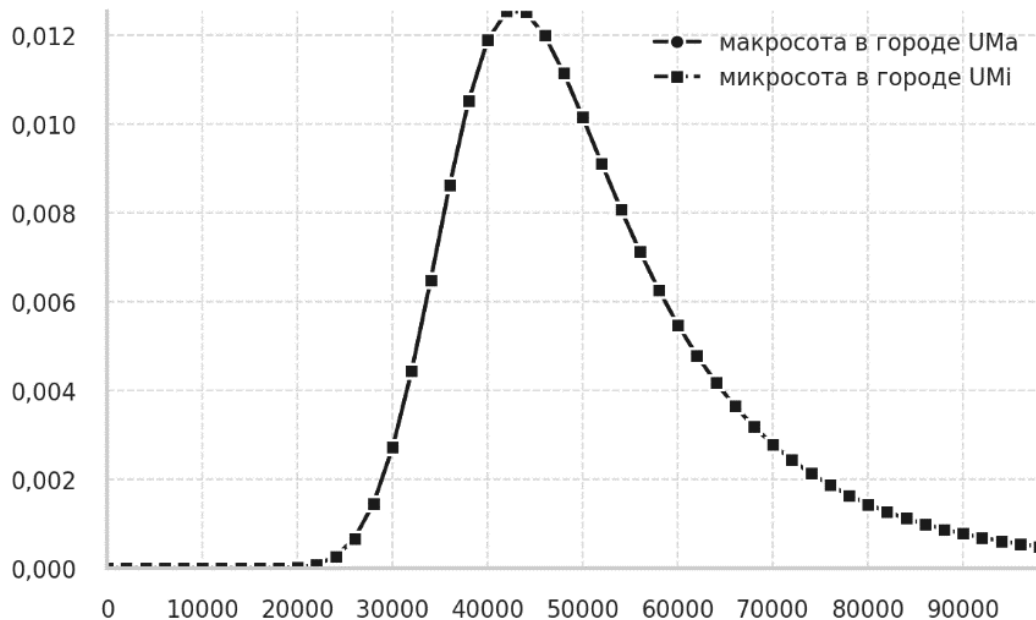


Рисунок 2.5 – Вероятность прерывания обслуживания сессии широкополосного трафика в зависимости от доли интенсивности потока запросов на передачу узкополосного трафика для разных моделей затухания сигнала

2.5 МОДЕЛЬ ДОСТУПА ДЛЯ ВИДЕО ПО ТРЕБОВАНИЮ И ВИДЕОКОНФЕРЕНЦИИ

2.5.1 Системная модель

В разделах 2.5–2.6 главы 2 исследованы модели приоритетного доступа двух классов потокового трафика – одноадресного и многоадресного, генерируемого при предоставлении услуг видео по требованию и видеоконференции соответственно. Приоритетными являются обе услуги, объем ресурса постоянный. Рассмотрим сценарии, где для предоставления услуги мультивещания видеоконференции сессии видео по запросу могут прерываться, а ресурс

освобождается после завершения обслуживания пользователя, инициировавшего сессию. Если скорость видеоконференции допускает снижение, то это может произойти в случае необходимости предоставления услуги видео по требованию. Рассмотрим также случай с дополнительной услугой мультивещания – телевидением и освобождением ресурса после завершения обслуживания последнего из присутствующих на сессии пользователей.

Рассмотрим несколько сценариев обслуживания многоадресного трафика. Скорость передачи может быть постоянной или изменяться по уровням скорости. Многоадресный трафик одного класса с завершением сессии по первому пользователю или двух классов дополнительно с завершением сессии по последнему пользователю. Для обеспечения приоритета многоадресного трафика применяется механизм прерывания обслуживания одноадресного трафика, а для обеспечения приоритета одноадресного трафика – снижение скорости многоадресного трафика.

В таблице 2.2 приведены основные обозначения, параметры и компоненты состояния системы. Для моделирования мультивещания используется дисциплина обслуживания «прозрачных» заявок – П1 для многоадресного трафика 1-класса и П2 для многоадресного трафика 2-класса.

Таблица 2.2 – Обозначения, используемые для моделей с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса (глава 2, разделы 2.5–2.6)

Обозначение	Параметр
λ_1	Интенсивность пуассоновского потока запросов на передачу многоадресного трафика с завершением сессии по первому пользователю (1-класса)
λ_2	Интенсивность пуассоновского потока запросов на передачу одноадресного трафика
λ_3	Интенсивность пуассоновского потока запросов на передачу многоадресного трафика с завершением сессии по последнему пользователю (2-класса)
L	Число уровней скорости, на которых может передаваться многоадресный трафик 1-класса

Обозначение	Параметр
$b = b_1 > \dots > b_l > \dots > b_L$	Скорости передачи многоадресного трафика 1-класса
$d > b$	Скорость передачи многоадресного трафика 2-класса
$1/\mu_1$	Среднее экспоненциальное время сессии многоадресного трафика 1-класса
$1/\mu_2$	Среднее экспоненциальное время передачи одноадресного трафика
$1/\mu_3$	Среднее экспоненциальное время сессии многоадресного трафика 2-класса
$n = (n_1, \dots, n_L)$	Число обслуживаемых сессий многоадресного трафика 1-класса
m	Число обслуживаемых сессий одноадресного трафика
k	Число обслуживаемых сессий многоадресного трафика 2-класса

2.5.2 Схемы приоритетного доступа

Рассмотрим следующие схемы приоритетного доступа для моделей с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса:

- (а) прерывание обслуживания одноадресного трафика и одна дисциплина обслуживания многоадресного трафика;
- (б) прерывание обслуживания одноадресного трафика и две дисциплины обслуживания многоадресного трафика;
- (в) прерывание обслуживания одноадресного трафика и снижение скорости передачи многоадресного трафика.

В этих схемах используются механизмы $p \in \{1, 3, 4\} \subset \mathcal{P}$ реализации приоритета и принципы $w \in \{2, 5\}$ выбора сессии для применения механизма из параграфа 1.6.3.

Опишем особенности исследуемых моделей с перечисленными схемами приоритетного доступа через матрицы приоритетов.

Схема (а) – базовая система с одним механизмом приоритетного доступа – прерыванием обслуживания одноадресного трафика. Многоадресный трафик при этом обслуживается по дисциплине П1 – с завершением сессии по первому пользователю. Это СМО $M_2 | M_2 | V | 0 | F_5$ с матрицей приоритетов

F_5		одноадресный трафик обслуживается	
		$\begin{matrix} p \\ e \end{matrix}$	
многоадресный трафик 1-класса	поступление	1	(1,5)
	завершение обслуживания	2	0
одноадресный трафик	поступление	3	0
	завершение обслуживания	4	0

Схема (б). В этой системе также применяется один механизм приоритетного доступа – прерывание обслуживания одноадресного трафика. Однако число классов многоадресного трафика два. Трафик первого класса также обслуживается по дисциплине П1 – сессия завершается по первому пользователю. Трафик второго класса – по дисциплине П2, когда сессия завершается по последнему из присутствующих на сессии пользователей. Это СМО $M_3 | M_3 | V | 0 | F_6$ с матрицей приоритетов

F_6		одноадресный трафик обслуживается	
		$\begin{matrix} p \\ e \end{matrix}$	
многоадресный трафик 1-класса	поступление	1	(1,5)
	завершение обслуживания	2	0
одноадресный трафик	поступление	3	0
	завершение обслуживания	4	0
многоадресный трафик 2- класса	поступление	5	(1,5)
	завершение обслуживания	6	0

Схема (в). Последняя система – с двумя механизмами приоритетного доступа – с прерыванием обслуживания одноадресного трафика и снижением скорости передачи многоадресного трафика. Это СМО $M_2 | M_2 | V | 0 | F_7$ с матрицей приоритетов

F_7			одноадресный трафик обслуживается	многоадресный трафик 1-класса обслуживается	
			$\begin{matrix} p \\ e \end{matrix}$	1	$\begin{matrix} 3 & 4 \end{matrix}$
многоадресный трафик 1-класса	поступление	1		(1,5)	$\begin{matrix} 0 & 0 \end{matrix}$
	завершение обслуживания	2		0	$\begin{matrix} 0 & 0 \end{matrix}$
одноадресный трафик	поступление	3		0	$\begin{matrix} (1,2) & 0 \end{matrix}$
	завершение обслуживания	4		0	$\begin{matrix} 0 & (1,2) \end{matrix}$

2.5.3 Модели математической теории телетрафика

Функционирование систем опишем при помощи СП $X(t)$ над пространствами состояний

модель (а)

$$\mathcal{X} = \{(n, m): n = 0, 1, m \geq 0, bn + m \leq V\}; \quad (2.56)$$

модель (б)

$$\mathcal{X} = \{(n, m, k): n = 0, 1, m \geq 0, k = 0, 1, bn + m + dk \leq V\}; \quad (2.57)$$

модель (в)

$$\mathcal{X} = \{(n, m): n = 0, 1, m \geq 0, b_L n + m \leq V\}. \quad (2.58)$$

Ненулевые элементы матриц интенсивностей переходов имеют вид

модель (а)

$$\begin{aligned}
 Q[(0, m), (1, m)] &= \lambda_1, \quad b + m \leq V, \\
 Q[(0, m), (1, V - b)] &= \lambda_1, \quad b + m > V, \\
 Q[(n, m), (n, m + 1)] &= \lambda_2, \quad bn + m + 1 \leq V, \\
 Q[(1, m), (0, m)] &= \mu_1, \\
 Q[(n, m), (n, m - 1)] &= m\mu_2, \quad m > 0;
 \end{aligned} \tag{2.59}$$

модель (б)

$$\begin{aligned}
 Q[(0, m, k), (1, m, k)] &= \lambda_1, \quad b + m + dk \leq V, \\
 Q[(0, m, k), (1, V - b - dk, k)] &= \lambda_1, \quad b + m + dk > V, \\
 Q[(n, m, k), (n, m + 1, k)] &= \lambda_2, \quad bn + m + 1 + dk \leq V, \\
 Q[(n, m, 0), (n, m, 1)] &= \lambda_3, \quad bn + m + d \leq V, \\
 Q[(n, m, 0), (n, V - bn - d, 1)] &= \lambda_3, \quad bn + m + d > V, \\
 Q[(1, m, k), (0, m, k)] &= \mu_1, \\
 Q[(n, m, k), (n, m - 1, k)] &= m\mu_2, \quad m > 0, \\
 Q[(n, m, 1), (n, m, 0)] &= \mu_3;
 \end{aligned} \tag{2.60}$$

модель (в)

$$\begin{aligned}
 Q[(0, m), (1, m)] &= \lambda_1, \quad b_L + m \leq V, \\
 Q[(0, m), (1, V - b)] &= \lambda_1, \quad b_L + m > V, \\
 Q[(n, m), (n, m + 1)] &= \lambda_2, \quad b_L n + m + 1 \leq V, \\
 Q[(1, m), (0, m)] &= \mu_1, \\
 Q[(n, m), (n, m - 1)] &= m\mu_2, \quad m > 0.
 \end{aligned} \tag{2.61}$$

В следующем разделе 2.6 описаны результаты исследования моделей схем приоритетного доступа с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса. Предложены в скалярном виде алгоритм расчета стационарного распределения для модели (а) и снижения размерности системы уравнений равновесия для модели (б). Обсуждается выбор уровней скорости многоадресного трафика для модели (в).

2.6 СИСТЕМА С ОДНОАДРЕСНЫМ И МНОГОАДРЕСНЫМ ТРАФИКОМ

2.6.1 Рекуррентный алгоритм расчета стационарного распределения

Для модели с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания одноадресного трафика и одной дисциплиной обслуживания многоадресного трафика, справедлива следующая теорема.

Теорема 2.6. Для модели с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания одноадресного трафика и одной дисциплиной обслуживания многоадресного трафика, стационарное распределение СП $X(t)$ рассчитывается по формуле

$$\pi(n, m) = \frac{p(n, m)}{\sum_{(i, j) \in \mathcal{X}} p(i, j)}, \quad (n, m) \in \mathcal{X}, \quad (2.62)$$

$$p(n, m) = A_0(n, m) + A_1(n, m) \cdot X, \quad (n, m) \in \mathcal{X}, \quad (2.63)$$

$$X = \frac{A_0(0, V-1) \cdot \lambda_2 - A_0(0, V) \cdot (\lambda_1 + V\mu_2)}{A_1(0, V) \cdot (\lambda_1 + V\mu_2) - A_1(0, V-1) \cdot \lambda_2}, \quad (2.64)$$

где коэффициенты $A_i(n, m)$ вычисляются по рекуррентным соотношениям

$$\begin{aligned} A_0(0, 0) &= 1, \quad A_1(0, 0) = 0, \\ A_0(1, 0) &= 0, \quad A_1(1, 0) = 1, \\ A_0(0, 1) &= \frac{\lambda_1 + \lambda_2}{\mu_2}, \quad A_1(0, 1) = -\frac{\mu_1}{\mu_2}, \\ A_0(1, 1) &= -\frac{\lambda_1}{\mu_2}, \quad A_1(1, 1) \cdot \mu_2 = \frac{\lambda_2 + \mu_1}{\mu_2}, \end{aligned} \quad (2.65)$$

$$\begin{aligned}
 A_i(0, m) \cdot m\mu_2 &= \\
 &= A_i(0, m-1) \cdot [\lambda_1 + \lambda_2 + (m-1)\mu_2] - A_i(1, m-1) \cdot \mu_1 - A_i(0, m-2) \cdot \lambda_2, \\
 m &= 2, \dots, V-b+1, \quad i=0,1,
 \end{aligned} \tag{2.66}$$

$$\begin{aligned}
 A_i(0, m) \cdot m\mu_2 &= \\
 &= A_i(0, m-1) \cdot [\lambda_1 + \lambda_2 + (m-1)\mu_2] - A_i(0, m-2) \cdot \lambda_2, \\
 m &= V-b+2, \dots, V, \quad i=0,1,
 \end{aligned}$$

$$\begin{aligned}
 A_i(1, m) \cdot m\mu_2 &= \\
 &= A_i(1, m-1) \cdot [\lambda_2 + \mu_1 + (m-1)\mu_2] - A_i(0, m-1) \cdot \lambda_1 - A_i(1, m-2) \cdot \lambda_2, \\
 m &= 2, \dots, V-b, \quad i=0,1.
 \end{aligned} \tag{2.67}$$

Доказательство.

а) Получение зависимости $\pi(n, m)$ от $\pi(i, m-1)$ и $\pi(i, m-2)$

Разобьем пространство $\mathcal{X}$ состояний (2.56) СП $\mathbf{X}(t)$ на подмножества

$$\mathcal{X} = \mathcal{X}_0 \cup \mathcal{X}_1 \cup \mathcal{X}_2,$$

$$\mathcal{X}_0 = \{(n, m) \in \mathcal{X} : m = 0\} = \{(n, 0) \in \mathcal{X}\} = \{(0, 0), (1, 0)\},$$

$$\mathcal{X}_1 = \{(n, m) \in \mathcal{X} : 0 < m < V - bn\},$$

$$\mathcal{X}_2 = \{(n, m) \in \mathcal{X} : m = V - bn\} = \{(n, V - bn) \in \mathcal{X}\} = \{(0, V), (1, V - b)\}.$$

Запишем СУР $\pi\mathbf{Q} = \mathbf{0}$ в виде СУГБ с учетом ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (2.59)

$$\begin{aligned}
 \pi(n, m) \cdot [\lambda_1 \cdot 1\{n=0\} + \lambda_2 \cdot 1\{bn+m+1 \leq V\} + \mu_1 \cdot 1\{n=1\} + m\mu_2] &= \\
 &= \pi(n-1, m) \cdot \lambda_1 \cdot 1\{n=1\} + \sum_{i=V-b+1}^V \pi(n-1, i) \cdot \lambda_1 \cdot 1\{(n, m) = (1, V-b)\} + \\
 &+ \pi(n, m-1) \cdot \lambda_2 \cdot 1\{m > 0\} + \\
 &+ \pi(n+1, m) \cdot \mu_1 \cdot 1\{n=0, b+m \leq V\} + \\
 &+ \pi(n, m+1) \cdot (m+1)\mu_2 \cdot 1\{bn+m+1 \leq V\}, \quad (n, m) \in \mathcal{X}.
 \end{aligned}$$

Выразим $\pi(n, m+1)$ из уравнений СУГБ при $(n, m) \in \mathcal{X}_0 \cup \mathcal{X}_1$

$$\begin{aligned} \pi(n, m+1) \cdot (m+1) \mu_2 = \\ = \pi(n, m) \cdot [\lambda_1 \cdot 1\{n=0\} + \lambda_2 + \mu_1 \cdot 1\{n=1\} + m\mu_2] - \\ - \pi(n-1, m) \cdot \lambda_1 \cdot 1\{n=1\} - \\ - \pi(n, m-1) \cdot \lambda_2 \cdot 1\{m>0\} - \\ - \pi(n+1, m) \cdot \mu_1 \cdot 1\{n=0, b+m \leq V\}, \quad (n, m) \in \mathcal{X}_0 \cup \mathcal{X}_1. \end{aligned}$$

Заменим $m+1$ на m , тогда уравнения будут справедливы для $(n, m-1) \in \mathcal{X}_0 \cup \mathcal{X}_1$,

то есть для $(n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2$, и $\pi(n, m)$ представляет собой линейную комбинацию

$$\pi(i, m-1) \text{ и } \pi(i, m-2)$$

$$\begin{aligned} \pi(n, m) \cdot m\mu_2 = \\ = \pi(n, m-1) \cdot [\lambda_1 \cdot 1\{n=0\} + \lambda_2 + \mu_1 \cdot 1\{n=1\} + (m-1)\mu_2] - \\ - \pi(n-1, m-1) \cdot \lambda_1 \cdot 1\{n=1\} - \\ - \pi(n+1, m-1) \cdot \mu_1 \cdot 1\{n=0, b+m-1 \leq V\} - \\ - \pi(n, m-2) \cdot \lambda_2 \cdot 1\{m-1>0\}, \quad (n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2. \end{aligned}$$

б) Выражение $\pi(n, m)$, $m > 0$ через $\pi(i, 0)$

Воспользовавшись рекуррентно полученным соотношением, можно представить $\pi(n, m)$ в виде линейной комбинации $\pi(i, 0)$

$$\pi(n, m) = A_0(n, m) \cdot \pi(0, 0) + A_1(n, m) \cdot \pi(1, 0), \quad (n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2.$$

Обозначим

$$X_n = \pi(n, 0), \quad n = 0, 1,$$

тогда

$$\pi(n, m) = A_0(n, m) \cdot X_0 + A_1(n, m) \cdot X_1, \quad (n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2.$$

Доопределим $A_i(n, 0)$ для $(n, m) \in \mathcal{X}_0$

$$A_0(0, 0) = 1, \quad A_1(0, 0) = 0,$$

$$A_0(1, 0) = 0, \quad A_1(1, 0) = 1,$$

тогда

$$\pi(n, m) = A_0(n, m) \cdot X_0 + A_1(n, m) \cdot X_1, \quad (n, m) \in \mathcal{X}.$$

Подставим полученный вид $\pi(n, m)$ в СУГБ, и соотнеся коэффициенты при X_i в левой и правой частях равенств, получим рекуррентные соотношения для расчета $A_i(n, m)$

$$\begin{aligned} A_i(n, m) \cdot m\mu_2 = & \\ = A_i(n, m-1) \cdot [\lambda_1 \cdot 1\{n=0\} + \lambda_2 + \mu_1 \cdot 1\{n=1\} + (m-1)\mu_2] - & \\ - A_i(n-1, m-1) \cdot \lambda_1 \cdot 1\{n=1\} - & \\ - A_i(n+1, m-1) \cdot \mu_1 \cdot 1\{n=0, b+m-1 \leq V\} - & \\ - A_i(n, m-2) \cdot \lambda_2 \cdot 1\{m>1\}, \quad (n, m) \in \mathcal{X}_1 \cup \mathcal{X}_2, \quad i=0,1. & \end{aligned}$$

Для различных комбинаций условий в функциях-индикаторах

$$n=0, \quad m=1,$$

$$n=0, \quad m>1, \quad m \leq V-b+1,$$

$$n=0, \quad m>1, \quad m > V-b+1,$$

$$n=1, \quad m=1,$$

$$n=1, \quad m>1,$$

распишем полученные соотношения

$$A_i(0,1) \cdot \mu_2 = A_i(0,0) \cdot [\lambda_1 + \lambda_2] - A_i(1,0) \cdot \mu_1, \quad i=0,1,$$

$$\begin{aligned} A_i(0, m) \cdot m\mu_2 = A_i(0, m-1) \cdot [\lambda_1 + \lambda_2 + (m-1)\mu_2] - & \\ - A_i(1, m-1) \cdot \mu_1 - A_i(0, m-2) \cdot \lambda_2, & \\ m=2, \dots, V-b+1, \quad i=0,1, & \end{aligned}$$

$$\begin{aligned} A_i(0, m) \cdot m\mu_2 = A_i(0, m-1) \cdot [\lambda_1 + \lambda_2 + (m-1)\mu_2] - & \\ - A_i(0, m-2) \cdot \lambda_2, & \\ m=V-b+2, \dots, V, \quad i=0,1, & \end{aligned}$$

$$A_i(1,1) \cdot \mu_2 = A_i(1,0) \cdot [\lambda_2 + \mu_1] - A_i(0,0) \cdot \lambda_1, \quad i=0,1,$$

$$\begin{aligned} A_i(1, m) \cdot m\mu_2 = A_i(1, m-1) \cdot [\lambda_2 + \mu_1 + (m-1)\mu_2] - & \\ - A_i(0, m-1) \cdot \lambda_1 - A_i(1, m-2) \cdot \lambda_2, & \\ m=2, \dots, V-b, \quad i=0,1. & \end{aligned}$$

в) Получение уравнения относительно $\pi(n, 0)$

Уравнения СУГБ при $(n, m) \in \mathcal{X}_2$ имеют вид

$$\begin{aligned} \pi(n, V-bn) \cdot [\lambda_1 \cdot 1\{n=0\} + \mu_1 \cdot 1\{n=1\} + (V-bn)\mu_2] - & \\ - \sum_{i=V-b}^V \pi(n-1, i) \cdot \lambda_1 \cdot 1\{n=1\} - & \\ - \pi(n, V-bn-1) \cdot \lambda_2 = 0, \quad n=0,1. & \end{aligned}$$

В уравнение при $(0, V)$

$$\pi(0, V) \cdot [\lambda_1 + V\mu_2] - \pi(0, V-1) \cdot \lambda_2 = 0$$

подставим $\pi(n, m)$ в виде линейной комбинации $\pi(i, 0)$

$$[A_0(0, V) \cdot X_0 + A_1(0, V) \cdot X] \cdot [\lambda_1 + V\mu_2] - [A_0(0, V-1) \cdot X_0 + A_1(0, V-1) \cdot X_1] \cdot \lambda_2 = 0.$$

Запишем условие нормировки $\pi \mathbf{1}^T = 1$

$$\sum_{(n, m) \in \mathcal{X}} \pi(n, m) = 1,$$

примем $X_0 = 1$ и получаем уравнение относительно X_1

$$[A_0(0, V) + A_1(0, V) \cdot X_1] \cdot [\lambda_1 + V\mu_2] = [A_0(0, V-1) + A_1(0, V-1) \cdot X_1] \cdot \lambda_2.$$

При этом по окончании вычислений, вероятности $\pi(n, m)$ надо будет отнормировать

$$\pi(n, m) \leftarrow \frac{\pi(n, m)}{\sum_{(i, j) \in \mathcal{X}} \pi(i, j)}, \quad (n, m) \in \mathcal{X}.$$

■

2.6.2 Алгоритм снижения размерности системы уравнений равновесия

Для модели с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания одноадресного трафика и двумя дисциплинами обслуживания многоадресного трафика, также справедлива следующая теорема.

Теорема 2.7. Для модели с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания одноадресного трафика и двумя дисциплинами обслуживания многоадресного трафика, стационарное распределение СП $X(t)$ рассчитывается по формуле

$$\pi(n, m, k) = \frac{p(n, m, k)}{\sum_{(i, l, j) \in \mathcal{X}} p(i, l, j)}, \quad (n, m, k) \in \mathcal{X}, \quad (2.68)$$

$$p(n, m, k) = A_{00}(n, m, k) + A_{10}(n, m, k) \cdot X_{10} + A_{01}(n, m, k) \cdot X_{01} + A_{11}(n, m, k) \cdot X_{11}, \quad (2.69)$$

$$(n, m, k) \in \mathcal{X},$$

где вектор (X_{10}, X_{01}, X_{11}) является единственным решением системы уравнений

$$B_{00}(l) + B_{10}(l) \cdot X_{10} + B_{01}(l) \cdot X_{01} + B_{11}(l) \cdot X_{11} = 0, \quad l = 1, 2, 3, \quad (2.70)$$

$$B_{ij}(1) = [\lambda_1 + \lambda_3] \cdot A_{ij}(0, V, 0) - \lambda_2 \cdot A_{ij}(0, V-1, 0), \quad i, j = 0, 1,$$

$$B_{ij}(2) = [\lambda_3 + \mu_1 + (V-b)\mu_2] \cdot A_{ij}(1, V-b, 0) - \lambda_2 \cdot A_{ij}(1, V-b-1, 0) -$$

$$- \lambda_1 \sum_{m=V-b}^V A_{ij}(0, m, 0), \quad i, j = 0, 1, \quad (2.71)$$

$$B_{ij}(3) = [\lambda_1 + (V-d)\mu_2 + \mu_3] \cdot A_{ij}(0, V-d, 1) - \lambda_2 \cdot A_{ij}(0, V-d-1, 1) -$$

$$- \lambda_3 \sum_{m=V-d}^V A_{ij}(0, m, 0), \quad i, j = 0, 1,$$

а коэффициенты $A_{ij}(n, m, k)$ вычисляются по рекуррентным соотношениям

$$A_{00}(0, 0, 0) = 1, \quad A_{10}(0, 0, 0) = 0, \quad A_{01}(0, 0, 0) = 0, \quad A_{11}(0, 0, 0) = 0,$$

$$A_{00}(1, 0, 0) = 0, \quad A_{10}(1, 0, 0) = 1, \quad A_{01}(1, 0, 0) = 0, \quad A_{11}(1, 0, 0) = 0,$$

$$A_{00}(0, 0, 1) = 0, \quad A_{10}(0, 0, 1) = 0, \quad A_{01}(0, 0, 1) = 1, \quad A_{11}(0, 0, 1) = 0,$$

$$A_{00}(1, 0, 1) = 0, \quad A_{10}(1, 0, 1) = 0, \quad A_{01}(1, 0, 1) = 0, \quad A_{11}(1, 0, 1) = 1, \quad (2.72)$$

$$A_{ij}(n, m, k) \cdot m\mu_2 =$$

$$= A_{ij}(n, m-1, k) \cdot [\lambda_1 \cdot 1\{n=0\} + \lambda_2 + \lambda_3 \cdot 1\{k=0\} +$$

$$+ \mu_1 \cdot 1\{n=1\} + (m-1)\mu_2 + \mu_3 \cdot 1\{k=1\}] -$$

$$- A_{ij}(n-1, m-1, k) \cdot \lambda_1 \cdot 1\{n=1\} -$$

$$- A_{ij}(n, m-1, k-1) \cdot \lambda_3 \cdot 1\{k=1\} -$$

$$- A_{ij}(n+1, m-1, k) \cdot \mu_1 \cdot 1\{n=0, b+m-1+dk \leq V\} -$$

$$- A_{ij}(n, m-1, k+1) \cdot \mu_3 \cdot 1\{k=0, bn+m-1+d \leq V\} -$$

$$- A_{ij}(n, m-2, k) \cdot \lambda_2 \cdot 1\{m-1 > 0\}, \quad (n, m, k) \in \mathcal{X}: m > 0, \quad i, j = 0, 1. \quad (2.73)$$

Доказательство.

а) Получение зависимости $\pi(n, m, k)$ от $\pi(i, m-1, j)$ и $\pi(i, m-2, j)$

Разобьем пространство $\mathcal{X}$ состояний (2.57) СП $\mathbf{X}(t)$ на подмножества

$$\mathcal{X} = \mathcal{X}_0 \cup \mathcal{X}_1 \cup \mathcal{X}_2,$$

$$\mathcal{X}_0 = \{(n, m, k) \in \mathcal{X} : m = 0\} = \{(n, 0, k) \in \mathcal{X}\} = \{(0, 0, 0), (1, 0, 0), (0, 0, 1), (1, 0, 1)\},$$

$$\mathcal{X}_1 = \{(n, m, k) \in \mathcal{X} : 0 < m < V - bn - dk\},$$

$$\begin{aligned} \mathcal{X}_2 &= \{(n, m, k) \in \mathcal{X} : m = V - bn - dk\} = \{(n, V - bn - dk, k) \in \mathcal{X}\} = \\ &= \{(0, V, 0), (1, V - b, 0), (0, V - d, 1), (1, V - b - d, 1)\}. \end{aligned}$$

Запишем СУР $\pi\mathbf{Q} = \mathbf{0}$ в виде СУГБ с учетом ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (2.60)

$$\begin{aligned} \pi(n, m, k) \cdot & \left[\lambda_1 \cdot 1\{n = 0\} + \lambda_2 \cdot 1\{bn + m + 1 + dk \leq V\} + \lambda_3 \cdot 1\{k = 0\} + \right. \\ & \left. + \mu_1 \cdot 1\{n = 1\} + m\mu_2 + \mu_3 \cdot 1\{k = 1\} \right] = \\ & = \pi(n-1, m, k) \cdot \lambda_1 \cdot 1\{n = 1\} + \\ & + \sum_{i=V-b+1}^V \pi(n-1, i, 0) \cdot \lambda_1 \cdot 1\{(n, m, k) = (1, V - b, 0)\} + \\ & + \sum_{i=V-b-d+1}^{V-d} \pi(n-1, i, 1) \cdot \lambda_1 \cdot 1\{(n, m, k) = (1, V - b - d, 1)\} + \\ & + \pi(n, m-1, k) \cdot \lambda_2 \cdot 1\{m > 0\} + \\ & + \pi(n, m, k-1) \cdot \lambda_3 \cdot 1\{k = 1\} + \\ & + \sum_{i=V-d+1}^V \pi(n, i, k-1) \cdot \lambda_3 \cdot 1\{(n, m, k) = (0, V - d, 1)\} + \\ & + \sum_{i=V-b-d+1}^{V-b} \pi(n, i, k-1) \cdot \lambda_3 \cdot 1\{(n, m, k) = (1, V - b - d, 1)\} + \\ & + \pi(n+1, m, k) \cdot \mu_1 \cdot 1\{n = 0, b + m + dk \leq V\} + \\ & + \pi(n, m+1, k) \cdot (m+1)\mu_2 \cdot 1\{bn + m + 1 + dk \leq V\} + \\ & + \pi(n, m, k+1) \cdot \mu_3 \cdot 1\{k = 0, bn + m + d \leq V\}, \quad (n, m, k) \in \mathcal{X}. \end{aligned}$$

Выразим $\pi(n, m+1, k)$ из уравнений СУГБ при $(n, m, k) \in \mathcal{X}_0 \cup \mathcal{X}_1$

$$\begin{aligned} \pi(n, m+1, k) \cdot (m+1) \mu_2 = & \\ = \pi(n, m, k) \cdot [\lambda_1 \cdot 1\{n=0\} + \lambda_2 + \lambda_3 \cdot 1\{k=0\} + \mu_1 \cdot 1\{n=1\} + m\mu_2 + \mu_3 \cdot 1\{k=1\}] - & \\ - \pi(n-1, m, k) \cdot \lambda_1 \cdot 1\{n=1\} - & \\ - \pi(n, m-1, k) \cdot \lambda_2 \cdot 1\{m>0\} - & \\ - \pi(n, m, k-1) \cdot \lambda_3 \cdot 1\{k=1\} - & \\ - \pi(n+1, m, k) \cdot \mu_1 \cdot 1\{n=0, b+m+dk \leq V\} - & \\ - \pi(n, m, k+1) \cdot \mu_3 \cdot 1\{k=0, bn+m+d \leq V\}, \quad (n, m, k) \in \mathcal{X}_0 \cup \mathcal{X}_1. & \end{aligned}$$

Заменим $m+1$ на m , тогда уравнения будут справедливы для $(n, m-1, k) \in \mathcal{X}_0 \cup \mathcal{X}_1$, то есть для $(n, m, k) \in \mathcal{X}_1 \cup \mathcal{X}_2$, и $\pi(n, m, k)$ представляет собой линейную комбинацию $\pi(i, m-1, j)$ и $\pi(i, m-2, j)$

$$\begin{aligned} \pi(n, m, k) \cdot m \mu_2 = & \\ = \pi(n, m-1, k) \cdot [\lambda_1 \cdot 1\{n=0\} + \lambda_2 + \lambda_3 \cdot 1\{k=0\} + & \\ + \mu_1 \cdot 1\{n=1\} + (m-1)\mu_2 + \mu_3 \cdot 1\{k=1\}] - & \\ - \pi(n-1, m-1, k) \cdot \lambda_1 \cdot 1\{n=1\} - & \\ - \pi(n, m-1, k-1) \cdot \lambda_3 \cdot 1\{k=1\} - & \\ - \pi(n+1, m-1, k) \cdot \mu_1 \cdot 1\{n=0, b+m-1+dk \leq V\} - & \\ - \pi(n, m-1, k+1) \cdot \mu_3 \cdot 1\{k=0, bn+m-1+d \leq V\} - & \\ - \pi(n, m-2, k) \cdot \lambda_2 \cdot 1\{m-1>0\}, \quad (n, m, k) \in \mathcal{X}_1 \cup \mathcal{X}_2. & \end{aligned}$$

б) Выражение $\pi(n, m, k)$, $m > 0$ через $\pi(i, 0, j)$

Воспользовавшись рекуррентно полученным соотношением, можно представить $\pi(n, m, k)$ в виде линейной комбинации $\pi(i, 0, j)$

$$\begin{aligned} \pi(n, m, k) = A_{00}(n, m, k) \cdot \pi(0, 0, 0) + A_{10}(n, m, k) \cdot \pi(1, 0, 0) + & \\ + A_{01}(n, m, k) \cdot \pi(0, 0, 1) + A_{11}(n, m, k) \cdot \pi(1, 0, 1), \quad (n, m, k) \in \mathcal{X}_1 \cup \mathcal{X}_2. & \end{aligned}$$

Обозначим

$$X_{nk} = \pi(n, 0, k), \quad n, k = 0, 1,$$

тогда

$$\pi(n, m, k) = A_{00}(n, m, k) \cdot X_{00} + A_{10}(n, m, k) \cdot X_{10} + A_{01}(n, m, k) \cdot X_{01} + A_{11}(n, m, k) \cdot X_{11},$$

$$(n, m, k) \in \mathcal{X}_1 \cup \mathcal{X}_2.$$

Доопределим $A_{ij}(n, 0, k)$ для $(n, m, k) \in \mathcal{X}_0$

$$\begin{aligned} A_{00}(0, 0, 0) &= 1, & A_{10}(0, 0, 0) &= 0, & A_{01}(0, 0, 0) &= 0, & A_{11}(0, 0, 0) &= 0, \\ A_{00}(1, 0, 0) &= 0, & A_{10}(1, 0, 0) &= 1, & A_{01}(1, 0, 0) &= 0, & A_{11}(1, 0, 0) &= 0, \\ A_{00}(0, 0, 1) &= 0, & A_{10}(0, 0, 1) &= 0, & A_{01}(0, 0, 1) &= 1, & A_{11}(0, 0, 1) &= 0, \\ A_{00}(1, 0, 1) &= 0, & A_{10}(1, 0, 1) &= 0, & A_{01}(1, 0, 1) &= 0, & A_{11}(1, 0, 1) &= 1, \end{aligned}$$

тогда

$$\begin{aligned} \pi(n, m, k) &= A_{00}(n, m, k) \cdot X_{00} + A_{10}(n, m, k) \cdot X_{10} + \\ &+ A_{01}(n, m, k) \cdot X_{01} + A_{11}(n, m, k) \cdot X_{11}, \quad (n, m, k) \in \mathcal{X}. \end{aligned}$$

Подставим полученный вид $\pi(n, m, k)$ в СУГБ, и соотнеся коэффициенты при X_{ij} в левой и правой частях равенств, получим рекуррентные соотношения для расчета $A_{ij}(n, m, k)$

$$\begin{aligned} A_{ij}(n, m, k) \cdot m\mu_2 &= \\ &= A_{ij}(n, m-1, k) \cdot [\lambda_1 \cdot 1\{n=0\} + \lambda_2 + \lambda_3 \cdot 1\{k=0\} + \\ &+ \mu_1 \cdot 1\{n=1\} + (m-1)\mu_2 + \mu_3 \cdot 1\{k=1\}] - \\ &- A_{ij}(n-1, m-1, k) \cdot \lambda_1 \cdot 1\{n=1\} - \\ &- A_{ij}(n, m-1, k-1) \cdot \lambda_3 \cdot 1\{k=1\} - \\ &- A_{ij}(n+1, m-1, k) \cdot \mu_1 \cdot 1\{n=0, b+m-1+dk \leq V\} - \\ &- A_{ij}(n, m-1, k+1) \cdot \mu_3 \cdot 1\{k=0, bn+m-1+d \leq V\} - \\ &- A_{ij}(n, m-2, k) \cdot \lambda_2 \cdot 1\{m-1 > 0\}, \quad (n, m, k) \in \mathcal{X}_1 \cup \mathcal{X}_2, \quad i, j = 0, 1. \end{aligned}$$

в) Получение системы уравнений относительно $\pi(n, 0, k)$

Уравнения СУГБ при $(n, m, k) \in \mathcal{X}_2$ имеют вид

$$\begin{aligned} & \pi(n, m, k) \cdot [\lambda_1 \cdot 1\{n=0\} + \lambda_3 \cdot 1\{k=0\} + \mu_1 \cdot 1\{n=1\} + m\mu_2 + \mu_3 \cdot 1\{k=1\}] - \\ & - \pi(n-1, m, k) \cdot \lambda_1 \cdot 1\{n=1\} - \\ & - \sum_{i=V-b+1}^V \pi(n-1, i, 0) \cdot \lambda_1 \cdot 1\{(n, m, k) = (1, V-b, 0)\} - \\ & - \sum_{i=V-b-d+1}^{V-d} \pi(n-1, i, 1) \cdot \lambda_1 \cdot 1\{(n, m, k) = (1, V-b-d, 1)\} - \\ & - \pi(n, m-1, k) \cdot \lambda_2 - \\ & - \pi(n, m, k-1) \cdot \lambda_3 \cdot 1\{k=1\} - \\ & - \sum_{i=V-d+1}^V \pi(n, i, k-1) \cdot \lambda_3 \cdot 1\{(n, m, k) = (0, V-d, 1)\} - \\ & - \sum_{i=V-b-d+1}^{V-b} \pi(n, i, k-1) \cdot \lambda_3 \cdot 1\{(n, m, k) = (1, V-b-d, 1)\} = 0, \quad (n, m, k) \in \mathcal{X}_2. \end{aligned}$$

В уравнения при $(0, V, 0), (1, V-b, 0), (0, V-d, 1)$

$$\pi(0, V, 0) \cdot [\lambda_1 + \lambda_3] - \pi(0, V-1, 0) \cdot \lambda_2 = 0,$$

$$\pi(1, V-b, 0) \cdot [\lambda_3 + \mu_1 + (V-b)\mu_2] - \pi(1, V-b-1, 0) \cdot \lambda_2 - \sum_{i=V-b}^V \pi(0, i, 0) \cdot \lambda_1 = 0,$$

$$\pi(0, V-d, 1) \cdot [\lambda_1 + (V-d)\mu_2 + \mu_3] - \pi(0, V-d-1, 1) \cdot \lambda_2 - \sum_{i=V-d}^V \pi(0, i, 0) \cdot \lambda_3 = 0.$$

подставим $\pi(n, m, k)$ в виде линейной комбинации $\pi(i, 0, j)$

$$[\lambda_1 + \lambda_3] \sum_{i,j=0,1} A_{ij}(0, V, 0) \cdot X_{ij} - \lambda_2 \sum_{i,j=0,1} A_{ij}(0, V-1, 0) \cdot X_{ij} = 0,$$

$$[\lambda_3 + \mu_1 + (V-b)\mu_2] \sum_{i,j=0,1} A_{ij}(1, V-b, 0) \cdot X_{ij} - \lambda_2 \sum_{i,j=0,1} A_{ij}(1, V-b-1, 0) \cdot X_{ij} -$$

$$- \lambda_1 \sum_{m=V-b+1}^V \sum_{i,j=0,1} A_{ij}(0, m, 0) \cdot X_{ij} = 0,$$

$$[\lambda_1 + (V-d)\mu_2 + \mu_3] \sum_{i,j=0,1} A_{ij}(0, V-d, 1) \cdot X_{ij} - \lambda_2 \sum_{i,j=0,1} A_{ij}(0, V-d-1, 1) \cdot X_{ij} -$$

$$- \lambda_3 \sum_{m=V-d+1}^V \sum_{i,j=0,1} A_{ij}(0, m, 0) \cdot X_{ij} = 0.$$

Обозначим выражения при X_{ij} через $B_{ij}(l)$

$$\begin{aligned} B_{ij}(1) &= [\lambda_1 + \lambda_3] \cdot A_{ij}(0, V, 0) - \lambda_2 \cdot A_{ij}(0, V-1, 0), \quad i, j = 0, 1, \\ B_{ij}(2) &= [\lambda_3 + \mu_1 + (V-b)\mu_2] \cdot A_{ij}(1, V-b, 0) - \lambda_2 \cdot A_{ij}(1, V-b-1, 0) - \\ &\quad - \lambda_1 \sum_{m=V-b}^V A_{ij}(0, m, 0), \quad i, j = 0, 1, \\ B_{ij}(3) &= [\lambda_1 + (V-d)\mu_2 + \mu_3] \cdot A_{ij}(0, V-d, 1) - \lambda_2 \cdot A_{ij}(0, V-d-1, 1) - \\ &\quad - \lambda_3 \sum_{m=V-d}^V A_{ij}(0, m, 0), \quad i, j = 0, 1, \end{aligned}$$

тогда

$$B_{00}(l) \cdot X_{00} + B_{10}(l) \cdot X_{10} + B_{01}(l) \cdot X_{01} + B_{11}(l) \cdot X_{11} = 0, \quad l = 1, 2, 3.$$

Запишем условие нормировки $\pi \mathbf{1}^T = 1$

$$\sum_{(n,m,k) \in \mathcal{X}} \pi(n, m, k) = 1,$$

примем $X_{00} = 1$ и получаем систему уравнений относительно (X_{10}, X_{01}, X_{11})

$$B_{00}(l) + B_{10}(l) \cdot X_{10} + B_{01}(l) \cdot X_{01} + B_{11}(l) \cdot X_{11} = 0, \quad l = 1, 2, 3.$$

При этом по окончании вычислений, вероятности $\pi(n, m, k)$ надо будет отнормировать

$$\pi(n, m, k) \leftarrow \frac{\pi(n, m, k)}{\sum_{(i,l,j) \in \mathcal{X}} \pi(i, l, j)}, \quad (n, m, k) \in \mathcal{X}.$$

■

2.6.3 Средняя скорость передачи многоадресного трафика

Для модели с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания одноадресного трафика и снижением скорости передачи многоадресного трафика, из решения СУР $\pi \mathbf{Q} = \mathbf{0}$ с условием нормировки $\pi \mathbf{1}^T = 1$, вытекает следующее следствие.

Следствие 2.5. Для модели с одноадресным и многоадресным потоковым трафиком и постоянным объемом ресурса, схемой с прерыванием обслуживания одноадресного трафика и снижением скорости передачи многоадресного трафика, следующие показатели эффективности рассчитываются по формулам:

средняя скорость передачи многоадресного трафика,

$$\bar{b} = \left[b_1 \sum_{m=0}^{V-b_1} \pi(1, m) + \sum_{l=2}^L b_l \sum_{m=V-b_{l-1}+1}^{V-b_l} \pi(1, m) \right] \cdot \left[\sum_{m=0}^{V-b_L} \pi(1, m) \right]^{-1}; \quad (2.74)$$

вероятность прерывания обслуживания сессии одноадресного трафика

$$P_{\text{in}} = \sum_{m=V-b_L+1}^{V-1} \frac{\lambda_1}{\lambda_1 + \lambda_2 + m\mu_2} \cdot \frac{m - V + b_L}{m} \cdot \pi(0, m) + \frac{\lambda_1}{\lambda_1 + V\mu_2} \cdot \frac{b_L}{V} \cdot \pi(0, V); \quad (2.75)$$

вероятность блокировки сессии одноадресного трафика

$$P_{\text{bl}} = \pi(0, V) + \pi(1, V - b_L). \quad (2.76)$$

Задачу максимизации средней скорости многоадресного трафика по уровням скорости и при ограничениях на среднюю скорость многоадресного трафика, вероятность прерывания и вероятность блокировки сессии одноадресного трафика запишем следующим образом

$$\begin{aligned} (L^*, \mathbf{b}^*) &= \arg \max_{L, \mathbf{b}} \bar{b}(L, \mathbf{b}), \\ 1 \leq L \leq \hat{L}, \quad b_l &\in \{\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_{\hat{L}}\}, \quad l = 1, \dots, L, \\ \bar{b}(L, \mathbf{b}) &\geq \hat{b}, \\ P_{\text{in}}(L, \mathbf{b}) &\leq \hat{P}_{\text{in}}, \\ P_{\text{bl}}(L, \mathbf{b}) &\leq \hat{P}_{\text{bl}}. \end{aligned} \quad (2.77)$$

Задача выбора уровней $\mathbf{b} = (b_1, b_2, \dots, b_L)$ скорости передачи многоадресного трафика схожа с рассмотренной аналогичной задачей для широкополосного трафика в разделе 2.4.

Проиллюстрируем численный анализ задачи максимизации средней скорости видеоконференции. На рисунке 2.6 представлены графики средней скорости передачи многоадресного трафика в зависимости от предложенной нагрузки для разных наборов уровней скорости. Анализируются разные наборы

скоростей для видеоконференции. Наивысшая средняя скорость достигается при использовании набора с максимальной начальной скоростью, но это сопровождается высокими вероятностями блокировки и прерывания видео по запросу. Снижение минимальной скорости в наборе уменьшает среднюю скорость видео конференции, но улучшает качество обслуживания для видео по запросу. График показывает компромисс между качеством многоадресной передачи и стабильностью обслуживания одноадресного сервиса.

На рисунке 2.6 показаны области возможных значений уровней скорости многоадресного трафика в зависимости от предложенной нагрузки при ограничениях на показатели эффективности: среднюю скорость, вероятности прерывания обслуживания и блокировки. Это схема выбора наборов уровней скорости. По радиусу круга отложены интервалы суммарной нагрузки, а сектора соответствуют различным ограничениям – одинарным и попарным. На отрезке приведен случай одновременного выполнения всех трех ограничений.

Различные наборы уровней скорости соответствуют разным уровням нагрузки. При умеренной нагрузке оптимальным становится набор, который обеспечивает баланс между высокой средней скоростью видеоконференции с соблюдением заданных ограничений. При увеличении нагрузки требования к качеству обслуживания вынуждают переходить к более стабильным, но менее гибким наборам. График подчеркивает важность адаптивности скоростей для эффективного управления ресурсами в условиях переменного трафика.

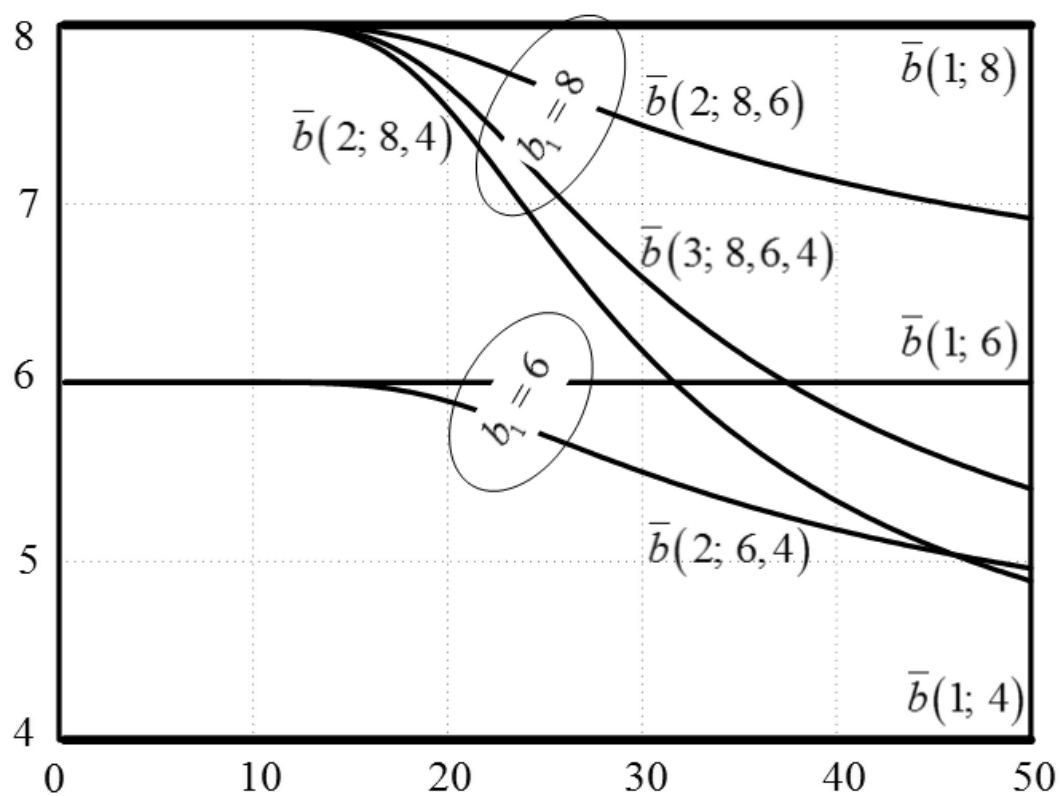


Рисунок 2.6 – Средняя скорость передачи многоадресного трафика (Мбит/с) в зависимости от предложенной нагрузки (Эрл) для разных наборов уровней скорости

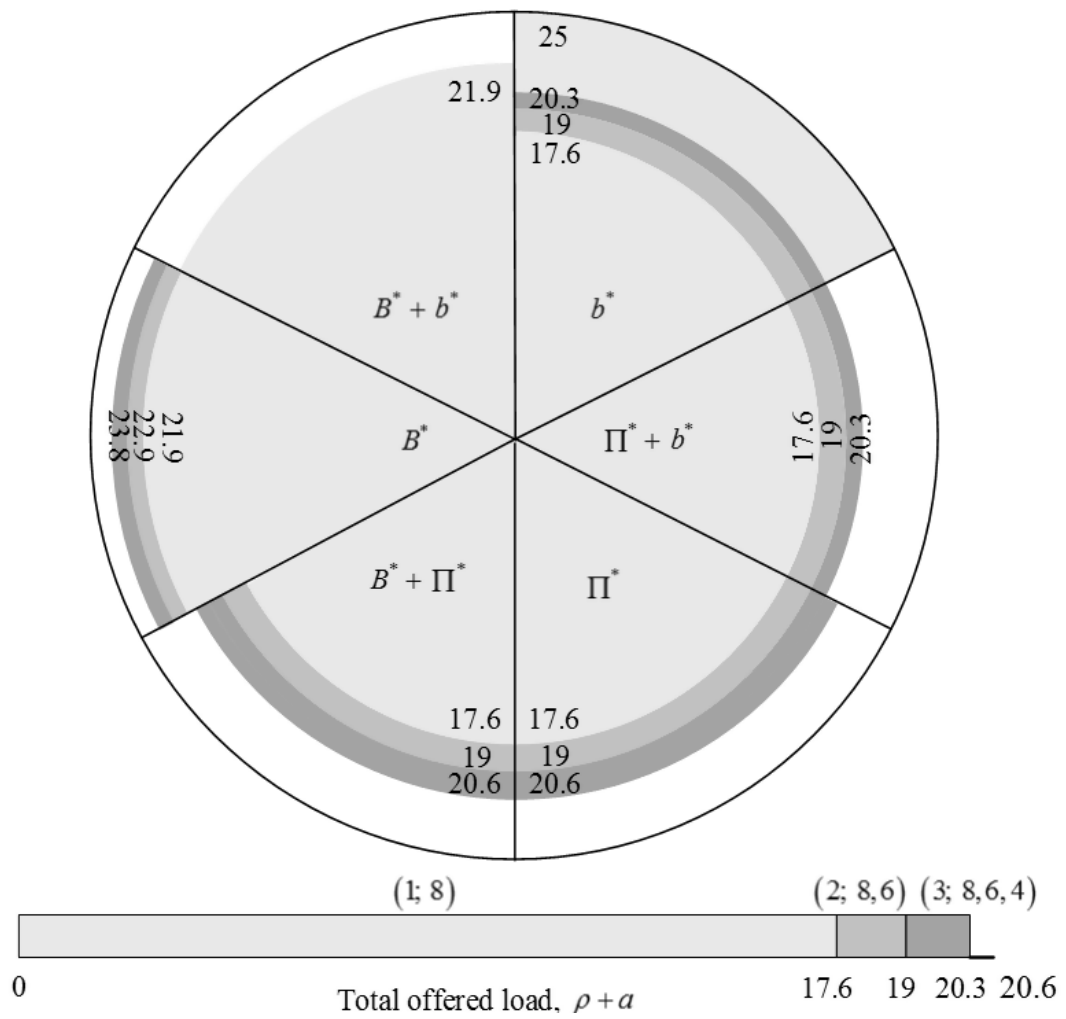


Рисунок 2.7 – Области возможных значений уровней скорости многоадресного трафика в зависимости от предложенной нагрузки (Эрл) при ограничениях на показатели эффективности: среднюю скорость, вероятности прерывания обслуживания и блокировки

Глава 3 МОДЕЛИ С ЭЛАСТИЧНЫМ ТРАФИКОМ И ПОСТОЯННЫМ ОБЪЕМОМ РЕСУРСА

3.1 МОДЕЛЬ ОБСЛУЖИВАНИЯ ТРАФИКА В НЕЛИЦЕНЗИРУЕМЫХ РАДИОЧАСТОТАХ

3.1.1 Системная модель

В разделах 3.1–3.2 главы 3 исследованы модели приоритетного доступа двух классов эластичного трафика – без требования и с требованием к минимальной скорости передачи. Трафик соответственно генерируется при передаче данных пользователей сети нелицензируемых радиочастот Wi-Fi и пользователей оператора 4G/5G – трафик, который был выгружен в сеть Wi-Fi. Приоритетным является трафик Wi-Fi, объем ресурса постоянный. В сети нелицензируемых радиочастот все пользователи не различимы вне зависимости от наличия требования к минимальной скорости передачи. Каждому пользователю выделяется равная доля от доступной пропускной способности.

Поскольку трафик является эластичным, любое событие, приводящее к изменению числа обслуживаемых сессий, также становится приоритетным. Если число сессий возрастает, скорость всех сессий снижается, а если уменьшается – повышается.

Рассмотрим несколько сценариев обслуживания эластичного трафика с требованием к минимальной скорости передачи. Для обеспечения приоритета трафика без требования минимальной скорости применяются два механизма доступа трафика с требованием к ней: снижение скорости передачи и прерывание обслуживания.

В таблице 3.1 приведены основные обозначения, параметры и компоненты состояния системы. Для моделирования эластичного трафика используется дисциплина эгалитарного разделения процессора EPS.

Таблица 3.1 – Обозначения, используемые для моделей с эластичным трафиком и полнодоступным ресурсом постоянного объема (глава 3, разделы 3.1–3.2)

Обозначение	Параметр
λ_1	Интенсивность пуассоновского потока запросов на передачу трафика без требования к минимальной скорости (1-класса)
λ_2	Интенсивность пуассоновского потока запросов на передачу трафика с требованием к минимальной скорости (2-класса)
b	Требование к минимальной скорости передачи трафика 2-класса
$1/\mu_1$	Средний экспоненциальный объем трафика 1-класса
$1/\mu_2$	Средний экспоненциальный объем трафика 2-класса
$M = \lfloor V/b \rfloor$	Максимальное число обслуживаемых сессий трафика 2-класса
n	Число обслуживаемых сессий трафика 1-класса
m	Число обслуживаемых сессий трафика 2-класса

3.1.2 Схемы приоритетного доступа

Рассмотрим следующие схемы приоритетного доступа для моделей с эластичным трафиком и полнодоступным ресурсом постоянного объема:

- (а) снижение скорости передачи трафика;
- (б) прерывание обслуживания трафика.

В этих схемах используются механизмы $p \in \{1, 3, 4\} \subset \mathcal{P}$ реализации приоритета и принципы $w \in \{1, 2\}$ выбора сессии для применения механизма из параграфа 1.6.3.

Опишем особенности исследуемых моделей с перечисленными схемами приоритетного доступа через матрицы приоритетов.

Схема (а) – базовая система с одним механизмом приоритетного доступа – снижением скорости передачи эластичного трафика. Это СМО $M_2 | M_2 | V | 0 | F_8$ с матрицей приоритетов

F_8		$\begin{matrix} p \\ e \end{matrix}$	эластичный трафик 1-класса обслуживается		эластичный трафик 2-класса обслуживается	
			3	4	3	4
эластичный трафик 1-класса	поступление	1	(1,2)	0	(1,2)	0
	завершение обслуживания	2	0	(1,2)	0	(1,2)
эластичный трафик 2-класса	поступление	3	(1,2)	0	(1,2)	0
	завершение обслуживания	4	0	(1,2)	0	(1,2)

Схема (б). В этой системе к базовому механизму снижения скорости передачи эластичного трафика добавляется еще один – прерывание обслуживания трафика с требованием к минимальной скорости (2-класса). Это СМО $M_2 | M_2 | V | 0 | F_9$ с матрицей приоритетов

F_9		$\begin{matrix} p \\ e \end{matrix}$	эластичный трафик 1-класса обслуживается		эластичный трафик 2-класса обслуживается		
			3	4	3	1	4
эластичный трафик 1-класса	поступление	1	(1,2)	0	(1,2)	(2,1)	0
	завершение обслуживания	2	0	(1,2)	0	0	(1,2)
эластичный трафик 2-класса	поступление	3	(1,2)	0	(1,2)	0	0
	завершение обслуживания	4	0	(1,2)	0	0	(1,2)

3.1.3 Модели математической теории телетрафика

Функционирование систем опишем при помощи СП $X(t)$ над пространствами состояний

модель (а)

$$\mathcal{X} = \{(n, m): n \geq 0, m \geq 0\}; \quad (3.1)$$

модель (б)

$$\mathcal{X} = \{(n, m): n \geq 0, m \geq 0, 1\{m > 0\} \cdot (n + m) \leq M\}. \quad (3.2)$$

Ненулевые элементы матриц интенсивностей переходов имеют вид
модель (а)

$$\begin{aligned} Q[(n, m), (n + 1, m)] &= \lambda_1, \\ Q[(n, m), (n, m + 1)] &= \lambda_2, \\ Q[(n, m), (n - 1, m)] &= \frac{n}{n + m} V \mu_1, \quad n > 0, \\ Q[(n, m), (n, m - 1)] &= \frac{m}{n + m} V \mu_2, \quad m > 0; \end{aligned} \quad (3.3)$$

модель (б)

$$\begin{aligned} Q[(n, 0), (n + 1, 0)] &= \lambda_1, \\ Q[(n, m), (n + 1, m)] &= \lambda_1, \quad m > 0, n + 1 + m \leq M, \\ Q[(n, m), (n + 1, m - 1)] &= \lambda_1, \quad m > 0, n + 1 + m > M, \\ Q[(n, m), (n, m + 1)] &= \lambda_2, \quad n + m + 1 \leq M, \\ Q[(n, m), (n - 1, m)] &= \frac{n}{n + m} V \mu_1, \quad n > 0, \\ Q[(n, m), (n, m - 1)] &= \frac{m}{n + m} V \mu_2, \quad m > 0. \end{aligned} \quad (3.4)$$

В следующем разделе 3.2 описаны результаты исследования моделей схем приоритетного доступа с эластичным трафиком и полнодоступным ресурсом постоянного объема. Приведены показатели эффективности приоритетного доступа, включая вероятность снижения скорости передачи эластичного трафика ниже минимального порога.

3.2 СИСТЕМА С МИНИМАЛЬНОЙ СКОРОСТЬЮ ПЕРЕДАЧИ

3.2.1 Мультипликативный вид стационарного распределения

Для модели с эластичным трафиком и полнодоступным ресурсом постоянного объема, схемой со снижением скорости передачи трафика, справедлива следующая теорема.

Теорема 3.1. Для модели с эластичным трафиком и полнодоступным ресурсом постоянного объема, схемой со снижением скорости передачи трафика, стационарное распределение СП $X(t)$ рассчитывается в мультипликативном виде по формуле

$$\begin{aligned}\pi(n, m) &= \left(\frac{\lambda_1}{\mu_1} \right)^n \left(\frac{\lambda_2}{\mu_2} \right)^m \frac{1}{V^{n+m}} \frac{(n+m)!}{n!m!} \cdot \pi(0, 0), \quad (n, m) \in \mathcal{X} \setminus \{(0, 0)\}, \\ \pi(0, 0) &= \left(\sum_{(n, m) \in \mathcal{X}} \left(\frac{\lambda_1}{\mu_1} \right)^n \left(\frac{\lambda_2}{\mu_2} \right)^m \frac{1}{V^{n+m}} \frac{(n+m)!}{n!m!} \right)^{-1}.\end{aligned}\tag{3.5}$$

Доказательство.

а) *Запись системы уравнений частичного баланса*

СП $X(t)$ является обратимым, поскольку для него выполнен критерий Колмогорова. Продемонстрируем это на примере последовательности четырех состояний «замкнутого контура»

$$\begin{aligned}(n, m) &\rightarrow (n-1, m) \rightarrow (n-1, m-1) \rightarrow (n, m-1) \rightarrow (n, m), \\ (n, m) &\rightarrow (n, m-1) \rightarrow (n-1, m-1) \rightarrow (n-1, m) \rightarrow (n, m).\end{aligned}$$

Произведения интенсивностей переходов между состояниями системы при движении в разных направлениях по контуру совпадают

$$\begin{aligned}&\mathbf{Q}[(n, m), (n-1, m)] \cdot \mathbf{Q}[(n-1, m), (n-1, m-1)] \times \\ &\quad \times \mathbf{Q}[(n-1, m-1), (n, m-1)] \cdot \mathbf{Q}[(n, m-1), (n, m)] = \\ &= \frac{n}{n+m} V \mu_1 \cdot \frac{m}{n+m-1} V \mu_2 \cdot \lambda_1 \cdot \lambda_2,\end{aligned}$$

$$\begin{aligned} & \mathbf{Q}[(n, m), (n, m-1)] \cdot \mathbf{Q}[(n, m-1), (n-1, m-1)] \times \\ & \quad \times \mathbf{Q}[(n-1, m-1), (n-1, m)] \cdot \mathbf{Q}[(n-1, m), (n, m)] = \\ & = \frac{m}{n+m} V \mu_2 \cdot \frac{n}{n+m-1} V \mu_1 \cdot \lambda_2 \cdot \lambda_1. \end{aligned}$$

Для обратимого СП $\mathbf{X}(t)$ запишем систему уравнений частичного баланса (СУЧБ) с учетом ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (3.3) при $(n, m) \in \mathcal{X} \setminus \{(0, 0)\}$

$$\frac{n}{n+m} V \mu_1 \cdot \pi(n, m) = \lambda_1 \cdot \pi(n-1, m), \quad n > 0,$$

$$\frac{m}{n+m} V \mu_2 \cdot \pi(n, m) = \lambda_2 \cdot \pi(n, m-1), \quad m > 0,$$

Следовательно, зависимость $\pi(n, m)$ от $\pi(n-1, m)$ и $\pi(n, m)$ от $\pi(n, m-1)$ имеет вид

$$\pi(n, m) = \frac{\lambda_1}{V \mu_1} \frac{n+m}{n} \cdot \pi(n-1, m), \quad n > 0,$$

$$\pi(n, m) = \frac{\lambda_2}{V \mu_2} \frac{n+m}{m} \cdot \pi(n, m-1), \quad m > 0.$$

б) Выражение $\pi(n, m)$ через $\pi(0, 0)$

Выразим $\pi(n, m)$ через $\pi(0, 0)$, рекуррентно подставляя полученные соотношения

$$\begin{aligned} \pi(n, m) &= \frac{\lambda_1}{V \mu_1} \frac{n+m}{n} \cdot \pi(n-1, m) = \left(\frac{\lambda_1}{V \mu_1} \right)^2 \frac{n+m}{n} \frac{n+m-1}{n-1} \cdot \pi(n-1, m) = \dots = \\ &= \left(\frac{\lambda_1}{V \mu_1} \right)^n \frac{n+m}{n} \frac{n-1+m}{n-1} \dots \frac{1+m}{1} \cdot \pi(0, m) = \\ &= \left(\frac{\lambda_1}{V \mu_1} \right)^n \frac{\lambda_2}{V \mu_2} \frac{n+m}{n} \frac{n-1+m}{n-1} \dots \frac{1+m}{1} \frac{m}{m} \cdot \pi(0, m-1) = \\ &= \left(\frac{\lambda_1}{V \mu_1} \right)^n \left(\frac{\lambda_2}{V \mu_2} \right)^2 \frac{n+m}{n} \frac{n-1+m}{n-1} \dots \frac{1+m}{1} \frac{m}{m} \frac{m-1}{m-1} \cdot \pi(0, m-2) = \dots = \end{aligned}$$

$$\begin{aligned}
 &= \left(\frac{\lambda_1}{V\mu_1} \right)^n \left(\frac{\lambda_2}{V\mu_2} \right)^m \frac{n+m}{n} \frac{n-1+m}{n-1} \dots \frac{1+m}{1} \frac{m}{m} \frac{m-1}{m-1} \dots \frac{1}{1} \cdot \pi(0,0) = \\
 &= \left(\frac{\lambda_1}{V\mu_1} \right)^n \left(\frac{\lambda_2}{V\mu_2} \right)^m \frac{(n+m)!}{n!m!} \cdot \pi(0,0), \quad (n,m) \in \mathcal{X} \setminus \{(0,0)\}.
 \end{aligned}$$

в) Нахождение $\pi(0,0)$

Запишем условие нормировки $\pi \mathbf{1}^T = 1$

$$\sum_{(n,m) \in \mathcal{X}} \pi(n,m) = 1$$

и подставим полученное выражение $\pi(n,m)$ через $\pi(0,0)$

$$\pi(0,0) \left[1 + \sum_{(n,m) \in \mathcal{X} \setminus \{(0,0)\}} \left(\frac{\lambda_1}{V\mu_1} \right)^n \left(\frac{\lambda_2}{V\mu_2} \right)^m \frac{(n+m)!}{n!m!} \right] = 1.$$

Таким образом, получаем вид $\pi(0,0)$

$$\pi(0,0) = \left(\sum_{(n,m) \in \mathcal{X}} \left(\frac{\lambda_1}{\mu_1} \right)^n \left(\frac{\lambda_2}{\mu_2} \right)^m \frac{1}{V^{n+m}} \frac{(n+m)!}{n!m!} \right)^{-1}.$$

■

3.2.2 Показатели эффективности передачи эластичного трафика

Для модели с эластичным трафиком и полнодоступным ресурсом постоянного объема, схемой со снижением скорости передачи трафика, из леммы 3.1 вытекает следующее следствие.

Следствие 3.1. Для модели с эластичным трафиком и полнодоступным ресурсом постоянного объема, схемой со снижением скорости передачи трафика, вероятность снижения скорости передачи трафика 2-класса ниже минимального порога рассчитывается по формуле

$$P_{\text{deg}} = 1 - \sum_{n=0}^M \sum_{m=0}^{M-n} \pi(n,m). \quad (3.6)$$

Для модели с эластичным трафиком и полнодоступным ресурсом постоянного объема, схемой с прерыванием обслуживания трафика, решения СУР $\pi Q = 0$ с условием нормировки $\pi \mathbf{1}^T = 1$, вытекает следующее следствие.

Следствие 3.2. Для модели с эластичным трафиком и полнодоступным ресурсом постоянного объема, схемой с прерыванием обслуживания трафика, следующие показатели эффективности обслуживания эластичного трафика 2-класса рассчитываются по формулам:

вероятность прерывания обслуживания сессии эластичного трафика 2-класса

$$P_{\text{in}} = \sum_{m=1}^M \frac{1}{m} \cdot \frac{\lambda_1}{\lambda_1 + \lambda_2 + \frac{M-m}{M} V \mu_1 + \frac{m}{M} V \mu_2} \cdot \pi(M-m, m); \quad (3.7)$$

средняя скорость передачи эластичного трафика 2-класса

$$\bar{b} = \left[\sum_{m=1}^M \sum_{n=0}^{M-m} \frac{V}{n+m} \cdot \pi(n, m) \right] \left[\sum_{m=1}^M \sum_{n=0}^{M-m} \pi(n, m) \right]^{-1}; \quad (3.8)$$

доля ресурса, занятая эластичным трафиком 2-класса

$$\text{UTIL}_2 = \sum_{m=1}^M \sum_{n=0}^{M-m} \frac{m}{n+m} \cdot \pi(n, m). \quad (3.9)$$

3.2.3 Анализ вероятности снижения скорости передачи

Рассмотрим сценарий использования технологии LAA для обслуживания трафика в нелицензируемых радиочастотах. В условиях нехватки лицензируемого спектра и растущего числа пользователей, критически важно использовать нелицензируемые радиочастоты, такие как 2.4 ГГц и 5 ГГц, для разгрузки трафика LTE. Однако эти частоты уже заняты Wi-Fi, что вызывает конкуренцию за ресурсы. Технология LAA объединяет LTE и нелицензируемый спектр через механизмы контроля и динамического распределения ресурсов. Это сохраняет минимальные гарантии пропускной способности и задержки, что важно для приложений, чувствительных к качеству обслуживания.

Рассмотрим два сценария выгрузки трафика. Первый сценарий соответствует схеме со снижением скорости передачи трафика – без использования LAA. Все сессии принимаются без ограничений, что увеличивает емкость системы. Однако при высокой нагрузке это приводит к резкому ухудшению качества обслуживания: повышается вероятность нарушения минимальной скорости.

Второй сценарий – с использованием LAA. Базовая станция с LAA управляет доступом к каналу и перераспределяет ресурсы между LTE и Wi-Fi сессиями. При появлении новых Wi-Fi-сессий, система адаптирует долю времени для каждого типа трафика, чтобы избежать ухудшения QoS для текущих LTE-сессий. Такой подход обеспечивает стабильность, но ограничивает количество подключений.

LAA лучше всего подходит для ситуаций с высокой нагрузкой и строгими требованиями к качеству обслуживания. В то же время, разгрузка на Wi-Fi эффективна при низком объеме трафика, когда допустимо балансировать между количеством подключений и качеством обслуживания.

Проиллюстрируем поведение вероятности снижения скорости передачи эластичного трафика ниже минимального порога для первого сценария. На рисунках 3.1, 3.2 и 3.3 представлены соответствующие графики. На рисунке 3.1 показана зависимость от интенсивности поступления запросов на передачу трафика с требованием к минимальной скорости для разной интенсивности поступления запросов на передачу трафика без требования к минимальной скорости. При низкой интенсивности риск нарушения QoS невелик. Но при увеличении нагрузки вероятность резко увеличивается. Это говорит о неспособности Wi-Fi справляться с высокой нагрузкой. Даже небольшой рост трафика LTE приводит к значительному ухудшению качества обслуживания.

Рисунок 3.2 демонстрирует зависимость от интенсивности поступления запросов на передачу трафика без требования к минимальной скорости для разной интенсивности поступления запросов на передачу трафика с требованием к минимальной скорости. Даже при низкой интенсивности Wi-Fi вероятность нарушения требований к скорости для LTE высока. Это связано с неконтролируемым доступом Wi-Fi к каналу. При увеличении нагрузки

вероятность нарушения QoS растет. Это подтверждает, что Wi-Fi доминирует в нелицензируемом спектре, конкурируя за ресурсы и делая выгрузку трафика LTE ненадежной даже при умеренной нагрузке.

Рисунок 3.3 показывает зависимость от требования к минимальной скорости для разной интенсивности поступления запросов на передачу трафика без требования к минимальной скорости. При увеличении требования вероятность возрастает. Высокие требования к скорости ужесточают условия для системы: даже незначительные колебания нагрузки приводят к снижению пропускной способности ниже заданного уровня.

Все графики демонстрируют основную проблему Wi-Fi-разгрузки: отсутствие механизмов резервирования ресурсов приводит к неконтролируемой конкуренции за канал. Даже умеренный рост нагрузки или повышение требований к скорости делают QoS ненадежным. Это ограничивает использование Wi-Fi для QoS-чувствительных сервисов. В таких случаях целесообразно применять технологию LAA, которая обеспечивает более стабильное качество обслуживания.

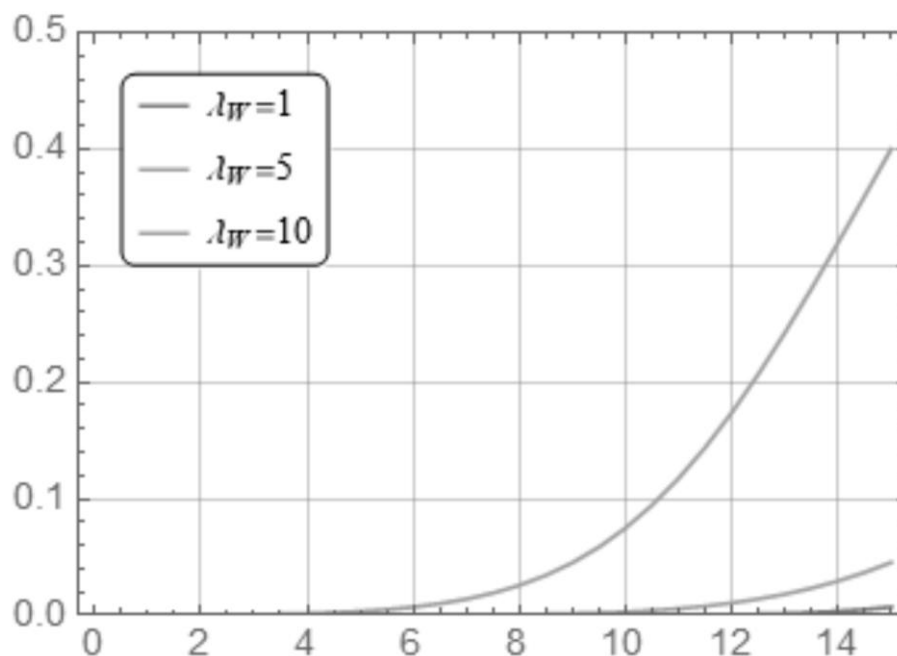


Рисунок 3.1 – Вероятность снижения скорости передачи эластичного трафика ниже минимального порога в зависимости от интенсивности поступления запросов на передачу трафика с требованием к минимальной скорости (1/с) для разной интенсивности поступления запросов на передачу трафика без требования к минимальной скорости (1/с)

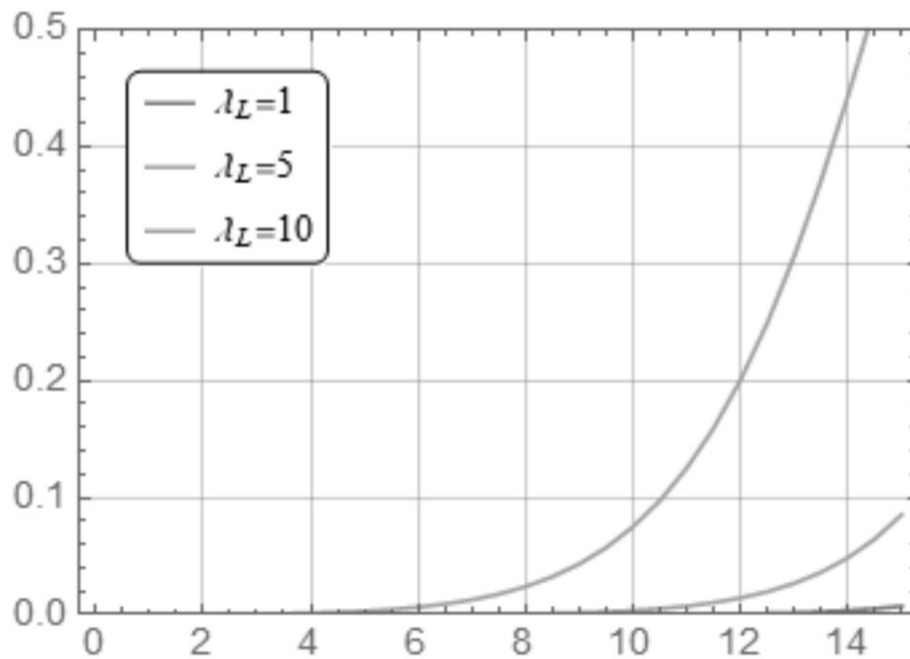


Рисунок 3.2 – Вероятность снижения скорости передачи эластичного трафика ниже минимального порога в зависимости от интенсивности поступления запросов на передачу трафика без требования к минимальной скорости (1/с) для разной интенсивности поступления запросов на передачу трафика с требованием к минимальной скорости (1/с)

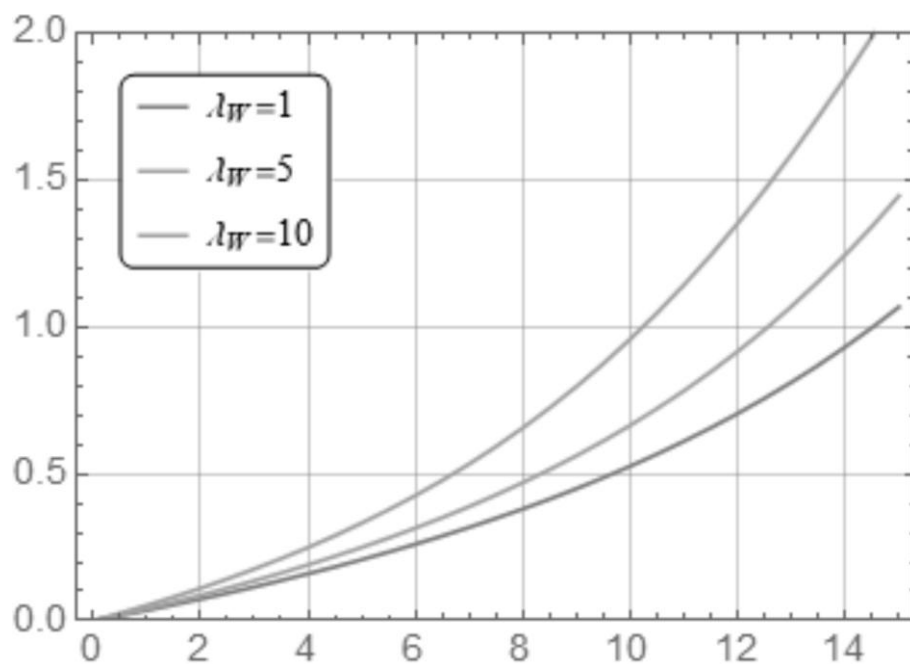


Рисунок 3.3 – Вероятность снижения скорости передачи эластичного трафика ниже минимального порога в зависимости от требования к минимальной скорости (Мбит/с) для разной интенсивности поступления запросов на передачу трафика без требования к минимальной скорости (1/с)

3.3 МОДЕЛЬ НАРЕЗКИ РАДИОРЕСУРСОВ

3.3.1 Системная модель

В разделах 3.3–3.6 главы 3 исследованы модели приоритетного доступа двух классов нетерпеливого эластичного трафика в условиях перераспределения ресурса между ними – между двумя сегментами сети с нарезкой радиоресурсов. При этом передача трафика допускает задержки, а скорость больше и равна минимальному порогу, одинаковому для обоих сегментов.

Приоритетными являются сообщения – сигналы контроллера для принятия решения о перераспределении ресурса. В интервалах времени между поступлениями сигналов распределение ресурса фиксированное. Когда поступает сигнал, перераспределение может как произойти, так и не произойти. Это зависит от ожидающих начала обслуживания сессий трафика. Перераспределение возможно только тогда, когда в одном из сегментов есть ожидающие сессии, а в другом их нет, и есть возможность обслуживать дополнительно хотя бы одну сессию. В противном случае перераспределение не осуществляется.

При приоритетном перераспределении ресурса по сигналу применяется механизм уменьшения пропускной способности у одного из сегментов, и ее увеличение у второго. Поскольку трафик является эластичным, любое событие, приводящее к изменению числа обслуживаемых сессий, также становится приоритетным. Если число сессий возрастает, скорость всех сессий в сегменте снижается, а если уменьшается – повышается.

Рассмотрим два сценария перераспределения ресурса. В первом случае применяется фиксированная политика, основанная на одном критерии – максимизации использования ресурса. Во втором случае политика является динамической и позволяет управлять перераспределением ресурса, основываясь на трех критериях. Помимо эффективного использования ресурса, она учитывает соответствие разделения пропускной способности соглашению SLA и

результативность сигналов, т.е. минимизирует ситуации, когда перераспределения не происходит, несмотря на наличие сигнала.

В таблице 3.2 приведены основные обозначения, параметры и компоненты состояния системы. Для моделирования сигналов контролера используется СМО с сигналами – потоком внешних событий, механизма приостановки обслуживания – СМО с двумя очередями, а для управления перераспределением ресурса – УСМО и МППР.

Таблица 3.2 – Обозначения, используемые для моделей с эластичным трафиком и перераспределением ресурса постоянного объема (глава 3, разделы 3.3–3.6)

Обозначение	Параметр
λ_1	Интенсивность пуассоновского потока запросов на передачу эластичного трафика 1-класса
λ_2	Интенсивность пуассоновского потока запросов на передачу эластичного трафика 2-класса
b	Требование к минимальной скорости передачи трафика
$1/\mu_1$	Средний экспоненциальный объем трафика 1-класса
$1/\mu_2$	Средний экспоненциальный объем трафика 2-класса
$N = \lfloor V/b \rfloor$	Максимальное число обслуживаемых сессий трафика
R_1	Число мест в очереди для ожидающих начала обслуживания сессий трафика 1-класса
R_2	Число мест в очереди для ожидающих начала обслуживания сессий трафика 2-класса
$1/\varepsilon_1$	Среднее экспоненциальное время на ожидание начала обслуживания нетерпеливого трафика 1-класса
$1/\varepsilon_2$	Среднее экспоненциальное время на ожидание начала обслуживания нетерпеливого трафика 2-класса
δ	Интенсивность пуассоновского потока сигналов на перераспределение ресурса
k	Максимальное число обслуживаемых сессий трафика 1-класса
u_1	Число обслуживаемых сессий трафика 1-класса
u_2	Число обслуживаемых сессий трафика 2-класса
r_1	Число ожидающих начала обслуживания сессий трафика 1-класса

Обозначение	Параметр
r_2	Число ожидающих начала обслуживания сессий трафика 2-класса
$n = u_1 + r_1$	Число сессий трафика 1-класса в системе – обслуживаемых и в очереди
$m = u_2 + r_2$	Число сессий трафика 2-класса в системе – обслуживаемых и в очереди

3.3.2 Схемы приоритетного доступа

Рассмотрим следующие схемы приоритетного доступа для моделей с эластичным трафиком и перераспределением ресурса постоянного объема:

- (а) фиксированная политика перераспределения ресурса между классами трафика;
- (б) политика управления перераспределением ресурса между классами трафика.

В этих схемах используются механизмы $p \in \{3,4,10,11\} \subset \mathcal{P}$ реализации приоритета и принципы $w=2$ выбора сессии для применения механизма из параграфа 1.6.3.

Опишем особенности исследуемых моделей с перечисленными схемами приоритетного доступа через матрицы приоритетов.

Схема (а) – система с фиксированной политикой перераспределения ресурса между классами эластичного трафика. Политика направлена на максимальное использование ресурса при поступлении сигнала. Она применяется, когда в одном из сегментов есть ожидающие начала обслуживания сессии трафика, в то время как в другом сегменте их нет, и есть возможность обслужить еще хотя бы одну сессию.

Если за счет передачи части ресурса от «простаивающего» сегмента к «перегруженному» можно обеспечить минимальную скорость для всех ожидающих сессий, то ресурс перераспределяется между сегментами на этот необходимый минимум. После такого перераспределения ожидающие сессии отсутствуют, а скорость всех сессий в ранее перегруженном сегменте становится минимальной b .

Если обеспечить минимальную скорость для всех ожидающих сессий невозможно, то ресурс перераспределяется между сегментами максимально. В результате скорость всех сессий в обоих сегментах становится минимальной b , а часть ожидающих сессий останется. Это СМО $M_2 | M_2 | V | \langle R_1, R_2 \rangle | F_{10}$ с матрицей приоритетов

$F_{10} = F_{11}$		$\begin{matrix} p \\ e \end{matrix}$	ресурс				эластичный трафик			
			1 сегмент		2 сегмент		1 сегмент обслуживается		2 сегмент обслуживается	
			10	11	10	11	3	4	3	4
сигнал поступает	умен. ресурс 1 сегм.	1	1	0	0	1	(1,2)	0	0	0
	увелич. ресурс 1 сегм.	2	0	1	1	0	0	0	(1,2)	0
эластичный трафик 1 сегмента	поступ.	3	0	0	0	0	(1,2)	0	0	0
	заверш. ождан.	4	0	0	0	0	0	0	0	0
	заверш. обслуж.	5	0	0	0	0	0	(1,2)	0	0
эластичный трафик 2 сегмента	поступ.	6	0	0	0	0	0	0	(1,2)	0
	заверш. ождан.	7	0	0	0	0	0	0	0	0
	заверш. обслуж.	8	0	0	0	0	0	0	0	(1,2)

Схема (б) – система с политикой управления перераспределением ресурса между классами трафика. В отличие от предыдущей модели, она стремится не только максимально использовать ресурс при поступлении сигнала, но и соблюдать равное разделение ресурса по SLA и высокую результативность сигналов, т.е. когда они приводят к перераспределению ресурса. Значимость каждого из критериев эффективности перераспределения ресурса определяется весовыми коэффициентами ω_i , $i=1,2,3$. Это УСМО $M_2 | M_2 | V | \langle R_1, R_2 \rangle | F_{11}$ с матрицей приоритетов, аналогичной схеме (а) $F_{11} = F_{10}$.

3.3.3 Модели математической теории телетрафика

Функционирование систем опишем при помощи СП $X(t)$ над пространствами состояний

модель (а)

$$\begin{aligned} \mathcal{X} = & \{(k, u_1, u_2, 0, 0): 0 \leq k \leq N, 0 \leq u_1 \leq k, 0 \leq u_2 \leq N - k\} \cup \\ & \cup \{(k, k, u_2, r_1, 0): 0 \leq k \leq N, 0 \leq u_2 \leq N - k, 0 < r_1 \leq R_1\} \cup \\ & \cup \{(k, u_1, N - k, 0, r_2): 0 \leq k \leq N, 0 \leq u_1 \leq k, 0 < r_2 \leq R_2\} \cup \\ & \cup \{(k, k, N - k, r_1, r_2): 0 \leq k \leq N, 0 < r_1 \leq R_1, 0 < r_2 \leq R_2\}; \end{aligned} \quad (3.10)$$

модель (б)

$$\mathcal{X} = \{(k, n, m): 0 \leq k \leq N, 0 \leq n \leq k + R_1, 0 \leq m \leq N - k + R_2\}. \quad (3.11)$$

Ненулевые элементы матриц интенсивностей переходов имеют вид

модель (а)

$$\begin{aligned} Q[(k, u_1, u_2, 0, r_2), (k, u_1 + 1, u_2, 0, r_2)] &= \lambda_1, \quad u_1 + 1 \leq k, \\ Q[(k, k, u_2, r_1, r_2), (k, k, u_2, r_1 + 1, r_2)] &= \lambda_1, \quad r_1 + 1 \leq R_1, \\ Q[(k, u_1, u_2, r_1, 0), (k, u_1, u_2 + 1, r_1, 0)] &= \lambda_2, \quad u_2 + 1 \leq N - k, \\ Q[(k, u_1, N - k, r_1, r_2), (k, k, u_2, r_1 + 1, r_2)] &= \lambda_2, \quad r_2 + 1 \leq R_2, \\ Q[(k, u_1, u_2, 0, r_2), (k, u_1 - 1, u_2, 0, r_2)] &= \frac{k}{N} V \mu_1, \quad u_1 > 0, \\ Q[(k, k, u_2, r_1, r_2), (k, k, u_2, r_1 - 1, r_2)] &= \frac{k}{N} V \mu_1, \quad r_1 > 0, \\ Q[(k, u_1, u_2, r_1, 0), (k, u_1, u_2 - 1, r_1, 0)] &= \frac{N - k}{N} V \mu_2, \quad u_2 > 0, \\ Q[(k, u_1, N - k, r_1, r_2), (k, u_1, N - k, r_1, r_2 - 1)] &= \frac{N - k}{N} V \mu_2, \quad r_2 > 0, \\ Q[(k, k, u_2, r_1, r_2), (k, k, u_2, r_1 - 1, r_2)] &= r_1 \varepsilon_1, \quad r_1 > 0, \\ Q[(k, u_1, N - k, r_1, r_2), (k, u_1, N - k, r_1, r_2 - 1)] &= r_2 \varepsilon_2, \quad r_2 > 0, \end{aligned} \quad (3.12)$$

$$\mathbf{Q}[(k, u_1, N - k, 0, r_2), (k - r_2, u_1, N - k + r_2, 0, 0)] = \delta,$$

$$u_1 < k, \quad 0 < r_2 \leq k - u_1,$$

$$\mathbf{Q}[(k, u_1, N - k, 0, r_2), (u_1, u_1, N - u_1, 0, r_2 - k + u_1)] = \delta,$$

$$u_1 < k, \quad r_2 > k - u_1,$$

$$\mathbf{Q}[(k, k, u_2, r_1, 0), (k + r_1, k + r_1, u_2, 0, 0)] = \delta,$$

$$u_2 < N - k, \quad 0 < r_1 \leq N - k - u_2,$$

$$\mathbf{Q}[(k, k, u_2, r_1, 0), (N - u_2, N - u_2, u_2, r_1 - N + k + u_2, 0)] = \delta,$$

$$u_2 < N - k, \quad r_1 > N - k - u_2;$$

модель (б)

$$\mathbf{Q}_f[(k, n, m), (k, n + 1, m)] = \lambda_1, \quad n + 1 \leq k + R_1,$$

$$\mathbf{Q}_f[(k, n, m), (k, n, m + 1)] = \lambda_2, \quad m + 1 \leq N - k + R_2,$$

$$\mathbf{Q}_f[(k, n, m), (k, n - 1, m)] = \frac{k}{N} V \mu_1 \cdot 1\{k > 0, n > 0\} + (n - k) \varepsilon_1 \cdot 1\{n > k\},$$

$$\mathbf{Q}_f[(k, n, m), (k, n, m - 1)] =$$

$$= \frac{N - k}{N} V \mu_2 \cdot 1\{N - k > 0, m > 0\} + (m - k + N) \varepsilon_2 \cdot 1\{m > N - k\},$$

$$\mathbf{Q}_f[(k, n, m), (a, n, m)] = \delta, \quad (n < k, m > N - k, n + m \leq N) \vee$$

$$\vee (n < k, m > N - k, n + m > N) \vee (n > k, m < N - k, n + m \leq N) \vee$$

$$\vee (n > k, m < N - k, n + m > N),$$

(3.13)

где f фиксированная политика, a конкретное действие при политике f .

Подробнее будет пояснено в разделе 3.5.

В следующих разделах 3.4–3.6 описаны результаты исследования моделей схем приоритетного доступа с эластичным трафиком и перераспределением ресурса постоянного объема. Для модели (а) матрица интенсивностей переходов представлена в блочно-трехдиагональном виде, а стационарное распределение рассчитывается с помощью матричного алгоритма (раздел 3.4).

Для модели (б) построен МППР с функцией вознаграждения, которая учитывает три ранее упомянутых критерия (раздел 3.5). Получен вид среднего вознаграждения для произвольной фиксированной политики и сформулирована задача оптимизации. Эта функция вознаграждения используется в системе

уравнений относительно вознаграждения за пребывание в различных состояниях системы при фиксированной политике (раздел 3.6).

Для поиска оптимальной политики использованы два подхода (раздел 3.6): метод простой итерации для решения системы линейных уравнений и с применением машинного обучения. Итерационный алгоритм Ховарда Р.А. является базовым для МППР и УМО, и для него выписаны формулы для расчета нового приближения политики на текущем шаге итерации. В качестве искусственной нейронной сети использован однослойный перцептрон, на вход которого подаются состояния системы, а на выходе генерируется оптимальная политика для каждого состояния.

3.4 СИСТЕМА С СИГНАЛАМИ ДЛЯ ПЕРЕРАСПРЕДЕЛЕНИЯ РЕСУРСА

3.4.1 Матрица интенсивностей переходов в блочно-трехдиагональном виде

Лемма 3.1. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с фиксированной политикой перераспределения ресурса между классами трафика, если на пространстве $\mathcal{X}$ состояний СП $X(t)$ введен лексикографический порядок

$$\begin{aligned} \mathbf{x} = (k, u_1, u_2, r_1, r_2) < (k', u'_1, u'_2, r'_1, r'_2) = \mathbf{x}' \Leftrightarrow \\ (u_1 + u_2 + r_1 + r_2 < u'_1 + u'_2 + r'_1 + r'_2) \vee \\ \vee (u_1 + u_2 + r_1 + r_2 = u'_1 + u'_2 + r'_1 + r'_2, k < k') \vee \\ \vee (u_1 + u_2 + r_1 + r_2 = u'_1 + u'_2 + r'_1 + r'_2, \\ k = k', u_1 + r_1 < u'_1 + r'_1), \quad \mathbf{x}, \mathbf{x}' \in \mathcal{X}, \end{aligned} \quad (3.14)$$

тогда матрица интенсивностей переходов СП $\mathbf{X}(t)$ представима в блочно-трехдиагональном виде

$$\mathbf{Q} = \text{diag}(\mathbf{Q}_{00}, \mathbf{Q}_{01}, \dots, \mathbf{Q}_{0L}) + \text{diag}^+(\mathbf{Q}_{11}, \mathbf{Q}_{12}, \dots, \mathbf{Q}_{1L}) + \text{diag}^-(\mathbf{Q}_{21}, \mathbf{Q}_{22}, \dots, \mathbf{Q}_{2L}), \quad (3.15)$$

где ненулевые недиагональные элементы блоков определяются по формулам

$$\begin{aligned} \mathbf{Q}_{0l} \left[(k, u_1, u_2, r_1, r_2), (k', u'_1, u'_2, r'_1, r'_2) \right] &= \delta, \\ (k' = k - r_2, u'_1 = u_1, u_2 = N - k, u'_2 = u_2 + r_2, \\ r'_1 = r_1 = 0, r'_2 = 0, u_1 < k, 0 < r_2 \leq k - u_1) \vee \\ \vee (k' = u_1, u'_1 = u_1, u_2 = N - k, u'_2 = u_2 + k - u_1, \\ r'_1 = r_1 = 0, r'_2 = r_2 - k + u_1, u_1 < k, r_2 > k - u_1) \vee \\ \vee (k' = k + r_1, u_1 = k, u'_1 = u_1 + r_1, u'_2 = u_2, \\ r'_1 = 0, r'_2 = r_2 = 0, u_2 < N - k, 0 < r_1 \leq N - k - u_2) \vee \\ \vee (k' = N - u_2, u_1 = k, u'_1 = u_1 + N - k - u_2, u'_2 = u_2, \\ r'_1 = r_1 - N + k + u_2, r'_2 = r_2 = 0, u_2 < N - k, r_1 > N - k - u_2), \end{aligned} \quad (3.16)$$

$$\dim \mathbf{Q}_{0l} = \sum_{k=0}^N \sum_{n=0}^{\min(l, k+R_1)} (l-n+1) \times \sum_{k=0}^N \sum_{n=0}^{\min(l, k+R_1)} (l-n+1), \quad l = 0, \dots, L,$$

$$\begin{aligned} \mathbf{Q}_{1l} \left[(k, u_1, u_2, r_1, r_2), (k', u'_1, u'_2, r'_1, r'_2) \right] &= \\ = \begin{cases} \lambda_1, & (k' = k, u'_1 = u_1 + 1, u'_2 = u_2, r'_1 = r_1 = 0, r'_2 = r_2) \vee \\ & \vee (k' = k, u'_1 = u_1 = k, u'_2 = u_2, r'_1 = r_1 + 1, r'_2 = r_2), \\ \lambda_2, & (k' = k, u'_1 = u_1, u'_2 = u_2 + 1, r'_1 = r_1 = 0, r'_2 = r_2) \vee \\ & \vee (k' = k, u'_1 = u_1, u'_2 = u_2 = N - k, r'_1 = r_1, r'_2 = r_2 + 1), \end{cases} \end{aligned} \quad (3.17)$$

$$\dim \mathbf{Q}_{1l} = \sum_{k=0}^N \sum_{n=0}^{\min(l-1, k+R_1)} (l-n) \times \sum_{k=0}^N \sum_{n=0}^{\min(l, k+R_1)} (l-n+1), \quad l = 1, \dots, L,$$

$$\begin{aligned}
 \mathbf{Q}_{2l} \left[(k, u_1, u_2, r_1, r_2), (k', u'_1, u'_2, r'_1, r'_2) \right] = \\
 = \begin{cases} \frac{k}{N} V \mu_1, & (k' = k, u'_1 = u_1 - 1, u'_2 = u_2, r'_1 = r_1 = 0, r'_2 = r_2) \vee \\ & \vee (k' = k, u'_1 = u_1 = k, u'_2 = u_2, r'_1 = r_1 - 1, r'_2 = r_2), \\ \frac{N-k}{N} V \mu_2, & (k' = k, u'_1 = u_1, u'_2 = u_2 - 1, r'_1 = r_1 = 0, r'_2 = r_2) \vee \\ & \vee (k' = k, u'_1 = u_1, u'_2 = u_2 = N - k, r'_1 = r_1, r'_2 = r_2 - 1), \\ r_1 \varepsilon_1, & k' = k, u'_1 = u_1 = k, u'_2 = u_2, r'_1 = r_1 - 1, r'_2 = r_2, \\ r_2 \varepsilon_2, & k' = k, u'_1 = u_1, u'_2 = u_2 = N - k, r'_1 = r_1, r'_2 = r_2 - 1, \end{cases} \quad (3.18) \\
 \dim \mathbf{Q}_{2l} = \sum_{k=0}^N \sum_{n=0}^{\min(l, k+R_1)} (l-n+1) \times \sum_{k=0}^N \sum_{n=0}^{\min(l-1, k+R_1)} (l-n), \quad l=1, \dots, L,
 \end{aligned}$$

и $L = N + R_1 + R_2$, $l = u_1 + u_2 + r_1 + r_2$ и $j = \min(k + R_1, N - k + R_2)$.

Доказательство.

а) Введение лексикографического порядка

Разобьем пространство $\mathcal{X}$ состояний (3.10) СП $X(t)$ на подмножества по общему числу $u_1 + u_2 + r_1 + r_2$ сессий трафика в системе

$$\begin{aligned}
 \mathcal{X} &= \bigcup_{l=0}^L \mathcal{X}(l), \\
 \mathcal{X}(l) &= \{(k, u_1, u_2, r_1, r_2) \in \mathcal{X} : u_1 + u_2 + r_1 + r_2 = l\}.
 \end{aligned}$$

Введем лексикографический порядок на множестве $\mathcal{X}$ в порядке объединения $\mathcal{X}(l)$, тогда с учетом ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (3.12), матрица $\mathbf{Q}$ имеет блочно-трехдиагональный вид (3.15).

б) Определение размерности блоков

Найдем мощность множества $\mathcal{X}(l)$

$$\begin{aligned}
 |\mathcal{X}(l)| &= \sum_{k=0}^N \sum_{n=0}^{\min(l, k+R_1)} (l-n+1) = \\
 &= \sum_{k=0}^N \left\{ (l+1) [\min(l, k+R_1) + 1] - \frac{\min(l, k+R_1) [\min(l, k+R_1) + 1]}{2} \right\}, \quad l=0, \dots, L.
 \end{aligned}$$

Для краткости, в формулировке леммы оставим вид размерности блоков, записанный через суммы.

■

3.4.2 Матричный алгоритм расчета стационарного распределения

Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с фиксированной политикой перераспределения ресурса между классами трафика, в соответствии с лексикографическим порядком (3.14), представим стационарное распределение СП $X(t)$ в виде $\pi = (\pi_0, \dots, \pi_L)$. Тогда с учетом леммы 3.1, справедлива следующая теорема.

Теорема 3.2. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с фиксированной политикой перераспределения ресурса между классами трафика, стационарное распределение СП $X(t)$ рассчитывается в матричном виде по формуле

$$\pi_l = \pi_L \prod_{i=1}^{L-l} \mathbf{A}_{L-i}, \quad l = 0, \dots, L-1, \quad (3.19)$$

где π_M является единственным решением системы уравнений

$$\begin{aligned} \pi_L (\mathbf{A}_{L-1} \mathbf{Q}_{1L} + \mathbf{Q}_{0L}) &= \mathbf{0}, \\ \pi_L \left(\sum_{l=0}^L \prod_{i=1}^{L-l} \mathbf{A}_{L-i} \mathbf{1}^T \right) &= 1, \end{aligned} \quad (3.20)$$

а матрицы $\mathbf{A}_l$ вычисляются по рекуррентным соотношениям

$$\begin{aligned} \mathbf{A}_0 &= -\mathbf{Q}_{21} \mathbf{Q}_{00}^{-1}, \\ \mathbf{A}_l &= -\mathbf{Q}_{2,l+1} (\mathbf{A}_{l-1} \mathbf{Q}_{1l} + \mathbf{Q}_{0l})^{-1}, \quad l = 1, \dots, L-1, \end{aligned} \quad (3.21)$$

и $L = N + R_1 + R_2$.

Доказательство.

Доказательство можно провести аналогично теореме 2.1, заменив $M \rightarrow L = N + R_1 + R_2$.

■

3.4.3 О влиянии интенсивности потока сигналов на показатели эффективности

Рассмотрим базовую характеристику, связанную с перераспределением ресурса. Под вероятностью перераспределения ресурса по сигналу будем понимать вероятность того, что система находится в состоянии, при котором в случае поступления сигнала произойдет перераспределение ресурса.

Из теоремы 3.2 вытекает следующее следствие.

Следствие 3.3. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с фиксированной политикой перераспределения ресурса между классами трафика, следующие показатели эффективности рассчитываются по формулам:

вероятность перераспределения ресурса по сигналу

$$P_r = \sum_{k=0}^N \sum_{u_1=0}^{k-1} \sum_{r_2=1}^{R_2} \pi(k, u_1, N-k, 0, r_2) + \sum_{k=0}^N \sum_{u_2=0}^{N-k-1} \sum_{r_1=1}^{R_1} \pi(k, k, u_2, r_1, 0); \quad (3.22)$$

вероятности блокировки сессии эластичного трафика

$$P_1^{bl} = \sum_{k=0}^N \sum_{u_2=0}^{N-k} \pi(k, k, u_2, R_1, 0) + \sum_{k=0}^N \sum_{r_2=1}^{R_2} \pi(k, k, N-k, R_1, r_2), \quad (3.23)$$

$$P_2^{bl} = \sum_{k=0}^N \sum_{u_1=0}^k \pi(k, u_1, N-k, 0, R_2) + \sum_{k=0}^N \sum_{r_1=1}^{R_1} \pi(k, k, N-k, r_1, R_2). \quad (3.24)$$

Проиллюстрируем, как вероятность перераспределения ресурса по сигналу и вероятности блокировки влияют на выбор интенсивности поступления сигналов. На рисунке 3.4 изображены области значений характеристик, которые соответствуют или не соответствуют пороговым значениям. По оси абсцисс в логарифмическом масштабе отложены интенсивности δ , а по оси ординат – коэффициент λ_2/λ_1 . В зависимости от целевой функции в задачи оптимизации, выбираем значение δ , как показано на рисунке.

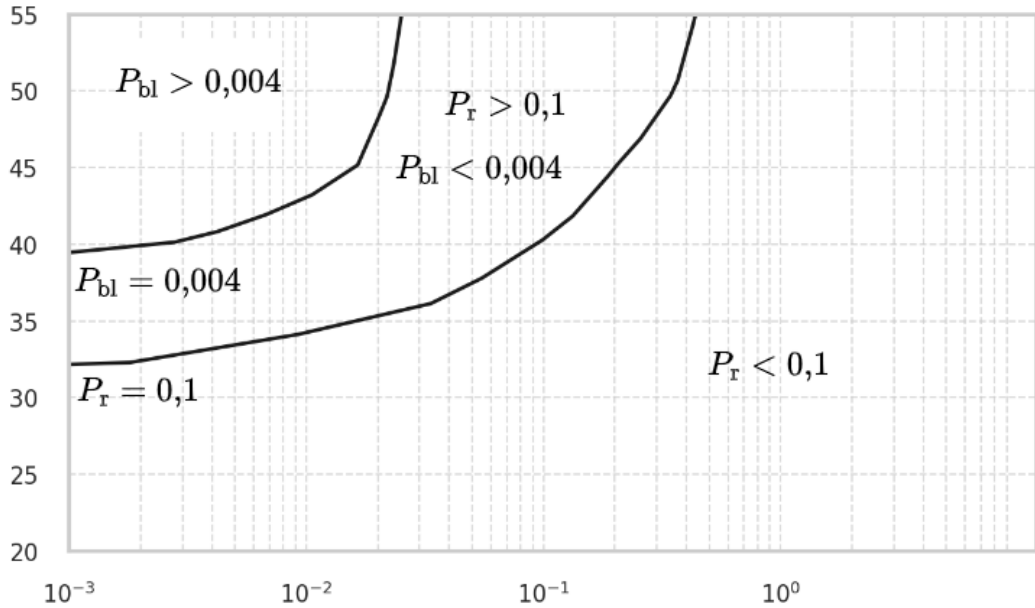


Рисунок 3.4 – Области возможных значений интенсивности поступления сигналов (1/с) в зависимости от коэффициента λ_2/λ_1 при ограничениях на показатели эффективности: вероятность перераспределения ресурса по сигналу и вероятность блокировки сессии

3.5 УПРАВЛЯЕМАЯ СИСТЕМА ПЕРЕРАСПРЕДЕЛЕНИЯ РЕСУРСА

3.5.1 Марковский процесс принятия решения

Функционирование системы опишем при помощи МППР в непрерывном времени

$$(\mathcal{X}, \mathcal{A}(\mathbf{x}), \mathbf{Q}_f, g(\mathbf{x})). \quad (3.25)$$

Первый элемент кортежа, представляющий пространство состояний, имеет вид

$$\mathcal{X} = \{(k, n, m): 0 \leq k \leq N, 0 \leq n \leq k + R_1, 0 \leq m \leq N - k + R_2\}. \quad (3.26)$$

Действие возникает при поступлении сигнала и означает новое разделение ресурса. Поскольку в системе два сегмента, достаточно задать объем ресурса только для одного из них. Выберем первый сегмент. Тогда множество действий в состоянии $\mathbf{x}$ будет выглядеть следующим образом

$$\mathcal{A}(\mathbf{x}) = \begin{cases} \{N-m, \dots, k\}, & n < k, m > N-k, n+m \leq N, \\ \{n, \dots, k\}, & n < k, m > N-k, n+m > N, \\ \{k, \dots, N-m\}, & n > k, m < N-k, n+m \leq N, \\ \{k, \dots, n\}, & n > k, m < N-k, n+m > N, \\ \emptyset, & (n \geq k) \vee (m \leq N-k), (n \leq k) \vee (m \geq N-k). \end{cases} \quad (3.27)$$

Ненулевые элементы матрицы интенсивностей переходов при фиксированной политике f имеют вид

$$\begin{aligned} \mathbf{Q}_f[(k, n, m), (k, n+1, m)] &= \lambda_1, \quad n+1 \leq k + R_1, \\ \mathbf{Q}_f[(k, n, m), (k, n, m+1)] &= \lambda_2, \quad m+1 \leq N-k + R_2, \\ \mathbf{Q}_f[(k, n, m), (k, n-1, m)] &= \frac{k}{N} V \mu_1 \cdot 1\{k > 0, n > 0\} + (n-k) \varepsilon_1 \cdot 1\{n > k\}, \\ \mathbf{Q}_f[(k, n, m), (k, n, m-1)] &= \\ &= \frac{N-k}{N} V \mu_2 \cdot 1\{N-k > 0, m > 0\} + (m-k+N) \varepsilon_2 \cdot 1\{m > N-k\}, \\ \mathbf{Q}_f[(k, n, m), (a, n, m)] &= \delta, \quad (n < k, m > N-k, n+m \leq N) \vee \\ &\vee (n < k, m > N-k, n+m > N) \vee (n > k, m < N-k, n+m \leq N) \vee \\ &\vee (n > k, m < N-k, n+m > N), \quad a \in \mathcal{A}(k, n, m). \end{aligned} \quad (3.28)$$

Последний элемент кортежа – функция вознаграждения. В рассматриваемой системе оно получается не за выполнение действия $a \in \mathcal{A}(\mathbf{x})$, а только за пребывание в состоянии $\mathbf{x}$. Функция вознаграждения определяется следующей леммой.

Лемма 3.2. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с политикой управления перераспределением ресурса между классами трафика, функция вознаграждения в единицу времени за пребывание в состояниях системы имеет вид

$$g(\mathbf{x}) = \omega_1 g_1(\mathbf{x}) + \omega_2 g_2(\mathbf{x}) + \omega_3 g_3(\mathbf{x}), \quad \mathbf{x} \in \mathcal{X}, \quad (3.29)$$

где вознаграждение за соответствие равному разделению ресурса

$$g_1(\mathbf{x}) = -\bar{g}_1(\mathbf{x}),$$

$$\bar{g}_1(\mathbf{x}) = \begin{cases} n - k, & k < N/2, n > k, n \leq N/2, \\ N/2 - k, & k < N/2, n > k, n > N/2, \\ n - (N - k), & k > N/2, m > N - k, m \leq N/2, \\ N/2 - (N - k), & k > N/2, m > N - k, m \leq N/2, \\ 0, & (k < N/2) \vee (n > k), \\ & (k > N/2) \vee (m > N - k), \end{cases} \quad (3.30)$$

вознаграждение за результативность сигналов

$$g_2(\mathbf{x}) = \frac{\delta}{A(\mathbf{x})} \cdot 1\{(n < k, m > N - k) \vee (n > k, m < N - k)\},$$

$$A(\mathbf{x}) = \lambda_1 \cdot 1\{n + 1 \leq k + R_1\} + \lambda_2 \cdot 1\{m + 1 \leq N - k + R_2\} +$$

$$+ \frac{k}{N} V \mu \cdot 1\{n > 0\} + \frac{N - k}{N} V \mu_2 \cdot 1\{m > 0\} +$$

$$+ (n - k) \varepsilon_1 \cdot 1\{n > k\} + (m - k + N) \varepsilon_2 \cdot 1\{m > N - k\} + \delta, \quad (3.31)$$

вознаграждение за использование ресурса

$$g_3(\mathbf{x}) = -\bar{g}_3(\mathbf{x}),$$

$$\bar{g}_3(\mathbf{x}) = \begin{cases} k - N + m, & n < k, m > N - k, n + m \leq N, \\ k - n, & n < k, m > N - k, n + m > N, \\ N - m - k, & n > k, m < N - k, n + m \leq N, \\ n - k, & n > k, m < N - k, n + m > N, \\ 0, & (n \geq k) \vee (m \leq N - k), \\ & (n \leq k) \vee (m \geq N - k). \end{cases} \quad (3.32)$$

Доказательство.

Первое слагаемое $g_1(\mathbf{x})$ в функции вознаграждения – это вознаграждение за соответствие равному разделению ресурса. Определим его через обратную величину стоимости за несоответствие равному разделению ресурса. Это число сессий, ожидающих начала обслуживания, которые могли бы обслуживаться, если бы ресурс был разделен между сегментами поровну.

Второе слагаемое $g_2(\mathbf{x})$ – вознаграждение за результативность сигналов. Это вероятность того, что сигнал поступает, и система находится в состоянии, при котором может произойти перераспределение ресурса. Исключаются состояния, когда множество действий пустое, но при этом из возможных действий может быть выбрано то, при котором разделение ресурса останется неизменным.

И наконец, третье слагаемое $g_3(\mathbf{x})$ – это вознаграждение за использование ресурса. Определим его через обратную величину стоимости за простой ресурса. Это число сессий, ожидающих начала обслуживания, которые могли бы обслуживаться, если бы ресурс был разделен между сегментами так, чтобы ресурс использовался максимально.

■

Таким образом, политика представляет собой функцию, которая ставит однозначное соответствие между состоянием системы и действием, и имеет вид

$$f = \mathbf{a} = [a(\mathbf{x})]_{\mathbf{x} \in \mathcal{X}}, \quad a(\mathbf{x}) \in \mathcal{A}(\mathbf{x}). \quad (3.33)$$

Тогда с учетом леммы 3.2, справедлива следующая теорема.

Теорема 3.3. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с политикой управления перераспределением ресурса между классами трафика, среднее вознаграждение в единицу времени рассчитывается по формуле

$$\bar{g}_f = \sum_{\mathbf{x} \in \mathcal{X}} g(\mathbf{x}) \pi_f(\mathbf{x}), \quad (3.34)$$

где $\pi_f(\mathbf{x})$ стационарное распределение при политике f .

Доказательство.

Среднее вознаграждение в единицу времени будем использовать как целевую функцию в задаче оптимизации для нахождения оптимальной политики.

■

Оптимальной политикой назовем политику, которая максимизирует среднее вознаграждение в единицу времени

$$f^* = \arg \max_f \bar{g}_f. \quad (3.35)$$

Стационарное распределение при оптимальной политике f^* обозначим

$$\pi^*(x) = \pi_{f^*}(x), \quad x \in \mathcal{X}. \quad (3.36)$$

3.5.2 Вероятность перераспределения ресурса по сигналу

Справедливо следующее следствие для вероятности перераспределения ресурса по сигналу, которая упоминалась в разделе 3.4.3.

Следствие 3.4. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с политикой управления перераспределением ресурса между классами трафика, вероятность перераспределения ресурса по сигналу рассчитывается по формуле

$$P_r = \sum_{k=0}^N \sum_{n=0}^{k-1} \sum_{m=N-k+1}^{N-k+R_2} \pi_f(k, n, m) + \sum_{k=0}^N \sum_{n=k+1}^{k+R_1} \sum_{m=0}^{N-k-1} \pi_f(k, n, m). \quad (3.37)$$

Рассмотрим как оптимальную политику f^* , так и случай фиксированной политики f_0 , направленной на максимальное использование ресурса

$$a_0(k, n, m) = \begin{cases} N - m, & (n < k, m > N - k, n + m \leq N) \vee \\ & \vee (n > k, m < N - k, n + m > N), \\ n, & (n < k, m > N - k, n + m > N) \vee \\ & \vee (n > k, m < N - k, n + m \leq N). \end{cases} \quad (3.38)$$

Проиллюстрируем поведение вероятности перераспределения ресурса между сегментами сети. На рисунках 3.5 и 3.6 представлены графики вероятности перераспределения ресурса по сигналу соответственно в зависимости от интенсивности поступления запросов на передачу эластичного трафика для фиксированной и оптимальной политик и в зависимости интенсивности поступления сигналов контроллера для фиксированной политики

Из рисунка 3.5 видно, что при равных весах оптимальная политика обеспечивает более высокое значение вероятности перераспределения ресурса по сравнению с фиксированной политикой, ориентированной только на максимальное

использование ресурса. Это объясняется тем, что оптимальная политика более активно реагирует на дисбалансы, даже если это снижает эффективность в краткосрочной перспективе.

С увеличением интенсивности поступления запросов на передачу эластичного трафика вероятность перераспределения ресурса снижается, поскольку свободные ресурсы становятся менее доступными для перераспределения. Буферы сегментов заполняются быстрее, сокращая количество состояний, где перераспределение возможно. Однако оптимальная политика демонстрирует более плавное снижение вероятности за счет учета нескольких принципов, включая результативность сигналов.

Рисунок 3.6 демонстрирует, что частота поступления сигналов также влияет на вероятность перераспределения ресурса между сегментами. При высокой частоте контроллер чаще проверяет систему, но ограниченность ресурсов приводит к тому, что многие сигналы не приводят к действиям, снижая вероятность перераспределения ресурса.

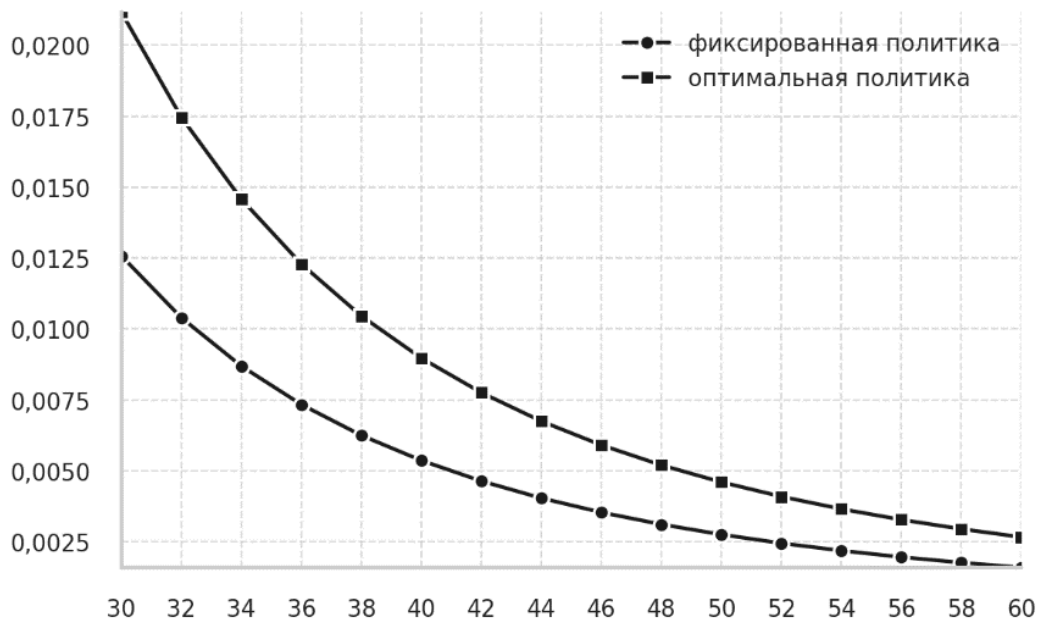


Рисунок 3.5 – Вероятность перераспределения ресурса по сигналу в зависимости от интенсивности поступления запросов на передачу эластичного трафика (1/с) для фиксированной и оптимальной политик

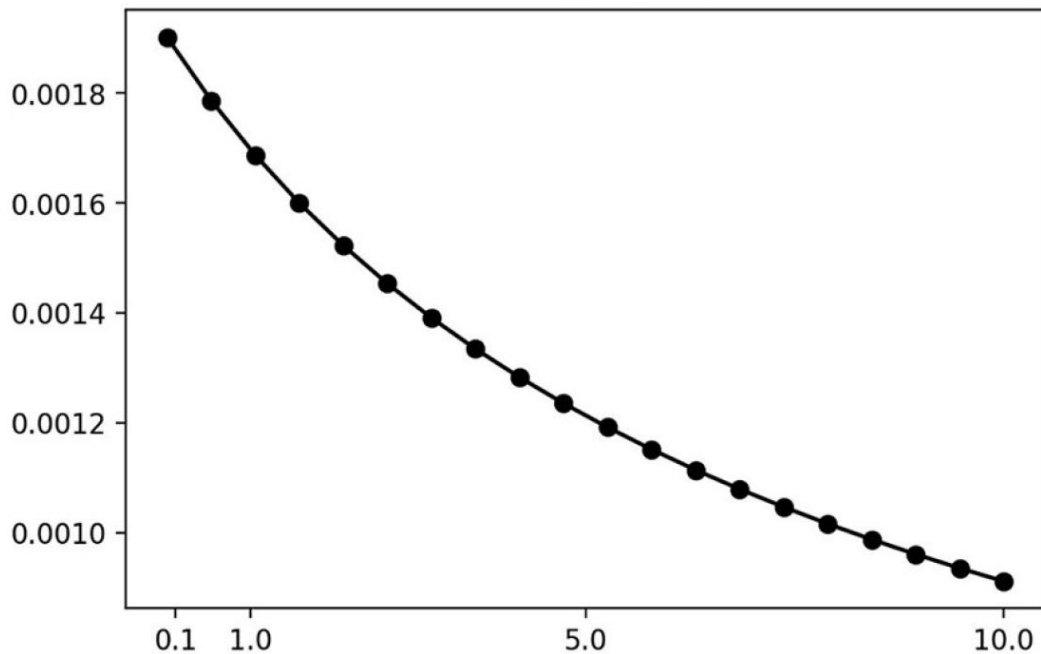


Рисунок 3.6 – Вероятность перераспределения ресурса по сигналу в зависимости от интенсивности поступления сигналов контроллера (1/с) для фиксированной политики

3.5.3 О влиянии весов в функции вознаграждения на политику

Весовые коэффициенты $\omega = (\omega_1, \omega_2, \omega_3)$ в функции вознаграждения определяют относительную значимость каждого из критериев эффективности перераспределения ресурса: ω_1 равное разделение ресурса, ω_2 результативность сигналов и ω_3 эффективное использование ресурса. При равных весах $\omega_1 = \omega_2 = \omega_3$ система стремится к балансу между этими принципами, что приводит к осторожной политике с редкими изменениями.

Операторы могут гибко управлять приоритетами с помощью настройки весов. Увеличение ω_3 смещает фокус на максимальное использование ресурса. Если важно снизить сигнальную нагрузку от контроллера, увеличение ω_2 уменьшает вероятность отправки ненужных сигналов. Наоборот, если приоритет – соблюдение SLA и равное разделение ресурса, доминирование ω_1 сохраняет изначальное разделение, даже если это снижает общую эффективность. Выбор

весов требует компромиссов между целями. Одновременная максимизация всех принципов невозможна.

Проиллюстрируем, как весовые коэффициенты ω_i , $i=1,2,3$ в функции $g(\mathbf{x})$ вознаграждения в единицу времени влияют на оптимальную политику f^* . В таблице 3.3 представлен фрагмент политики f^* для различных наборов весов при $N=8$. По строкам показаны состояния системы (k,n,m) , а по столбцам – веса $\omega=(\omega_1,\omega_2,\omega_3)$. Например, при $\omega=(1,1,1)$ все критерии перераспределения ресурса равнозначны, в то время как при $\omega=(1000,1,1)$ наиболее важным является равное разделение ресурса. Так, если в состоянии $(k,n,m)=(7,8,0)$ поступает сигнал и $\omega=(1,1,1)$, то объем ресурса первого сегмента будет увеличен, а если $\omega=(1000,1,1)$, то останется прежним.

Таблица 3.3 – Фрагмент оптимальной политики перераспределения ресурса

$(\omega_1,\omega_2,\omega_3) \backslash (k,n,m)$	$(5,6,2)$	$(6,4,3)$	$(7,5,3)$	$(7,8,0)$	$(8,5,1)$	$(8,7,1)$
$(1,1,1)$	6	6	6	8	8	8
$(1000,1,1)$	6	6	6	7	7	7
$(1,1000,1)$	6	6	6	8	8	8
$(1,1,1000)$	5	5	5	8	8	8
$(1000,1000,1)$	6	6	6	7	7	7
$(1000,1,1000)$	6	6	6	8	8	8
$(1,1000,1000)$	5	5	5	8	8	8

3.6 НАХОЖДЕНИЕ ПОЛИТИКИ ПЕРЕРАСПРЕДЕЛЕНИЯ РЕСУРСА

3.6.1 Система уравнений относительно вознаграждения

Чтобы решить задачу оптимизации (3.35) в том виде, как она сформулирована, можно использовать метод перебора. Для этого нужно рассмотреть все возможные политики и для каждой из них составить матрицу интенсивностей переходов, используя формулы (3.28). Затем необходимо решить СУР и по полученному стационарному распределению вычислить среднее вознаграждение по формуле (3.34). Однако этот подход отличается высокой вычислительной сложностью.

Более эффективным решением является метод итерации по политикам, который позволяет увеличивать среднее вознаграждение $\bar{g}_f$ в единицу времени на каждом шаге. Этот метод, предложенный Ховардом Р.А., основан на решении системы уравнений непосредственно относительно вознаграждения на каждом шаге, без необходимости решения СУР относительно стационарного распределения.

Обозначим $v(x, t)$ вознаграждение, получаемое системой после выхода из состояния x за время t . Это означает, что в некоторый момент времени t_0 система находится в состоянии x . Далее в течение интервала времени t система изменяет свои состояния, получая вознаграждения, и в итоге накапливает их в объеме $v_f(x, t)$. СП $X_f(t)$ является эргодическим, тогда

$$v_f(x) = \lim_{t \rightarrow \infty} v_f(x, t), \quad x \in \mathcal{X}, \quad (3.39)$$

является предельным вознаграждением, получаемым системой после выхода из состояния x .

Система уравнений относительно вознаграждения описывает баланс средней «скорости получения вознаграждения» $\bar{g}_f$ вне зависимости от состояния системы. Она складывается из «скорости получения вознаграждения» $g(x)$ за пребывание в

состоянии $\mathbf{x}$ и суммы вознаграждений $v_f(\mathbf{x}')$ при переходе в состояние $\mathbf{x}'$ с заданными интенсивностями. Эта система уравнений представлена в следующей лемме.

Лемма 3.3. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с политикой управления перераспределением ресурса между классами трафика, система уравнений относительно вознаграждения имеет вид

$$\bar{g}_f = g(\mathbf{x}) + \sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{0}\}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}'), \quad \mathbf{x} \in \mathcal{X}. \quad (3.40)$$

Доказательство.

а) Получение системы уравнений относительно $v_f(\mathbf{x}, t)$

Динамику приращения вознаграждения за малый интервал времени Δ можно описать через соотношение между вознаграждением за время $t + \Delta$ и вознаграждением за время t . Вознаграждение за время $t + \Delta$ сложим из двух частей: первая – это вознаграждение за время Δ , а вторая – за время t . Вознаграждение за время Δ получим, если система оставалась в состоянии $\mathbf{x}$, в противном случае вознаграждения не будет

$$v_f(\mathbf{x}, t + \Delta) = \Pr(\mathbf{X}_f(t_0 + \Delta) = \mathbf{x} \mid \mathbf{X}_f(t_0) = \mathbf{x}) \cdot [g(\mathbf{x})\Delta + v_f(\mathbf{x}, t)] + \\ + \sum_{\mathbf{x}' \in \mathcal{X}: \mathbf{x}' \neq \mathbf{x}} \Pr(\mathbf{X}_f(t_0 + \Delta) = \mathbf{x}' \mid \mathbf{X}_f(t_0) = \mathbf{x}) \cdot v_f(\mathbf{x}', t), \quad \mathbf{x} \in \mathcal{X}, t > 0,$$

где $\mathbf{X}_f(t)$ СП для фиксированной политики f .

Подставим в систему уравнений выражение переходных вероятностей через интенсивности переходов между состояниями СП

$$\Pr(\mathbf{X}_f(t_0 + \Delta) = \mathbf{x}' \mid \mathbf{X}_f(t_0) = \mathbf{x}) = 1\{\mathbf{x}' = \mathbf{x}\} + \mathbf{Q}_f[\mathbf{x}, \mathbf{x}']\Delta + o(\Delta), \quad \mathbf{x}, \mathbf{x}' \in \mathcal{X},$$

и получим

$$v_f(\mathbf{x}, t + \Delta) = [1 + \mathbf{Q}_f[\mathbf{x}, \mathbf{x}]\Delta + o(\Delta)] \cdot [g(\mathbf{x})\Delta + v_f(\mathbf{x}, t)] + \\ + \sum_{\mathbf{x}' \in \mathcal{X}: \mathbf{x}' \neq \mathbf{x}} [\mathbf{Q}_f[\mathbf{x}, \mathbf{x}']\Delta + o(\Delta)] \cdot v_f(\mathbf{x}', t) = \\ = g(\mathbf{x})\Delta + v_f(\mathbf{x}, t) + \sum_{\mathbf{x}' \in \mathcal{X}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}']\Delta \cdot v_f(\mathbf{x}', t) + o(\Delta), \quad \mathbf{x} \in \mathcal{X}, t > 0.$$

б) Получение системы дифференциальных уравнений относительно $v_f(\mathbf{x}, t)$

Таким образом, из системы уравнений относительно вознаграждения $v_f(\mathbf{x}, t)$, получаемого системой после выхода из состояния $\mathbf{x}$ за время t ,

$$\frac{v_f(\mathbf{x}, t + \Delta) - v_f(\mathbf{x}, t)}{\Delta} = g(\mathbf{x}) + \sum_{\mathbf{x}' \in \mathcal{X}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}', t) + \frac{o(\Delta)}{\Delta}, \quad \mathbf{x} \in \mathcal{X}, \quad t > 0,$$

получим соответствующую систему дифференциальных уравнений

$$v'_f(\mathbf{x}, t) = \lim_{\Delta \rightarrow 0} \frac{v_f(\mathbf{x}, t + \Delta) - v_f(\mathbf{x}, t)}{\Delta} = g(\mathbf{x}) + \sum_{\mathbf{x}' \in \mathcal{X}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}', t), \quad \mathbf{x} \in \mathcal{X}, \quad t > 0.$$

Производная $v'_f(\mathbf{x}, t)$ представляет собой вознаграждение, получаемое системой в единицу времени после выхода из состояния $\mathbf{x}$ за время t .

в) Получение системы уравнений относительно $v_f(\mathbf{x})$

Перейдем к предельному режиму – устремим $t \rightarrow \infty$

$$\lim_{t \rightarrow \infty} v'_f(\mathbf{x}, t) = \lim_{t \rightarrow \infty} \left[g(\mathbf{x}) + \sum_{\mathbf{x}' \in \mathcal{X}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}', t) \right], \quad \mathbf{x} \in \mathcal{X},$$

откуда

$$v'_f(\mathbf{x}) = g(\mathbf{x}) + \sum_{\mathbf{x}' \in \mathcal{X}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}'), \quad \mathbf{x} \in \mathcal{X}.$$

Вознаграждение, получаемое системой в единицу времени, также является стационарным – не зависит от начального состояния $\mathbf{x}$

$$v'_f(\mathbf{x}) = \bar{g}_f, \quad \mathbf{x} \in \mathcal{X},$$

$$\text{поэтому } \bar{g}_f = g(\mathbf{x}) + \sum_{\mathbf{x}' \in \mathcal{X}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}'), \quad \mathbf{x} \in \mathcal{X}.$$

В полученной системе $|\mathcal{X}|$ уравнений и $|\mathcal{X}| + 1$ неизвестных $v_f(\mathbf{x})$, $\mathbf{x} \in \mathcal{X}$ и $\bar{g}_f$. Чтобы уменьшить число неизвестных, приравняем одну из величин $v_f(\mathbf{x})$ к константе, например, $v_f(\mathbf{0}) = 0$. Вероятностный смысл величин $v_f(\mathbf{x})$ поменяется – они станут относительными друг к другу. Однако в рамках данной задачи это несущественно, поскольку вознаграждения $v_f(\mathbf{x})$ являются промежуточными переменными в процессе итерационного поиска оптимальной политики.

■

3.6.2 Итерационный алгоритм Ховарда

Из леммы 3.3 вытекает следующее следствие.

Следствие 3.5. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с политикой управления перераспределением ресурса между классами трафика, систему уравнений относительно вознаграждения для ее решения методом простых итераций можно записать в следующем виде

$$\bar{g}_f = g(\mathbf{0}) + \sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{0}\}} \mathbf{Q}_f[\mathbf{0}, \mathbf{x}'] \cdot v_f(\mathbf{x}'), \quad (3.41)$$

$$v_f(\mathbf{x}) = \left[\sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{x}\}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \right]^{-1} \cdot \left[\sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{x}, \mathbf{0}\}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}') + g(\mathbf{x}) - \bar{g}_f \right], \quad (3.42)$$

$\mathbf{x} \in \mathcal{X} \setminus \{\mathbf{0}\}.$

Доказательство.

В каждом уравнении системы (3.40) выделим из суммы слагаемое при $\mathbf{x}' = \mathbf{x}$

$$\begin{aligned} \bar{g}_f &= g(\mathbf{x}) + \mathbf{Q}_f[\mathbf{x}, \mathbf{x}] \cdot v_f(\mathbf{x}) \cdot 1\{\mathbf{x} \neq \mathbf{0}\} + \sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{x}, \mathbf{0}\}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}') = \\ &= g(\mathbf{x}) - v_f(\mathbf{x}) \cdot 1\{\mathbf{x} \neq \mathbf{0}\} \cdot \sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{x}\}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] + \sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{x}, \mathbf{0}\}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}'), \quad \mathbf{x} \in \mathcal{X}. \end{aligned}$$

Приведем систему к виду, чтобы ее можно было решить методом итераций

$$\begin{aligned} \bar{g}_f &= g(\mathbf{0}) + \sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{0}\}} \mathbf{Q}_f[\mathbf{0}, \mathbf{x}'] \cdot v_f(\mathbf{x}'), \\ \bar{g}_f &= g(\mathbf{x}) - v_f(\mathbf{x}) \cdot \sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{x}\}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] + \sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{x}, \mathbf{0}\}} \mathbf{Q}_f[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}'), \quad \mathbf{x} \in \mathcal{X} \setminus \{\mathbf{0}\}. \end{aligned}$$

В результате получаем систему уравнений (3.41).

■

Таким образом, для фиксированной политики f перераспределения ресурса решение системы уравнений (3.41) позволяет определить среднее вознаграждение $\bar{g}_f$ в единицу времени (3.34) и предельные вознаграждения $v_f(\mathbf{x})$, $\mathbf{x} \in \mathcal{X} \setminus \{\mathbf{0}\}$ (3.39).

Алгоритм Ховарда Р.А. совершает итерации по политикам. Чтобы вычислить новую политику для следующей итерации, максимизируем среднее вознаграждение $\bar{g}_f$ в единицу времени в каждом состоянии системы, выбирая возможные действия из множества (3.27). Из этого вытекает следующее следствие.

Следствие 3.6. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с политикой управления перераспределением ресурса между классами трафика, при решении системы уравнений относительно вознаграждения политика для новой итерации рассчитывается по формуле

$$a_{\text{new}}(k, n, m) = \arg \max_{a \in \mathcal{A}(k, n, m)} v_f(a, n, m), \quad (k, n, m) \in \mathcal{A}(k, n, m) \neq \emptyset. \quad (3.43)$$

Доказательство.

Прежде всего, отметим, что политика определена только для тех состояний $\mathbf{x}$, в которых множество действий не пустое $\mathcal{A}(\mathbf{x}) \neq \emptyset$.

Далее, среднее вознаграждение $\bar{g}_f$ в единицу времени при известных предельных вознаграждениях $v_f(\mathbf{x})$, $\mathbf{x} \in \mathcal{X} \setminus \{\mathbf{0}\}$ вычисляется по формуле (3.40).

Для новой итерации зададим политику $f_{\text{new}} = \mathbf{a}_{\text{new}}$ как аргумент максимизации среднего вознаграждения по возможным действиям из множества (3.27)

$$a_{\text{new}}(\mathbf{x}) = a_{\text{new}}(k, n, m) = \arg \max_{a \in \mathcal{A}(\mathbf{x})} \left(g(\mathbf{x}) + \sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{0}\}} \mathbf{Q}_{f(a)}[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}') \right),$$

$$\mathbf{x} \in \mathcal{X} : \mathcal{A}(\mathbf{x}) \neq \emptyset,$$

где интенсивность $\mathbf{Q}_{f(a)}[\mathbf{x}, \mathbf{x}']$ перехода с индексом $f(a)$ означает, что в состоянии $\mathbf{x}$ происходит действие a и система переходит в состояние $\mathbf{x}' = \mathbf{x}'(a)$.

Вознаграждение $g(\mathbf{x})$ в единицу времени за пребывание в состоянии $\mathbf{x}$ не зависит от политики, поэтому

$$a_{\text{new}}(\mathbf{x}) = \arg \max_{a \in \mathcal{A}(\mathbf{x})} \left(\sum_{\mathbf{x}' \in \mathcal{X} \setminus \{\mathbf{0}\}} \mathbf{Q}_{f(a)}[\mathbf{x}, \mathbf{x}'] \cdot v_f(\mathbf{x}') \right), \quad \mathbf{x} \in \mathcal{X} : \mathcal{A}(\mathbf{x}) \neq \emptyset.$$

В сумме только одно слагаемое зависит от политики $f(a)$ – это слагаемое, в котором действие a влияет на переход из состояния $\mathbf{x}$ в состояние $\mathbf{x}'$. Это происходит, когда поступает сигнал. С учетом вида (3.28) получим

$$a_{\text{new}}(\mathbf{x}) = \arg \max_{a \in \mathcal{A}(\mathbf{x})} \left(\mathbf{Q}_{f(a)}[(k, n, m), (a, n, m)] \cdot v_f(a, n, m) \times \right. \\ \left. \times 1\{(n < k, m > N - k) \vee (n > k, m < N - k)\} \cdot 1\{\mathbf{x} \neq \mathbf{0}\} \right), \quad \mathbf{x} \in \mathcal{X}: \mathcal{A}(\mathbf{x}) \neq \emptyset.$$

Поскольку $\mathbf{x} \in \mathcal{X}: \mathcal{A}(\mathbf{x}) \neq \emptyset$, условия в индикаторных функциях всегда выполняются, поэтому

$$a_{\text{new}}(\mathbf{x}) = \arg \max_{a \in \mathcal{A}(\mathbf{x})} (\delta \cdot v_f(a, n, m)), \quad \mathbf{x} \in \mathcal{X}: \mathcal{A}(\mathbf{x}) \neq \emptyset.$$

Из это следует формула (3.43). ■

Для итерационного алгоритма в качестве начальной f_0 можно использовать любую политику. Однако чтобы ускорить процесс поиска оптимальной политики, выберем такую f_0 , которая в какой-то степени соответствует целевой функции $\bar{g}_f$ (см. формулы (3.34), (3.29)–(3.32)). В качестве такой политики возьмем направленную на максимальное использование ресурса. Эта фиксированная политика уже использовалась в разделе 5.2.

Таким образом, алгоритм расчета оптимальной политики можно представить следующим образом.

Алгоритм 3.1. Для модели с эластичным трафиком и перераспределением ресурса постоянного объема, схемой с политикой управления перераспределением ресурса между классами трафика, расчет оптимальной политики по итерационному алгоритму Ховарда Р.А. записывается следующим образом:

- 1: $i \leftarrow 0$, задать начальную политику $f_0 = \mathbf{a}_0$ по формулам (3.38)
- 2: рассчитать матрицу $\mathbf{Q}_{f_i}$ интенсивностей переходов по формулам (3.28)
- 3: решить систему уравнений относительно $\mathbf{v}_{f_i}$ по формулам (3.41)
- 4: найти новую политику $f_{\text{new}} = \mathbf{a}_{\text{new}}$ по формулам (3.43)
- 5: если $\mathbf{a}_{\text{new}} = \mathbf{a}_i$
- 6: то вернуть оптимальную политику $f^* \leftarrow \mathbf{a}_{\text{new}}$
- 7: иначе $\mathbf{a}_i = \mathbf{a}_{\text{new}}$, $i \leftarrow i + 1$ и вернуться к шагу 2

3.6.3 Применение искусственной нейронной сети

В предыдущем разделе описан итерационный алгоритм Ховарда Р.А., который позволяет рассчитать оптимальную политику f^* как решение задачи оптимизации (3.35). Однако существует и другой подход к решению этой задачи – использование методов машинного обучения. Определение политики f^* можно представить как задачу классификации, где пространство состояний $\mathcal{X}$ (3.26) необходимо разделить на классы, соответствующие новому разделению ресурса при поступлении сигнала. Множество классов – это множество действий $\mathcal{A}(x)$ (3.27).

Проиллюстрируем решение задачи классификации на примере однослойного персептрона – простейшей искусственной нейронной сети. На рисунке 3.7 показана ее архитектура. Во входном слое находится три нейрона, каждый из которых соответствует компоненте вектора $x = (k, n, m)$ состояния системы. В скрытом слое число L нейронов не фиксировано и настраивается. В выходном слое число нейронов $N + 1$ и соответствует числу возможных значений нового максимального числа обслуживаемых сессий трафика первого класса.

$$\begin{aligned} x &= (x_1, x_2, x_3) = (k, n, m), \\ y &= (y_1, y_2, \dots, y_L), \\ z &= (z_0, z_1, \dots, z_N). \end{aligned} \tag{3.44}$$

Система уравнений, описывающая взаимосвязь между входным и скрытым слоями, выглядит следующим образом

$$y^T = Vx^T + \alpha^T, \quad V = \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ \vdots & \vdots & \vdots \\ v_{L1} & v_{L2} & v_{L3} \end{bmatrix}, \quad \alpha^T = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_L \end{bmatrix}. \tag{3.45}$$

Для активации скрытого слоя скрытого слоя используем функцию положительной срезки ReLU (rectified linear unit)

$$\tilde{y}_l = \phi(y_l) = \text{ReLU}(y_l) = \max(0, y_l), \quad l = 1, \dots, L. \quad (3.46)$$

Система уравнений, характеризующая взаимосвязь между скрытым и выходным слоями, имеет вид

$$\mathbf{z}^T = \mathbf{W}\tilde{\mathbf{y}}^T + \boldsymbol{\beta}^T, \quad \mathbf{W} = \begin{bmatrix} w_{01} & w_{02} & \cdots & w_{0L} \\ w_{11} & w_{12} & \cdots & w_{1L} \\ \vdots & \vdots & & \vdots \\ w_{N1} & w_{N2} & \cdots & w_{NL} \end{bmatrix}, \quad \boldsymbol{\beta}^T = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_N \end{bmatrix}. \quad (3.47)$$

В качестве функции активации выходного слоя используем многопеременную логистическую функцию softmax (normalized exponential function)

$$\tilde{z}_k = \phi(z_k) = \text{softmax}(\mathbf{z})_k = \frac{\exp(z_k)}{\sum_{i=0}^N \exp(z_i)}, \quad k = 0, \dots, N. \quad (3.48)$$

Для решения применим два итерационных алгоритма: метод адаптивной оценки моментов Adam (adaptive moment estimation) и алгоритм Бройдена – Флетчера – Гольдфарба – Шанно с ограниченным использованием памяти L-BFGS (Limited-memory Broyden – Fletcher – Goldfarb – Shanno algorithm). В качестве функции потерь используем

$$L(\tilde{z}_k) = -\log(\tilde{z}_k), \quad k = 0, \dots, N. \quad (3.49)$$

Наконец, действие определим по формуле

$$a = \arg \max_{k=0, \dots, N} (z_k). \quad (3.50)$$

На рисунке 3.8 проиллюстрирована точность классификации, вычисляемая по формуле

$$\text{Accuracy}(a, \hat{a}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} 1\{\hat{a}_i = a_i\}. \quad (3.51)$$

По оси абсцисс отложено число нейронов в скрытом слое. Графики отображают результаты работы различных алгоритмов как для обучающей, так и для тестовой выборок.

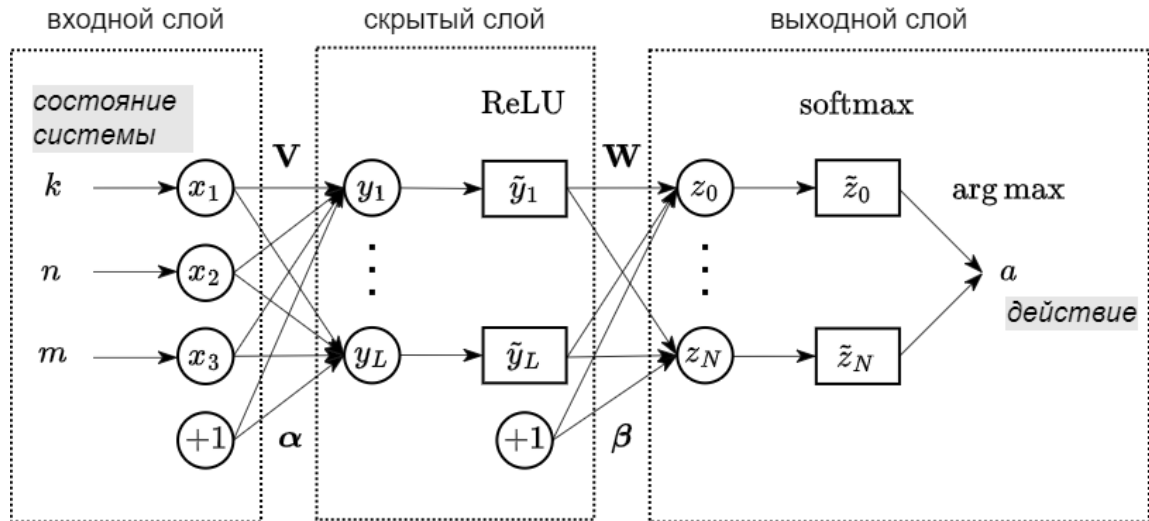


Рисунок 3.7 – Однослойный перцептрон для вычисления оптимальной политики перераспределения ресурса

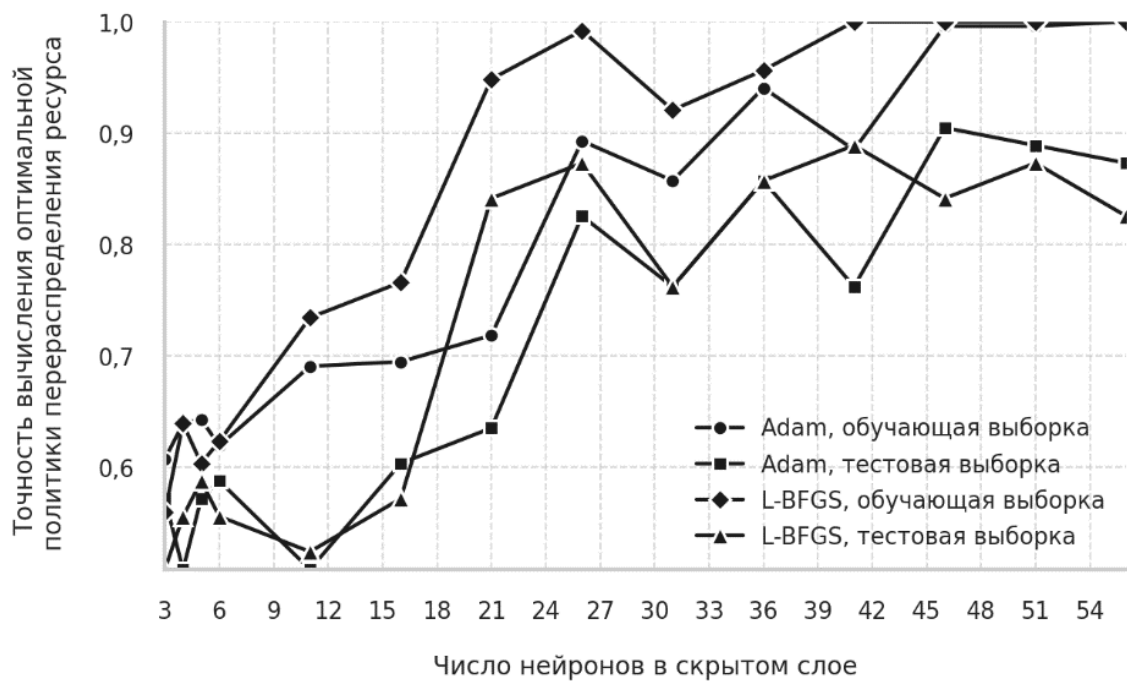


Рисунок 3.8 – Точность вычисления оптимальной политики перераспределения ресурса

Глава 4 МОДЕЛИ С ЭЛАСТИЧНЫМ ТРАФИКОМ И ИЗМЕНЯЕМЫМ ОБЪЕМОМ РЕСУРСА

4.1 МОДЕЛЬ ПЕРЕДАЧИ ДАННЫХ В МИЛЛИМЕТРОВОМ ДИАПАЗОНЕ

4.1.1 Системная модель

В главе 4 исследованы модели приоритетного доступа эластичного трафика в условиях блокировок пути прямой видимости в сети миллиметрового диапазона. Приоритетными являются блокировки пути прямой видимости. Диапазон находится в двух состояниях: во время блокировки – от момента ее возникновения до исчезновения – и между блокировками. В каждом таком состоянии объем ресурса фиксирован. Когда происходит блокировка, объем ресурса изменяется – может полностью отказать и стать недоступным, или частично отказать – объем уменьшится. Потом объем ресурса восстанавливается.

В качестве услуг рассмотрим загрузку веб-страниц, передачу критически важных данных. Поскольку трафик является эластичным, любое событие, приводящее к изменению числа обслуживаемых сессий, также становится приоритетным. Если число сессий возрастает, скорость всех сессий снижается, а если уменьшается – повышается.

Рассмотрим два сценария влияния блокировки на доступность ресурса. В первом случае ресурс отказывает полностью, во втором – частично. Передача эластичного трафика может допускать задержки или их не происходит. При этом запросы на передачу трафика могут поступать независимо друг от друга в виде пуассоновского потока, либо поток коррелированный. Для обеспечения приоритета возникновения блокировок, применяются два механизма доступа эластичного трафика: приостановка обслуживания и снижение скорости передачи.

В таблице 4.1 приведены основные обозначения, параметры и компоненты состояния системы. Для моделирования полного отказа ресурса используется СМО

с ненадежными приборами, а для частичного – СМО с двумя группами приборов – надежными и ненадежными. Для моделирования механизма приостановки обслуживания используется очередь, а для коррелированного характера потока запросов на передачу трафика – МАР поток.

Таблица 4.1 – Обозначения, используемые для моделей с эластичным трафиком и изменяемым объемом ресурса (глава 4)

Обозначение	Параметр
λ	Интенсивность пуассоновского потока запросов на передачу трафика
$\mathbf{D} = \mathbf{D}_0 + \mathbf{D}_1$	Матрица интенсивностей переходов управляющей цепи Маркова
θ	Стационарное распределение управляющей цепи Маркова
$\lambda = \theta \mathbf{D}_1 \mathbf{1}^T$	Интенсивность марковского потока запросов на передачу трафика
b	Требование к минимальной скорости передачи трафика
$M = \lfloor C/b \rfloor$	Максимальное число обслуживаемых сессий трафика
$1/\mu$	Средний экспоненциальный объем трафика
R	Число мест в очереди для приостановленных и ожидающих начала обслуживания сессий трафика
m	Число обслуживаемых сессий трафика
q	Число приостановленных сессий трафика
r	Число ожидающих начала обслуживания сессий трафика
$k = q + r$	Число сессий трафика в очереди – приостановленных и ожидающих начала обслуживания
$n = m + k$	Число сессий трафика в системе – обслуживаемых и в очереди
u	Состояние ненадежного ресурса
w	Состояние управляющей цепи Маркова
ω	Поглощающее состояние

4.1.2 Схемы приоритетного доступа

Рассмотрим следующие схемы приоритетного доступа для моделей с эластичным трафиком и полным отказом изменяемого объема ресурса:

- (а) приостановка обслуживания трафика и пуассоновский входящий поток;
- (б) приостановка обслуживания трафика и марковский входящий поток;

и следующую схему приоритетного доступа для моделей с эластичным трафиком и частичным отказом изменяемого объема ресурса:

- (в) снижение скорости передачи трафика и марковский входящий поток.

В этих схемах используются механизмы $p \in \{2, 3, 4, 9, 10, 11, 12, 13\} \subset \mathcal{P}$ реализации приоритета и принципы $w=2$ выбора сессии для применения механизма из параграфа 1.6.3.

Опишем особенности исследуемых моделей с перечисленными схемами приоритетного доступа через матрицы приоритетов.

Схемы (а) и (б) – системы с полным отказом ресурса и приостановкой обслуживания эластичного трафика. В случае отказа ресурса вся пропускная способность становится недоступной, и обслуживание трафика приостанавливается. Сессии отправляются в очередь для ожидания возобновления обслуживания. Ни одна сессия не будет прервана, так как в очереди зарезервировано S мест для приостановленных сессий. После восстановления ресурса все приостановленные сессии возобновляют свое обслуживание.

Таким образом, трафик может находиться в трех состояниях: обслуживания, ожидания возобновления обслуживания и ожидания начала обслуживания. При этом попытки возобновить обслуживание имеют относительный приоритет перед попытками начать обслуживание. Рассмотрим два случая – схему (а) с пуассоновским входящим потоком и схему (б) с марковским входящим потоком MAP.

Для схемы (а) – это СМО $M | M | C | R | F_{12}$ с матрицей приоритетов

$F_{12} = F_{13}$			ресурс		эластичный трафик			
					обслуживание			ожидание возобновления обслуживания
$\begin{matrix} p \\ e \end{matrix}$			12	13	2	3	4	9
ресурс	отказ	1	1	0	(1,2)	0	0	0
	восстановление	2	0	1	0	0	0	(1,2)
эластичный трафик	поступление	3	0	0	0	(1,2)	0	0
	завершение обслуживания	4	0	0	0	0	(1,2)	0

для схемы (б) – это СМО $MAP | M | C | R | F_{13}$ с матрицей приоритетов, аналогичной схеме (а) $F_{13} = F_{12}$.

Схема (в) – система с частичным отказом ресурса, снижением скорости передачи трафика и потоком MAP. Отличие от предыдущих моделей в том, что ресурс отказывает не полностью – пропускная способность V доступна всегда. Когда ненадежный ресурс отказывает, сессии эластичного трафика «стягиваются», т.е. перемещаются, на надежный ресурс, и скорость передачи снижается. Это СМО $MAP | M | \langle V, C \rangle | 0 | F_{14}$ с матрицей приоритетов

F_{14}		$\begin{matrix} p \\ e \end{matrix}$	ресурс		эластичный трафик	
			10	11	обслуживается	
					3	4
ресурс	отказ	1	1	0	(1,2)	0
	восстановление	2	0	1	0	(1,2)
эластичный трафик	поступление	3	0	0	(1,2)	0
	завершение обслуживания	4	0	0	0	(1,2)

4.1.3 Модели математической теории телетрафика

Функционирование систем опишем при помощи СП $X(t)$ над пространствами состояний

модель (а)

$$\mathcal{X} = \{(n, u): 0 \leq n \leq R, u = 0, 1\}; \quad (4.1)$$

модель (б)

$$\begin{aligned} \mathcal{X} = & \{(n, w): 0 \leq n \leq R, 1 \leq w \leq W\} \cup \\ & \cup \{(n, q, w): 0 \leq n \leq R, n \leq q \leq M, 1 \leq w \leq W\} = \\ = & \{(n, w): 0 \leq n \leq R, 1 \leq w \leq W\} \cup \\ & \cup \{(n, q, w): q \leq n \leq R, 0 \leq q \leq M, 1 \leq w \leq W\}; \end{aligned} \quad (4.2)$$

модель (в)

$$\mathcal{X} = \{(n, u, w): 0 \leq n \leq R, u = 0, 1, 1 \leq w \leq W\}. \quad (4.3)$$

Обозначим $x(n) = ((n, 1), (n, 2), \dots, (n, w), \dots, (n, W))$ и

$x(n, q) = ((n, q, 1), (n, q, 2), \dots, (n, q, w), \dots, (n, q, W))$, тогда ненулевые элементы

матриц интенсивностей переходов имеют вид

модель (а)

$$\begin{aligned} Q[(n, u), (n+1, u)] &= \lambda, \quad n+1 \leq R, \\ Q[(n, 1), (n-1, 1)] &= C\mu, \quad n > 0, \\ Q[(n, 1), (n, 0)] &= \alpha, \\ Q[(n, 0), (n, 1)] &= \beta; \end{aligned} \quad (4.4)$$

модель (б)

$$\begin{aligned} Q[x(n), x(n+1)] &= \mathbf{D}_1, \quad n+1 \leq R, \\ Q[x(n, q), x(n+1, q)] &= \mathbf{D}_1, \quad n+1 \leq R, \\ Q[x(n), x(n-1)] &= C\mu\mathbf{I}, \quad n > 0, \\ Q[x(n), x(n, q)] &= \alpha\mathbf{I}, \\ Q[x(n, q), x(n)] &= \beta\mathbf{I}; \end{aligned} \quad (4.5)$$

модель (в)

$$\begin{aligned}
 Q[x(n, u), x(n+1, u)] &= D_1, \quad n+1 \leq R, \\
 Q[x(n, 0), x(n-1, 0)] &= C\mu I, \quad n > 0, \\
 Q[x(n, 1), x(n-1, 1)] &= (V+C)\mu I, \quad n > 0, \\
 Q[x(n, 1), x(n, 0)] &= \alpha I, \\
 Q[x(n, 0), x(n, 1)] &= \beta I.
 \end{aligned} \tag{4.6}$$

В следующих разделах 4.2–4.6 описаны результаты исследования моделей схем приоритетного доступа с эластичным трафиком, полным и частичным отказом изменяемого объема ресурса. Матрицы интенсивностей переходов представлены в блочно-трехдиагональном виде. Для расчета стационарного распределения предложены алгоритм в скалярном виде для модели (а) (раздел 4.2), и матричный алгоритм для модели (б) – (раздел 4.3). Проиллюстрирован вид показателя приоритетного доступа – вероятности блокировки запроса на передачу эластичного трафика.

Для анализа другого показателя приоритетного доступа – среднего времени передачи эластичного трафика, необходимо учесть тот факт, что не всегда передача происходит непрерывно. Она может приостанавливаться, если ресурс полностью отказывает, а затем возобновляться при его восстановлении (раздел 4.4). Или может замедляться, если ресурс частично отказывает, а затем ускоряться при его восстановлении (раздел 4.6). Поэтому для нахождения среднего времени передачи трафика вводится поглощающий СП, вложенный по моментам времени поступления запроса на передачу и до момента завершения его передачи.

Матрицы интенсивностей переходов для поглощающих СП представлены в блочно-трехдиагональном виде. Записан вид начального распределения вероятностей состояний поглощающих СП. Для расчета не только математического ожидания, но и других моментов времени передачи эластичного трафика записано соответствующее ПЛС. Получен матричный алгоритм расчета ожидаемого времени пребывания в состояниях СП до поглощения и моментов

времени передачи трафика. Проиллюстрирован вид математического ожидания и дисперсии времени передачи эластичного трафика.

4.2 СИСТЕМА С ПОЛНЫМ ОТКАЗОМ РЕСУРСА

4.2.1 Матрица интенсивностей переходов в блочном виде

Лемма 4.1. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, если на пространстве $\mathcal{X}$ состояний СП $X(t)$ введен лексикографический порядок

$$\begin{aligned} \mathbf{x} = (n, u) < (n', u') = \mathbf{x}' \Leftrightarrow \\ (u > u') \vee \\ \vee (u = u', n < n'), \quad \mathbf{x}, \mathbf{x}' \in \mathcal{X}, \end{aligned} \quad (4.7)$$

тогда матрица интенсивностей переходов СП $X(t)$ представима в блочном виде из четырех блоков

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{10} \\ \mathbf{Q}_{01} & \mathbf{Q}_{00} \end{bmatrix}, \quad (4.8)$$

где блоки определяются по формулам

$$\begin{aligned} \mathbf{Q}_{11} = \text{diag}(-(\lambda + \alpha), -(\lambda + C\mu + \alpha), \dots, -(\lambda + C\mu + \alpha), -(C\mu + \alpha)) + \\ + \text{diag}^+(\lambda, \dots, \lambda) + \text{diag}^-(C\mu, \dots, C\mu), \end{aligned} \quad (4.9)$$

$$\mathbf{Q}_{00} = \text{diag}(-(\lambda + \beta), \dots, -(\lambda + \beta), -\beta) + \text{diag}^+(\lambda, \dots, \lambda), \quad (4.10)$$

$$\mathbf{Q}_{10} = \alpha \mathbf{I}, \quad (4.11)$$

$$\begin{aligned} \mathbf{Q}_{01} = \beta \mathbf{I}, \\ \dim \mathbf{Q}_{11} = \dim \mathbf{Q}_{00} = \dim \mathbf{Q}_{10} = \dim \mathbf{Q}_{01} = (R+1) \times (R+1). \end{aligned} \quad (4.12)$$

Доказательство.

а) Введение лексикографического порядка

Разобьем пространство $\mathcal{X}$ состояний (4.1) СП $X(t)$ на подмножества по состоянию u ненадежного ресурса

$$\mathcal{X} = \mathcal{X}(1) \cup \mathcal{X}(0),$$

$$\mathcal{X}(u) = \{(n, i) \in \mathcal{X} : i = u\} = \{(n, u) \in \mathcal{X}\} = \{(0, u), (1, u), \dots, (n, u), \dots, (R, u)\}.$$

Введем лексикографический порядок на множестве $\mathcal{X}$ в порядке объединения $\mathcal{X}(u)$

$$\mathcal{X} = \{(0, 1), (1, 1), \dots, (n, 1), \dots, (R, 1), \\ (0, 0), (1, 0), \dots, (n, 0), \dots, (R, 0)\}.$$

б) Описание блоков матрицы

С учетом введенного лексикографического порядка и ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (4.4), матрица $\mathbf{Q}$ имеет блочный вид из четырех блоков

$$\mathbf{Q} = \begin{bmatrix} \begin{array}{cccc|cccc} -(\lambda + \alpha) & \lambda & & & & & & \\ C\mu & \ddots & \ddots & & & & & \\ & \ddots & \ddots & \ddots & & & & \\ & & C\mu & \ddots & \lambda & & & \\ & & C\mu & - (C\mu + \alpha) & & & & \alpha \end{array} & \begin{array}{cccc} \alpha & & & \\ & \alpha & & \\ & & \ddots & \\ & & & \alpha \end{array} \\ \hline \begin{array}{cccc|cccc} \beta & & & & -(\lambda + \beta) & \lambda & & \\ & \beta & & & & \ddots & \ddots & \\ & & \ddots & & & \ddots & \ddots & \\ & & & \beta & & & -(\lambda + \beta) & \lambda \\ & & & & \beta & & & -\beta \end{array} \end{bmatrix}$$

■

4.2.2 Рекуррентный алгоритм расчета стационарного распределения

Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, в соответствии с лексикографическим порядком (4.7), представим стационарное распределение СП $X(t)$ в виде

$$\begin{aligned}\pi &= (\pi_1, \pi_0), \\ \pi_u &= (\pi(0, u), \pi(1, u), \dots, \pi(n, u), \dots, \pi(R, u)).\end{aligned}\tag{4.13}$$

Тогда с учетом леммы 4.1, справедлива следующая теорема.

Теорема 4.1. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, стационарное распределение СП $X(t)$ рассчитывается по рекуррентным формулам

$$\begin{aligned}\pi(1, 1) &= \frac{1}{C\mu} \left[(\lambda + \alpha) \pi(0, 1) - \frac{\alpha\beta}{\lambda + \beta} \pi(0, 1) \right], \\ \pi(n, 1) &= \\ &= \frac{1}{C\mu} \left[(\lambda + C\mu + \alpha) \pi(n-1, 1) - \lambda \pi(n-2, 1) - \frac{\alpha\beta}{\lambda + \beta} \sum_{i=0}^{n-1} \left(\frac{\lambda}{\lambda + \beta} \right)^{n-1-i} \pi(i, 1) \right], \\ n &= 2, \dots, R,\end{aligned}\tag{4.14}$$

$$\pi(n, 0) = \frac{\alpha}{\lambda + \beta} \sum_{i=0}^n \left(\frac{\lambda}{\lambda + \beta} \right)^{n-i} \pi(i, 1), \quad n = 0, \dots, R-1,$$

$$\pi(R, 0) = \frac{\alpha}{\beta} \sum_{i=0}^R \left(\frac{\lambda}{\lambda + \beta} \right)^{R-i} \pi(i, 1).$$

Доказательство.

а) Запись СУР

Запишем СУР $\pi Q = 0$ с учетом ненулевых элементов матрицы Q интенсивностей переходов (4.4)

$$\begin{aligned} -(\lambda + \alpha)\pi(0,1) + C\mu\pi(1,1) + \beta\pi(0,0) &= 0, \\ \lambda\pi(n-1,1) - (\lambda + C\mu + \alpha)\pi(n,1) + C\mu\pi(n+1,1) + \beta\pi(n,0) &= 0, \quad n=1, \dots, R-1, \\ \lambda\pi(R-1,1) - (C\mu + \alpha)\pi(R,1) + \beta\pi(R,0) &= 0, \\ \alpha\pi(0,1) - (\lambda + \beta)\pi(0,0) &= 0, \\ \alpha\pi(n,1) + \lambda\pi(n-1,0) - (\lambda + \beta)\pi(n,0) &= 0, \quad n=1, \dots, R-1, \\ \alpha\pi(R,1) + \lambda\pi(R-1,0) - \beta\pi(R,0) &= 0. \end{aligned}$$

б) Выражение $\pi(0,0)$ от $\pi(0,1)$

Выразим $\pi(0,0)$ через $\pi(0,1)$ из уравнения СУР при $(0,0)$

$$\pi(0,0) = \frac{\alpha}{\lambda + \beta} \pi(0,1).$$

Подставим полученное соотношение в уравнение СУР при $(0,1)$

$$-(\lambda + \alpha)\pi(0,1) + C\mu\pi(1,1) + \beta \frac{\alpha}{\lambda + \beta} \pi(0,1) = 0$$

и выразим $\pi(1,1)$ через $\pi(0,1)$

$$\pi(1,1) = \frac{1}{C\mu} \left[(\lambda + \alpha)\pi(0,1) - \frac{\alpha\beta}{\lambda + \beta} \pi(0,1) \right].$$

в) Получение зависимости $\pi(n,0)$ от $\pi(i,1)$

Получим зависимость $\pi(n,0)$ от $\pi(i,1)$. Уравнение СУР при $(1,0)$ имеет вид

$$\alpha\pi(1,1) + \lambda\pi(0,0) - (\lambda + \beta)\pi(1,0) = 0.$$

Подставим выражение $\pi(0,0)$ через $\pi(0,1)$ и выразим $\pi(1,0)$

$$\pi(1,0) = \frac{\alpha}{\lambda + \beta} \left(\pi(1,1) + \frac{\lambda}{\lambda + \beta} \pi(0,1) \right) = \frac{\alpha}{\lambda + \beta} \sum_{i=0}^1 \left(\frac{\lambda}{\lambda + \beta} \right)^{1-i} \pi(i,1).$$

Предположим, что соотношение верно для $n-1$

$$\pi(n-1,0) = \frac{\alpha}{\lambda + \beta} \sum_{i=0}^{n-1} \left(\frac{\lambda}{\lambda + \beta} \right)^{n-1-i} \pi(i,1).$$

Подставим его в уравнение СУР при $(n,0)$

$$\begin{aligned} \pi(n,0) &= \frac{1}{\lambda + \beta} \left[\alpha \pi(n,1) + \lambda \frac{\alpha}{\lambda + \beta} \sum_{i=0}^{n-1} \left(\frac{\lambda}{\lambda + \beta} \right)^{n-1-i} \pi(i,1) \right] = \\ &= \frac{\alpha}{\lambda + \beta} \sum_{i=0}^n \left(\frac{\lambda}{\lambda + \beta} \right)^{n-i} \pi(i,1), \quad n=1, \dots, R-1, \end{aligned}$$

и объединим с соотношением для $\pi(0,0)$. Следовательно, сделанное предположение о зависимости $\pi(n,0)$ от $\pi(i,1)$ верно, таким образом

$$\pi(n,0) = \frac{\alpha}{\lambda + \beta} \sum_{i=0}^n \left(\frac{\lambda}{\lambda + \beta} \right)^{n-i} \pi(i,1), \quad n=0, \dots, R-1.$$

г) Выражение $\pi(R,0)$ через $\pi(i,1)$

Подставим выражение $\pi(R-1,0)$ в уравнение СУР при $(R,0)$ и выразим $\pi(R,0)$

$$\pi(R,0) = \frac{\alpha}{\beta} \left[\pi(R,1) + \frac{\lambda}{\lambda + \beta} \sum_{i=0}^{R-1} \left(\frac{\lambda}{\lambda + \beta} \right)^{R-1-i} \pi(i,1) \right] = \frac{\alpha}{\beta} \sum_{i=0}^R \left(\frac{\lambda}{\lambda + \beta} \right)^{R-i} \pi(i,1).$$

д) Выражение $\pi(n,1)$ через $\pi(i,1)$

Заменим $n+1$ на n в уравнениях СУР при $(n,1)$

$$\lambda \pi(n-2,1) - (\lambda + C\mu + \alpha) \pi(n-1,1) + C\mu \pi(n,1) + \beta \pi(n-1,0) = 0, \quad n=2, \dots, R,$$

выразим $\pi(n,1)$

$$\pi(n,1) = \frac{1}{C\mu} [(\lambda + C\mu + \alpha) \pi(n-1,1) - \lambda \pi(n-2,1) - \beta \pi(n-1,0)], \quad n=2, \dots, R,$$

и подставим выражение $\pi(n-1,0)$ через $\pi(i,1)$

$$\pi(n,1) = \frac{1}{C\mu} \left[(\lambda + C\mu + \alpha)\pi(n-1,1) - \lambda\pi(n-2,1) - \frac{\alpha\beta}{\lambda + \beta} \sum_{i=0}^{n-1} \left(\frac{\lambda}{\lambda + \beta} \right)^{n-1-i} \pi(i,1) \right],$$

$$n = 2, \dots, R.$$

■

Замечание. Рекуррентные соотношения в теореме 4.1 выражают $\pi(n,u)$, $(n,u) \in \mathcal{X} \setminus \{(0,1)\}$ через $\pi(0,1)$. Вероятность $\pi(0,1)$ находится из условия нормировки $\pi \mathbf{1}^T = 1$.

4.2.3 Анализ вероятности блокировки

Из теоремы 4.1 вытекает следующее следствие.

Следствие 4.1. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, вероятность блокировки запроса на передачу эластичного трафика рассчитывается по формуле

$$P_{bl} = \pi(R,0) + \pi(R,1) \quad (4.15)$$

Проиллюстрируем поведение вероятности блокировки запроса на передачу эластичного трафика на примере просмотра веб-страниц. Рассмотрим сценарий веб-серфинга – основы цифровой жизни. Он включает в себя поиск информации, покупки, общение в соцсетях и потребление мультимедийного контента. От скорости и стабильности загрузки страниц зависит, насколько комфортно пользователю. Даже небольшие задержки могут вызвать разочарование и подорвать доверие к ресурсу. В мобильных сетях с ограниченной пропускной способностью и сбоями связи оптимизация загрузки страниц становится ключевой. Современные технологии, такие как кэширование, адаптивный дизайн и сжатие данных, помогают ускорить загрузку и снизить зависимость от качества соединения.

На рисунках 4.1 и 4.2 представлены графики вероятности блокировки запроса на передачу эластичного трафика в зависимости от числа пользователей для разных схем ММО и ширины радиочастот, а также объема трафика соответственно. Чем выше пропускная способность, тем меньше вероятность блокировки. Это связано с тем, что система может обслуживать больше пользователей одновременно, сокращая очередь. Чем больше объем контента, тем выше вероятность блокировки запросов. Обработка крупных файлов занимает больше времени, что заполняет очередь и ограничивает прием новых пользователей. Однако кэширование помогает решить эту проблему. Оно сокращает объем передаваемой информации за счет использования ранее загруженных элементов, снижая нагрузку на канал и освобождая ресурсы для новых запросов. Это особенно важно для контента с мультимедийными элементами.

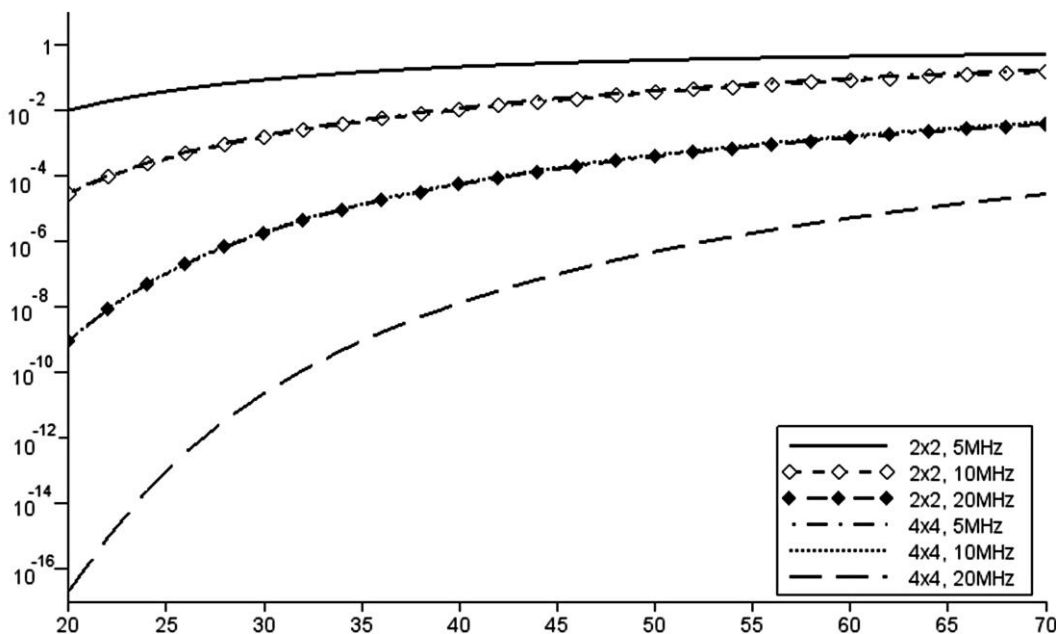


Рисунок 4.1 – Вероятность блокировки запроса на передачу эластичного трафика в зависимости от числа пользователей для разных схем ММО и ширины радиочастот (МГц)

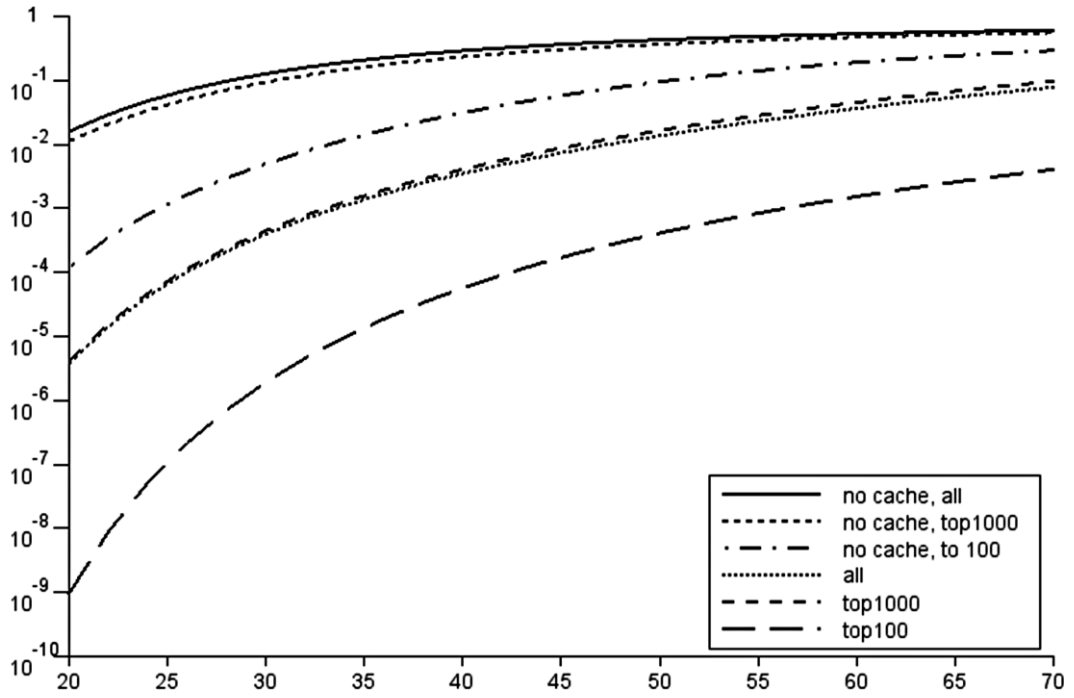


Рисунок 4.2 – Вероятность блокировки запроса на передачу эластичного трафика в зависимости от числа пользователей для разного объема трафика

4.3 СИСТЕМА С МАРКОВСКИМ ВХОДЯЩИМ ПОТОКОМ

4.3.1 Матрица интенсивностей переходов в блочном виде

Лемма 4.2. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и марковским входящим потоком, если на пространстве $\mathcal{X}$ состояний СП $X(t)$ введен лексикографический порядок

$$\begin{aligned}
 x = (n, w) &< (n, q, w) = x', \\
 x = (n, w) &< (n', w') = x' \Leftrightarrow \\
 &(n < n') \vee \\
 &\vee (n = n', w < w'),
 \end{aligned} \tag{4.16}$$

$$\begin{aligned} \mathbf{x} = (n, q, w) < (n', q', w') = \mathbf{x}' &\Leftrightarrow \\ (q < q') \vee \\ \vee (q = q', n < n') \vee \\ \vee (q = q', n = n', w < w'), \quad \mathbf{x}, \mathbf{x}' \in \mathcal{X}, \end{aligned}$$

тогда матрица интенсивностей переходов СП $\mathbf{X}(t)$ представима в блочном виде

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{B}_0 & \mathbf{B}_1 & \cdots & \mathbf{B}_M \\ \mathbf{C}_0 & \mathbf{A}_0 & & & \\ \mathbf{C}_1 & & \mathbf{A}_1 & & \\ \vdots & & & \ddots & \\ \mathbf{C}_M & & & & \mathbf{A}_M \end{bmatrix}, \quad (4.17)$$

где блоки определяются по формулам

$$\begin{aligned} \mathbf{Q}_{11} &= \text{diag}(\mathbf{D}_0 - \alpha \mathbf{I}, \mathbf{D}_0 - (\alpha + C\mu) \mathbf{I}, \dots, \mathbf{D}_0 - (\alpha + C\mu) \mathbf{I}, \mathbf{D} - (\alpha + C\mu) \mathbf{I}) + \\ &\quad + \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1) + \text{diag}^-(C\mu \mathbf{I}, \dots, C\mu \mathbf{I}), \\ \dim \mathbf{Q}_{11} &= (R+1)W \times (R+1)W, \end{aligned} \quad (4.18)$$

$$\begin{aligned} \mathbf{A}_q &= \text{diag}(\mathbf{D}_0 - \beta \mathbf{I}, \dots, \mathbf{D}_0 - \beta \mathbf{I}, \mathbf{D} - \beta \mathbf{I}) + \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1), \\ \dim \mathbf{A}_q &= (R-q+1)W \times (R-q+1)W, \quad q = 0, \dots, M, \end{aligned} \quad (4.19)$$

$$\mathbf{B}_q = \begin{bmatrix} \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \hline \alpha \mathbf{I} & & & \\ & \mathbf{0} & & \\ & & \ddots & \\ & & & \mathbf{0} \end{bmatrix}, \quad (4.20)$$

$$\dim \mathbf{B}_q = (R+1)W \times (R-q+1)W, \quad q = 0, \dots, M-1,$$

$$\mathbf{B}_M = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} \\ \hline \alpha \mathbf{I} & & \\ & \ddots & \\ & & \alpha \mathbf{I} \end{bmatrix}, \quad (4.21)$$

$$\dim \mathbf{B}_M = (R+1)W \times (R-M+1)W,$$

$$\mathbf{C}_q = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \beta \mathbf{I} & & \\ \vdots & & \vdots & & \ddots & \\ \mathbf{0} & \cdots & \mathbf{0} & \vdots & & \beta \mathbf{I} \end{bmatrix}, \quad (4.22)$$

$$\dim \mathbf{C}_q = (R-q+1)W \times (R+1)W, \quad q = 0, \dots, M.$$

Доказательство.

а) Введение лексикографического порядка

Разобьем пространство $\mathcal{X}$ состояний (4.2) СП $\mathbf{X}(t)$ на подмножества по состоянию ненадежного ресурса, числу q приостановленных сессий трафика и числу n сессий трафика в системе

$$\mathcal{X} = \mathcal{X}(1) \cup \mathcal{X}(0),$$

$$\mathcal{X}(1) = \{(n, w) : 0 \leq n \leq R, 1 \leq w \leq W\},$$

$$\mathcal{X}(0) = \{(n, q, w) : q \leq n \leq R, 0 \leq q \leq M, 1 \leq w \leq W\},$$

$$\mathcal{X}(1) = \bigcup_{n=0}^R \mathcal{X}(1, n),$$

$$\mathcal{X}(1, n) = \{(i, w) \in \mathcal{X}(1) : i = n\} = \{(n, w) \in \mathcal{X}(1)\},$$

$$\mathcal{X}(0) = \bigcup_{q=0}^M \mathcal{X}(0, q),$$

$$\mathcal{X}(0, q) = \{(n, i, w) \in \mathcal{X}(0) : i = q\} = \{(n, q, w) \in \mathcal{X}(0)\},$$

$$\mathcal{X}(0, q) = \bigcup_{n=q}^R \mathcal{X}(0, q, n),$$

$$\mathcal{X}(0, q, n) = \{(i, q, w) \in \mathcal{X}(0, q) : i = n\} = \{(n, q, w) \in \mathcal{X}(0, q)\}.$$

Введем лексикографический порядок на множестве $\mathcal{X}$ в порядке объединения $\mathcal{X}(u)$, $\mathcal{X}(0, q)$ и $\mathcal{X}(0, q, n)$

$$\begin{aligned} \mathcal{X} = & \{ \mathbf{x}(0), \mathbf{x}(1), \dots, \mathbf{x}(n), \dots, \mathbf{x}(R), \\ & \mathbf{x}(0,0), \mathbf{x}(1,0), \dots, \mathbf{x}(n,0), \dots, \mathbf{x}(R,0), \\ & \mathbf{x}(1,1), \mathbf{x}(2,1), \dots, \mathbf{x}(n,1), \dots, \mathbf{x}(R,1), \\ & \dots \\ & \mathbf{x}(q,q), \mathbf{x}(q+1,q), \dots, \mathbf{x}(n,q), \dots, \mathbf{x}(R,q), \\ & \dots \\ & \mathbf{x}(M,M), \mathbf{x}(M+1,M), \dots, \mathbf{x}(n,M), \dots, \mathbf{x}(R,M) \}. \end{aligned}$$

б) Описание блоков матрицы

С учетом введенного лексикографического порядка и ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (4.5), матрица $\mathbf{Q}$ имеет блочный вид из четырех блоков

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{10} \\ \mathbf{Q}_{01} & \mathbf{Q}_{00} \end{bmatrix}.$$

Блоки матрицы $\mathbf{Q}$ имеют следующую структуру

$$\mathbf{Q}_{10} = [\mathbf{B}_0 \quad \mathbf{B}_1 \quad \dots \quad \mathbf{B}_M],$$

$$\mathbf{Q}_{01} = \begin{bmatrix} \mathbf{C}_0 \\ \mathbf{C}_1 \\ \vdots \\ \mathbf{C}_M \end{bmatrix},$$

$$\mathbf{Q}_{11} = \begin{bmatrix} \mathbf{D}_0 - \alpha \mathbf{I} & \mathbf{D}_1 & & & \\ C\mu \mathbf{I} & \mathbf{D}_0 - (\alpha + C\mu) \mathbf{I} & \mathbf{D}_1 & & \\ & \ddots & \ddots & \ddots & \\ & & C\mu \mathbf{I} & \mathbf{D}_0 - (\alpha + C\mu) \mathbf{I} & \mathbf{D}_1 \\ & & & C\mu \mathbf{I} & \mathbf{D} - (\alpha + C\mu) \mathbf{I} \end{bmatrix},$$

$$\mathbf{Q}_{00} = \begin{bmatrix} \mathbf{A}_0 & & & \\ & \mathbf{A}_1 & & \\ & & \ddots & \\ & & & \mathbf{A}_M \end{bmatrix},$$

$$\mathbf{A}_q = \begin{bmatrix} \mathbf{D}_0 - \beta \mathbf{I} & \mathbf{D}_1 & & & \\ & \ddots & \ddots & & \\ & & \mathbf{D}_0 - \beta \mathbf{I} & \mathbf{D}_1 & \\ & & & \mathbf{D} - \beta \mathbf{I} & \\ & & & & \end{bmatrix}, \quad q = 0, \dots, M.$$

■

4.3.2 Матричный алгоритм расчета стационарного распределения

Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и марковским входящим потоком, в соответствии с лексикографическим порядком (4.16), представим стационарное распределение СП $\mathbf{X}(t)$ в виде

$$\begin{aligned} \boldsymbol{\pi} &= (\boldsymbol{\pi}^{11}, \boldsymbol{\pi}^0, \boldsymbol{\pi}^1, \dots, \boldsymbol{\pi}^q, \dots, \boldsymbol{\pi}^M), \\ \boldsymbol{\pi}^{11} &= (\pi_0, \pi_1, \dots, \pi_n, \dots, \pi_R), \\ \boldsymbol{\pi}^q &= (\pi_{qq}, \pi_{q+1,q}, \dots, \pi_{nq}, \dots, \pi_{Rq}), \quad q = 0, \dots, M, \\ \pi_n &= \pi(\mathbf{x}(n)), \quad n = 0, \dots, R, \\ \pi_{nq} &= (\pi(\mathbf{x}(q, q)), \pi(\mathbf{x}(q+1, q)), \dots, \pi(\mathbf{x}(n, q)), \dots, \pi(\mathbf{x}(R, q))), \\ & \quad q = 0, \dots, M, \quad n = q, \dots, R. \end{aligned} \tag{4.23}$$

Тогда с учетом леммы 4.2, справедлива следующая теорема.

Теорема 4.2. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и марковским входящим потоком, стационарное распределение СП $\mathbf{X}(t)$ рассчитывается в матричном виде по рекуррентным формулам

$$\begin{aligned}
 \pi_{nq} &= \alpha \cdot \pi_q \mathbf{XY}^{n-q}, \quad n = q, \dots, R-1, \quad q = 0, \dots, M-1, \\
 \pi_{Rq} &= \alpha \cdot \pi_q \mathbf{XY}^{R-1-q} \mathbf{W}, \quad q = 0, \dots, M-1, \\
 \pi_{nM} &= \alpha \sum_{i=M}^n \pi_i \mathbf{XY}^{n-i}, \quad n = M, \dots, R-1, \\
 \pi_{RM} &= \alpha \sum_{i=M}^{R-1} \pi_i \mathbf{XY}^{R-1-i} \mathbf{W} + \alpha \pi_R \mathbf{V},
 \end{aligned} \tag{4.24}$$

где вектор $(\pi_0, \pi_1, \dots, \pi_n, \dots, \pi_R)$ является единственным решением системы уравнений

$$\begin{aligned}
 \pi_0 [\mathbf{D}_0 - \alpha \mathbf{I}] + C\mu \cdot \pi_1 + \alpha\beta \cdot \pi_0 \mathbf{X} &= \mathbf{0}, \\
 \pi_{n-1} \mathbf{D}_1 + \pi_n [\mathbf{D}_0 - (\alpha + C\mu) \mathbf{I}] + C\mu \cdot \pi_{n+1} + \alpha\beta \sum_{q=0}^n \pi_q \mathbf{XY}^{n-q} &= \mathbf{0}, \quad n = 1, \dots, R-1, \\
 \pi_{R-1} \mathbf{D}_1 + \pi_R [\mathbf{D} - (\alpha + C\mu) \mathbf{I}] + \alpha\beta \left(\sum_{q=0}^{R-1} \pi_q \mathbf{XY}^{R-1-q} \mathbf{W} + \pi_R \mathbf{V} \right) &= \mathbf{0},
 \end{aligned} \tag{4.25}$$

а матрицы $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{V}$ и $\mathbf{W}$ имеют вид

$$\begin{aligned}
 \mathbf{X} &= -[\mathbf{D}_0 - \beta \mathbf{I}]^{-1}, \\
 \mathbf{Y} &= -\mathbf{D}_1 [\mathbf{D}_0 - \beta \mathbf{I}]^{-1}, \\
 \mathbf{V} &= -[\mathbf{D} - \beta \mathbf{I}]^{-1}, \\
 \mathbf{W} &= -\mathbf{D}_1 [\mathbf{D} - \beta \mathbf{I}]^{-1}.
 \end{aligned} \tag{4.26}$$

Доказательство.

а) Запись СУР в матричном виде

Запишем СУР $\pi \mathbf{Q} = \mathbf{0}$ с учетом блочно-трехдиагонального вида матрицы $\mathbf{Q}$ интенсивностей переходов (4.17)

$$\begin{aligned}
 \pi^{11} \mathbf{Q}_{11} + \sum_{q=0}^M \pi^q \mathbf{C}_q &= \mathbf{0}, \\
 \pi^{11} \mathbf{B}_q + \pi^q \mathbf{A}_q &= \mathbf{0}, \quad q = 0, \dots, M.
 \end{aligned}$$

Подставим вид матриц $\mathbf{Q}_{11}$, $\mathbf{A}_q$, $\mathbf{B}_q$ и $\mathbf{C}_q$ (4.18)–(4.22) в полученный вид СУР

$$\pi_0[\mathbf{D}_0 - \alpha\mathbf{I}] + C\mu \cdot \pi_1 + \beta \cdot \pi_{00} = \mathbf{0},$$

$$\pi_{n-1}\mathbf{D}_1 + \pi_n[\mathbf{D}_0 - (\alpha + C\mu)\mathbf{I}] + C\mu \cdot \pi_{n+1} + \beta \sum_{q=0}^n \pi_{nq} = \mathbf{0}, \quad n = 1, \dots, M,$$

$$\pi_{n-1}\mathbf{D}_1 + \pi_n[\mathbf{D}_0 - (\alpha + C\mu)\mathbf{I}] + C\mu \cdot \pi_{n+1} + \beta \sum_{q=0}^M \pi_{nq} = \mathbf{0}, \quad n = M + 1, \dots, R - 1,$$

$$\pi_{R-1}\mathbf{D}_1 + \pi_R[\mathbf{D} - (\alpha + C\mu)\mathbf{I}] + \beta \sum_{q=0}^M \pi_{Rq} = \mathbf{0},$$

$$\alpha \cdot \pi_q + \pi_{qq}[\mathbf{D}_0 - \beta\mathbf{I}] = \mathbf{0}, \quad q = 0, \dots, M - 1,$$

$$\pi_{n-1,q}\mathbf{D}_1 + \pi_{nq}[\mathbf{D}_0 - \beta\mathbf{I}] = \mathbf{0}, \quad n = q + 1, \dots, R - 1, \quad q = 0, \dots, M - 1,$$

$$\pi_{R-1,q}\mathbf{D}_1 + \pi_{Rq}[\mathbf{D} - \beta\mathbf{I}] = \mathbf{0}, \quad q = 0, \dots, M - 1,$$

$$\alpha \cdot \pi_M + \pi_{MM}[\mathbf{D}_0 - \beta\mathbf{I}] = \mathbf{0},$$

$$\alpha \cdot \pi_n + \pi_{n-1,M}\mathbf{D}_1 + \pi_{nM}[\mathbf{D}_0 - \beta\mathbf{I}] = \mathbf{0}, \quad n = M + 1, \dots, R - 1,$$

$$\alpha \cdot \pi_R + \pi_{R-1,M}\mathbf{D}_1 + \pi_{RM}[\mathbf{D} - \beta\mathbf{I}] = \mathbf{0}.$$

б) Получение зависимости π_{nq} от π_q для $q < M$

Выразим π_{qq} через π_q из матричного уравнения СУР при $\mathbf{x}(q, q)$, $q < M$

$$\pi_{qq} = -\alpha \cdot \pi_q[\mathbf{D}_0 - \beta\mathbf{I}]^{-1}, \quad q = 0, \dots, M - 1,$$

и обозначим $\mathbf{X} = -[\mathbf{D}_0 - \beta\mathbf{I}]^{-1}$, тогда

$$\pi_{qq} = \alpha \cdot \pi_q \mathbf{X}, \quad q = 0, \dots, M - 1.$$

Получим зависимость π_{nq} от π_q для $0 \leq q < M$, $q < n < R$. Матричное уравнение СУР при $\mathbf{x}(q + 1, q)$ имеет вид

$$\pi_{qq}\mathbf{D}_1 + \pi_{q+1,q}[\mathbf{D}_0 - \beta\mathbf{I}] = \mathbf{0}, \quad q = 0, \dots, M - 1,$$

Подставим полученное соотношение для π_{qq}

$$\pi_{q+1,q} = -\alpha \cdot \pi_q \mathbf{X} \mathbf{D}_1[\mathbf{D}_0 - \beta\mathbf{I}]^{-1}, \quad q = 0, \dots, M - 1,$$

и обозначим $\mathbf{Y} = -\mathbf{D}_1[\mathbf{D}_0 - \beta\mathbf{I}]^{-1}$, тогда

$$\pi_{q+1,q} = \alpha \cdot \pi_q \mathbf{X} \mathbf{Y} = \alpha \cdot \pi_q \mathbf{X} \mathbf{Y}^1, \quad q = 0, \dots, M - 1.$$

Предположим, что соотношение верно для $n-1$

$$\pi_{n-1,q} = \alpha \cdot \pi_q \mathbf{XY}^{n-1-q}, \quad q = 0, \dots, M-1.$$

Подставим его в матричное уравнение СУР при $x(n, q)$

$$\alpha \cdot \pi_q \mathbf{XY}^{n-1-q} \mathbf{D}_1 + \pi_{nq} [\mathbf{D}_0 - \beta \mathbf{I}] = \mathbf{0}, \quad q = 0, \dots, M-1,$$

выразим π_{nq}

$$\pi_{nq} = -\alpha \cdot \pi_q \mathbf{XY}^{n-1-q} \mathbf{D}_1 [\mathbf{D}_0 - \beta \mathbf{I}]^{-1} = \alpha \cdot \pi_q \mathbf{XY}^{n-q}, \quad q = 0, \dots, M-1,$$

и объединим с соотношением для π_{qq} . Следовательно, сделанное предположение о зависимости π_{nq} от π_q верно, таким образом

$$\pi_{nq} = \alpha \cdot \pi_q \mathbf{XY}^{n-q}, \quad n = q, \dots, R-1, \quad q = 0, \dots, M-1,$$

$$\mathbf{X} = -[\mathbf{D}_0 - \beta \mathbf{I}]^{-1},$$

$$\mathbf{Y} = -\mathbf{D}_1 [\mathbf{D}_0 - \beta \mathbf{I}]^{-1}.$$

в) Выражение π_{Rq} через π_q для $q < M$

Подставим выражение $\pi_{R-1,q}$ в матричное уравнение СУР при $x(R, q)$,
 $0 \leq q < M$

$$\alpha \cdot \pi_q \mathbf{XY}^{R-1-q} \mathbf{D}_1 + \pi_{Rq} [\mathbf{D} - \beta \mathbf{I}] = \mathbf{0}, \quad q = 0, \dots, M-1,$$

и выразим π_{Rq}

$$\pi_{Rq} = -\alpha \cdot \pi_q \mathbf{XY}^{R-1-q} \mathbf{D}_1 [\mathbf{D} - \beta \mathbf{I}]^{-1}, \quad q = 0, \dots, M-1.$$

Обозначим $\mathbf{W} = -\mathbf{D}_1 [\mathbf{D} - \beta \mathbf{I}]^{-1}$, тогда

$$\pi_{Rq} = \alpha \cdot \pi_q \mathbf{XY}^{R-1-q} \mathbf{W}, \quad q = 0, \dots, M-1.$$

г) Получение зависимости π_{nM} от π_i для $M < n < R$

Выразим π_{MM} через π_M из матричного уравнения СУР при $x(M, M)$

$$\pi_{MM} = -\alpha \cdot \pi_M [\mathbf{D}_0 - \beta \mathbf{I}]^{-1} = \alpha \cdot \pi_M \mathbf{F}.$$

Получим зависимость π_{nM} от π_i для $M < n < R$. Матричное уравнение СУР при $x(M+1, M)$ имеет вид

$$\alpha \cdot \pi_{M+1} + \pi_{MM} \mathbf{D}_1 + \pi_{M+1, M} [\mathbf{D}_0 - \beta \mathbf{I}] = \mathbf{0}.$$

Подставим полученное соотношение для π_{MM}

$$\begin{aligned} \pi_{M+1, M} &= -(\alpha \cdot \pi_{M+1} + \alpha \cdot \pi_M \mathbf{X} \mathbf{D}_1) [\mathbf{D}_0 - \beta \mathbf{I}]^{-1} = \\ &= -\alpha \left(\pi_{M+1} [\mathbf{D}_0 - \beta \mathbf{I}]^{-1} + \pi_M \mathbf{X} \mathbf{D}_1 [\mathbf{D}_0 - \beta \mathbf{I}]^{-1} \right) = \\ &= \alpha (\pi_{M+1} \mathbf{X} + \pi_M \mathbf{X} \mathbf{Y}) = \alpha \sum_{i=M}^{M+1} \pi_n \mathbf{X} \mathbf{Y}^{M+1-i}. \end{aligned}$$

Предположим, что соотношение верно для $n-1$

$$\pi_{n-1, M} = \alpha \sum_{i=M}^{n-1} \pi_i \mathbf{X} \mathbf{Y}^{n-1-i}.$$

Подставим его в матричное уравнение СУР при $x(n, M)$

$$\alpha \cdot \pi_n + \alpha \sum_{i=M}^{n-1} \pi_i \mathbf{X} \mathbf{Y}^{n-1-i} \mathbf{D}_1 + \pi_{nM} [\mathbf{D}_0 - \beta \mathbf{I}] = \mathbf{0},$$

выразим π_{nM}

$$\begin{aligned} \pi_{nM} &= -\left(\alpha \cdot \pi_n + \alpha \sum_{i=M}^{n-1} \pi_i \mathbf{X} \mathbf{Y}^{n-1-i} \mathbf{D}_1 \right) [\mathbf{D}_0 - \beta \mathbf{I}]^{-1} = \\ &= -\alpha \left(\pi_n [\mathbf{D}_0 - \beta \mathbf{I}]^{-1} + \sum_{i=M}^{n-1} \pi_i \mathbf{X} \mathbf{Y}^{n-1-i} \mathbf{D}_1 [\mathbf{D}_0 - \beta \mathbf{I}]^{-1} \right) = \\ &= \alpha \left(\pi_n \mathbf{X} + \sum_{i=M}^{n-1} \pi_i \mathbf{X} \mathbf{Y}^{n-1-i} \mathbf{Y} \right) = \alpha \sum_{i=M}^n \pi_i \mathbf{X} \mathbf{Y}^{n-i} \end{aligned}$$

и объединим с соотношением для π_{MM} . Следовательно, сделанное предположение о зависимости π_{nM} от π_i верно, таким образом

$$\pi_{nM} = \alpha \sum_{i=M}^n \pi_i \mathbf{X} \mathbf{Y}^{n-i}, \quad n = M, \dots, R-1.$$

д) Выражение π_{RM} через π_i

Подставим выражение $\pi_{R-1,M}$ в матричное уравнение СУР при $x(R, M)$

$$\alpha \cdot \pi_R + \alpha \sum_{i=M}^{R-1} \pi_i \mathbf{X} \mathbf{Y}^{R-1-i} \mathbf{D}_1 + \pi_{RM} [\mathbf{D} - \beta \mathbf{I}] = \mathbf{0}$$

и выразим π_{RM}

$$\begin{aligned} \pi_{RM} &= -\alpha \left(\pi_R + \sum_{i=M}^{R-1} \pi_i \mathbf{X} \mathbf{Y}^{R-1-i} \mathbf{D}_1 \right) [\mathbf{D} - \beta \mathbf{I}]^{-1} = \\ &= -\alpha \left(\pi_R [\mathbf{D} - \beta \mathbf{I}]^{-1} + \sum_{i=M}^{R-1} \pi_i \mathbf{X} \mathbf{Y}^{R-1-i} \mathbf{D}_1 [\mathbf{D} - \beta \mathbf{I}]^{-1} \right). \end{aligned}$$

Обозначим $\mathbf{V} = -[\mathbf{D} - \beta \mathbf{I}]^{-1}$, тогда

$$\pi_{RM} = \alpha \left(\pi_R \mathbf{V} + \sum_{i=M}^{R-1} \pi_i \mathbf{X} \mathbf{Y}^{R-1-i} \mathbf{W} \right) = \alpha \sum_{i=M}^{R-1} \pi_i \mathbf{X} \mathbf{Y}^{R-1-i} \mathbf{W} + \alpha \pi_R \mathbf{V}.$$

е) Получение системы уравнений относительно π_n

Подставим выражение π_{00} в матричное уравнение СУР при $x(0)$

$$\pi_0 [\mathbf{D}_0 - \alpha \mathbf{I}] + C\mu \cdot \pi_1 + \alpha\beta \cdot \pi_0 \mathbf{X} = \mathbf{0}.$$

Подставим выражение π_{nq} в матричное уравнение СУР при $x(n)$, $0 < n < M$

$$\pi_{n-1} \mathbf{D}_1 + \pi_n [\mathbf{D}_0 - (\alpha + C\mu) \mathbf{I}] + C\mu \cdot \pi_{n+1} + \alpha\beta \sum_{q=0}^n \pi_q \mathbf{X} \mathbf{Y}^{n-q} = \mathbf{0}, \quad n = 1, \dots, M-1.$$

Подставим выражение π_{nq} в матричное уравнение СУР при $x(n)$, $M \leq n < R$

$$\begin{aligned} \pi_{n-1} \mathbf{D}_1 + \pi_n [\mathbf{D}_0 - (\alpha + C\mu) \mathbf{I}] + C\mu \cdot \pi_{n+1} + \alpha\beta \left(\sum_{q=0}^{M-1} \pi_q \mathbf{X} \mathbf{Y}^{n-q} + \sum_{i=M}^n \pi_i \mathbf{X} \mathbf{Y}^{n-i} \right) &= \mathbf{0}, \\ n &= M, \dots, R-1. \end{aligned}$$

Объединим с соотношением для $0 < n < M$

$$\pi_{n-1} \mathbf{D}_1 + \pi_n [\mathbf{D}_0 - (\alpha + C\mu) \mathbf{I}] + C\mu \cdot \pi_{n+1} + \alpha\beta \sum_{q=0}^n \pi_q \mathbf{X} \mathbf{Y}^{n-q} = \mathbf{0}, \quad n = 1, \dots, R-1.$$

Подставим выражение π_{Rq} в матричное уравнение СУР при $x(R)$

$$\pi_{R-1}\mathbf{D}_1 + \pi_R[\mathbf{D} - (\alpha + C\mu)\mathbf{I}] + \\ + \alpha\beta \left(\sum_{q=0}^{M-1} \pi_q \mathbf{X}\mathbf{Y}^{R-1-q}\mathbf{W} + \sum_{i=M}^{R-1} \pi_i \mathbf{X}\mathbf{Y}^{R-1-i}\mathbf{W} + \pi_R \mathbf{V} \right) = \mathbf{0},$$

т.е.

$$\pi_{R-1}\mathbf{D}_1 + \pi_R[\mathbf{D} - (\alpha + C\mu)\mathbf{I}] + \alpha\beta \left(\sum_{q=0}^{R-1} \pi_q \mathbf{X}\mathbf{Y}^{R-1-q}\mathbf{W} + \pi_R \mathbf{V} \right) = \mathbf{0}.$$

■

Замечание. При нахождении вектора $\pi^{11} = (\pi_0, \pi_1, \dots, \pi_n, \dots, \pi_R)$ из системы уравнений (4.25) в теореме 4.2 потребуется условие нормировки

$$\sum_{n=0}^R \pi_n + \sum_{q=0}^M \sum_{n=q}^R \pi_{nq} = \theta. \quad (4.27)$$

4.3.3 Анализ вероятности блокировки

Из теоремы 4.2 вытекает следующее следствие.

Следствие 4.2. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и марковским входящим потоком, вероятность блокировки запроса на передачу эластичного трафика рассчитывается по формуле

$$P_{bl} = \frac{1}{\lambda} \left(\pi_R + \sum_{q=0}^M \pi_{Rq} \right) \mathbf{D}_1 \mathbf{1}^T. \quad (4.28)$$

Проиллюстрируем поведение среднего времени передачи эластичного трафика на примере просмотра веб-страниц. На рисунке Рисунок 4.3 представлены графики вероятности блокировки запроса на передачу эластичного трафика в зависимости от интенсивности отказа для разных интенсивностей восстановления.

Вероятность блокировки растет с увеличением интенсивности возникновения блокировок пути прямой видимости. Они не только прерывают

текущие сессии, но и ограничивают прием новых запросов. Когда очередь заполняется, система отклоняет новые запросы пользователей. Это особенно важно в сценариях, где физические помехи, такие как движение объектов или погодные условия, приводят к частым потерям сигнала. Например, в беспроводных сетях с высокой плотностью пользователей каждый новый сбой сокращает доступные ресурсы, увеличивая риск перегрузки. Рост параметра интенсивности возникновения блокировок пути прямой видимости усиливает эту тенденцию, так как путь чаще становится недоступным, а очередь – переполненной.

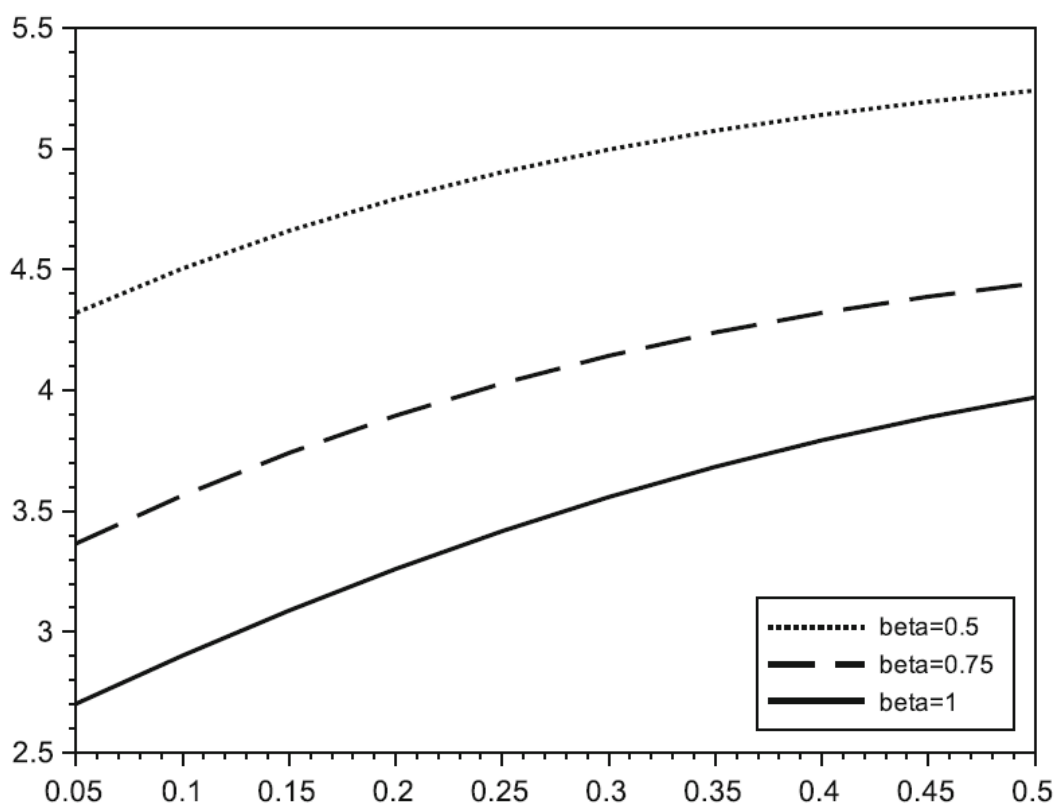


Рисунок 4.3 – Вероятность блокировки запроса на передачу эластичного трафика в зависимости от интенсивности отказа ресурса (1/c) для разной интенсивности восстановления ресурса (1/c)

4.4 АНАЛИЗ ВЛОЖЕННОГО ПОГЛОЩАЮЩЕГО СЛУЧАЙНОГО ПРОЦЕССА

4.4.1 Построение поглощающего случайного процесса

Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, построим поглощающий СП $Y(t)$, описывающий траекторию сессии эластичного трафика в системе – с момента поступления запроса на передачу трафика и до момента завершения его передачи. Чтобы отслеживать траекторию сессии, введем дополнительную переменную r к состояниям $(n, u) \in \mathcal{X}: n > 0$ СП $X(t)$:

если $r = 0$ при $u = 1$, то сессия находится на ресурсе и обслуживается,

если $r = 0$ при $u = 0$, то после восстановления ресурса сессия сразу окажется на нем и начнет обслуживаться,

если $r > 0$ при $u = 1$, то сессия находится в очереди на r месте и ожидает начала обслуживания,

если $r > 0$ при $u = 0$, то после восстановления ресурса сессия окажется в очереди на r месте.

Добавим также поглощающее состояние ω .

Таким образом, соответствие между состояниями СП $X(t)$ и состояниями СП $Y(t)$ выглядит следующим образом

$$\begin{aligned} \{(0,0), (0,1)\} &\rightarrow \emptyset, \\ \{(n,u)\} &\rightarrow \{(0,n,u)\}, \quad n=1,\dots,M, \quad u=0,1, \\ \{(n,u)\} &\rightarrow \{(0,n,u), (1,n,u), (2,n,u), \dots, (n-M,n,u)\}, \quad n=M+1,\dots,R, \quad u=0,1, \end{aligned} \quad (4.29)$$

а пространство состояний СП $Y(t)$ будет иметь вид

$$\begin{aligned} \mathcal{Y} = & \{(0, n, u): 0 < n \leq M, u = 0, 1\} \cup \\ & \cup \{(r, n, u): M < n \leq R, u = 0, 1, 0 \leq r \leq n - M\} \cup \{\omega\} = \\ = & \{(0, n, u): 1 \leq n \leq R, u = 0, 1\} \cup \\ & \cup \{(r, n, u): M + r \leq n \leq R, u = 0, 1, 1 \leq r \leq R - M\} \cup \{\omega\}. \end{aligned} \quad (4.30)$$

Ненулевые элементы матрицы интенсивностей переходов имеют вид

$$\begin{aligned} G[(r, n, u), (r, n+1, u)] &= \lambda, \quad n+1 \leq R, \\ G[(0, n, 1), \omega] &= \frac{1}{n} C\mu, \quad 0 < n \leq M, \\ G[(0, n, 1), (0, n-1, 1)] &= \frac{n-1}{n} C\mu, \quad 0 < n \leq M, \\ G[(0, n, 1), \omega] &= \frac{1}{M} C\mu, \quad n > M, \\ G[(0, n, 1), (0, n-1, 1)] &= \frac{M-1}{M} C\mu, \quad n > M, \\ G[(r, n, 1), (r-1, n-1, 1)] &= C\mu, \quad n > M, \quad r > 0, \\ G[(r, n, 1), (r, n, 0)] &= \alpha, \\ G[(r, n, 0), (r, n, 1)] &= \beta. \end{aligned} \quad (4.31)$$

Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, построим поглощающий СП $Y(t)$, описывающий траекторию сессии эластичного трафика в системе – с момента поступления запроса на передачу трафика и до момента завершения его передачи. Чтобы отслеживать траекторию сессии, состояний $(n, w) \in \mathcal{X}: n > 0$ и $(n, q, w) \in \mathcal{X}: n > 0$ СП $X(t)$ достаточно. Добавим также поглощающее состояние ω .

Таким образом, соответствие между состояниями СП $X(t)$ и состояниями СП $Y(t)$ выглядит следующим образом

$$\begin{aligned} \{(0, w)\} &\rightarrow \emptyset, \quad w = 1, \dots, W, \\ \{(n, w)\} &\rightarrow \{(n, w)\}, \quad n = 1, \dots, R, \quad w = 1, \dots, W, \\ \{(0, q, w)\} &\rightarrow \emptyset, \quad q = 0, \dots, M, \quad w = 1, \dots, W, \\ \{(n, q, w)\} &\rightarrow \{(n, q, w)\}, \quad n > 0, \quad n = q, \dots, R, \quad q = 0, \dots, M, \quad w = 1, \dots, W, \end{aligned} \quad (4.32)$$

а пространство состояний СП $Y(t)$ будет иметь вид

$$\begin{aligned} \mathcal{Y} &= \{(n, w): 0 < n \leq R, 1 \leq w \leq W\} \cup \\ &\cup \{(n, q, w): 0 < n \leq R, n \leq q \leq M, 1 \leq w \leq W\} \cup \{\omega\} = \\ &= \{(n, w): 0 < n \leq R, 1 \leq w \leq W\} \cup \\ &\cup \{(n, q, w): n > 0, q \leq n \leq R, 0 \leq q \leq M, 1 \leq w \leq W\} \cup \{\omega\}. \end{aligned} \quad (4.33)$$

Ненулевые элементы матрицы интенсивностей переходов имеют вид

$$\begin{aligned} G[x(n), x(n+1)] &= D_1, \quad n+1 \leq R, \\ G[x(n, q), x(n+1, q)] &= D_1, \quad n+1 \leq R, \\ G[x(n), \omega] &= \frac{1}{n} C \mu I, \quad n > 0, \\ G[x(n), x(n-1)] &= \frac{n-1}{n} C \mu I, \quad n > 0, \\ G[x(n), x(n, q)] &= \alpha I, \\ G[x(n, q), x(n)] &= \beta I. \end{aligned} \quad (4.34)$$

4.4.2 Матрица интенсивностей переходов для пуассоновского потока

Лемма 4.3. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, если на пространстве $\mathcal{Y}$ состояний СП $Y(t)$ введен лексикографический порядок

$$\begin{aligned}
 \mathbf{y} &= (r, n, u) < \omega = \mathbf{y}', \\
 \mathbf{y} &= (r, n, u) < (r', n', u') = \mathbf{y}' \Leftrightarrow \\
 & (u > u') \vee \\
 & \vee (u = u', r < r') \vee \\
 & \vee (u = u', r = r', n < n'), \quad \mathbf{y}, \mathbf{y}' \in \mathcal{Y},
 \end{aligned} \tag{4.35}$$

тогда матрица интенсивностей переходов СП $\mathbf{Y}(t)$ представима в блочном виде

$$\mathbf{G} = \begin{bmatrix} \mathbf{S} & \mathbf{s}^T \\ \mathbf{0} & 0 \end{bmatrix}, \tag{4.36}$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{10} \\ \mathbf{S}_{01} & \mathbf{S}_{00} \end{bmatrix}, \tag{4.37}$$

$$\mathbf{s} = (\mathbf{s}_1, \mathbf{0}), \tag{4.38}$$

где блоки $\mathbf{S}_{11}$ и $\mathbf{S}_{00}$ также являются блочными матрицами

$$\begin{aligned}
 \mathbf{S}_{10} &= \alpha \mathbf{I}, \\
 \mathbf{S}_{01} &= \beta \mathbf{I}, \\
 \mathbf{S}_{11} &= \text{diag}(\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_{R-M}) + \text{diag}^-(\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_{R-M}), \\
 \mathbf{S}_{00} &= \text{diag}(\mathbf{C}_0, \mathbf{C}_1, \dots, \mathbf{C}_{R-M})
 \end{aligned} \tag{4.39}$$

с блоками, определяемыми по формулам

$$\begin{aligned}
 \mathbf{A}_0 &= \text{diag}(-(\lambda + C\mu + \alpha), \dots, -(\lambda + C\mu + \alpha), -(C\mu + \alpha)) + \text{diag}^+(\lambda, \dots, \lambda) + \\
 &+ \text{diag}^-\left(\frac{1}{2}C\mu, \frac{2}{3}C\mu, \dots, \frac{M-2}{M-1}C\mu, \frac{M-1}{M}C\mu, \underbrace{\frac{M-1}{M}C\mu, \dots, \frac{M-1}{M}C\mu}_{R-M+1}\right),
 \end{aligned} \tag{4.40}$$

$$\dim \mathbf{A}_0 = R \times R,$$

$$\begin{aligned}
 \mathbf{A}_r &= \text{diag}(-(\lambda + C\mu + \alpha), \dots, -(\lambda + C\mu + \alpha), -(C\mu + \alpha)) + \text{diag}^+(\lambda, \dots, \lambda), \\
 \dim \mathbf{A}_r &= (R - M - r + 1) \times (R - M - r + 1), \quad r = 1, \dots, R - M,
 \end{aligned} \tag{4.41}$$

$$\mathbf{B}_1 = \begin{bmatrix} 0 & \dots & 0 & \vdots & C\mu & 0 \\ \vdots & & \vdots & & \ddots & \ddots \\ 0 & \dots & 0 & \vdots & C\mu & 0 \end{bmatrix}, \tag{4.42}$$

$$\dim \mathbf{B}_1 = (R - M) \times R,$$

$$\mathbf{B}_r = \begin{bmatrix} C\mu & 0 & & \\ & \ddots & \ddots & \\ & & C\mu & 0 \end{bmatrix}, \quad (4.43)$$

$$\dim \mathbf{B}_r = (R - M - r + 1) \times (R - M - r + 2), \quad r = 2, \dots, R - M,$$

$$\mathbf{C}_r = \text{diag}(-(\lambda + \beta), \dots, -(\lambda + \beta), -\beta) + \text{diag}^+(\lambda, \dots, \lambda),$$

$$\dim \mathbf{C}_0 = R \times R, \quad (4.44)$$

$$\dim \mathbf{C}_r = (R - M - r + 1) \times (R - M - r + 1), \quad r = 1, \dots, R - M,$$

$$\mathbf{s}_1 = \left(C\mu, \frac{1}{2}C\mu, \dots, \frac{1}{M-1}C\mu, \frac{1}{M}C\mu, \underbrace{\frac{1}{M}C\mu, \dots, \frac{1}{M}C\mu}_{R-M} \right), \quad (4.45)$$

$$\dim \mathbf{s}_1 = R.$$

Доказательство.

а) Введение лексикографического порядка

Разобьем пространство $\mathcal{Y}$ состояний (4.30) СП $\mathbf{Y}(t)$ на подмножества по состоянию u ненадежного ресурса и параметру r

$$\mathcal{Y} = \mathcal{Y}(1) \cup \mathcal{Y}(0) \cup \{\omega\},$$

$$\mathcal{Y}(u) = \{(r, n, i) \in \mathcal{Y} : i = u\} = \{(r, n, u) \in \mathcal{Y}\},$$

$$\mathcal{Y}(u) = \bigcup_{r=0}^{R-M} \mathcal{Y}(r, u),$$

$$\mathcal{Y}(0, u) = \{(0, n, u) \in \mathcal{Y}(u)\} = \{(0, 1, u), (0, 2, u), \dots, (0, n, u), \dots, (0, R, u)\},$$

$$\mathcal{Y}(r, u) = \{(i, n, u) \in \mathcal{Y} : i = r\} = \{(r, n, u) \in \mathcal{Y}(u)\} =$$

$$= \{(r, M + r, u), (r, M + r + 1, u), \dots, (r, n, u), \dots, (r, R, u)\}, \quad r = 1, \dots, R - M.$$

Введем лексикографический порядок на множестве $\mathcal{Y}$ в порядке объединения $\mathcal{Y}(u)$, $\{\omega\}$ и $\mathcal{Y}(r, u)$

$$\mathcal{Y} = \{ (0,1,1), (0,2,1), \dots, (0,n,1), \dots, (0,R,1), \\ (1,M+1,1), (1,M+2,1), \dots, (1,n,1), \dots, (1,R,1), \\ (2,M+2,1), (2,M+3,1), \dots, (2,n,1), \dots, (2,R,1), \\ \dots, \\ (r,M+r,1), (r,M+r+1,1), \dots, (r,n,1), \dots, (r,R,1), \\ \dots, \\ (R-M-1, R-1,1), (R-M-1, R,1), \\ (R-M, R,1), \\ (0,1,0), (0,2,0), \dots, (0,n,0), \dots, (0,R,0), \\ (1,M+1,0), (1,M+2,0), \dots, (1,n,0), \dots, (1,R,0), \\ (2,M+2,0), (2,M+3,0), \dots, (2,n,0), \dots, (2,R,0), \\ \dots, \\ (r,M+r,0), (r,M+r+1,0), \dots, (r,n,0), \dots, (r,R,0), \\ \dots, \\ (R-M-1, R-1,0), (R-M-1, R,0), \\ (R-M, R,0), \\ \omega \}.$$

б) Описание блоков матрицы

С учетом введенного лексикографического порядка, матрица $\mathbf{G}$ интенсивностей переходов имеет блочный вид, где последние строка и столбец соответствуют поглощающему состоянию ω

$$\mathbf{G} = \left[\begin{array}{cccc|cccc|c} \mathbf{A}_0 & & & & \alpha \mathbf{I} & & & & \mathbf{s}_1^T \\ \mathbf{B}_1 & \mathbf{A}_1 & & & & \alpha \mathbf{I} & & & \mathbf{0}^T \\ & \ddots & \ddots & & & & \ddots & & \vdots \\ & & \mathbf{B}_{R-M} & \mathbf{A}_{R-M} & & & & \alpha \mathbf{I} & \mathbf{0}^T \\ \hline \beta \mathbf{I} & & & & \mathbf{C}_0 & & & & \mathbf{0}^T \\ & \beta \mathbf{I} & & & & \mathbf{C}_1 & & & \mathbf{0}^T \\ & & \ddots & & & & \ddots & & \vdots \\ & & & \beta \mathbf{I} & & & & \mathbf{C}_{R-M} & \mathbf{0}^T \\ \hline \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & 0 \end{array} \right],$$

$$\mathbf{A}_0 = \begin{bmatrix} -(\lambda + C\mu + \alpha) & \lambda & & & & \\ \frac{1}{2}C\mu & \ddots & \lambda & & & \\ & \frac{2}{3}C\mu & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \frac{M-1}{M}C\mu & \ddots & \lambda \\ & & & & \ddots & \ddots & \lambda \\ & & & & & \frac{M-1}{M}C\mu & -(C\mu + \alpha) \end{bmatrix},$$

$$\mathbf{A}_r = \begin{bmatrix} -(\lambda + C\mu + \alpha) & \lambda & & & \\ & \ddots & \ddots & & \\ & & -(\lambda + C\mu + \alpha) & \lambda & \\ & & & \ddots & \\ & & & & -(C\mu + \alpha) \end{bmatrix}, \quad r > 0,$$

$$\mathbf{C}_r = \begin{bmatrix} -(\lambda + \beta) & \lambda & & \\ & \ddots & \ddots & \\ & & -(\lambda + \beta) & \lambda \\ & & & -\beta \end{bmatrix}.$$

■

4.4.3 Матрица интенсивностей переходов для марковского потока

Лемма 4.4. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и марковским входящим потоком, если на пространстве $\mathcal{Y}$ состояний СП $\mathbf{Y}(t)$ введен лексикографический порядок

$$\begin{aligned}
 y &= (n, w) < \omega = y', \\
 y &= (n, q, w) < \omega = y', \\
 y &= (n, w) < (n, q, w) = y', \\
 y &= (n, w) < (n', w') = y' \Leftrightarrow \\
 &\quad (n < n') \vee \\
 &\quad \vee (n = n', w < w'), \\
 y &= (n, q, w) < (n', q', w') = y' \Leftrightarrow \\
 &\quad (q < q') \vee \\
 &\quad \vee (q = q', n < n') \vee \\
 &\quad \vee (q = q', n = n', w < w'), \quad y, y' \in \mathcal{Y},
 \end{aligned} \tag{4.46}$$

тогда матрица интенсивностей переходов СП $\mathbf{Y}(t)$ представима в блочном виде

$$\mathbf{G} = \begin{bmatrix} \mathbf{S} & \mathbf{s}^T \\ \mathbf{0} & 0 \end{bmatrix}, \tag{4.47}$$

$$\mathbf{S} = \begin{bmatrix} \tilde{\mathbf{Q}}_{11} & \tilde{\mathbf{B}}_0 & \tilde{\mathbf{B}}_1 & \cdots & \tilde{\mathbf{B}}_M \\ \tilde{\mathbf{C}}_0 & \tilde{\mathbf{A}}_0 & & & \\ \tilde{\mathbf{C}}_1 & & \tilde{\mathbf{A}}_1 & & \\ \vdots & & & \ddots & \\ \tilde{\mathbf{C}}_M & & & & \tilde{\mathbf{A}}_M \end{bmatrix}, \tag{4.48}$$

$$\mathbf{s} = (\mathbf{s}_1, \mathbf{0}), \tag{4.49}$$

где блоки определяются по формулам

$$\begin{aligned}
 \tilde{\mathbf{Q}}_{11} &= \text{diag}(\mathbf{D}_0 - (\alpha + C\mu)\mathbf{I}, \dots, \mathbf{D}_0 - (\alpha + C\mu)\mathbf{I}, \mathbf{D} - (\alpha + C\mu)\mathbf{I}) + \\
 &\quad + \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1) + \\
 &\quad + \text{diag}^-\left(\frac{1}{2}C\mu\mathbf{I}, \frac{2}{3}C\mu\mathbf{I}, \dots, \frac{M-2}{M-1}C\mu\mathbf{I}, \frac{M-1}{M}C\mu\mathbf{I}, \underbrace{\frac{M-1}{M}C\mu\mathbf{I}, \dots, \frac{M-1}{M}C\mu\mathbf{I}}_{R-M+1}\right),
 \end{aligned} \tag{4.50}$$

$$\dim \tilde{\mathbf{Q}}_{11} = RW \times RW,$$

$$\begin{aligned}
 \tilde{\mathbf{A}}_q &= \text{diag}(\mathbf{D}_0 - \beta\mathbf{I}, \dots, \mathbf{D}_0 - \beta\mathbf{I}, \mathbf{D} - \beta\mathbf{I}) + \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1), \\
 \dim \tilde{\mathbf{A}}_0 &= (R - q)W \times (R - q)W, \\
 \dim \tilde{\mathbf{A}}_q &= (R - q + 1)W \times (R - q + 1)W, \quad q = 1, \dots, M,
 \end{aligned} \tag{4.51}$$

$$\tilde{\mathbf{B}}_q = \begin{bmatrix} \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \hline \alpha \mathbf{I} & & & \\ & \mathbf{0} & & \\ & & \ddots & \\ & & & \mathbf{0} \end{bmatrix},$$

$$\dim \tilde{\mathbf{B}}_0 = RW \times (R - q)W, \quad (4.52)$$

$$\dim \tilde{\mathbf{B}}_q = RW \times (R - q + 1)W, \quad q = 1, \dots, M - 1,$$

$$\tilde{\mathbf{B}}_M = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} \\ \hline \alpha \mathbf{I} & & \\ & \ddots & \\ & & \alpha \mathbf{I} \end{bmatrix}, \quad (4.53)$$

$$\dim \tilde{\mathbf{B}}_M = RW \times (R - M)W,$$

$$\tilde{\mathbf{C}}_q = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \vdots & \beta \mathbf{I} \\ \vdots & & \vdots & \ddots & \\ \mathbf{0} & \cdots & \mathbf{0} & \vdots & \beta \mathbf{I} \end{bmatrix},$$

$$\dim \tilde{\mathbf{C}}_0 = (R - q)W \times RW, \quad (4.54)$$

$$\dim \tilde{\mathbf{C}}_q = (R - q + 1)W \times RW, \quad q = 1, \dots, M,$$

$$\mathbf{s}_1 = \left(C\mu \mathbf{I}, \frac{1}{2}C\mu \mathbf{I}, \dots, \frac{1}{M-1}C\mu \mathbf{I}, \frac{1}{M}C\mu \mathbf{I}, \underbrace{\frac{1}{M}C\mu \mathbf{I}, \dots, \frac{1}{M}C\mu \mathbf{I}}_{R-M} \right), \quad (4.55)$$

$$\dim \mathbf{s}_1 = RW.$$

Доказательство.

Схема доказательства аналогична лемме 4.3.

■

4.5 АЛГОРИТМ РАСЧЕТА МОМЕНТОВ ВРЕМЕНИ ПЕРЕДАЧИ ТРАФИКА

4.5.1 Преобразование Лапласа-Стилтьеса времени передачи трафика

Для анализа времени передачи трафика воспользуемся поглощающим СП $Y(t)$, описывающим траекторию сессии эластичного трафика в системе – с момента поступления запроса на передачу трафика и до момента завершения его передачи. Процесс начинается в момент поступления запроса и завершается при завершении передачи трафика. Все промежуточные состояния отражают возможные сценарии работы системы: активное передачи, приостановка обслуживания, ожидание возобновления передачи и т.д. Время передачи трафика эквивалентно времени, за которое процесс достигает поглощающего состояния. Это позволяет представить распределение времени передачи как фазовое распределение, для которого естественно применяются методы анализа на основе преобразований Лапласа-Стилтьеса.

$$\phi(s) = p(sI - S)^{-1} s^T,$$

где p начальное распределение

$$\begin{aligned} p(0, n, u) &= \frac{\pi(n-1, u)}{1 - P_{bl}}, \quad n = 1, \dots, M, \quad u = 0, 1, \\ p(n - M, n, u) &= \frac{\pi(n-1, u)}{1 - P_{bl}}, \quad n = M + 1, \dots, R, \quad u = 0, 1, \\ p(r, n, u) &= 0, \quad r = 0, \dots, n - M - 1, \quad n = M + 1, \dots, R, \quad u = 0, 1, \\ p(\omega) &= 0. \end{aligned} \tag{4.56}$$

Это стандартное выражение для фазовых распределений, но прямое вычисление обратной матрицы при больших размерах системы становится вычислительно сложным. Элементы обратной матрицы соответствуют ожидаемому времени пребывания в состояниях до поглощения.

$$St^T = -1^T$$

4.5.2 Матричный алгоритм расчета моментов времени

Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, справедлива следующая теорема.

Теорема 4.3. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, ожидаемое время пребывания в состояниях СП $Y(t)$ до поглощения рассчитывается в матричном виде по формуле

$$\begin{aligned} \mathbf{t}_{01}^T &= -(\mathbf{A}_0 + \alpha\beta\mathbf{X}_0)^{-1}(\mathbf{1}^T + \alpha\mathbf{X}_0\mathbf{1}^T), \\ \mathbf{t}_{r1}^T &= -(\mathbf{A}_r + \alpha\beta\mathbf{X}_r)^{-1}(\mathbf{1}^T - \alpha\mathbf{X}_r\mathbf{1}^T + \mathbf{B}_r\mathbf{t}_{r-1,1}^T), \quad r = 1, \dots, R-M, \\ \mathbf{t}_{r0}^T &= \mathbf{X}_r(\mathbf{1}^T + \beta\mathbf{t}_{r1}^T), \quad r = 0, \dots, R-M, \end{aligned} \quad (4.57)$$

где матрица $\mathbf{X}_r$ имеет вид

$$\mathbf{X}_r = -\mathbf{C}_r^{-1}. \quad (4.58)$$

Доказательство.

а) Запись системы уравнений относительно $\mathbf{t}_{ru}^T$ в матричном виде

Запишем систему $\mathbf{S}\mathbf{t}^T = -\mathbf{1}^T$ с учетом блочно-трехдиагонального вида матрицы $\mathbf{S}$ интенсивностей переходов

$$\begin{aligned} \mathbf{A}_0\mathbf{t}_{01}^T + \alpha\mathbf{t}_{00}^T &= -\mathbf{1}^T, \\ \mathbf{B}_r\mathbf{t}_{r-1,1}^T + \mathbf{A}_r\mathbf{t}_{r1}^T + \alpha\mathbf{t}_{r0}^T &= -\mathbf{1}^T, \quad r = 1, \dots, R-M, \\ \beta\mathbf{t}_{r1}^T + \mathbf{C}_r\mathbf{t}_{r0}^T &= -\mathbf{1}^T, \quad r = 0, \dots, R-M \end{aligned}$$

б) Выражение $\mathbf{t}_{r0}^T$ через $\mathbf{t}_{r1}^T$

Выразим $\mathbf{t}_{r0}^T$ через $\mathbf{t}_{r1}^T$ из матричного уравнения при $u = 0$

$$\mathbf{t}_{r0}^T = -\mathbf{C}_r^{-1}(\mathbf{1}^T + \beta\mathbf{t}_{r1}^T), \quad r = 0, \dots, R-M$$

и обозначим $\mathbf{X}_r = -\mathbf{C}_r^{-1}$, тогда

$$\mathbf{t}_{r0}^T = \mathbf{X}_r(\mathbf{1}^T + \beta\mathbf{t}_{r1}^T), \quad r = 0, \dots, R-M$$

в) Получение $\mathbf{t}_{01}^T$

Подставим соотношение при $r = 0$ в матричное уравнение при $u = 1$ и $r = 0$

$$\mathbf{A}_0 \mathbf{t}_{01}^T + \alpha \mathbf{X}_0 (\mathbf{1}^T + \beta \mathbf{t}_{01}^T) = -\mathbf{1}^T$$

и найдем $\mathbf{t}_{01}^T$

$$\mathbf{t}_{01}^T = -(\mathbf{A}_0 + \alpha \beta \mathbf{X}_0)^{-1} (\mathbf{1}^T + \alpha \mathbf{X}_0 \mathbf{1}^T).$$

г) Выражение $\mathbf{t}_{r1}^T$ через $\mathbf{t}_{r-1,1}^T$

Подставим соотношения при $0 < r \leq R - M$ в матричное уравнение при $u = 1$ и $0 < r \leq R - M$

$$\mathbf{B}_r \mathbf{t}_{r-1,1}^T + \mathbf{A}_r \mathbf{t}_{r1}^T + \alpha \mathbf{X}_r (\mathbf{1}^T + \beta \mathbf{t}_{r1}^T) = -\mathbf{1}^T, \quad r = 1, \dots, R - M,$$

и выразим $\mathbf{t}_{r1}^T$ через $\mathbf{t}_{r-1,1}^T$

$$\mathbf{t}_{r1}^T = -(\mathbf{A}_r + \alpha \beta \mathbf{X}_r)^{-1} (\mathbf{1}^T - \alpha \mathbf{X}_r \mathbf{1}^T + \mathbf{B}_r \mathbf{t}_{r-1,1}^T), \quad r = 1, \dots, R - M.$$

■

Следствие 4.3. Для модели с эластичным трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и пуассоновским входящим потоком, среднее время передачи эластичного трафика рассчитывается по формуле

$$T = \mathbf{p} \mathbf{t}^T. \quad (4.59)$$

4.5.3 Анализ среднего времени передачи

Проиллюстрируем поведение среднего времени передачи эластичного трафика на примере просмотра веб-страниц. На рисунках 4.4 и 4.5 представлены графики для пуассоновского входящего потока, а на рисунке 4.6 – марковского потока.

Из рисунка 4.4 следует, что увеличение пропускной способности канала сокращает время загрузки веб-страницы. Это происходит потому, что высокая

скорость позволяет системе быстрее обрабатывать запросы и распределять ресурсы между пользователями без задержек. При низкой пропускной способности система быстро достигает своих пределов, что приводит к увеличению времени ожидания. Это особенно заметно при высокой активности пользователей, когда даже небольшое увеличение нагрузки резко ухудшает производительность.

По рисунку 4.5 видно, что чем больше объем данных, тем дольше загрузка. Это связано с тем, что обработка крупных файлов требует больше времени, особенно при ограниченной пропускной способности. Однако кэширование помогает смягчить эту проблему. Даже при больших размерах страниц эффективная оптимизация контента может приблизить время загрузки к приемлемому уровню.

Рисунок 4.6 демонстрирует, что частые блокировки пути прямой видимости замедляют загрузку. Это происходит из-за недоступности канала, что прерывает загрузку контента. Пользователи остаются в системе, ожидая своей очереди, что удлиняет время ожидания. Кроме того, частые блокировки приводят к перегрузке системы, вызывая накопление очереди и дополнительные задержки.

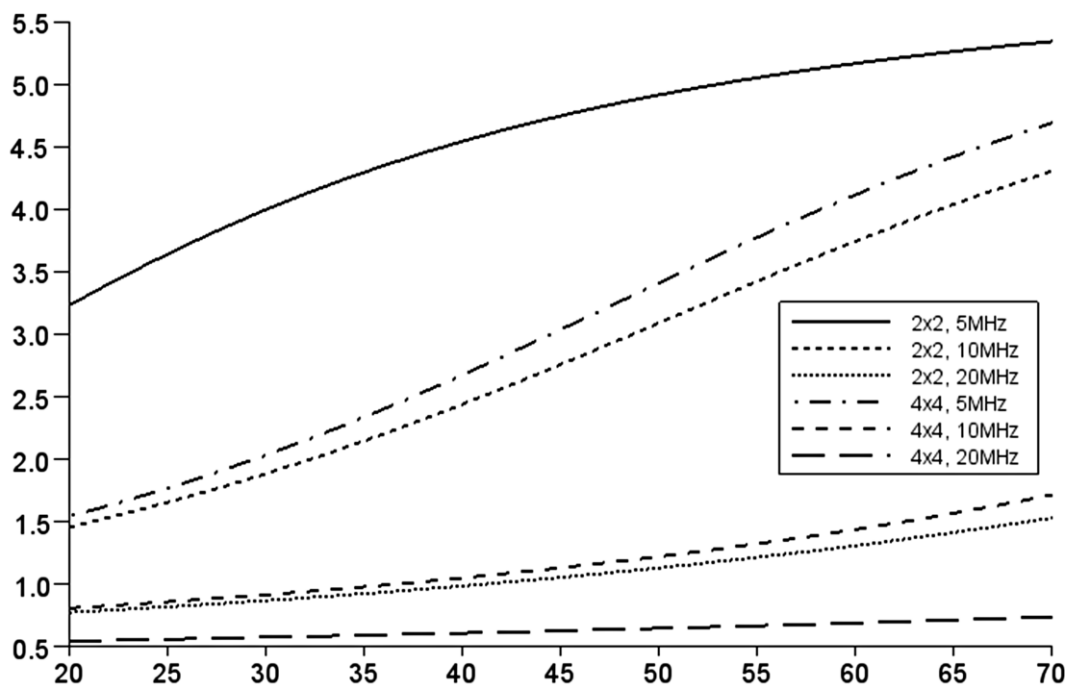


Рисунок 4.4 – Среднее время передачи эластичного трафика (с) в зависимости от числа пользователей для разных схем MIMO и ширины радиочастот (МГц)

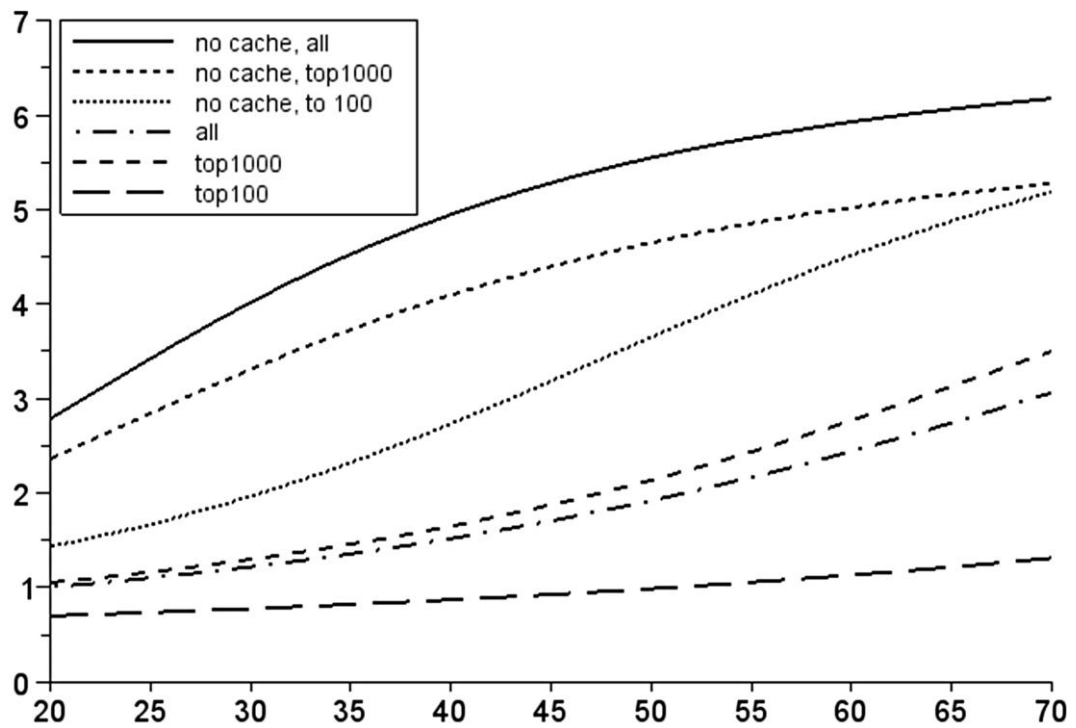


Рисунок 4.5 – Среднее время передачи эластичного трафика (с) в зависимости от числа пользователей для разного объема трафика

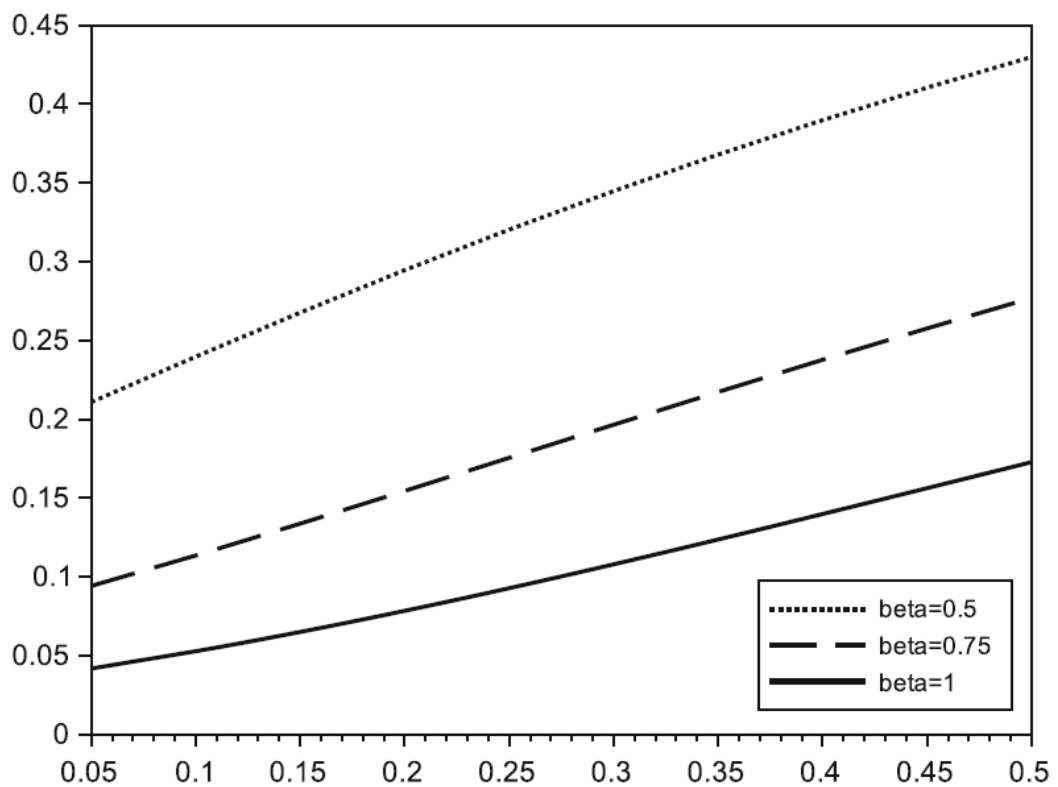


Рисунок 4.6 – Среднее время передачи эластичного трафика (с) в зависимости от интенсивности отказа ресурса (1/с) для разной интенсивности восстановления ресурса (1/с)

4.6 СИСТЕМА С ЧАСТИЧНЫМ ОТКАЗОМ РЕСУРСА

4.6.1 Матрица интенсивностей переходов в блочном виде

Лемма 4.5. Для модели с эластичным трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением скорости передачи трафика и марковским входящим потоком, если на пространстве $\mathcal{X}$ состояний СП $X(t)$ введен лексикографический порядок

$$\begin{aligned} x = (n, u, w) < (n', u', w') = x' &\Leftrightarrow \\ &(u > u') \vee \\ &\vee (u = u', n < n') \vee \\ &\vee (u = u', n = n', w < w'), \quad x, x' \in \mathcal{X}, \end{aligned} \quad (4.60)$$

тогда матрица интенсивностей переходов СП $X(t)$ представима в блочном виде

$$Q = \begin{bmatrix} Q_{11} & Q_{10} \\ Q_{01} & Q_{00} \end{bmatrix}, \quad (4.61)$$

где блоки определяются по формулам

$$\begin{aligned} Q_{11} = & \text{diag}(\mathbf{D}_0 - \alpha \mathbf{I}, \mathbf{D}_0 - (\alpha + (V + C)\mu) \mathbf{I}, \dots, \mathbf{D}_0 - (\alpha + (V + C)\mu) \mathbf{I}, \\ & \mathbf{D} - (\alpha + (V + C)\mu) \mathbf{I}) + \\ & + \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1) + \text{diag}^-((V + C)\mu \mathbf{I}, \dots, (V + C)\mu \mathbf{I}), \end{aligned} \quad (4.62)$$

$$\begin{aligned} Q_{00} = & \text{diag}(\mathbf{D}_0 - \beta \mathbf{I}, \mathbf{D}_0 - (\beta + V\mu) \mathbf{I}, \dots, \mathbf{D}_0 - (\beta + V\mu) \mathbf{I}, \mathbf{D} - (\beta + V\mu) \mathbf{I}) + \\ & + \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1) + \text{diag}^-(V\mu \mathbf{I}, \dots, V\mu \mathbf{I}), \end{aligned} \quad (4.63)$$

$$Q_{10} = \alpha \mathbf{I}, \quad (4.64)$$

$$Q_{01} = \beta \mathbf{I}, \quad (4.65)$$

$$\dim Q_{11} = \dim Q_{00} = \dim Q_{10} = \dim Q_{01} = (R + 1)W \times (R + 1)W.$$

Доказательство.

а) Введение лексикографического порядка

Разобьем пространство $\mathcal{X}$ состояний (4.3) СП $X(t)$ на подмножества по состоянию u ненадежного ресурса и числу n сессий трафика в системе

$$\mathcal{X} = \mathcal{X}(1) \cup \mathcal{X}(0),$$

$$\mathcal{X}(u) = \{(n, i, w) \in \mathcal{X} : i = u\} = \{(n, u, w) \in \mathcal{X}\},$$

$$\mathcal{X}(u) = \bigcup_{n=0}^R \mathcal{X}(u, n),$$

$$\mathcal{X}(u, n) = \{(i, u, w) \in \mathcal{X}(u) : i = n\} = \{(n, u, w) \in \mathcal{X}(u)\}.$$

Введем лексикографический порядок на множестве $\mathcal{X}$ в порядке объединения $\mathcal{X}(u)$ и $\mathcal{X}(u, n)$

$$\mathcal{X} = \{x(0,1), x(1,1), \dots, x(n,1), \dots, x(R,1), \\ x(0,0), x(1,0), \dots, x(n,0), \dots, x(R,0)\}.$$

б) Описание блоков матрицы

С учетом введенного лексикографического порядка и ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (4.6), матрица $\mathbf{Q}$ имеет блочный вид из четырех блоков

$$\mathbf{Q} = \begin{bmatrix} \mathbf{D}_0 - \alpha \mathbf{I} & \mathbf{D}_1 & & & \alpha \mathbf{I} & & & \\ (V+C)\mu \mathbf{I} & \ddots & \ddots & & & \ddots & & \\ & \ddots & \mathbf{D}_0 - X \mathbf{I} & \mathbf{D}_1 & & & \ddots & \\ & & (V+C)\mu \mathbf{I} & \mathbf{D} - X \mathbf{I} & & & & \alpha \mathbf{I} \\ \hline & \beta \mathbf{I} & & & \mathbf{D}_0 - \beta \mathbf{I} & \mathbf{D}_1 & & \\ & & \ddots & & V\mu \mathbf{I} & \ddots & \ddots & \\ & & & \ddots & & \ddots & \mathbf{D}_0 - Y \mathbf{I} & \mathbf{D}_1 \\ & & & & & & V\mu \mathbf{I} & \mathbf{D} - Y \mathbf{I} \\ & & & & & & & \beta \mathbf{I} \end{bmatrix},$$

где $X = \alpha + (V+C)\mu$ и $Y = \beta + V\mu$.

■

4.6.2 Поглощающий случайный процесс

Для модели с эластичным трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением скорости передачи трафика и марковским входящим потоком, построим поглощающий СП $Y(t)$, описывающий траекторию сессии эластичного трафика в системе – с момента поступления запроса на передачу трафика и до момента завершения его передачи. Чтобы отслеживать траекторию сессии, введем дополнительную переменную r к состояниям $(n, u, w) \in \mathcal{X}: n > 0$ СП $X(t)$:

если $r = 0$ при $u = 1$, то сессия находится на ресурсе и обслуживается,

если $r = 0$ при $u = 0$, то после восстановления ресурса сессия сразу окажется на нем и начнет обслуживаться,

если $r > 0$ при $u = 1$, то сессия находится в очереди на r месте и ожидает начала обслуживания,

если $r > 0$ при $u = 0$, то после восстановления ресурса сессия окажется в очереди на r месте.

Добавим также поглощающее состояние ω .

Таким образом, соответствие между состояниями СП $X(t)$ и состояниями СП $Y(t)$ выглядит следующим образом

$$\begin{aligned} \{(0, u, w)\} &\rightarrow \emptyset, \quad u = 0, 1, \quad w = 1, \dots, W, \\ \{(n, u, w)\} &\rightarrow \{(0, n, u, w)\}, \quad n = 1, \dots, M, \quad u = 0, 1, \quad w = 1, \dots, W, \\ \{(n, u, w)\} &\rightarrow \{(0, n, u, w), (1, n, u, w), (2, n, u, w), \dots, (n - M, n, u, w)\}, \\ &\quad n = M + 1, \dots, R, \quad u = 0, 1, \quad w = 1, \dots, W, \end{aligned} \quad (4.66)$$

а пространство состояний СП $Y(t)$ будет иметь вид

$$\begin{aligned} \mathcal{Y} &= \{(0, n, u, w): 0 < n \leq M, u = 0, 1, 1 \leq w \leq W\} \cup \\ &\quad \cup \{(r, n, u, w): M < n \leq R, u = 0, 1, 1 \leq w \leq W, 0 \leq r \leq n - M\} \cup \{\omega\} = \\ &= \{(0, n, u, w): 1 \leq n \leq R, u = 0, 1, 1 \leq w \leq W\} \cup \\ &\quad \cup \{(r, n, u, w): M + r \leq n \leq R, u = 0, 1, 1 \leq w \leq W, 1 \leq r \leq R - M\} \cup \{\omega\}. \end{aligned} \quad (4.67)$$

Ненулевые элементы матрицы интенсивностей переходов имеют вид

$$\begin{aligned}
 \mathbf{G}[\mathbf{x}(r, n, u), \mathbf{x}(r, n+1, u)] &= \mathbf{D}_1, \quad n+1 \leq R, \\
 \mathbf{G}[\mathbf{x}(0, n, 0), \omega] &= \frac{1}{n} C \mu \mathbf{I}, \quad 0 < n \leq M, \\
 \mathbf{G}[\mathbf{x}(0, n, 1), \omega] &= \frac{1}{n} (V + C) \mu \mathbf{I}, \quad 0 < n \leq M, \\
 \mathbf{G}[\mathbf{x}(0, n, 0), \mathbf{x}(0, n-1, 0)] &= \frac{n-1}{n} C \mu \mathbf{I}, \quad 0 < n \leq M, \\
 \mathbf{G}[\mathbf{x}(0, n, 1), \mathbf{x}(0, n-1, 1)] &= \frac{n-1}{n} (V + C) \mu \mathbf{I}, \quad 0 < n \leq M, \\
 \mathbf{G}[\mathbf{x}(0, n, 0), \omega] &= \frac{1}{M} C \mu \mathbf{I}, \quad n > M, \\
 \mathbf{G}[\mathbf{x}(0, n, 1), \omega] &= \frac{1}{M} (V + C) \mu \mathbf{I}, \quad n > M, \\
 \mathbf{G}[\mathbf{x}(0, n, 0), \mathbf{x}(0, n-1, 0)] &= \frac{M-1}{M} C \mu \mathbf{I}, \quad n > M, \\
 \mathbf{G}[\mathbf{x}(0, n, 1), \mathbf{x}(0, n-1, 1)] &= \frac{M-1}{M} (V + C) \mu \mathbf{I}, \quad n > M, \\
 \mathbf{G}[\mathbf{x}(r, n, 0), \mathbf{x}(r-1, n-1, 0)] &= C \mu \mathbf{I}, \quad n > M, \quad r > 0, \\
 \mathbf{G}[\mathbf{x}(r, n, 1), \mathbf{x}(r-1, n-1, 1)] &= (V + C) \mu \mathbf{I}, \quad n > M, \quad r > 0, \\
 \mathbf{G}[\mathbf{x}(r, n, 1), \mathbf{x}(r, n, 0)] &= \alpha \mathbf{I}, \\
 \mathbf{G}[\mathbf{x}(r, n, 0), \mathbf{x}(r, n, 1)] &= \beta \mathbf{I}.
 \end{aligned} \tag{4.68}$$

Лемма 4.6. Для модели с эластичным трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением скорости передачи трафика и марковским входящим потоком, если на пространстве $\mathcal{Y}$ состояний СП $\mathbf{Y}(t)$ введен лексикографический порядок

$$\begin{aligned}
 \mathbf{y} = (r, n, u, w) < (r', n', u', w') = \mathbf{y}' &\Leftrightarrow \\
 (u > u') \vee \\
 \vee (u = u', r < r') \vee \\
 \vee (u = u', r = r', n < n') \vee \\
 \vee (u = u', r = r', n = n', w < w'), \quad \mathbf{y}, \mathbf{y}' \in \mathcal{Y},
 \end{aligned} \tag{4.69}$$

тогда матрица интенсивностей переходов СП $\mathbf{Y}(t)$ представима в блочном виде

$$\mathbf{G} = \begin{bmatrix} \mathbf{S} & \mathbf{s}^T \\ \mathbf{0} & 0 \end{bmatrix}, \quad (4.70)$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{10} \\ \mathbf{S}_{01} & \mathbf{S}_{00} \end{bmatrix}, \quad (4.71)$$

$$\mathbf{s} = (\mathbf{s}_1, \mathbf{0}, \mathbf{s}_0, \mathbf{0}), \quad (4.72)$$

где блоки $\mathbf{S}_{11}$ и $\mathbf{S}_{00}$ также являются блочными матрицами

$$\begin{aligned} \mathbf{S}_{10} &= \alpha \mathbf{I}, \\ \mathbf{S}_{01} &= \beta \mathbf{I}, \end{aligned} \quad (4.73)$$

$$\mathbf{S}_{uu} = \text{diag}(\mathbf{A}_{u0}, \mathbf{A}_{u1}, \dots, \mathbf{A}_{u, R-M}) + \text{diag}^-(\mathbf{B}_{u1}, \mathbf{B}_{u2}, \dots, \mathbf{B}_{u, R-M}), \quad u = 0, 1$$

с блоками, определяемыми по формулам

$$\begin{aligned} \mathbf{A}_{00} &= \text{diag}(\mathbf{D}_0 - (\beta + C\mu)\mathbf{I}, \dots, \mathbf{D}_0 - (\beta + C\mu)\mathbf{I}, \mathbf{D} - (\beta + C\mu)\mathbf{I}) + \\ &+ \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1) + \\ &+ \text{diag}^-\left(\frac{1}{2}C\mu\mathbf{I}, \frac{2}{3}C\mu\mathbf{I}, \dots, \frac{M-2}{M-1}C\mu\mathbf{I}, \frac{M-1}{M}C\mu\mathbf{I}, \underbrace{\frac{M-1}{M}C\mu\mathbf{I}, \dots, \frac{M-1}{M}C\mu\mathbf{I}}_{R-M+1}\right), \end{aligned} \quad (4.74)$$

$$\dim \mathbf{A}_{00} = RW \times RW,$$

$$\begin{aligned} \mathbf{A}_{10} &= \\ &= \text{diag}(\mathbf{D}_0 - (\alpha + (V + C)\mu)\mathbf{I}, \dots, \mathbf{D}_0 - (\alpha + (V + C)\mu)\mathbf{I}, \mathbf{D} - (\alpha + (V + C)\mu)\mathbf{I}) + \\ &+ \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1) + \\ &+ \text{diag}^-\left(\frac{1}{2}(V + C)\mu\mathbf{I}, \frac{2}{3}(V + C)\mu\mathbf{I}, \dots, \frac{M-2}{M-1}(V + C)\mu\mathbf{I}, \frac{M-1}{M}(V + C)\mu\mathbf{I}, \right. \\ &\quad \left. \underbrace{\frac{M-1}{M}(V + C)\mu\mathbf{I}, \dots, \frac{M-1}{M}(V + C)\mu\mathbf{I}}_{R-M+1}\right), \end{aligned} \quad (4.75)$$

$$\dim \mathbf{A}_{10} = RW \times RW,$$

$$\begin{aligned} \mathbf{A}_{0r} &= \text{diag}(\mathbf{D}_0 - (\beta + C\mu)\mathbf{I}, \dots, \mathbf{D}_0 - (\beta + C\mu)\mathbf{I}, \mathbf{D} - (\beta + C\mu)\mathbf{I}) + \\ &+ \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1), \end{aligned} \quad (4.76)$$

$$\dim \mathbf{A}_{0r} = (R - M - r + 1)W \times (R - M - r + 1)W, \quad r = 1, \dots, R - M,$$

$$\begin{aligned} \mathbf{A}_{1r} &= \\ &= \text{diag}(\mathbf{D}_0 - (\alpha + (V + C)\mu)\mathbf{I}, \dots, \mathbf{D}_0 - (\alpha + (V + C)\mu)\mathbf{I}, \mathbf{D} - (\alpha + (V + C)\mu)\mathbf{I}) + \\ &\quad + \text{diag}^+(\mathbf{D}_1, \dots, \mathbf{D}_1), \end{aligned} \quad (4.77)$$

$$\dim \mathbf{A}_{0r} = (R - M - r + 1)W \times (R - M - r + 1)W, \quad r = 1, \dots, R - M,$$

$$\mathbf{B}_{01} = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & C\mu\mathbf{I} & \mathbf{0} \\ \vdots & & \vdots & & \ddots & \ddots \\ \mathbf{0} & \cdots & \mathbf{0} & & C\mu\mathbf{I} & \mathbf{0} \end{bmatrix}, \quad (4.78)$$

$$\dim \mathbf{B}_{01} = (R - M)W \times RW,$$

$$\mathbf{B}_{11} = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & (V + C)\mu\mathbf{I} & \mathbf{0} \\ \vdots & & \vdots & & \ddots & \ddots \\ \mathbf{0} & \cdots & \mathbf{0} & & (V + C)\mu\mathbf{I} & \mathbf{0} \end{bmatrix}, \quad (4.79)$$

$$\dim \mathbf{B}_{11} = (R - M)W \times RW,$$

$$\mathbf{B}_{0r} = \begin{bmatrix} C\mu\mathbf{I} & \mathbf{0} \\ & \ddots & \ddots \\ & & C\mu\mathbf{I} & \mathbf{0} \end{bmatrix}, \quad (4.80)$$

$$\dim \mathbf{B}_{0r} = (R - M - r + 1)W \times (R - M - r + 2)W, \quad r = 2, \dots, R - M,$$

$$\mathbf{B}_{1r} = \begin{bmatrix} (V + C)\mu\mathbf{I} & \mathbf{0} \\ & \ddots & \ddots \\ & & (V + C)\mu\mathbf{I} & \mathbf{0} \end{bmatrix}, \quad (4.81)$$

$$\dim \mathbf{B}_{1r} = (R - M - r + 1)W \times (R - M - r + 2)W, \quad r = 2, \dots, R - M,$$

$$\mathbf{s}_0 = \left(C\mu\mathbf{I}, \frac{1}{2}C\mu\mathbf{I}, \dots, \frac{1}{M-1}C\mu\mathbf{I}, \frac{1}{M}C\mu\mathbf{I}, \underbrace{\frac{1}{M}C\mu\mathbf{I}, \dots, \frac{1}{M}C\mu\mathbf{I}}_{R-M} \right), \quad (4.82)$$

$$\dim \mathbf{s}_0 = RW,$$

$$\begin{aligned} \mathbf{s}_1 &= \left((V + C)\mu\mathbf{I}, \frac{1}{2}(V + C)\mu\mathbf{I}, \dots, \frac{1}{M-1}(V + C)\mu\mathbf{I}, \frac{1}{M}(V + C)\mu\mathbf{I}, \right. \\ &\quad \left. \underbrace{\frac{1}{M}(V + C)\mu\mathbf{I}, \dots, \frac{1}{M}(V + C)\mu\mathbf{I}}_{R-M} \right), \end{aligned} \quad (4.83)$$

$$\dim \mathbf{s}_1 = RW.$$

Доказательство.

а) Введение лексикографического порядка

Разобьем пространство $\mathcal{Y}$ состояний (4.67) СП $Y(t)$ на подмножества по состоянию u ненадежного ресурса, параметру r и числу n сессий трафика в системе

$$\mathcal{Y} = \mathcal{Y}(1) \cup \mathcal{Y}(0) \cup \{\omega\},$$

$$\mathcal{Y}(u) = \{(r, n, i, w) \in \mathcal{Y} : i = u\} = \{(r, n, u, w) \in \mathcal{Y}\},$$

$$\mathcal{Y}(u) = \bigcup_{r=0}^{R-M} \mathcal{Y}(r, u),$$

$$\mathcal{Y}(r, u) = \{(i, n, u, w) \in \mathcal{Y}(u) : i = r\} = \{(r, n, u, w) \in \mathcal{Y}(u)\},$$

$$\mathcal{Y}(r, u) = \bigcup_{n=1}^{R-M-r} \mathcal{Y}(r, n, u),$$

$$\mathcal{Y}(r, n, u) = \{(r, i, u, w) \in \mathcal{Y}(r, u) : i = n\} = \{(r, n, u, w) \in \mathcal{Y}(r, u)\}.$$

Введем лексикографический порядок на множестве $\mathcal{Y}$ в порядке объединения

$\mathcal{Y}(u)$, $\{\omega\}$, $\mathcal{Y}(r, u)$ и $\mathcal{Y}(r, n, u)$

$$\begin{aligned} \mathcal{Y} = & \{y(0, 1, 1), y(0, 2, 1), \dots, y(0, n, 1), \dots, y(0, R, 1), \\ & y(1, M + 1, 1), y(1, M + 2, 1), \dots, y(1, n, 1), \dots, y(1, R, 1), \\ & y(2, M + 2, 1), y(2, M + 3, 1), \dots, y(2, n, 1), \dots, y(2, R, 1), \\ & \dots, \\ & y(r, M + r, 1), y(r, M + r + 1, 1), \dots, y(r, n, 1), \dots, y(r, R, 1), \\ & \dots, \\ & y(R - M - 1, R - 1, 1), y(R - M - 1, R, 1), \\ & y(R - M, R, 1), \end{aligned}$$

$$\begin{aligned}
 & y(0,1,0), y(0,2,0), \dots, y(0,n,0), \dots, y(0,R,0), \\
 & y(1,M+1,0), y(1,M+2,0), \dots, y(1,n,0), \dots, y(1,R,0), \\
 & y(2,M+2,0), y(2,M+3,0), \dots, y(2,n,0), \dots, y(2,R,0), \\
 & \dots, \\
 & y(r,M+r,0), y(r,M+r+1,0), \dots, y(r,n,0), \dots, y(r,R,0), \\
 & \dots, \\
 & y(R-M-1,R-1,0), y(R-M-1,R,0), \\
 & y(R-M,R,0), \\
 & \omega \}.
 \end{aligned}$$

б) Описание блоков матрицы

С учетом введенного лексикографического порядка, матрица $\mathbf{G}$ интенсивностей переходов имеет блочный вид, где последние строка и столбец соответствуют поглощающему состоянию ω

$$\mathbf{G} = \left[\begin{array}{cccc|cccc|c}
 \mathbf{A}_{10} & & & & \alpha \mathbf{I} & & & & s_1^T \\
 \mathbf{B}_{11} & \mathbf{A}_{11} & & & & \alpha \mathbf{I} & & & \mathbf{0}^T \\
 & \ddots & \ddots & & & & \ddots & & \vdots \\
 & & \mathbf{B}_{1,R-M} & \mathbf{A}_{1,R-M} & & & & \alpha \mathbf{I} & \mathbf{0}^T \\
 \hline
 \beta \mathbf{I} & & & & \mathbf{A}_{00} & & & & s_0^T \\
 & \beta \mathbf{I} & & & \mathbf{B}_{01} & \mathbf{A}_{01} & & & \mathbf{0}^T \\
 & & \ddots & & & \ddots & \ddots & & \vdots \\
 & & & \beta \mathbf{I} & & & \mathbf{B}_{0,R-M} & \mathbf{A}_{0,R-M} & \mathbf{0}^T \\
 \hline
 \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & 0
 \end{array} \right].$$

■

4.6.3 Моменты времени передачи трафика

Для модели с эластичным трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением скорости передачи трафика и марковским входящим потоком, ПЛС времени передачи эластичного трафика рассчитывается по формуле

$$\phi(s) = p(sI - S)^{-1} s^T, \quad (4.84)$$

где p начальное распределение поглощающего СП $Y(t)$.

Вытекает следующее следствие.

Следствие 4.4. Для модели с эластичным трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением скорости передачи трафика и марковским входящим потоком, начальный момент i -порядка времени передачи эластичного трафика рассчитывается по формуле

$$T_i = i! p(-S^{-1})^i, \quad i = 1, 2, \dots \quad (4.85)$$

Проиллюстрируем поведение среднего значения и дисперсии времени передачи эластичного трафика на примере передачи критически важных данных. В случае стихийных бедствий или аварий оперативная связь между службами, такими как полиция, спасатели и медики, помогает координировать действия и передавать информацию о пострадавших, маршрутах эвакуации и угрозах. Например, очки дополненной реальности или дроны требуют передачи видео в реальном времени, что невозможно без mmWave-технологий. Однако даже короткие сбои в сети из-за блокировок прямой видимости могут вызвать задержки и дезорганизацию команд. Непрерывность связи в экстремальных условиях крайне важна.

На рисунках 4.7 и 4.8 представлены графики среднего значения и дисперсии времени передачи критически важных данных соответственно от коэффициента $1 - \alpha/\beta$ для разных среднего времени полной доступности диапазона радиочастот и пропускной способности. Когда блокировки прямой видимости происходят часто, среднее время передачи данных увеличивается из-за высокой нагрузки, но дисперсия остается низкой. При редких блокировках среднее время снижается, но дисперсия возрастает. Низкая пропускная способность обеспечивает устойчивость системы даже при редких блокировках, в то время как высокая пропускная способность делает систему уязвимой к перегрузкам.

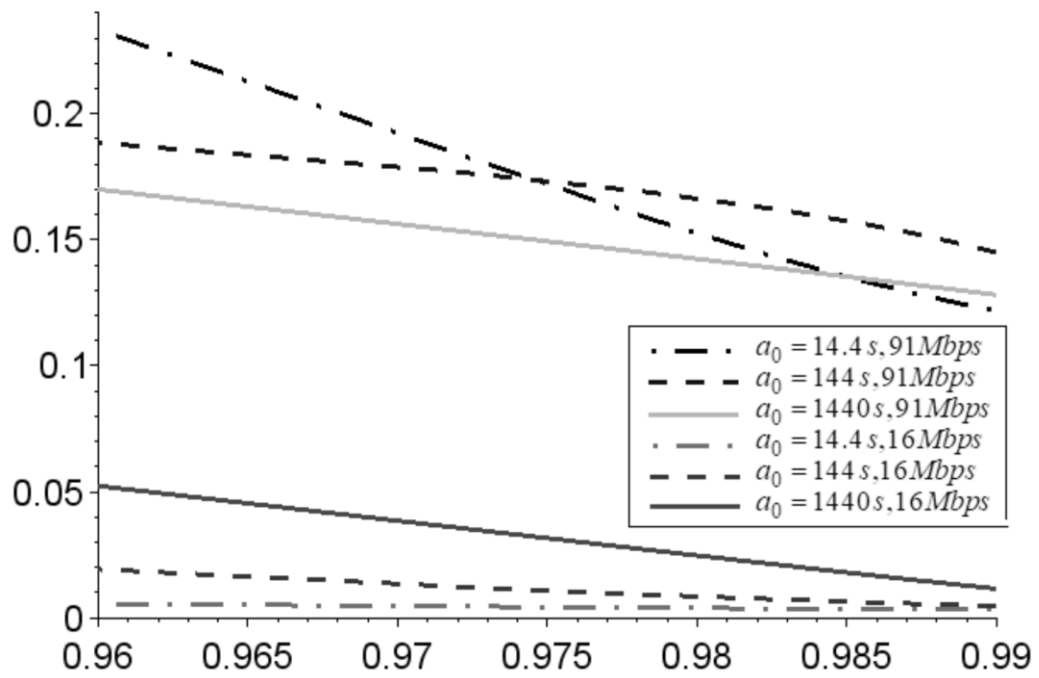


Рисунок 4.7 – Среднее время передачи эластичного трафика (с) в зависимости от коэффициента $1 - \alpha/\beta$ для разных среднего времени полной доступности диапазона радиочастот (с) и пропускной способности (Мбит/с)

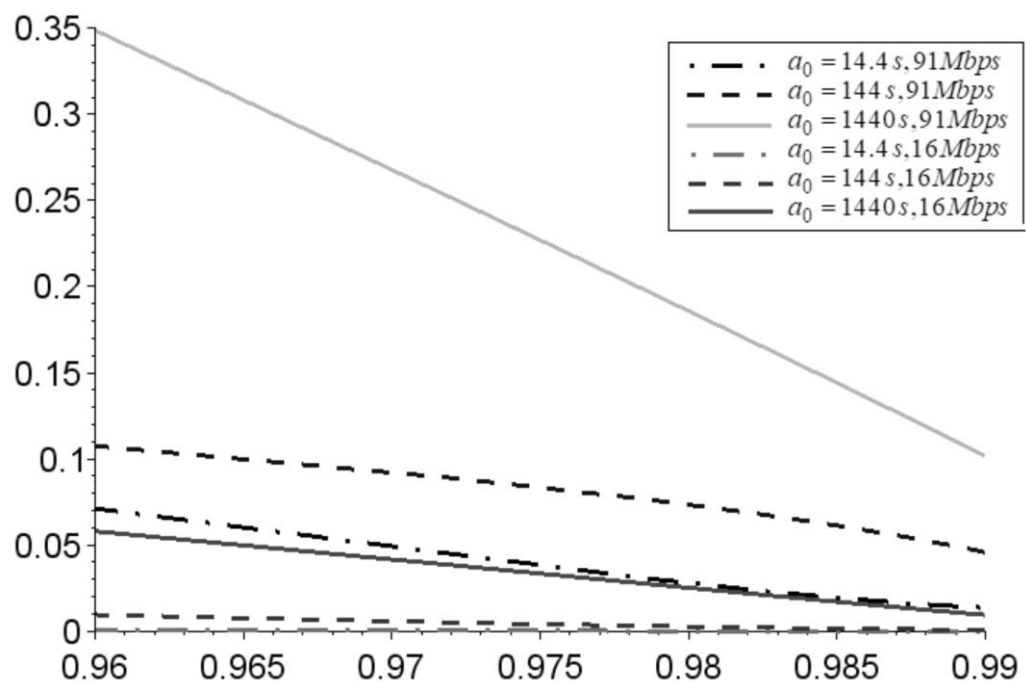


Рисунок 4.8 – Дисперсия времени передачи эластичного трафика (с) в зависимости от коэффициента $1 - \alpha/\beta$ для разных среднего времени полной доступности диапазона радиочастот (с) и пропускной способности (Мбит/с)

Глава 5 МОДЕЛИ С ПОТОКОВЫМ ТРАФИКОМ И ИЗМЕНЯЕМЫМ ОБЪЕМОМ РЕСУРСА

5.1 МОДЕЛЬ СОВМЕСТНОГО ИСПОЛЬЗОВАНИЯ РАДИОЧАСТОТ

5.1.1 Системная модель

В главе 5 исследованы модели приоритетного доступа потокового трафика в условиях совместного использования лицензируемых радиочастот. Приоритетными являются запросы владельца на занятие радиочастот, что делает их недоступными для пользователей оператора. Однако, когда владелец завершает доступ к частотам, оператор снова может их использовать.

Рассмотрим два сценария организации доступа к радиочастотам. В первом случае оператор имеет доступ только к диапазону, который используется совместно с владельцем, в то время как во втором случае у него есть возможность использовать дополнительно индивидуальный диапазон. Рассмотрим также случай, когда оператор самостоятельно управляет доступом к совместно используемому диапазону. Если число сессий трафика меньше порогового значения, доступ к диапазону выключен, а если больше – включен. Управление гистерезисное – первый порог N_1 отвечает за включение диапазона, а второй порог $N_2 < N_1$ – за его выключение.

Передача потокового трафика может допускать задержки или их не происходит. Система с полным отказом ресурса рассмотрена для двух случаев – когда возможны отказы на периоде простоя системы в отсутствии сессий, и когда отказы отсутствуют. Для обеспечения приоритета запросов владельца совместно используемого диапазона применяются различные механизмы доступа потокового трафика: прерывание и приостановка обслуживания, снижение скорости передачи и мощности сигнала.

В таблице 5.1 приведены основные обозначения, параметры и компоненты состояния системы. Для моделирования полного отказа ресурса используется СМО с ненадежными приборами, а для частичного – СМО с двумя группами приборов – надежными и ненадежными. Для моделирования механизма приостановки обслуживания используется очередь, а для снижения мощности сигнала – РеСМО с двумя типами ресурсов.

Таблица 5.1 – Обозначения, используемые для моделей с потоковым трафиком и изменяемым объемом ресурса (глава 5)

Обозначение	Параметр
λ	Интенсивность пуассоновского потока запросов на передачу трафика
b	Скорость передачи трафика
$1/\mu$	Среднее экспоненциальное время передачи трафика
$M_0 = \lfloor C/b \rfloor$	Максимальное число обслуживаемых сессий трафика на ненадежном ресурсе
$M_1 = M = \lfloor (V + C)/b \rfloor$	Максимальное число обслуживаемых сессий трафика на надежном и ненадежном ресурсе
$R > C$	Число мест в очереди для приостановленных и ожидающих начала обслуживания сессий трафика
N_1	Порог гистерезисного управления на включение ненадежного ресурса
$N_2 < N_1$	Порог гистерезисного управления на выключение ненадежного ресурса
p	Вероятность того, что после включения ненадежного ресурса он будет доступен
P_0	Максимальная мощность передающей антенны, когда ненадежный ресурс выключен
$P_1 > P_0$	Максимальная мощность передающей антенны, когда ненадежный ресурс включен
$PL_{\min}$	Минимальная мощность затухания сигнала
$PL_{\max}$	Максимальная мощность затухания сигнала
$d_{\min}$	Минимальное расстояние от устройства пользователя до базовой станции

Обозначение	Параметр
$d_{\max}$	Радиус зоны покрытия передающей антенны базовой станции
l	Число обслуживаемых сессий трафика на надежном ресурсе
m	Число обслуживаемых сессий трафика на ненадежном ресурсе
q	Число приостановленных сессий трафика
r	Число ожидающих начала обслуживания сессий трафика
$k = q + r$	Число сессий трафика в очереди – приостановленных и ожидающих начала обслуживания
$n = m + l + k$	Число сессий трафика в системе – обслуживаемых на ненадежном и надежном ресурсах и в очереди
u	Состояние ненадежного ресурса

5.1.2 Схемы приоритетного доступа

Рассмотрим следующие схемы приоритетного доступа для моделей с потоковым трафиком и полным отказом изменяемого объема ресурса:

- (а) приостановка обслуживания трафика без отказов ресурса на периоде простоя системы;
- (б) приостановка обслуживания трафика с отказами ресурса на периоде простоя системы;

и следующие схемы приоритетного доступа для моделей с потоковым трафиком и частичным отказом изменяемого объема ресурса:

- (в) приостановка обслуживания трафика;
- (г) приостановка обслуживания трафика и гистерезисный доступ;
- (д) снижение мощности и скорости передачи трафика;
- (е) постоянная мощность и прерывание обслуживания трафика.

В этих схемах используются механизмы $p \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\} = \mathcal{P}$ реализации приоритета и принципы $w \in \{2, 5, 6, 7\}$ выбора сессии для применения механизма из параграфа 1.6.3.

Опишем особенности исследуемых моделей с перечисленными схемами приоритетного доступа через матрицы приоритетов.

Схемы (а) и (б) – системы с полным отказом ресурса и приостановкой обслуживания потокового трафика. В случае отказа ресурса вся пропускная способность становится недоступной, и обслуживание трафика приостанавливается. Сессии отправляются в очередь для ожидания возобновления обслуживания. Ни одна сессия не будет прервана, так как в очереди зарезервировано C мест для приостановленных сессий. После восстановления ресурса все приостановленные сессии возобновляют свое обслуживание.

Таким образом, трафик может находиться в трех состояниях: обслуживания, ожидания возобновления обслуживания и ожидания начала обслуживания. При этом попытки возобновить обслуживание имеют относительный приоритет перед попытками начать обслуживание. Отметим, что из-за того, что система построена в экспоненциальных предположениях – входящий поток, поток отказов и восстановлений пуассоновские, обслуживание экспоненциальное, возобновление обслуживания можно также интерпретировать как начало обслуживания заново.

Рассмотрим два случая – схему (а) без отказов на периоде простоя системы и схему (б) с отказами на периоде простоя системы. Для схемы (а) – это СМО $M | M | C | R | F_{15}$ с матрицей приоритетов

				ресурс		потоковый трафик	
						обслуживается	ожидание возобновления обслуживания
ресурс		$\begin{matrix} p \\ e \end{matrix}$	12	13		2	9
	отказ	1	1	0		(1,2)	0
	восстановление	2	0	1		0	(1,2)
	поступление	3	0	0		0	0
	завершение обслуживания	4	0	0		0	0

для схемы (б) – это СМО $M | M | C | R | F_{16}$ с матрицей приоритетов, аналогичной схеме (а) $F_{16} = F_{15}$.

Схемы (в) и (г) – системы с частичным отказом ресурса, приостановкой обслуживания потокового трафика. Отличие от предыдущих моделей в том, что ресурс отказывает не полностью – пропускная способность V доступна всегда. Когда ненадежный ресурс отказывает, а он использовался для обслуживания сессий потокового трафика, эти сессии «стягиваются», т.е. перемещаются, на надежный ресурс. Если пропускной способности надежного ресурса недостаточно для всех таких сессий, то оставшиеся будут ожидать возобновления обслуживания в очереди.

Рассмотрим два случая – базовую описанную схему (в) и схему (г) с гистерезисным доступом. Для схемы (в) – это СМО $M | M | \langle V, C \rangle | R | F_{17}$ с матрицей приоритетов

F_{17}			ресурс		потоковый трафик		
					обслуживается	ожидание возобновления обслуживания	
		$\begin{matrix} p \\ e \end{matrix}$	10	11	14	2	9
ресурс	отказ	1	1	0	(1,6)	(2,2)	0
	восстановление	2	0	1	0	0	(1,2)
потоковый трафик	поступление	3	0	0	0	0	0
	завершение обслуживания	4	0	0	0	0	0

В схеме (г) ненадежная часть ресурса C помимо двух состояний – доступен и недоступен – может параллельно находиться в двух дополнительных состояниях – доступ включен или выключен. Этим процессом управляет оператор, учитывая общее число n сессий в системе. Управление доступом гистерезисное – с уменьшением числа сессий до $n < N_1$ доступ выключается, а с увеличением до $n > N_2 < N_1$ доступ включается.

Таким образом, ненадежный ресурс может находиться в четырех состояниях: включен и доступен, включен и недоступен, выключен и доступен, выключен и недоступен. С точки зрения передачи трафика, последние два случая схожи – в обоих обслуживание приостановлено. Поэтому объединим их в один – ресурс выключен, и при его включении с вероятностью p он окажется доступным, а с вероятностью $1 - p$ – недоступным. Это СМО $M | M | \langle V, C \rangle | \infty | F_{18}$ с матрицей приоритетов

F_{18}		$\begin{matrix} p \\ e \end{matrix}$	ресурс		поточковый трафик		
			10	11	обслуживается	ожидание возобновления обслуживания	
ресурс	выключение	1	1	0	(1,6)	(2,2)	0
	включение	2	0	(1,7)	0	0	(1,8)
	отказ	3	1	0	(1,6)	(2,2)	0
	восстановление	4	0	1	0	0	(1,2)
поточковый трафик	поступление	5	0	0	0	0	0
	завершение обслуживания	6	0	0	0	0	0

Схема (д). Для системы со снижением мощности и скорости передачи потокового трафика в качестве модели требования к ресурсу выберем политику снижения мощности сигнала RR (см. раздел 1.3). Эта политика предполагает равное деление кадра между всеми сессиями и без деления ресурса по частоте. Ненадежный ресурс доступен, когда максимальная мощность передающей антенны сигнала равна P_1 . В этом случае мощность сигнала для каждой сессии устанавливается таким образом, чтобы обеспечить необходимую скорость b передачи потокового трафика. Ненадежный ресурс недоступен, когда максимальная мощность передающей антенны сигнала становится равной $P_0 < P_1$. В этом случае мощность сигнала устанавливается таким образом, чтобы обеспечить равную скорость передачи потокового трафика для каждой сессии. Это СМО $M | M | V, \langle T, P_0, P_1 \rangle | 0 | F_{19}$ с матрицей приоритетов

F ₁₉		ресурс		поточковый трафик обслуживается						
		$\begin{matrix} p \\ e \end{matrix}$	7	8	$\begin{matrix} 7 \\ \Downarrow \\ 3 \end{matrix}$	$\begin{matrix} 8 \\ \Downarrow \\ 4 \end{matrix}$	$\begin{matrix} 5 \\ \Downarrow \\ 8 \end{matrix}$	$\begin{matrix} 5 \\ \Downarrow \\ 8, 3 \end{matrix}$	$\begin{matrix} 6 \\ \Downarrow \\ 7 \end{matrix}$	$\begin{matrix} 6 \\ \Downarrow \\ 7, 4 \end{matrix}$
ресурс	отказ	1	1	0	(1,2)	0	0	0		
	восстановление	2	0	1	0	(1,2)	0	0		
поточковый трафик	поступление	3	0	0	0	0	(1,2)	(2,2)		
	завершение обслуживания	4	0	0	0	0	0	0	(1,2)	(2,2)

Схема (е). Последняя система – с постоянной мощностью и прерыванием обслуживания потокового трафика. Применена политика постоянной мощности сигнала FP без деления ресурса по частоте. Мощность сигнала устанавливается на максимальном уровне P_1 или P_0 в зависимости от состояния ресурса. Доля кадра, выделяемая каждому пользователю, определяется таким образом, чтобы обеспечить необходимую скорость b передачи для каждой сессии. Это СМО $M \mid M \mid \langle V, C \rangle, \langle T, P_0, P_1 \rangle \mid 0 \mid F_{20}$ с матрицей приоритетов

F ₂₀		ресурс		поточковый трафик обслуживается			
		$\begin{matrix} p \\ \diagdown \\ e \end{matrix}$	7	8	$\begin{matrix} 7 \\ \downarrow \\ 6 \end{matrix}$	1	$\begin{matrix} 8 \\ \downarrow \\ 5 \end{matrix}$
ресурс	отказ	1	1	0	(1,2)	(2,5)	0
	восстановление	2	0	1	0	0	(1,2)
поточковый трафик	поступление	3	0	0	0	0	0
	завершение обслуживания	4	0	0	0	0	0

Автор исследовал и другие схемы приоритетного доступа для моделей с потоковым трафиком, полным и частичным отказом изменяемого объема ресурса. Однако чтобы сохранить относительную соразмерность глав диссертации, эти схемы не включены в текст работы:

- (ж) прерывание обслуживания трафика без стягивания на надежном ресурсе;
- (з) прерывание обслуживания трафика со стягиванием на надежном ресурсе;
- (и) снижение скорости передачи трафика с восстановлением ненадежного ресурса на периоде простоя;

- (к) снижение скорости передачи трафика без восстановления ненадежного ресурса на периоде простоя.

5.1.3 Модели математической теории телетрафика

Функционирование систем опишем при помощи СП $X(t)$ над пространствами состояний

модель (а)

$$\mathcal{X} = \{(0, k): 1 \leq k \leq R\} \cup \{(m, 0): 0 \leq m \leq C\} \cup \{(C, k): 1 \leq k \leq R - C\}; \quad (5.1)$$

модель (б)

$$\mathcal{X} = \{(n, u): 0 \leq n \leq R, u = 0, 1\}; \quad (5.2)$$

модель (в)

$$\begin{aligned} \mathcal{X} = & \{(l, 0, 0, 0): 0 \leq l \leq V\} \cup \{(V, 0, k, 0): 1 \leq k \leq R\} \cup \\ & \cup \{(l, m, 0, 1): 0 \leq l \leq V, 0 \leq m \leq C\} \cup \{(V, C, k, 1): 1 \leq k \leq R - C\}; \end{aligned} \quad (5.3)$$

модель (г)

$$\mathcal{X} = \{(n, 0): 0 \leq n \leq N_1 - 1\} \cup \{(n, u): n \geq N_2 + 1, u = 1, 2\}, \quad (5.4)$$

где $u = 0$ ресурс выключен, $u = 1$ ресурс включен и доступен, $u = 2$ ресурс включен и недоступен;

модель (д)

$$\mathcal{X} = \{(n, u): 0 \leq n \leq M, u = 0, 1\}; \quad (5.5)$$

модель (е)

$$\mathcal{X} = \{(n, u): 0 \leq n \leq M_u, u = 0, 1\}. \quad (5.6)$$

Отметим, что в моделях (а)–(г) скорость передачи трафика $b = 1$.

Ненулевые элементы матриц интенсивностей переходов имеют вид

модель (а)

$$\begin{aligned}
 Q[(0, k), (0, k+1)] &= \lambda, \quad k < R, \\
 Q[(m, 0), (m+1, 0)] &= \lambda, \quad m < C, \\
 Q[(C, k), (C, k+1)] &= \lambda, \quad 0 < k < R - C, \\
 Q[(m, 0), (m-1, 0)] &= m\mu, \quad m > 0, \\
 Q[(C, k), (C, k-1)] &= C\mu, \quad k > 0, \\
 Q[(m, 0), (0, m)] &= \alpha, \quad 0 < m \leq C, \\
 Q[(C, k), (0, C+k)] &= \alpha, \quad 0 < k \leq R - C, \\
 Q[(0, k), (k, 0)] &= \beta, \quad 0 < k \leq C, \\
 Q[(0, k), (C, k-C)] &= \beta, \quad C < k \leq R;
 \end{aligned} \tag{5.7}$$

модель (б)

$$\begin{aligned}
 Q[(n, u), (n+1, u)] &= \lambda, \quad n+1 \leq R, \\
 Q[(n, 1), (n-1, 1)] &= n\mu, \quad 0 < n \leq C, \\
 Q[(n, 1), (n-1, 1)] &= C\mu, \quad n > C, \\
 Q[(n, 1), (n, 0)] &= \alpha, \\
 Q[(n, 0), (n, 1)] &= \beta;
 \end{aligned} \tag{5.8}$$

модель (в)

$$\begin{aligned}
 Q[(l, m, 0, u), (l+1, m, 0, u)] &= \lambda, \quad l+1 \leq V, \\
 Q[(V, m, 0, 1), (V, m+1, 0, 1)] &= \lambda, \quad m+1 \leq C, \\
 Q[(V, 0, k, 0), (V, 0, k+1, 0)] &= \lambda, \quad k+1 \leq R, \\
 Q[(V, C, k, 1), (V, C, k+1, 1)] &= \lambda, \quad k+1 \leq R - C, \\
 Q[(l, m, 0, u), (l-1, m, 0, u)] &= l\mu, \quad 0 < l \leq V, \\
 Q[(l, m, 0, 1), (l, m-1, 0, 1)] &= m\mu, \quad 0 < m \leq C, \\
 Q[(V, 0, k, 0), (V, 0, k-1, 0)] &= V\mu, \quad k > 0, \\
 Q[(V, C, k, 1), (V, C, k-1, 1)] &= (V+C)\mu, \quad k > 0,
 \end{aligned} \tag{5.9}$$

$$\begin{aligned} Q[(l, m, 0, 1), (l + m, 0, 0, 0)] &= \alpha, \quad l + m \leq V, \\ Q[(l, m, k, 1), (V, 0, k + V - l, 0)] &= \alpha, \quad l + m > V, \\ Q[(l, 0, k, 0), (l, k, 0, 1)] &= \beta, \quad k \leq C, \\ Q[(V, 0, k, 0), (V, C, k - C, 1)] &= \beta, \quad k > C; \end{aligned}$$

модель (г)

$$\begin{aligned} Q[(n, u), (n + 1, u)] &= \lambda, \quad (n, u) \notin (N_1 - 1, 0), \\ Q[(N_1 - 1, 0), (N_1, 1)] &= \lambda p, \\ Q[(N_1 - 1, 0), (N_1, 2)] &= \lambda(1 - p), \\ Q[(n, 0), (n - 1, 0)] &= \min(n, V) \mu, \quad n > 0, \\ Q[(N_2 + 1, 1), (N_2, 0)] &= \min(N_2 + 1, V + C) \mu, \\ Q[(n, 1), (n - 1, 1)] &= \min(n, V + C) \mu, \quad n > N_2 + 1, \\ Q[(N_2 + 1, 2), (N_2, 0)] &= \min(N_2 + 1, V) \mu, \\ Q[(n, 2), (n - 1, 2)] &= \min(n, V) \mu, \quad n > N_2 + 1, \\ Q[(n, 1), (n, 2)] &= \alpha, \\ Q[(n, 2), (n, 1)] &= \beta; \end{aligned} \tag{5.10}$$

модель (д)

$$\begin{aligned} Q[(n, u), (n + 1, u)] &= \lambda \cdot p(n, u), \quad n + 1 \leq M, \quad u = 0, 1, \\ Q[(n, u), (n - 1, u)] &= n \mu, \quad n > 0, \quad u = 0, 1, \\ Q[(n, 1), (n, 0)] &= \alpha, \\ Q[(n, 0), (n, 1)] &= \beta; \end{aligned} \tag{5.11}$$

модель (е)

$$\begin{aligned} Q[(n, u), (n + 1, u)] &= \lambda \cdot p(n, u), \quad n + 1 \leq M_u, \quad u = 0, 1, \\ Q[(n, u), (n - 1, u)] &= n \mu, \quad n > 0, \quad u = 0, 1, \\ Q[(n, 1), (n, 0)] &= \alpha, \quad n \leq M_0, \\ Q[(n, 1), (M_0, 0)] &= \alpha, \quad n > M_0, \\ Q[(n, 0), (n, 1)] &= \beta. \end{aligned} \tag{5.12}$$

В следующих разделах 5.2–5.6 описаны результаты исследования моделей схем приоритетного доступа с эластичным трафиком, полным и частичным

отказом изменяемого объема ресурса. Для моделей (б), (в) и (г) матрицы интенсивностей переходов представлены в блочно-трехдиагональном виде (разделы 5.2 и 5.4). Для моделей (б) и (г) стационарное распределение рассчитывается с помощью матричного алгоритма (разделы 5.2 и 5.4).

Для модели (а) предложен алгоритм расчета распределения вероятностей числа ожидающих начала обслуживания сессий трафика и среднего числа приостановленных сессий (раздел 5.2). В качестве исходных данных для алгоритма используется стационарное распределение $\pi(m, k)$, где k суммарное число сессий в очереди, как приостановленных, так и ожидающих начала обслуживания. Алгоритм позволяет в скалярном виде рассчитать характеристику приоритетного доступа – среднее число приостановленных сессий – на базе двухмерного СП с состояниями (m, k) вместо того, чтобы использовать трехмерный СП с состояниями (m, q, r) .

Для модели (б) проведен анализ длительности периода занятости, который начинается в момент, когда поступивший запрос на передачу трафика застаёт систему в свободном состоянии, а заканчивается в момент, когда последняя из сессий трафика завершается (раздел 5.3). Получены матричные алгоритмы расчета ПЛС времени первого достижения пустого состояния СП и ПЛС длительности периода занятости. Проиллюстрирован вид среднего времени, в течение которого ресурс находится в состоянии отказа в цикле регенерации, и среднего времени, в течение которого ресурс доступен и полностью занят.

Для модели (г) поставлена задача выбора порогов гистерезисного доступа как политики управления (раздел 5.5). Получен вид средней стоимости использования ненадежного ресурса для произвольной фиксированной политики и сформулирована задача оптимизации. Для поиска оптимальной политики использован подход с применением машинного обучения – применен генетический алгоритм. Политика – тройка чисел: объем ненадежного ресурса и два порога гистерезисного доступа – представлены в виде битовой строки, состоящей из трех

соответствующих бинарных сегментов. Проиллюстрирован вид функции стоимости использования ненадежного ресурса.

Для ресурсных систем (д) и (е) найдены вероятности занятия ресурса и условные вероятности выделения ресурса (раздел 5.6). Приведены показатели эффективности приоритетного доступа: вероятности прерывания обслуживания и средней скорости передачи потокового трафика.

5.2 СИСТЕМА С ПОЛНЫМ ОТКАЗОМ РЕСУРСА

5.2.1 Матричный алгоритм расчета стационарного распределения

Лемма 5.1. Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и отказами ресурса на периоде простоя системы, если на пространстве $\mathcal{X}$ состояний СП $X(t)$ введен лексикографический порядок

$$\begin{aligned} \mathbf{x} = (n, u) < (n', u') = \mathbf{x}' &\Leftrightarrow \\ (n < n') \vee & \\ \vee (n < n', u > u'), & \quad \mathbf{x}, \mathbf{x}' \in \mathcal{X}, \end{aligned} \quad (5.13)$$

тогда матрица интенсивностей переходов СП $X(t)$ представима в блочно-трехдиагональном виде

$$\begin{aligned} \mathbf{Q} = \text{diag}(\mathbf{Q}_{00}, \mathbf{Q}_{01}, \dots, \mathbf{Q}_{0C}, \mathbf{Q}_{0C}, \dots, \mathbf{Q}_{0C}, \mathbf{Q}_{0,C+1}) + \\ + \text{diag}^+(\mathbf{Q}_{11}, \dots, \mathbf{Q}_{11}) + \text{diag}^-(\mathbf{Q}_{21}, \mathbf{Q}_{22}, \dots, \mathbf{Q}_{2C}, \mathbf{Q}_{2C}, \dots, \mathbf{Q}_{2C}), \end{aligned} \quad (5.14)$$

где блоки определяются по формулам

$$\mathbf{Q}_{00} = \begin{bmatrix} -(\lambda + \alpha) & \alpha \\ \beta & -(\lambda + \beta) \end{bmatrix}, \quad (5.15)$$

$$\mathbf{Q}_{0n} = \begin{bmatrix} -(\lambda + n\mu + \alpha) & \alpha \\ \beta & -(\lambda + \beta) \end{bmatrix}, \quad n=1, \dots, C,$$

$$\mathbf{Q}_{0,C+1} = \begin{bmatrix} -(C\mu + \alpha) & \alpha \\ \beta & -\beta \end{bmatrix},$$

$$\mathbf{Q}_{11} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}, \tag{5.16}$$

$$\mathbf{Q}_{2n} = \begin{bmatrix} n\mu & 0 \\ 0 & 0 \end{bmatrix}, \quad n=1, \dots, C. \tag{5.17}$$

Доказательство.

а) Введение лексикографического порядка

Разобьем пространство $\mathcal{X}$ состояний (5.2) СП $\mathbf{X}(t)$ на подмножества по числу n сессий трафика в системе

$$\mathcal{X} = \bigcup_{n=0}^R \mathcal{X}(n),$$

$$\mathcal{X}(n) = \{(i, u) \in \mathcal{X} : i = n\} = \{(n, u) \in \mathcal{X}\} = \{(n, 1), (n, 0)\}.$$

Введем лексикографический порядок на множестве $\mathcal{X}$ в порядке объединения $\mathcal{X}(n)$

$$\mathcal{X} = \{(0, 1), (1, 0),$$

$$(1, 1), (1, 0),$$

$$\dots$$

$$(n, 1), (n, 0),$$

$$\dots$$

$$(R, 1), (R, 0)\}.$$

б) Описание блоков матрицы

С учетом введенного лексикографического порядка и ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (5.8), матрица $\mathbf{Q}$ имеет блочно-трехдиагональный вид

$$Q = \begin{bmatrix} Q_{00} & Q_{11} & & & & & \\ Q_{21} & Q_{01} & Q_{11} & & & & \\ & Q_{22} & \ddots & \ddots & & & \\ & & \ddots & \ddots & Q_{11} & & \\ & & & Q_{2C} & Q_{0C} & Q_{11} & \\ & & & & Q_{2C} & Q_{0C} & \ddots \\ & & & & & \ddots & \ddots & Q_{11} \\ & & & & & & Q_{2C} & Q_{0C} & Q_{11} \\ & & & & & & & Q_{2C} & Q_{0,C+1} \end{bmatrix}.$$

■

Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и отказами ресурса на периоде простоя системы, в соответствии с лексикографическим порядком (5.13), представим стационарное распределение СП $X(t)$ в виде

$$\begin{aligned} \pi &= (\pi_0, \pi_1, \dots, \pi_n, \dots, \pi_R), \\ \pi_n &= (\pi(n, 1), \pi(n, 0)), \quad n = 0, \dots, R. \end{aligned} \quad (5.18)$$

Тогда с учетом леммы 5.1, справедлива следующая теорема.

Теорема 5.1. Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и отказами ресурса на периоде простоя системы, стационарное распределение СП $X(t)$ рассчитывается в матричном виде по формуле

$$\pi_n = \pi_R \prod_{i=1}^{R-n} A_{R-i}, \quad n = 0, \dots, R-1, \quad (5.19)$$

где π_R является единственным решением системы уравнений

$$\begin{aligned} \pi_R (A_{R-1} Q_{11} + Q_{0,C+1}) &= 0, \\ \pi_R \left(\sum_{n=0}^R \prod_{i=1}^{R-n} A_{R-i} \mathbf{1}^T \right) &= 1, \end{aligned} \quad (5.20)$$

а матрицы A_n вычисляются по рекуррентным соотношениям

$$\begin{aligned} A_0 &= -Q_{21}Q_{00}^{-1}, \\ A_n &= -Q_{2,n+1}(A_{n-1}Q_{11} + Q_{0n})^{-1}, \quad n=1, \dots, C-1, \\ A_n &= -Q_{2C}(A_{n-1}Q_{11} + Q_{0C})^{-1}, \quad n=C, \dots, R-1. \end{aligned} \quad (5.21)$$

Доказательство.

Доказательство можно провести аналогично теореме 2.1, заменив $M \rightarrow R$ и матрицы

$$\begin{aligned} Q_{0n} &\rightarrow Q_{0C}, \quad n=C, \dots, R, \\ Q_{1n} &\rightarrow Q_{11}, \quad n=1, \dots, R, \\ Q_{2n} &\rightarrow Q_{2C}, \quad n=C, \dots, R. \end{aligned}$$

■

5.2.2 Алгоритм расчета среднего числа приостановленных сессий

Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика без отказов ресурса на периоде простоя системы, справедлива следующая теорема.

Теорема 5.2. Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика без отказов ресурса на периоде простоя системы, распределение вероятностей числа ожидающих начала обслуживания сессий трафика рассчитывается по формуле

$$p_r = \sum_{(m,k) \in \mathcal{X}} A_r(m,k) \cdot \pi(m,k), \quad r=0, \dots, R-C, \quad (5.22)$$

где вероятности $A_r(m,k)$ того, что число ожидающих начала обслуживания сессий равно r при условии того, что система находится в состоянии (m,k) , вычисляются по формулам

$$\begin{aligned}
 A_r(0, k) &= B_0(k-r) \cdot \prod_{i=k-r+1}^k B_1(i), \quad k=2, \dots, C, \quad r=0, \dots, k-1, \\
 A_{k-C}(0, k) &= \sum_{i=0}^{k-C} B_0(C+i) \cdot \prod_{i=C+i+1}^k B_1(i), \quad k=C+1, \dots, R-C,
 \end{aligned} \tag{5.23}$$

$$A_r(0, k) = B_0(k-r) \cdot \prod_{i=k-r+1}^k B_1(i), \quad k=C+1, \dots, R-C, \quad r=k-C+1, \dots, k-1,$$

а вероятности $B_0(n)$ перехода в состояние $(0, n)$ из-за отказа ресурса и $B_1(n)$ перехода в состояние $(0, n)$ из-за поступления запроса на передачу трафика вычисляются соответственно по формулам

$$B_0(1) = 0, \quad B_0(n) = \frac{\alpha\pi(n, 0)}{\alpha\pi(n, 0) + \lambda\pi(0, n-1)}, \quad n=1, \dots, C, \tag{5.24}$$

$$B_0(n) = \frac{\alpha\pi(C, n-C)}{\alpha\pi(C, n-C) + \lambda\pi(0, n-1)}, \quad n=C+1, \dots, R,$$

$$B_1(1) = 1,$$

$$B_1(n) = \frac{\lambda\pi(0, n-1)}{\alpha\pi(n, 0) + \lambda\pi(0, n-1)}, \quad n=2, \dots, C, \tag{5.25}$$

$$B_1(n) = \frac{\lambda\pi(0, n-1)}{\alpha\pi(C, n-C) + \lambda\pi(0, n-1)}, \quad n=C+1, \dots, R.$$

Доказательство.

а) Выражение p_r через $A_r(m, k)$

По формуле полной вероятности, вероятность p_r того, что число ожидающих начала обслуживания сессий равно r , запишем в виде

$$p_r = \Pr(r) = \sum_{(m, k) \in \mathcal{X}} \Pr(r \mid (m, k)) \cdot \Pr((m, k)).$$

Обозначим вероятность $A_r(m, k) = \Pr(r \mid (m, k))$ того, что число ожидающих начала обслуживания сессий равно r при условии того, что система находится в состоянии (m, k) , тогда

$$p_r = \sum_{(m, k) \in \mathcal{X}} A_r(m, k) \cdot \pi(m, k), \quad r=0, \dots, R-C.$$

Определим область значений $A_r(m, k)$

$$A_0(0, 1) = 1,$$

$$A_r(0, 1) = 0, \quad r = 1, \dots, R - C,$$

$$A_r(0, k) > 0, \quad k = 2, \dots, C, \quad r = 0, \dots, k - 1,$$

$$A_r(0, k) = 0, \quad k = 2, \dots, C, \quad r = k, \dots, R - C,$$

$$A_r(0, k) > 0, \quad k = C + 1, \dots, R - C, \quad r = k - C, \dots, k - 1,$$

$$A_r(0, k) = 0, \quad k = C + 1, \dots, R - C, \quad r = 0, \dots, k - C - 1, k, \dots, R - C,$$

$$A_0(m, 0) = 1, \quad m = 0, \dots, C,$$

$$A_r(m, 0) = 0, \quad m = 0, \dots, C, \quad r = 1, \dots, R - C,$$

$$A_k(C, k) = 1, \quad k = 1, \dots, R - C,$$

$$A_r(C, k) = 0, \quad k = 1, \dots, R - C, \quad r = 1, \dots, k - 1, k + 1, \dots, R - C.$$

б) Выражение $A_r(m, k)$ через $B_0(n)$ и $B_1(n)$

Проанализируем вероятности $A_r(m, k)$, отличные от 0 и 1, а именно три случая

$$A_r(0, k) > 0, \quad k = 2, \dots, C, \quad r = 0, \dots, k - 1,$$

$$A_r(0, k) > 0, \quad k = C + 1, \dots, R - C, \quad r = k - C,$$

$$A_r(0, k) > 0, \quad k = C + 1, \dots, R - C, \quad r = k - C + 1, \dots, k - 1,$$

с точки зрения возможных траекторий СП $X(t)$, которые заканчиваются в состоянии $(m, k) = (0, k)$ с числом ожидающих начала обслуживания сессий r , а начинаются в состояниях с однозначной интерпретацией, сколько из $k' = q' + r'$ сессий в очереди приостановлено q' , а сколько ожидает начала обслуживания r' .

Первое состояние в минимальной по длине траектории должно быть такое, когда ресурс доступен, поскольку только в этом случае все сессии в очереди ожидают начала обслуживания $k' = r'$. Первое событие, приводящее к переходу во второе состояние будет отказ ресурса. Длина траектории будет определяться числом событий – поступлений запросов на передачу трафика, – таким, чтобы в состоянии $(m, k) = (0, k)$ число ожидающих начала обслуживания сессий стало r .

Таким образом, вероятности $A_r(0, k)$ можно интерпретировать как вероятности соответствующих траекторий СП. Найдем их для каждого из трех указанных выше случаев

$$\begin{aligned} A_r(0, k) &= \Pr(r \mid (0, k)) = \\ &= \Pr((k-r, 0) \rightarrow (0, k-r) \rightarrow (0, k-r+1) \rightarrow \dots \rightarrow (0, k-1) \rightarrow (0, k)) = \\ &= \Pr((k-r, 0) \rightarrow (0, k-r)) \cdot \prod_{i=k-r+1}^k \Pr((0, i-1) \rightarrow (0, i)), \\ k &= 2, \dots, C, \quad r = 0, \dots, k-1, \end{aligned}$$

$$\begin{aligned} A_{k-C}(0, k) &= \Pr(r = k-C \mid (0, k)) = \\ &= \sum_{i=0}^{k-C} \Pr((C, i) \rightarrow (0, C+i) \rightarrow (0, C+i+1) \rightarrow \dots \rightarrow (0, k-1) \rightarrow (0, k)) = \\ &= \sum_{i=0}^{k-C} \Pr((C, i) \rightarrow (0, C+i)) \cdot \prod_{i=C+i+1}^k \Pr((0, i-1) \rightarrow (0, i)), \\ k &= C+1, \dots, R-C, \end{aligned}$$

$$\begin{aligned} A_r(0, k) &= \Pr(r \mid (0, k)) = \\ &= \Pr((k-r, 0) \rightarrow (0, k-r) \rightarrow (0, k-r+1) \rightarrow \dots \rightarrow (0, k-1) \rightarrow (0, k)) = \\ &= \Pr((k-r, 0) \rightarrow (0, k-r)) \cdot \prod_{i=k-r+1}^k \Pr((0, i-1) \rightarrow (0, i)), \\ k &= C+1, \dots, R-C, \quad r = k-C+1, \dots, k-1. \end{aligned}$$

Обозначим

$$\begin{aligned} B_0(n) &= \Pr((n, 0) \rightarrow (0, n)), \quad n = 1, \dots, C, \\ B_0(n) &= \Pr((C, n-C) \rightarrow (0, n)), \quad n = C+1, \dots, R, \\ B_1(n) &= \Pr((0, n-1) \rightarrow (0, n)), \quad n = 1, \dots, R, \end{aligned}$$

при этом $B_0(n)$ можно интерпретировать как вероятность перехода в состояние $(0, n)$ из-за отказа ресурса, а $B_1(n)$ как вероятность перехода в состояние $(0, n)$ из-за поступления запроса на передачу трафика, причем $B_0(n) + B_1(n) = 1$. Тогда

$$A_r(0, k) = B_0(k - r) \cdot \prod_{i=k-r+1}^k B_1(i), \quad k = 2, \dots, C, \quad r = 0, \dots, k - 1,$$

$$A_{k-C}(0, k) = \sum_{i=0}^{k-C} B_0(C + i) \cdot \prod_{i=C+i+1}^k B_1(i), \quad k = C + 1, \dots, R - C,$$

$$A_r(0, k) = B_0(k - r) \cdot \prod_{i=k-r+1}^k B_1(i), \quad k = C + 1, \dots, R - C, \quad r = k - C + 1, \dots, k - 1.$$

в) Выражение $B_0(n)$ и $B_1(n)$ через $\pi(m, k)$

Найдем вероятности $B_0(n)$ и $B_1(n)$. Система построена в экспоненциальных предположениях – входящий поток, поток отказов и восстановлений пуассоновские, обслуживание экспоненциальное, поэтому $B_0(n)$ можно интерпретировать как вероятность того, что минимум независимых экспоненциально распределенных СВ – времени до поступления запроса на передачу трафика и времени до отказа ресурса – будет отказ ресурса, а $B_1(n)$ – вероятность дополнительного события. Согласно соответствующему свойству экспоненциального распределения и с учетом стационарных вероятностей, получим

$$B_0(1) = \Pr(\emptyset) = 0,$$

$$B_0(n) = \Pr((n, 0) \rightarrow (0, n)) = \frac{\alpha\pi(n, 0)}{\alpha\pi(n, 0) + \lambda\pi(0, n - 1)}, \quad n = 1, \dots, C,$$

$$B_0(n) = \Pr((C, n - C) \rightarrow (0, n)) = \frac{\alpha\pi(C, n - C)}{\alpha\pi(C, n - C) + \lambda\pi(0, n - 1)}, \quad n = C + 1, \dots, R,$$

$$B_1(1) = \Pr((0, 0) \rightarrow (0, 1)) = 1,$$

$$B_1(n) = \Pr((0, n - 1) \rightarrow (0, n)) = \frac{\lambda\pi(0, n - 1)}{\alpha\pi(n, 0) + \lambda\pi(0, n - 1)}, \quad n = 2, \dots, C,$$

$$B_1(n) = \Pr((0, n - 1) \rightarrow (0, n)) = \frac{\lambda\pi(0, n - 1)}{\alpha\pi(C, n - C) + \lambda\pi(0, n - 1)}, \quad n = C + 1, \dots, R.$$

■

5.2.3 Анализ среднего числа приостановленных сессий

Из теоремы 5.2 вытекает следующее следствие.

Следствие 5.1. Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика без отказов ресурса на периоде простоя системы, среднее число приостановленных сессий потокового трафика рассчитывается по формуле

$$\bar{q} = \sum_{(m,k) \in \mathcal{X}} k \cdot \pi(m,k) - \sum_{r=0}^{R-C} r \cdot p_r. \quad (5.26)$$

Проиллюстрируем поведение показателей эффективности обслуживания потокового трафика на примере сценария радиотелеметрии с самолетами. Радиотелеметрия играет ключевую роль в безопасности и эффективности авиаперевозок. Она передает данные о скорости, высоте, состоянии двигателей и навигации в реальном времени с самолета на землю. Система LSA позволяет мобильным операторам временно использовать частоты радиотелеметрии, когда они не заняты авиацией. Это расширяет пропускную способность сети в районах с высокой нагрузкой вокруг аэропортов. При этом приоритет всегда остается за авиацией: при необходимости радиотелеметрии оператор должен освободить LSA-диапазон. Такая модель выгодна для всех: операторы получают дополнительный спектр, а владельцы частот – доход. Однако для успешной реализации требуется точная координация, динамическое управление ресурсами и надежные алгоритмы переключения.

На рисунке 5.1 представлены графики среднего числа приостановленных сессий потокового трафика в зависимости от предложенной нагрузки для разного среднего времени доступности диапазона радиочастот. С увеличением нагрузки растет и среднее число приостановленных сессий потокового трафика. Это происходит из-за ограниченной пропускной способности системы и частой потребности использования частот для радиотелеметрии. Чем меньше среднее время доступности диапазона для оператора, тем выше число приостановленных сессий.

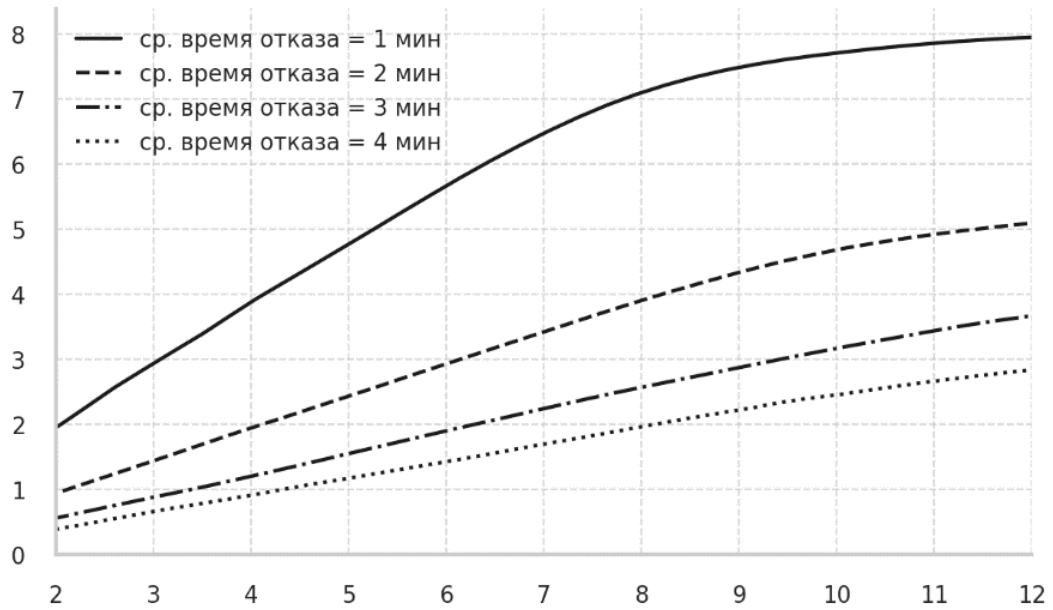


Рисунок 5.1 – Среднее число приостановленных сессий потокового трафика от предложенной нагрузки (Эрл) для разного среднего времени доступности диапазона радиочастот (мин)

5.3 АНАЛИЗ ПЕРИОДА ЗАНЯТОСТИ СИСТЕМЫ

5.3.1 Предварительные обозначения

Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и отказами ресурса на периоде простоя системы, рассмотрим период занятости. Обозначим ξ_b СВ длительности периода занятости – времени от момента поступления в свободную систему и принятия запроса на передачу трафика до момента полного освобождения системы. Период занятости начинается в момент, когда поступивший запрос на передачу трафика застаёт систему в состоянии $x \in \mathcal{X}(0)$, мгновенно принимается, и система переходит в состояние $x \in \mathcal{X}(1)$, а заканчивается в момент, когда последняя из сессий трафика завершается, и система переходит в состояние $x \in \mathcal{X}(0)$.

Пусть $F_{\xi_b}(y) = \Pr(\xi_b \leq y)$ ФР СВ ξ_b , а $f_{\xi_b}(y) = dF_{\xi_b}(y)/dy$ плотность распределения вероятностей СВ ξ_b . Тогда преобразование Лапласа-Стилтьеса (ПЛС) СВ ξ_b имеет вид

$$\varphi(s) = \left\{ \mathcal{L}^* F_{\xi_b} \right\}(s) = E \left[\exp(-s \cdot \xi_b) \right] = \int_0^{\infty} \exp(-sy) dF_{\xi_b}(y) = \int_0^{\infty} \exp(-sy) f_{\xi_b}(y) dy.$$

Обозначим $\xi_b(x) = \xi_b(n, u)$ СВ времени первого достижения пустого состояния из состояния x . Этот интервал времени начинается в момент, когда система находится в состоянии $x = (n, u)$, а заканчивается в момент, когда последняя из сессий трафика завершается, и система переходит в состояние $x \in \{(0, 1), (0, 0)\}$.

Пусть $F_{\xi_b(x)}(y) = \Pr(\xi_b(x) \leq y)$ ФР СВ $\xi_b(x)$, а $f_{\xi_b(x)}(y) = dF_{\xi_b(x)}(y)/dy$ плотность распределения вероятностей СВ $\xi_b(x)$. Тогда ПЛС СВ $\xi_b(x)$ имеет вид

$$\begin{aligned} \varphi_x(s) &= \left\{ \mathcal{L}^* F_{\xi_b(x)} \right\}(s) = E \left[\exp(-s \cdot \xi_b(x)) \right] = \int_0^{\infty} \exp(-sy) dF_{\xi_b(x)}(y) = \\ &= \int_0^{\infty} \exp(-sy) f_{\xi_b(x)}(y) dy. \end{aligned}$$

Пусть также $\varphi_n(s) = (\varphi_{(n,1)}(s), \varphi_{(n,0)}(s))$ ПЛС времени первого достижения пустого состояния из состояния, когда в системе n сессий трафика.

Обозначим ξ_{id} СВ длительности периода простоя, т.е. времени от момента полного освобождения системы до момента поступления и принятия запроса на передачу трафика в свободную систему. Длительность периода простоя распределена экспоненциально $\xi_{id} \sim \text{Exp}(\lambda)$. Тогда $\xi_c = \xi_b + \xi_{id}$ СВ длительности цикла регенерации.

Следствие 5.2. Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и отказами ресурса на периоде простоя системы, следующие показатели эффективности рассчитываются по формулам:

среднее время, в течение которого ресурс находится в состоянии отказа в цикле регенерации

$$E[T_0] = \frac{1}{\lambda \pi_0 \mathbf{1}^T} \sum_{n=0}^{R-1} \pi_n e_2^T; \quad (5.27)$$

среднее время, в течение которого ресурс доступен и полностью занят

$$E[T_1] = \frac{1}{\lambda \pi_0 \mathbf{1}^T} \sum_{n=C}^{R-1} \pi_n e_1^T. \quad (5.28)$$

5.3.2 Преобразование Лапласа-Стилтьеса времени первого достижения пустого состояния

Лемма 5.3. Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и отказами ресурса на периоде простоя системы, ПЛС времени первого достижения пустого состояния СП $X(t)$ рассчитывается в матричном виде по рекуррентным формулам

$$\begin{aligned} \varphi_0^T(s) &= e, \\ \varphi_n^T(s) &= X_n(s) \varphi_{n+1}^T(s) + Y_n(s), \quad n=1, \dots, R-1, \\ \varphi_R^T(s) &= Y_R(s), \end{aligned} \quad (5.29)$$

где матрицы $X_n(s)$ и $Y_n(s)$ вычисляются по рекуррентным соотношениям

$$\begin{aligned} X_1(s) &= -(\mathbf{Q}_{01} - s\mathbf{I})^{-1} \mathbf{Q}_{11}, \\ X_n(s) &= -[\mathbf{Q}_{0n} - s\mathbf{I} + \mathbf{Q}_{2n} X_{n-1}(s)]^{-1} \mathbf{Q}_{11}, \quad n=2, \dots, C, \\ X_n(s) &= -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C} X_{n-1}(s)]^{-1} \mathbf{Q}_{11}, \quad n=C+1, \dots, R-1, \end{aligned} \quad (5.30)$$

$$\begin{aligned} Y_1(s) &= -(\mathbf{Q}_{01} - s\mathbf{I})^{-1} \mathbf{Q}_{21} \mathbf{1}^T, \\ Y_n(s) &= -[\mathbf{Q}_{0n} - s\mathbf{I} + \mathbf{Q}_{2n} X_{n-1}(s)]^{-1} \mathbf{Q}_{2n} Y_{n-1}(s), \quad n=2, \dots, C, \\ Y_n(s) &= -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C} X_{n-1}(s)]^{-1} \mathbf{Q}_{2C} Y_{n-1}(s), \quad n=C+1, \dots, R-1, \\ Y_R(s) &= -[\mathbf{Q}_{0,C+1} - s\mathbf{I} + \mathbf{Q}_{2C} X_{R-1}(s)]^{-1} \mathbf{Q}_{2C} Y_{R-1}(s). \end{aligned} \quad (5.31)$$

Доказательство.

а) Запись системы уравнений относительно $\varphi_{(n,u)}(s)$

Воспользуемся подходом, который на английском языке звучит как «first step analysis», и составим уравнения для ПЛС $\varphi_{(n,u)}(s)$

$$\begin{aligned}(s + \lambda + n\mu + \alpha)\varphi_{(n,1)}(s) &= \lambda\varphi_{(n+1,1)}(s) + n\mu\varphi_{(n-1,1)}(s) + \alpha\varphi_{(n,0)}(s), \quad n = 1, \dots, C, \\(s + \lambda + C\mu + \alpha)\varphi_{(n,1)}(s) &= \lambda\varphi_{(n+1,1)}(s) + C\mu\varphi_{(n-1,1)}(s) + \alpha\varphi_{(n,0)}(s), \quad n = C + 1, \dots, R - 1, \\(s + C\mu + \alpha)\varphi_{(R,1)}(s) &= C\mu\varphi_{(R-1,1)}(s) + \alpha\varphi_{(R,0)}(s), \\(s + \lambda + \beta)\varphi_{(n,0)}(s) &= \lambda\varphi_{(n+1,0)}(s) + \beta\varphi_{(n,1)}(s), \quad n = 1, \dots, R - 1, \\(s + \beta)\varphi_{(R,0)}(s) &= \beta\varphi_{(R,1)}(s).\end{aligned}$$

б) Получение системы уравнений относительно $\boldsymbol{\varphi}_n^T(s)$ в матричном виде

Сгруппируем уравнения по числу сессий

$$\begin{aligned}(s + \lambda + \mu + \alpha)\varphi_{(1,1)}(s) &= \lambda\varphi_{(2,1)}(s) + \mu + \alpha\varphi_{(1,0)}(s), \\(s + \lambda + \beta)\varphi_{(1,0)}(s) &= \lambda\varphi_{(2,0)}(s) + \beta\varphi_{(1,1)}(s), \\(s + \lambda + n\mu + \alpha)\varphi_{(n,1)}(s) &= \lambda\varphi_{(n+1,1)}(s) + n\mu\varphi_{(n-1,1)}(s) + \alpha\varphi_{(n,0)}(s), \quad n = 2, \dots, C, \\(s + \lambda + \beta)\varphi_{(n,0)}(s) &= \lambda\varphi_{(n+1,0)}(s) + \beta\varphi_{(n,1)}(s), \quad n = 2, \dots, C, \\(s + \lambda + C\mu + \alpha)\varphi_{(n,1)}(s) &= \lambda\varphi_{(n+1,1)}(s) + C\mu\varphi_{(n-1,1)}(s) + \alpha\varphi_{(n,0)}(s), \quad n = C + 1, \dots, R - 1, \\(s + \lambda + \beta)\varphi_{(n,0)}(s) &= \lambda\varphi_{(n+1,0)}(s) + \beta\varphi_{(n,1)}(s), \quad n = C + 1, \dots, R - 1, \\(s + C\mu + \alpha)\varphi_{(R,1)}(s) &= C\mu\varphi_{(R-1,1)}(s) + \alpha\varphi_{(R,0)}(s), \\(s + \beta)\varphi_{(R,0)}(s) &= \beta\varphi_{(R,1)}(s)\end{aligned}$$

и запишем получившуюся систему уравнений следующим образом

$$\begin{aligned}(s + \lambda + \mu + \alpha)\varphi_{(1,1)}(s) - \alpha\varphi_{(1,0)}(s) &= \lambda\varphi_{(2,1)}(s) + \mu, \\-\beta\varphi_{(1,1)}(s) + (s + \lambda + \beta)\varphi_{(1,0)}(s) &= \lambda\varphi_{(2,0)}(s), \\(s + \lambda + n\mu + \alpha)\varphi_{(n,1)}(s) - \alpha\varphi_{(n,0)}(s) &= \lambda\varphi_{(n+1,1)}(s) + n\mu\varphi_{(n-1,1)}(s), \quad n = 2, \dots, C, \\-\beta\varphi_{(n,1)}(s) + (s + \lambda + \beta)\varphi_{(n,0)}(s) &= \lambda\varphi_{(n+1,0)}(s), \quad n = 2, \dots, C, \\(s + \lambda + C\mu + \alpha)\varphi_{(n,1)}(s) - \alpha\varphi_{(n,0)}(s) &= \lambda\varphi_{(n+1,1)}(s) + C\mu\varphi_{(n-1,1)}(s), \quad n = C + 1, \dots, R - 1, \\-\beta\varphi_{(n,1)}(s) + (s + \lambda + \beta)\varphi_{(n,0)}(s) &= \lambda\varphi_{(n+1,0)}(s), \quad n = C + 1, \dots, R - 1,\end{aligned}$$

$$(s + C\mu + \alpha)\varphi_{(R,1)}(s) - \alpha\varphi_{(R,0)}(s) = C\mu\varphi_{(R-1,1)}(s),$$

$$-\beta\varphi_{(R,1)}(s) + (s + \beta)\varphi_{(R,0)}(s) = 0.$$

Тогда с учетом вида матриц в лемме 5.1, получим

$$(-\mathbf{Q}_{01} + s\mathbf{I})\boldsymbol{\varphi}_1^T(s) = \mathbf{Q}_{11}\boldsymbol{\varphi}_2^T(s) + \mathbf{Q}_{21}\mathbf{1}^T,$$

$$(-\mathbf{Q}_{0n} + s\mathbf{I})\boldsymbol{\varphi}_n^T(s) = \mathbf{Q}_{11}\boldsymbol{\varphi}_{n+1}^T(s) + \mathbf{Q}_{2n}\boldsymbol{\varphi}_{n-1}^T(s), \quad n = 2, \dots, C,$$

$$(-\mathbf{Q}_{0C} + s\mathbf{I})\boldsymbol{\varphi}_n^T(s) = \mathbf{Q}_{11}\boldsymbol{\varphi}_{n+1}^T(s) + \mathbf{Q}_{2C}\boldsymbol{\varphi}_{n-1}^T(s), \quad n = C + 1, \dots, R - 1,$$

$$(-\mathbf{Q}_{0,C+1} + s\mathbf{I})\boldsymbol{\varphi}_R^T(s) = \mathbf{Q}_{2C}\boldsymbol{\varphi}_{R-1}^T(s).$$

в) Выражение $\boldsymbol{\varphi}_n^T(s)$ через $\boldsymbol{\varphi}_{n+1}^T(s)$ для $n \leq C$

Выразим $\boldsymbol{\varphi}_1^T(s)$ через $\boldsymbol{\varphi}_2^T(s)$ из матричного уравнения при $n = 1$

$$\boldsymbol{\varphi}_1^T(s) = -(\mathbf{Q}_{01} - s\mathbf{I})^{-1}\mathbf{Q}_{11}\boldsymbol{\varphi}_2^T(s) - (\mathbf{Q}_{01} - s\mathbf{I})^{-1}\mathbf{Q}_{21}\mathbf{1}^T,$$

и обозначим $\mathbf{X}_1(s) = -(\mathbf{Q}_{01} - s\mathbf{I})^{-1}\mathbf{Q}_{11}$, $\mathbf{Y}_1(s) = -(\mathbf{Q}_{01} - s\mathbf{I})^{-1}\mathbf{Q}_{21}\mathbf{1}^T$, тогда

$$\boldsymbol{\varphi}_1^T(s) = \mathbf{X}_1(s)\boldsymbol{\varphi}_2^T(s) + \mathbf{Y}_1(s).$$

Получим зависимость $\boldsymbol{\varphi}_n^T(s)$ от $\boldsymbol{\varphi}_{n+1}^T(s)$ для $1 < n \leq C$. Матричное уравнение при $n = 2$ имеет вид

$$-(\mathbf{Q}_{02} - s\mathbf{I})\boldsymbol{\varphi}_2^T(s) = \mathbf{Q}_{11}\boldsymbol{\varphi}_3^T(s) + \mathbf{Q}_{22}\boldsymbol{\varphi}_1^T(s).$$

Подставим полученное соотношение для $\boldsymbol{\varphi}_1^T(s)$

$$\boldsymbol{\varphi}_2^T(s) = -[\mathbf{Q}_{02} - s\mathbf{I} + \mathbf{Q}_{22}\mathbf{X}_1(s)]^{-1}\mathbf{Q}_{11}\boldsymbol{\varphi}_3^T(s) - [\mathbf{Q}_{02} - s\mathbf{I} + \mathbf{Q}_{22}\mathbf{X}_1(s)]^{-1}\mathbf{Q}_{22}\mathbf{Y}_1(s)$$

и обозначим $\mathbf{X}_2(s) = -[\mathbf{Q}_{02} - s\mathbf{I} + \mathbf{Q}_{22}\mathbf{X}_1(s)]^{-1}\mathbf{Q}_{11}$,

$$\mathbf{Y}_2(s) = -[\mathbf{Q}_{02} - s\mathbf{I} + \mathbf{Q}_{22}\mathbf{X}_1(s)]^{-1}\mathbf{Q}_{22}\mathbf{Y}_1(s), \text{ тогда}$$

$$\boldsymbol{\varphi}_2^T(s) = \mathbf{X}_2(s)\boldsymbol{\varphi}_3^T(s) + \mathbf{Y}_2(s).$$

Предположим, что соотношение верно для $n - 1$

$$\boldsymbol{\varphi}_{n-1}^T(s) = \mathbf{X}_{n-1}(s)\boldsymbol{\varphi}_n^T(s) + \mathbf{Y}_{n-1}(s),$$

$$\mathbf{X}_{n-1}(s) = -[\mathbf{Q}_{0,n-1} - s\mathbf{I} + \mathbf{Q}_{2,n-1}\mathbf{X}_{n-2}(s)]^{-1}\mathbf{Q}_{11},$$

$$\mathbf{Y}_{n-1}(s) = -[\mathbf{Q}_{0,n-1} - s\mathbf{I} + \mathbf{Q}_{2,n-1}\mathbf{X}_{n-2}(s)]^{-1}\mathbf{Q}_{2,n-1}\mathbf{Y}_{n-2}(s).$$

Подставим его в матричное уравнение при n

$$\begin{aligned} \varphi_n^T(s) = \\ = -[\mathbf{Q}_{0n} - s\mathbf{I} + \mathbf{Q}_{2n}\mathbf{X}_{n-1}(s)]^{-1} \mathbf{Q}_{11}\varphi_{n+1}^T(s) + [\mathbf{Q}_{0n} - s\mathbf{I} + \mathbf{Q}_{2n}\mathbf{X}_{n-1}(s)]^{-1} \mathbf{Q}_{2n}\mathbf{Y}_{n-1}(s) \end{aligned}$$

и обозначим $\mathbf{X}_n(s) = -[\mathbf{Q}_{0n} - s\mathbf{I} + \mathbf{Q}_{2n}\mathbf{X}_{n-1}(s)]^{-1} \mathbf{Q}_{11}$ и

$$\mathbf{Y}_n(s) = -[\mathbf{Q}_{0n} - s\mathbf{I} + \mathbf{Q}_{2n}\mathbf{X}_{n-1}(s)]^{-1} \mathbf{Q}_{2n}\mathbf{Y}_{n-1}(s), \text{ тогда}$$

$$\varphi_n^T(s) = \mathbf{X}_n(s)\varphi_{n+1}^T(s) + \mathbf{Y}_n(s).$$

Следовательно, сделанное предположение о зависимости $\varphi_n^T(s)$ от $\varphi_{n+1}^T(s)$ верно, таким образом

$$\varphi_n^T(s) = \mathbf{X}_n(s)\varphi_{n+1}^T(s) + \mathbf{Y}_n(s), \quad n = 2, \dots, C,$$

$$\mathbf{X}_n(s) = -[\mathbf{Q}_{0n} - s\mathbf{I} + \mathbf{Q}_{2n}\mathbf{X}_{n-1}(s)]^{-1} \mathbf{Q}_{11}, \quad n = 2, \dots, C,$$

$$\mathbf{Y}_n(s) = -[\mathbf{Q}_{0n} - s\mathbf{I} + \mathbf{Q}_{2n}\mathbf{X}_{n-1}(s)]^{-1} \mathbf{Q}_{2n}\mathbf{Y}_{n-1}(s), \quad n = 2, \dots, C.$$

г) Выражение $\varphi_n^T(s)$ через $\varphi_{n+1}^T(s)$ для $n > C$

Получим зависимость $\varphi_n^T(s)$ от $\varphi_{n+1}^T(s)$ для $C < n < R$. Матричное уравнение при $n = C + 1$ имеет вид

$$-(\mathbf{Q}_{0C} - s\mathbf{I})\varphi_{C+1}^T(s) = \mathbf{Q}_{11}\varphi_{C+2}^T(s) + \mathbf{Q}_{2C}\varphi_C^T(s).$$

Подставим полученное соотношение для $\varphi_C^T(s)$

$$\begin{aligned} \varphi_{C+1}^T(s) = \\ = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_C(s)]^{-1} \mathbf{Q}_{11}\varphi_{C+2}^T(s) - [\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_C(s)]^{-1} \mathbf{Q}_{2C}\mathbf{Y}_C(s) \end{aligned}$$

и обозначим $\mathbf{X}_{C+1}(s) = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_C(s)]^{-1} \mathbf{Q}_{11}$,

$$\mathbf{Y}_{C+1}(s) = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_C(s)]^{-1} \mathbf{Q}_{2C}\mathbf{Y}_C(s), \text{ тогда}$$

$$\varphi_{C+1}^T(s) = \mathbf{X}_{C+1}(s)\varphi_{C+2}^T(s) + \mathbf{Y}_{C+1}(s)$$

Предположим, что соотношение верно для $n-1$

$$\boldsymbol{\varphi}_{n-1}^T(s) = \mathbf{X}_{n-1}(s)\boldsymbol{\varphi}_n^T(s) + \mathbf{Y}_{n-1}(s),$$

$$\mathbf{X}_{n-1}(s) = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{n-2}(s)]^{-1}\mathbf{Q}_{11},$$

$$\mathbf{Y}_{n-1}(s) = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{n-2}(s)]^{-1}\mathbf{Q}_{2C}\mathbf{Y}_{n-2}(s).$$

Подставим его в матричное уравнение при n

$$\begin{aligned} \boldsymbol{\varphi}_n^T(s) = \\ = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{n-1}(s)]^{-1}\mathbf{Q}_{11}\boldsymbol{\varphi}_{n+1}^T(s) - [\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{n-1}(s)]^{-1}\mathbf{Q}_{2C}\mathbf{Y}_{n-1}(s) \end{aligned}$$

и обозначим $\mathbf{X}_n(s) = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{n-1}(s)]^{-1}\mathbf{Q}_{11}$ и

$$\mathbf{Y}_n(s) = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{n-1}(s)]^{-1}\mathbf{Q}_{2C}\mathbf{Y}_{n-1}(s), \text{ тогда}$$

$$\boldsymbol{\varphi}_n^T(s) = \mathbf{X}_n(s)\boldsymbol{\varphi}_{n+1}^T(s) + \mathbf{Y}_n(s).$$

Следовательно, сделанное предположение о зависимости $\boldsymbol{\varphi}_n^T(s)$ от $\boldsymbol{\varphi}_{n+1}^T(s)$ верно, таким образом

$$\boldsymbol{\varphi}_n^T(s) = \mathbf{X}_n(s)\boldsymbol{\varphi}_{n+1}^T(s) + \mathbf{Y}_n(s), \quad n = C+1, \dots, R-1,$$

$$\mathbf{X}_n(s) = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{n-1}(s)]^{-1}\mathbf{Q}_{11}, \quad n = C+1, \dots, R-1,$$

$$\mathbf{Y}_n(s) = -[\mathbf{Q}_{0C} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{n-1}(s)]^{-1}\mathbf{Q}_{2C}\mathbf{Y}_{n-1}(s), \quad n = C+1, \dots, R-1.$$

д) *Получение выражения для $\boldsymbol{\varphi}_R^T(s)$*

Получим выражение для $\boldsymbol{\varphi}_R^T(s)$, подставив соотношение для $\boldsymbol{\varphi}_{R-1}^T(s)$ в матричное уравнения при $n = R$

$$\boldsymbol{\varphi}_R^T(s) = -[\mathbf{Q}_{0,C+1} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{R-1}(s)]^{-1}\mathbf{Q}_{2C}\mathbf{Y}_{R-1}(s)$$

и обозначим $\mathbf{Y}_R(s) = -[\mathbf{Q}_{0,C+1} - s\mathbf{I} + \mathbf{Q}_{2C}\mathbf{X}_{R-1}(s)]^{-1}\mathbf{Q}_{2C}\mathbf{Y}_{R-1}(s)$, тогда

$$\boldsymbol{\varphi}_R^T(s) = \mathbf{Y}_R(s).$$

■

5.3.3 Матричный алгоритм расчета преобразования Лапласа-Стилтьеса длительности периода занятости

С учетом леммы 5.3, справедлива следующая теорема.

Теорема 5.3. Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и отказами ресурса на периоде простоя системы, ПЛС времени первого достижения пустого состояния рассчитывается в матричном виде по формулам

$$\begin{aligned}\varphi_0^T(s) &= \mathbf{e}, \\ \varphi_n^T(s) &= \sum_{i=n}^R \prod_{j=n}^{i-1} \mathbf{X}_j(s) \mathbf{Y}_i(s), \quad n=1, \dots, R-1, \\ \varphi_R^T(s) &= \mathbf{Y}_R(s),\end{aligned}\tag{5.32}$$

где матрицы $\mathbf{X}_n(s)$ и $\mathbf{Y}_n(s)$ вычисляются по рекуррентным соотношениям леммы 5.3.

Доказательство.

Если $n=0$, тогда время $\xi_b(0,1) = \xi_b(0,0) = 0$, поэтому $\varphi_0(s) = \mathbf{1}$.

Выражение для $\varphi_R^T(s)$ уже есть в матричном уравнении при $n=R$

$$\varphi_R^T(s) = \mathbf{Y}_R(s),$$

Матричное уравнение при $n=R-1$ имеет вид

$$\varphi_{R-1}^T(s) = \mathbf{X}_{R-1}(s) \varphi_R^T(s) + \mathbf{Y}_{R-1}(s).$$

Подставим соотношение для $\varphi_R^T(s)$

$$\varphi_{R-1}^T(s) = \mathbf{X}_{R-1}(s) \mathbf{Y}_R(s) + \mathbf{Y}_{R-1}(s) = \sum_{i=R-1}^R \prod_{j=R-1}^{i-1} \mathbf{X}_j(s) \mathbf{Y}_i(s)$$

Предположим, что соотношение верно для $n + 1$

$$\boldsymbol{\varphi}_{n+1}^T(s) = \sum_{i=n+1}^R \prod_{j=n+1}^{i-1} \mathbf{X}_j(s) \mathbf{Y}_i(s)$$

Подставим его в матричное уравнение при n

$$\begin{aligned} \boldsymbol{\varphi}_n^T(s) &= \mathbf{X}_n(s) \sum_{i=n+1}^R \prod_{j=n+1}^{i-1} \mathbf{X}_j(s) \mathbf{Y}_i(s) + \mathbf{Y}_n(s) = \sum_{i=n+1}^R \prod_{j=n+1}^{i-1} \mathbf{X}_n(s) \mathbf{X}_j(s) \mathbf{Y}_i(s) + \mathbf{Y}_n(s) = \\ &= \sum_{i=n+1}^R \prod_{j=n}^{i-1} \mathbf{X}_j(s) \mathbf{Y}_i(s) + \mathbf{Y}_n(s) = \sum_{i=n}^R \prod_{j=n}^{i-1} \mathbf{X}_j(s) \mathbf{Y}_i(s) \end{aligned}$$

Следовательно, сделанное предположение о виде $\boldsymbol{\varphi}_n^T(s)$ верно. ■

Из теоремы 5.3 вытекает следующее следствие.

Следствие 5.2. Для модели с потоковым трафиком и полным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и отказами ресурса на периоде простоя системы, ПЛС длительности периода занятости рассчитывается по формуле

$$\varphi(s) = \frac{\boldsymbol{\pi}_0}{\boldsymbol{\pi}_0 \mathbf{1}^T} \boldsymbol{\varphi}_1^T(s), \quad (5.33)$$

где время $\boldsymbol{\varphi}_1^T(s)$ первого достижения пустого состояния из состояния, когда в системе 1 сессия потокового трафика вычисляется в матричном виде по формуле

$$\boldsymbol{\varphi}_1^T(s) = \sum_{i=1}^R \prod_{j=1}^{i-1} \mathbf{X}_j(s) \mathbf{Y}_i(s), \quad (5.34)$$

а матрицы $\mathbf{X}_n(s)$ и $\mathbf{Y}_n(s)$ вычисляются по рекуррентным соотношениям леммы 5.3.

5.4 СИСТЕМА С ЧАСТИЧНЫМ ОТКАЗОМ РЕСУРСОМ

5.4.1 Матрица интенсивностей переходов в блочно-трехдиагональном виде

Лемма 5.4. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика, если на пространстве $\mathcal{X}$ состояний СП $X(t)$ введен лексикографический порядок

$$\begin{aligned} x = (l, m, k, u) < (l', m', k', u') = x' \Leftrightarrow \\ (l + m + k < l' + m' + k') \vee \\ \vee (l + m + k = l' + m' + k', u < u') \vee \\ \vee (l + m + k = l' + m' + k', u = u', l > l'), \quad x, x' \in \mathcal{X}, \end{aligned} \quad (5.35)$$

тогда матрица интенсивностей переходов СП $X(t)$ представима в блочно-трехдиагональном виде

$$\begin{aligned} Q = \text{diag}(Q_{00}, Q_{01}, \dots, Q_{0, V+R}) + \\ + \text{diag}^+(Q_{11}, Q_{12}, \dots, Q_{1, V+R}) + \text{diag}^-(Q_{21}, Q_{22}, \dots, Q_{2, V+R}), \end{aligned} \quad (5.36)$$

где блоки определяются по формулам

$$Q_{00} = \begin{bmatrix} -(\lambda + \beta) & \beta \\ \alpha & -(\lambda + \alpha) \end{bmatrix}, \quad (5.37)$$

$$\dim Q_{00} = 2 \times 2,$$

$$Q_{0n} = \begin{bmatrix} -(\lambda + n\mu + \beta) & \beta & 0 & \dots & 0 \\ \alpha & -(\lambda + n\mu + \alpha) & & & \\ \alpha & & -(\lambda + n\mu + \alpha) & & \\ \vdots & & & \ddots & \\ \alpha & & & & -(\lambda + n\mu + \alpha) \end{bmatrix},$$

$$\dim Q_{0n} = (n + 2) \times (n + 2), \quad n = 1, \dots, C, \quad (5.38)$$

$$\dim Q_{0n} = (C + 2) \times (C + 2), \quad n = C + 1, \dots, V,$$

$$\mathbf{Q}_{0n} = \left[\begin{array}{c|cccc} -(\lambda + V\mu + \beta) & \beta & 0 & \cdots & 0 \\ \hline \alpha & -(\lambda + n\mu + \alpha) & & & \\ \alpha & & -(\lambda + n\mu + \alpha) & & \\ \vdots & & & \ddots & \\ \alpha & & & & -(\lambda + n\mu + \alpha) \end{array} \right], \quad (5.39)$$

$$\dim \mathbf{Q}_{0n} = (V + C - n + 2) \times (V + C - n + 2), \quad n = V + 1, \dots, V + C,$$

$$\mathbf{Q}_{0n} = \left[\begin{array}{c|c} -(\lambda + V\mu + \beta) & \beta \\ \hline \alpha & -(\lambda + (V + C)\mu + \alpha) \end{array} \right], \quad (5.40)$$

$$\dim \mathbf{Q}_{0n} = 2 \times 2, \quad n = V + C + 1, \dots, V + R - 1,$$

$$\mathbf{Q}_{0,V+R} = \left[\begin{array}{c|c} -(V\mu + \beta) & \beta \\ \hline \alpha & -((V + C)\mu + \alpha) \end{array} \right], \quad (5.41)$$

$$\dim \mathbf{Q}_{0,V+R} = 2 \times 2,$$

$$\mathbf{Q}_{1n} = \left[\begin{array}{c|ccc} \lambda & & & \\ \hline & \lambda & 0 & \\ & & \ddots & \ddots \\ & & & \lambda & 0 \end{array} \right], \quad (5.42)$$

$$\dim \mathbf{Q}_{1n} = (n + 1) \times (n + 2), \quad n = 1, \dots, C,$$

$$\mathbf{Q}_{1n} = \left[\begin{array}{c|ccc} \lambda & & & \\ \hline & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{array} \right], \quad (5.43)$$

$$\dim \mathbf{Q}_{1n} = (C + 2) \times (C + 2), \quad n = C + 1, \dots, V,$$

$$\mathbf{Q}_{1n} = \left[\begin{array}{c|cccc} \lambda & & & & \\ \hline & \lambda & & & \\ & \lambda & 0 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 0 \\ & & & & \lambda \end{array} \right], \quad (5.44)$$

$$\dim \mathbf{Q}_{1n} = (V + C - n + 2) \times (V + C - n + 1), \quad n = V + 1, \dots, V + C,$$

$$\mathbf{Q}_{1n} = \begin{bmatrix} \lambda & \\ & \lambda \end{bmatrix}, \quad (5.45)$$

$$\dim \mathbf{Q}_{1n} = 2 \times 2, \quad n = V + C + 1, \dots, V + R,$$

$$\mathbf{Q}_{2n} = \begin{bmatrix} n\mu & & & & & \\ & n\mu & & & & \\ & \mu & (n-1)\mu & & & \\ & & 2\mu & \ddots & & \\ & & & \ddots & 2\mu & \\ & & & & (n-1)\mu & \mu \\ & & & & & n\mu \end{bmatrix}, \quad (5.46)$$

$$\dim \mathbf{Q}_{2n} = (n+2) \times (n+1), \quad n = 1, \dots, C,$$

$$\mathbf{Q}_{2n} = \begin{bmatrix} n\mu & & & & & \\ & n\mu & & & & \\ & \mu & (n-1)\mu & & & \\ & & 2\mu & (n-2)\mu & & \\ & & & \ddots & \ddots & \\ & & & & C\mu & (n-C)\mu \end{bmatrix}, \quad (5.47)$$

$$\dim \mathbf{Q}_{2n} = (C+2) \times (C+2), \quad n = C+1, \dots, V,$$

$$\mathbf{Q}_{2n} = \begin{bmatrix} V\mu & & & & & \\ & (n-V)\mu & V\mu & & & \\ & & (n-V+1)\mu & (V-1)\mu & & \\ & & & \ddots & \ddots & \\ & & & & C\mu & (n-C)\mu \end{bmatrix}, \quad (5.48)$$

$$\dim \mathbf{Q}_{2n} = (V+C-n+1) \times (V+C-n+2), \quad n = V+1, \dots, V+C,$$

$$\mathbf{Q}_{2n} = \begin{bmatrix} V\mu & \\ & (V+C)\mu \end{bmatrix}, \quad (5.49)$$

$$\dim \mathbf{Q}_{2n} = 2 \times 2, \quad n = V + C + 1, \dots, V + R.$$

Доказательство.

Разобьем пространство $\mathcal{X}$ состояний (5.3) СП $\mathbf{X}(t)$ на подмножества по числу $l + m + k$ сессий трафика в системе и состоянию u ненадежного ресурса

$$\mathcal{X} = \bigcup_{n=0}^{V+R} \mathcal{X}(n),$$

$$\mathcal{X}(n) = \{(l, m, k, u) \in \mathcal{X} : l + m + k = n\},$$

$$\mathcal{X}(n) = \mathcal{X}(n, 1) \cup \mathcal{X}(n, 0),$$

$$\mathcal{X}(n, u) = \{(l, m, k, i) \in \mathcal{X}(n) : i = u\} = \{(l, m, k, u) \in \mathcal{X}(n)\}.$$

Введем лексикографический порядок на множестве $\mathcal{X}$ в порядке объединения $\mathcal{X}(n)$ и $\mathcal{X}(n, u)$

$$\begin{aligned} \mathcal{X} = & \{(0, 0, 0, 0), (0, 0, 0, 1), \\ & (1, 0, 0, 0), (1, 0, 0, 1), (0, 1, 0, 1), \\ & (2, 0, 0, 0), (2, 0, 0, 1), (1, 1, 0, 1), (0, 2, 0, 1), \\ & \dots, \\ & (n, 0, 0, 0), (n, 0, 0, 1), (n-1, 1, 0, 1), \dots, (l, n-l, 0, 1), \dots, (0, n, 0, 1), \\ & \dots, \\ & (C, 0, 0, 0), (C, 0, 0, 1), (1, C-1, 0, 1), \dots, (l, C-l, 0, 1), \dots, (0, C, 0, 1), \\ & (C+1, 0, 0, 0), (C+1, 0, 0, 1), (C, 1, 0, 1), \dots, (l, C+1-l, 0, 1), \dots, (1, C, 0, 1), \\ & (C+2, 0, 0, 0), (C+2, 0, 0, 1), (C+1, 1, 0, 1), \dots, (l, C+2-l, 0, 1), \dots, (2, C, 0, 1), \\ & \dots, \\ & (n, 0, 0, 0), (n, 0, 0, 1), (n-1, 1, 0, 1), \dots, (l, n-l, 0, 1), \dots, (n-C, C, 0, 1), \\ & \dots, \\ & (V, 0, 0, 0), (V, 0, 0, 1), (V-1, 1, 0, 1), \dots, (l, V-l, 0, 1), \dots, (V-C, C, 0, 1), \\ & (V, 0, 1, 0), (V, 1, 0, 1), (V-1, 2, 0, 1), \dots, (l, V+1-l, 0, 1), \dots, (V-C+1, C, 0, 1), \\ & (V, 0, 2, 0), (V, 2, 0, 1), (V-1, 3, 0, 1), \dots, (l, V+2-l, 0, 1), \dots, (V-C+2, C, 0, 1), \\ & \dots, \\ & (V, 0, n-V, 0), (V, n-V, 0, 1), \dots, (l, n-l, 0, 1), \dots, (n-C, C, 0, 1), \\ & \dots, \\ & (V, 0, C-1, 0), (V, C-1, 0, 1), (V-1, C, 0, 1), \\ & (V, 0, C, 0), (V, C, 0, 0), \end{aligned}$$

где блоки определяются по формулам

$$\mathbf{Q}_{00} = \text{diag}\left(-\lambda, -(\lambda + \mu), -(\lambda + 2\mu), \dots, -(\lambda + \min(N_1 - 1, V)\mu)\right) + \\ + \text{diag}^+(\lambda, \dots, \lambda) + \text{diag}^-(\mu, 2\mu, \dots, \min(N_1 - 1, V)\mu), \quad (5.52)$$

$$\dim \mathbf{Q}_{00} = N_1 \times N_1,$$

$$\mathbf{A}_{0n} = \begin{bmatrix} -(\lambda + \min(n, V + C)\mu + \alpha) & \alpha \\ \beta & -(\lambda + \min(n, V)\mu + \beta) \end{bmatrix}, \quad (5.53)$$

$$\dim \mathbf{A}_{0n} = 2 \times 2, \quad n = N_2 + 1, \dots, N - 1,$$

$$\mathbf{A}_{0N} = \begin{bmatrix} -(\lambda + (V + C)\mu + \alpha) & \alpha \\ \beta & -(\lambda + V\mu + \beta) \end{bmatrix}, \quad (5.54)$$

$$\dim \mathbf{A}_{0, N+1} = 2 \times 2, \quad n = N, N + 1, \dots,$$

$$\mathbf{A}_1 = \begin{bmatrix} \lambda & \\ & \lambda \end{bmatrix} = \lambda \mathbf{I}, \quad (5.55)$$

$$\dim \mathbf{A}_{1n} = 2 \times 2,$$

$$\mathbf{A}_{2n} = \begin{bmatrix} \min(n, V + C)\mu & \\ & \min(n, V)\mu \end{bmatrix}, \quad (5.56)$$

$$\dim \mathbf{A}_{2n} = 2 \times 2, \quad n = N_2 + 2, \dots, N - 1,$$

$$\mathbf{A}_{2N} = \begin{bmatrix} (V + C)\mu & \\ & V\mu \end{bmatrix}, \quad (5.57)$$

$$\dim \mathbf{A}_{2n} = 2 \times 2, \quad n = N, N + 1, \dots,$$

$$\mathbf{B} = \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ \lambda p & \lambda(1 - p) \end{bmatrix}, \quad (5.58)$$

$$\dim \mathbf{B} = N_1 \times 2,$$

$$\mathbf{C} = \begin{bmatrix} 0 & \dots & 0 & \overset{N_2+1 \text{ столбец}}{\min(N_2 + 1, V + C)\mu} & 0 & \dots & 0 \\ 0 & \dots & 0 & \min(N_2 + 1, V)\mu & 0 & \dots & 0 \end{bmatrix}, \quad (5.59)$$

$$\dim \mathbf{C} = 2 \times N_1$$

и $N = \max(N_1 + 1, V + C)$.

Доказательство.

а) Введение лексикографического порядка

Разобьем пространство $\mathcal{X}$ состояний (5.4) СП $X(t)$ на подмножества по состоянию u ненадежного ресурса и числу n сессий трафика в системе

$$\mathcal{X} = \mathcal{X}(0) \cup \bar{\mathcal{X}}(0),$$

$$\mathcal{X}(0) = \{(n, u) \in \mathcal{X} : u = 0\} = \{(n, 0) \in \mathcal{X}\} = \{(0, 0), (1, 0), \dots, (n, 0), \dots, (N_1 - 1, 0)\},$$

$$\bar{\mathcal{X}}(0) = \{(n, u) \in \mathcal{X} : u = 1, 2\},$$

$$\bar{\mathcal{X}}(0) = \bigcup_{n=N_2+1}^{\infty} \mathcal{X}(n),$$

$$\mathcal{X}(n) = \{(i, u) \in \bar{\mathcal{X}}(0) : i = n\} = \{(n, u) \in \bar{\mathcal{X}}(0)\} = \{(n, 1), (n, 2)\}.$$

Введем лексикографический порядок на множестве $\mathcal{X}$ в порядке объединения $\mathcal{X}(0)$ и $\mathcal{X}(n)$

$$\begin{aligned} \mathcal{X} = & \{(0, 0), (1, 0), \dots, (n, 0), \dots, (N_2, 0), \dots, (n, 0), \dots, (N_1 - 1, 0), \\ & (N_2 + 1, 1), (N_2 + 1, 2), \\ & (N_2 + 2, 1), (N_2 + 2, 2), \\ & \dots, \\ & (n, 1), (n, 2), \\ & \dots, \\ & (N_1, 1), (N_1, 2), \\ & \dots, \\ & (n, 1), (n, 2), \\ & \dots\}. \end{aligned}$$

б) Описание блоков матрицы

С учетом введенного лексикографического порядка и ненулевых элементов матрицы $\mathbf{Q}$ интенсивностей переходов (5.10), матрица $\mathbf{Q}$ имеет блочный вид из четырех блоков

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{00} & \mathbf{Q}_{01} \\ \mathbf{Q}_{10} & \mathbf{Q}_{11} \end{bmatrix}.$$

Блоки матрицы $\mathbf{Q}$ имеют следующую структуру

$$\mathbf{Q}_{00} = \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & \\ & 2\mu & -(\lambda + 2\mu) & \ddots & \\ & & \ddots & \ddots & \lambda \\ & & & \min(N_1 - 1, V)\mu & -(\lambda + \min(N_1 - 1, V)\mu) \end{bmatrix},$$

$$\mathbf{Q}_{11} = \begin{bmatrix} \mathbf{A}_{0, N_2+1} & \mathbf{A}_{1, N_2+2} & & \\ \mathbf{A}_{2, N_2+2} & \mathbf{A}_{0, N_2+2} & \mathbf{A}_{1, N_2+3} & \\ & \mathbf{A}_{2, N_2+3} & \ddots & \ddots \\ & & \ddots & \ddots \end{bmatrix},$$

$$\mathbf{Q}_{01} = [\mathbf{0} \quad \cdots \quad \mathbf{0} \quad \overset{N_1 - N_2 + 1}{\text{столбец}} \mathbf{B} \quad \mathbf{0} \quad \mathbf{0} \quad \cdots],$$

$$\mathbf{Q}_{10} = \begin{bmatrix} \mathbf{C} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \end{bmatrix}.$$

■

5.4.2 Матричный алгоритм расчета стационарного распределения

Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и гистерезисным доступом, в соответствии с лексикографическим порядком (5.50), представим стационарное распределение СП $\mathbf{X}(t)$ в виде

$$\begin{aligned} \boldsymbol{\pi} &= (\boldsymbol{\pi}^0, \boldsymbol{\pi}^1), \\ \boldsymbol{\pi}^0 &= (\pi(0, 0), \pi(1, 0), \dots, \pi(n, 0), \dots, \pi(N_2, 0), \dots, \pi(n, 0), \dots, \pi(N_1 - 1, 0)), \\ \boldsymbol{\pi}^1 &= (\pi_0, \pi_{N_2+1}, \pi_{N_2+2}, \dots), \\ \pi_n &= (\pi(n, 1), \pi(n, 2)), \quad n = N_2 + 1, N_2 + 2, \dots \end{aligned} \tag{5.60}$$

Тогда с учетом леммы 5.5, справедлива следующая теорема.

Теорема 5.4. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и гистерезисным доступом, стационарное распределение СП $X(t)$ рассчитывается в матричном виде по рекуррентным формулам

$$\begin{aligned}
 \pi(n,0) &= \frac{\lambda^n}{\mu^n} \frac{1}{n!} \pi(0,0), \quad n=1,\dots,\min(N_2,V), \\
 \pi(n,0) &= \frac{\lambda^n}{\mu^n} \frac{1}{V!V^{n-V}} \pi(0,0), \quad n=V+1,\dots,N_2, \quad N_2 > V, \\
 \pi(n,0) &= \frac{A(N_1-n-1)}{A(N_1-N_2-1)} \cdot \pi(N_2,0), \quad n=N_2+1,\dots,N_1-1, \\
 \pi_n &= \pi(N_1-1,0) \sum_{j=N_1}^N \mathbf{Y}_j \prod_{i=0}^{j-n-1} \mathbf{X}_{j-i-1}, \quad n=N_2+1,\dots,N_1-1, \quad N_1 > N_2+1, \\
 \pi_n &= \pi(N_1-1,0) \sum_{j=n}^N \mathbf{Y}_j \prod_{i=0}^{j-n-1} \mathbf{X}_{j-i-1}, \quad n=N_1,\dots,N, \\
 \pi_n &= \pi_N \mathbf{R}^{n-N}, \quad n=N+1, N+2, \dots
 \end{aligned} \tag{5.61}$$

где коэффициенты $A(n)$, матрицы $\mathbf{X}_n$ и $\mathbf{Y}_n$ вычисляются по рекуррентным соотношениям

$$\begin{aligned}
 A(0) &= 0, \\
 A(n) &= 1 + \frac{\min(N_1-n,V)\mu}{\lambda} A(n-1), \quad n=1,\dots,N_1-N_2-1,
 \end{aligned} \tag{5.62}$$

$$\mathbf{X}_{N_2+1} = -\mathbf{A}_{2,N_2+2} \mathbf{A}_{0,N_2+1}^{-1}, \tag{5.63}$$

$$\begin{aligned}
 \mathbf{X}_n &= -\mathbf{A}_{2,n+1} (\lambda \mathbf{X}_{n-1} + \mathbf{A}_{0n})^{-1}, \quad n=N_2+2,\dots,N-1, \\
 \mathbf{Y}_{N_1} &= -\lambda(p,1-p) (\lambda \mathbf{X}_{N_1-1} + \mathbf{A}_{0N_1})^{-1}, \quad N_1 > N_2+1, \\
 \mathbf{Y}_{N_1} &= -\lambda(p,1-p) \mathbf{A}_{0N_1}^{-1}, \quad N_1 = N_2+1, \\
 \mathbf{Y}_n &= -\lambda \mathbf{Y}_{n-1} (\lambda \mathbf{X}_{n-1} + \mathbf{A}_{0n})^{-1}, \quad n=N_1+1,\dots,N-1, \\
 \mathbf{Y}_{N-N_2} &= -\lambda \mathbf{Y}_{N-1-N_2} (\lambda \mathbf{X}_{N-1-N_2} + \mathbf{A}_{0,N+1} + \mathbf{R} \mathbf{A}_{2,N+1})^{-1}.
 \end{aligned} \tag{5.64}$$

а матрица $\mathbf{R}$ является минимальным неотрицательным решением квадратичного матричного уравнения

$$\mathbf{R}^2 \mathbf{A}_{2N} + \mathbf{R} \mathbf{A}_{0N} + \lambda \mathbf{I} = \mathbf{0} \quad (5.65)$$

и $N = \max(N_1 + 1, V + C)$.

Доказательство.

а) *Запись СУР*

Запишем СУР $\pi \mathbf{Q} = \mathbf{0}$ с учетом блочно-трехдиагонального вида матрицы $\mathbf{Q}$ интенсивностей переходов (5.51).

Если $N_1 > N_2 + 1$, то

$$\begin{aligned} \pi^0 \mathbf{Q}_{00} + \pi_{N_2+1} \mathbf{C} &= \mathbf{0}, \\ \pi_{N_2+1} \mathbf{A}_{0,N_2+1} + \pi_{N_2+2} \mathbf{A}_{2,N_2+2} &= \mathbf{0}, \\ \pi_{n-1} \mathbf{A}_1 + \pi_n \mathbf{A}_{0n} + \pi_{n+1} \mathbf{A}_{2,n+1} &= \mathbf{0}, \quad n = N_2 + 2, \dots, N_1 - 1, \\ \pi^0 \mathbf{B} + \pi_{N_1-1} \mathbf{A}_1 + \pi_{N_1} \mathbf{A}_{0N_1} + \pi_{N_1+1} \mathbf{A}_{2,N_1+1} &= \mathbf{0}, \\ \pi_{n-1} \mathbf{A}_1 + \pi_n \mathbf{A}_{0n} + \pi_{n+1} \mathbf{A}_{2,n+1} &= \mathbf{0}, \quad n = N_1 + 1, \dots, N - 1, \\ \pi_{n-1} \mathbf{A}_1 + \pi_n \mathbf{A}_{0,N+1} + \pi_{n+1} \mathbf{A}_{2,N+1} &= \mathbf{0}, \quad n = N, N + 1, \dots \end{aligned}$$

Если $N_1 = N_2 + 1$, то

$$\begin{aligned} \pi^0 \mathbf{Q}_{00} + \pi_{N_2+1} \mathbf{C} &= \mathbf{0}, \\ \pi^0 \mathbf{B} + \pi_{N_1} \mathbf{A}_{0N_1} + \pi_{N_1+1} \mathbf{A}_{2,N_1+1} &= \mathbf{0}, \\ \pi_{n-1} \mathbf{A}_1 + \pi_n \mathbf{A}_{0n} + \pi_{n+1} \mathbf{A}_{2,n+1} &= \mathbf{0}, \quad n = N_1 + 1, \dots, N - 1, \\ \pi_{n-1} \mathbf{A}_1 + \pi_n \mathbf{A}_{0,N+1} + \pi_{n+1} \mathbf{A}_{2,N+1} &= \mathbf{0}, \quad n = N, N + 1, \dots \end{aligned}$$

Подставим вид матриц $\mathbf{Q}_{00}$ (5.52) и $\mathbf{C}$ (5.59) в полученное матричное уравнение СУР при $((0,0), (1,0), \dots, (N_1-1,0))$

$$\begin{aligned} -\lambda \pi(0,0) + \mu \pi(1,0) &= 0, \\ \lambda \pi(n-1,0) - [\lambda + \min(n,V)\mu] \pi(n,0) + \min(n+1,V)\mu \pi(n+1,0) &= 0, \\ n &= 1, \dots, N_2 - 1, \end{aligned}$$

$$\begin{aligned} & \lambda \pi(N_2 - 1, 0) - [\lambda + \min(N_2, V) \mu] \pi(N_2, 0) + \min(N_2 + 1, V) \mu \pi(N_2 + 1, 0) + \\ & \quad + \min(N_2 + 1, V + C) \mu \pi(N_2 + 1, 1) + \min(N_2 + 1, V) \mu \pi(N_2 + 1, 2) = 0, \\ & \lambda \pi(n - 1, 0) - [\lambda + \min(n, V) \mu] \pi(n, 0) + \min(n + 1, V) \mu \pi(n + 1, 0) = 0, \\ & \quad n = N_2 + 1, \dots, N_1 - 2, \\ & \lambda \pi(N_1 - 2, 0) - [\lambda + \min(N_1 - 1, V) \mu] \pi(N_1 - 1, 0) = 0. \end{aligned}$$

Отметим также, что

$$\pi^0 \mathbf{B} = \lambda(p, 1 - p) \pi(N_1 - 1, 0).$$

б) Получение зависимости $\pi(n, 0)$ от $\pi(n - 1, 0)$ для $0 < n \leq N_2$

Зависимость $\pi(1, 0)$ и $\pi(0, 0)$ видна из уравнения СУР при $(0, 0)$

$$\lambda \pi(0, 0) = \mu \pi(1, 0).$$

Получим зависимость $\pi(n, 0)$ и $\pi(n - 1, 0)$ для $0 < n \leq N_2$. Уравнение СУР при $(1, 0)$ имеет вид

$$\lambda \pi(0, 0) - [\lambda + \min(1, V) \mu] \pi(1, 0) + \min(2, V) \mu \pi(2, 0) = 0.$$

Подставим выражение $\pi(0, 0)$ через $\pi(1, 0)$

$$\lambda \pi(1, 0) = \min(2, V) \mu \pi(2, 0).$$

Предположим, что соотношение верно для $n - 1$

$$\lambda \pi(n - 1, 0) = \min(n, V) \mu \pi(n, 0).$$

Подставим его в уравнение СУР при $(n, 0)$

$$\lambda \pi(n, 0) = \min(n + 1, V) \mu \pi(n + 1, 0)$$

и объединим с соотношением для $\pi(0, 0)$. Следовательно, сделанное предположение о зависимости $\pi(n, 0)$ и $\pi(n - 1, 0)$ верно, таким образом

$$\pi(n, 0) = \frac{\lambda}{\mu \min(n, V)} \pi(n - 1, 0), \quad n = 1, \dots, N_2.$$

в) Выражение $\pi(n,0)$ через $\pi(0,0)$ для $0 < n \leq N_2$

Выразим $\pi(n,0)$ через $\pi(0,0)$, рекуррентно подставляя полученное соотношение

$$\begin{aligned}\pi(n,0) &= \frac{\lambda}{\mu} \frac{1}{\min(n,V)} \pi(n-1,0) = \frac{\lambda^2}{\mu^2} \frac{1}{\min(n,V)\min(n-1,V)} \pi(n-2,0) = \dots = \\ &= \frac{\lambda^n}{\mu^n} \frac{1}{\min(n,V) \times \dots \times \min(1,V)} \pi(0,0) = \frac{\lambda^n}{\mu^n} \frac{1}{\prod_{i=1}^n \min(i,V)} \pi(0,0), \quad n=1, \dots, N_2.\end{aligned}$$

Таким образом, в зависимости от соотношения N_2 и V

$$\pi(n,0) = \frac{\lambda^n}{\mu^n} \frac{1}{n!} \pi(0,0), \quad n=1, \dots, N_2, \quad N_2 \leq V,$$

$$\pi(n,0) = \frac{\lambda^n}{\mu^n} \frac{1}{n!} \pi(0,0), \quad n=1, \dots, V, \quad N_2 > V,$$

$$\pi(n,0) = \frac{\lambda^n}{\mu^n} \frac{1}{V! V^{n-V}} \pi(0,0), \quad n=V+1, \dots, N_2, \quad N_2 > V.$$

г) Получение зависимости $\pi(n,0)$ от $\pi(n-1,0)$ для $N_2 < n < N_1$

Выразим $\pi(N_1-1,0)$ через $\pi(N_1-2,0)$ из уравнения СУР при $(N_1-1,0)$

$$\begin{aligned}\pi(N_1-1,0) &= \frac{\lambda}{\lambda + \min(N_1-1,V)\mu} \pi(N_1-2,0) = \\ &= \left(\frac{\lambda + \min(N_1-1,V)\mu}{\lambda} \right)^{-1} \pi(N_1-2,0) = \left(1 + \frac{\min(N_1-1,V)\mu}{\lambda} \right)^{-1} \pi(N_1-2,0).\end{aligned}$$

и обозначим $A(1) = 1 + \frac{\min(N_1-1,V)\mu}{\lambda}$, тогда

$$\pi(N_1-1,0) = \frac{1}{A(1)} \cdot \pi(N_1-2,0).$$

Выразим $\pi(n,0)$ через $\pi(n-1,0)$ для $N_2 < n < N_1$. Уравнение СУР при $(N_1-2,0)$ имеет вид

$$\lambda \pi(N_1-3,0) - [\lambda + \min(N_1-2,V)\mu] \pi(N_1-2,0) + \min(N_1-1,V)\mu \pi(N_1-1,0) = 0.$$

Подставим полученное соотношение для $(N_1 - 1, 0)$

$$\begin{aligned}
 \pi(N_1 - 2, 0) &= \frac{\lambda}{\lambda + \min(N_1 - 2, V)\mu - \min(N_1 - 1, V)\mu \frac{1}{A(1)}} \pi(N_1 - 3, 0) = \\
 &= \frac{A(1)}{A(1) + \frac{\min(N_1 - 2, V)\mu}{\lambda} A(1) - \frac{\min(N_1 - 1, V)\mu}{\lambda}} \pi(N_1 - 3, 0) = \\
 &= \frac{A(1)}{A(1) + \frac{\min(N_1 - 2, V)\mu}{\lambda} A(1) - A(1) + 1} \pi(N_1 - 3, 0) = \\
 &= \frac{A(1)}{1 + \frac{\min(N_1 - 2, V)\mu}{\lambda} A(1)} \pi(N_1 - 3, 0)
 \end{aligned}$$

и обозначим $A(2) = 1 + \frac{\min(N_1 - 2, V)\mu}{\lambda} A(1)$, тогда

$$\pi(N_1 - 2, 0) = \frac{A(1)}{A(2)} \cdot \pi(N_1 - 3, 0).$$

Предположим, что соотношение верно для $n + 1$

$$\pi(n + 1, 0) = \frac{A(N_1 - n - 2)}{A(N_1 - n - 1)} \cdot \pi(n, 0),$$

$$A(N_1 - n - 1) = 1 + \frac{\min(n + 1, V)\mu}{\lambda} A(N_1 - n - 2).$$

Подставим его в уравнение СУР при $(n, 0)$

$$\begin{aligned}
 \pi(n, 0) &= \frac{\lambda}{\lambda + \min(n, V)\mu - \min(n + 1, V)\mu \frac{A(N_1 - n - 2)}{A(N_1 - n - 1)}} \pi(n - 1, 0) = \\
 &= \frac{A(N_1 - n - 1)}{A(N_1 - n - 1) + \frac{\min(n, V)\mu}{\lambda} A(N_1 - n - 1) - \frac{\min(n + 1, V)\mu}{\lambda} A(N_1 - n - 2)} \times \\
 &\quad \times \pi(n - 1, 0) =
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{A(N_1 - n - 1)}{A(N_1 - n - 1) + \frac{\min(n, V)\mu}{\lambda} A(N_1 - n - 1) - A(N_1 - n - 1) + 1} \pi(n - 1, 0) = \\
 &= \frac{A(N_1 - n - 1)}{1 + \frac{\min(n, V)\mu}{\lambda} A(N_1 - n - 1)} \pi(n - 1, 0)
 \end{aligned}$$

и обозначим $A(N_1 - n) = 1 + \frac{\min(n, V)\mu}{\lambda} A(N_1 - n - 1)$, тогда

$$\pi(n, 0) = \frac{A(N_1 - n - 1)}{A(N_1 - n)} \cdot \pi(n - 1, 0).$$

Следовательно, сделанное предположение о зависимости $\pi(n, 0)$ и $\pi(n - 1, 0)$ верно, таким образом

$$\pi(n, 0) = \frac{A(N_1 - n - 1)}{A(N_1 - n)} \cdot \pi(n - 1, 0), \quad n = N_2 + 1, \dots, N_1 - 1,$$

$$A(1) = 1 + \frac{\min(N_1 - 1, V)\mu}{\lambda},$$

$$A(n) = 1 + \frac{\min(N_1 - n, V)\mu}{\lambda} A(n - 1), \quad n = 2, \dots, N_1 - N_2 - 1.$$

Доопределим $A(0) = 0$, тогда

$$A(n) = 1 + \frac{\min(N_1 - n, V)\mu}{\lambda} A(n - 1), \quad n = 1, \dots, N_1 - N_2 - 1.$$

д) *Выражение $\pi(n, 0)$ через $\pi(N_2, 0)$ для $N_2 < n < N_1$*

Выразим $\pi(n, 0)$ через $\pi(N_2, 0)$, рекуррентно подставляя полученное соотношение

$$\begin{aligned}
 \pi(n, 0) &= \frac{A(N_1 - n - 1)}{A(N_1 - n)} \cdot \pi(n - 1, 0) = \frac{A(N_1 - n - 1)}{A(N_1 - n)} \frac{A(N_1 - n)}{A(N_1 - n + 1)} \cdot \pi(n - 2, 0) = \dots = \\
 &= \frac{A(N_1 - n - 1)}{A(N_1 - n)} \frac{A(N_1 - n)}{A(N_1 - n + 1)} \dots \frac{A(N_1 - N_2)}{A(N_1 - N_2 - 1)} \cdot \pi(N_2, 0) = \\
 &= \frac{A(N_1 - n - 1)}{A(N_1 - N_2 - 1)} \cdot \pi(N_2, 0), \quad n = N_2 + 1, \dots, N_1 - 1.
 \end{aligned}$$

е) Получение зависимости π_n от π_{n+1} для $N_2 < n < N_1$ при $N_1 > N_2 + 1$

Выразим π_{N_2+1} через π_{N_2+2} из матричного уравнения СУР при $((N_2 + 1, 1), (N_2 + 1, 2))$

$$\pi_{N_2+1} = -\pi_{N_2+2} \mathbf{A}_{2, N_2+2} \mathbf{A}_{0, N_2+1}^{-1}$$

и обозначим $\mathbf{X}_{N_2+1} = -\mathbf{A}_{2, N_2+2} \mathbf{A}_{0, N_2+1}^{-1}$, тогда

$$\pi_{N_2+1} = \pi_{N_2+2} \mathbf{X}_{N_2+1}.$$

Получим зависимость π_n от π_{n+1} . Матричное уравнение СУР при $((N_2 + 2, 1), (N_2 + 2, 2))$ имеет вид

$$\pi_{N_2+1} \mathbf{A}_1 + \pi_{N_2+2} \mathbf{A}_{0, N_2+2} + \pi_{N_2+3} \mathbf{A}_{2, N_2+3} = \mathbf{0}.$$

Подставим полученное соотношение для π_{N_2+1}

$$\pi_{N_2+2} = -\pi_{N_2+3} \mathbf{A}_{2, N_2+3} \left(\mathbf{X}_{N_2+1} \mathbf{A}_1 + \mathbf{A}_{0, N_2+2} \right)^{-1}$$

и обозначим $\mathbf{X}_{N_2+2} = -\mathbf{A}_{2, N_2+3} \left(\mathbf{X}_{N_2+1} \mathbf{A}_1 + \mathbf{A}_{0, N_2+2} \right)^{-1}$, тогда

$$\pi_{N_2+2} = \pi_{N_2+3} \mathbf{X}_{N_2+2}.$$

Предположим, что соотношение верно для $n - 1$

$$\pi_{n-1} = \pi_n \mathbf{X}_{n-1},$$

$$\mathbf{X}_{n-1} = -\mathbf{A}_{2n} \left(\mathbf{X}_{n-2} \mathbf{A}_1 + \mathbf{A}_{0, n-1} \right)^{-1}.$$

Подставим его в матричное уравнение СУР при $((n, 1), (n, 2))$

$$\pi_n = -\pi_{n+1} \mathbf{A}_{2, n+1} \left(\mathbf{X}_{n-1} \mathbf{A}_1 + \mathbf{A}_{0n} \right)^{-1}$$

и обозначим $\mathbf{X}_n = -\mathbf{A}_{2, n+1} \left(\mathbf{X}_{n-1} \mathbf{A}_1 + \mathbf{A}_{0n} \right)^{-1}$, тогда

$$\pi_n = \pi_{n+1} \mathbf{X}_n.$$

Следовательно, сделанное предположение о зависимости π_n от π_{n+1} верно, таким образом

$$\pi_n = \pi_{n+1} \mathbf{X}_n, \quad n = N_2 + 1, \dots, N_1 - 1,$$

$$\mathbf{X}_{N_2+1} = -\mathbf{A}_{2,N_2+2} \mathbf{A}_{0,N_2+1}^{-1},$$

$$\mathbf{X}_n = -\mathbf{A}_{2,n+1} (\mathbf{X}_{n-1} \mathbf{A}_1 + \mathbf{A}_{0n})^{-1}, \quad n = N_2 + 2, \dots, N_1 - 1.$$

ж) Выражение π_n через π_{N_1} для $N_2 < n < N_1$ при $N_1 > N_2 + 1$

Выразим π_n через π_{N_1} , рекуррентно подставляя полученное соотношение

$$\pi_n = \pi_{n+1} \mathbf{X}_n = \pi_{n+2} \mathbf{X}_{n+1} \mathbf{X}_n = \dots = \pi_{N_1} \mathbf{X}_{N_1-1} \dots \mathbf{X}_{n+1} \mathbf{X}_n = \pi_{N_1} \prod_{i=0}^{N_1-n-1} \mathbf{X}_{N_1-i-1},$$

$$n = N_2 + 1, \dots, N_1 - 1.$$

з) Получение зависимости π_n от π_{n+1} и $\pi(N_1 - 1, 0)$ для $N_1 \leq n < N$ при $N_1 > N_2 + 1$

Подставим полученное соотношение для π_{N_1-1} в матричное уравнение СУР при $((N_1, 1), (N_1, 2))$

$$\begin{aligned} \pi_{N_1} &= -(\pi_{N_1+1} \mathbf{A}_{2,N_1+1} + \lambda(p, 1-p) \pi(N_1 - 1, 0)) (\mathbf{X}_{N_1-1} \mathbf{A}_1 + \mathbf{A}_{0N_1})^{-1} = \\ &= -\pi_{N_1+1} \mathbf{A}_{2,N_1+1} (\mathbf{X}_{N_1-1} \mathbf{A}_1 + \mathbf{A}_{0N_1})^{-1} - \lambda(p, 1-p) \pi(N_1 - 1, 0) (\mathbf{X}_{N_1-1} \mathbf{A}_1 + \mathbf{A}_{0N_1})^{-1}. \end{aligned}$$

Обозначим $\mathbf{X}_{N_1} = -\mathbf{A}_{2,N_1+1} (\mathbf{X}_{N_1-1} \mathbf{A}_1 + \mathbf{A}_{0N_1})^{-1}$ и

$$\mathbf{Y}_{N_1} = -\lambda(p, 1-p) (\mathbf{X}_{N_1-1} \mathbf{A}_1 + \mathbf{A}_{0N_1})^{-1}, \text{ тогда}$$

$$\pi_{N_1} = \pi_{N_1+1} \mathbf{X}_{N_1} + \pi(N_1 - 1, 0) \mathbf{Y}_{N_1}.$$

Получим зависимость π_n от π_{n+1} и $\pi(N_1 - 1, 0)$. Матричное уравнение СУР при $((N_1 + 1, 1), (N_1 + 1, 2))$ имеет вид

$$\pi_{N_1} \mathbf{A}_1 + \pi_{N_1+1} \mathbf{A}_{0,N_1+1} + \pi_{N_1+2} \mathbf{A}_{2,N_1+2} = \mathbf{0}.$$

Подставим полученное соотношение для π_{N_1}

$$\begin{aligned}\pi_{N_1+1} &= -(\pi_{N_1+2}\mathbf{A}_{2,N_1+2} + \pi(N_1-1,0)\mathbf{Y}_{N_1}\mathbf{A}_1)(\mathbf{X}_{N_1}\mathbf{A}_1 + \mathbf{A}_{0,N_1+1})^{-1} = \\ &= -\pi_{N_1+2}\mathbf{A}_{2,N_1+2}(\mathbf{X}_{N_1}\mathbf{A}_1 + \mathbf{A}_{0,N_1+1})^{-1} - \pi(N_1-1,0)\mathbf{Y}_{N_1}\mathbf{A}_1(\mathbf{X}_{N_1}\mathbf{A}_1 + \mathbf{A}_{0,N_1+1})^{-1}\end{aligned}$$

и обозначим $\mathbf{X}_{N_1+1} = -\mathbf{A}_{2,N_1+2}(\mathbf{X}_{N_1}\mathbf{A}_1 + \mathbf{A}_{0,N_1+1})^{-1}$ и

$$\mathbf{Y}_{N_1+1} = -\mathbf{Y}_{N_1}\mathbf{A}_1(\mathbf{X}_{N_1}\mathbf{A}_1 + \mathbf{A}_{0,N_1+1})^{-1}, \text{ тогда}$$

$$\pi_{N_1+1} = \pi_{N_1+2}\mathbf{X}_{N_1+1} + \pi(N_1-1,0)\mathbf{Y}_{N_1+1}.$$

Предположим, что соотношение верно для $n-1$

$$\pi_{n-1} = \pi_n\mathbf{X}_{n-1} + \pi(N_1-1,0)\mathbf{Y}_{n-1},$$

$$\mathbf{X}_{n-1} = -\mathbf{A}_{2n}(\mathbf{X}_{n-2}\mathbf{A}_1 + \mathbf{A}_{0,n-1})^{-1},$$

$$\mathbf{Y}_{n-1} = -\mathbf{Y}_{n-2}\mathbf{A}_1(\mathbf{X}_{n-2}\mathbf{A}_1 + \mathbf{A}_{0,n-1})^{-1}.$$

Подставим его в матричное уравнение СУР при $((n,1),(n,2))$

$$\begin{aligned}\pi_n &= -(\pi_{n+1}\mathbf{A}_{2,n+1} + \pi(N_1-1,0)\mathbf{Y}_{n-1}\mathbf{A}_1)(\mathbf{X}_{n-1}\mathbf{A}_1 + \mathbf{A}_{0n})^{-1} = \\ &= -\pi_{n+1}\mathbf{A}_{2,n+1}(\mathbf{X}_{n-1}\mathbf{A}_1 + \mathbf{A}_{0n})^{-1} - \pi(N_1-1,0)\mathbf{Y}_{n-1}\mathbf{A}_1(\mathbf{X}_{n-1}\mathbf{A}_1 + \mathbf{A}_{0n})^{-1}\end{aligned}$$

и обозначим $\mathbf{X}_n = -\mathbf{A}_{2,n+1}(\mathbf{X}_{n-1}\mathbf{A}_1 + \mathbf{A}_{0n})^{-1}$ и $\mathbf{Y}_n = -\mathbf{Y}_{n-1}\mathbf{A}_1(\mathbf{X}_{n-1}\mathbf{A}_1 + \mathbf{A}_{0n})^{-1}$, тогда

$$\pi_n = \pi_{n+1}\mathbf{X}_n + \pi(N_1-1,0)\mathbf{Y}_n.$$

Следовательно, сделанное предположение о зависимости π_n от π_{n+1} и $\pi(N_1-1,0)$

верно, таким образом

$$\pi_n = \pi_{n+1}\mathbf{X}_n + \pi(N_1-1,0)\mathbf{Y}_n, \quad n = N_1, \dots, N-1,$$

$$\mathbf{X}_n = -\mathbf{A}_{2,n+1}(\mathbf{X}_{n-1}\mathbf{A}_1 + \mathbf{A}_{0n})^{-1}, \quad n = N_1, \dots, N-1,$$

$$\mathbf{Y}_{N_1} = -\lambda(p, 1-p)(\mathbf{X}_{N_1-1}\mathbf{A}_1 + \mathbf{A}_{0N_1})^{-1},$$

$$\mathbf{Y}_n = -\mathbf{Y}_{n-1}\mathbf{A}_1(\mathbf{X}_{n-1}\mathbf{A}_1 + \mathbf{A}_{0n})^{-1}, \quad n = N_1+1, \dots, N-1.$$

и) Выражение π_{N_1} через π_{N_1+1} и $\pi(N_1-1,0)$ при $N_1 = N_2 + 1$

Из матричного уравнения СУР при $((N_1,1),(N_1,2))$ получим

$$\begin{aligned}\pi_{N_1} &= -(\pi_{N_1+1}\mathbf{A}_{2,N_1+1} + \lambda(p,1-p)\pi(N_1-1,0))\mathbf{A}_{0N_1}^{-1} = \\ &= -\pi_{N_1+1}\mathbf{A}_{2,N_1+1}\mathbf{A}_{0N_1}^{-1} - \lambda(p,1-p)\pi(N_1-1,0)\mathbf{A}_{0N_1}^{-1} = \\ &= \pi_{N_1+1}\mathbf{X}_{N_1} - \lambda(p,1-p)\pi(N_1-1,0)\mathbf{A}_{0N_1}^{-1}.\end{aligned}$$

Обозначим $\mathbf{Y}_{N_1} = -\lambda(p,1-p)\mathbf{A}_{0N_1}^{-1}$, тогда

$$\pi_{N_1} = \pi_{N_1+1}\mathbf{X}_{N_1} + \pi(N_1-1,0)\mathbf{Y}_{N_1}.$$

к) Выражение π_n через π_N и $\pi(N_1-1,0)$ для $N_1 \leq n < N$

Получим зависимость π_n от π_N и $\pi(N_1-1,0)$. Выразим π_{N-1} через π_N

$$\pi_{N-1} = \pi_N\mathbf{X}_{N-1} + \pi(N_1-1,0)\mathbf{Y}_{N-1}$$

и π_{N-2} через π_N

$$\begin{aligned}\pi_{N-2} &= \pi_{N-1}\mathbf{X}_{N-2} + \pi(N_1-1,0)\mathbf{Y}_{N-2} = \\ &= (\pi_N\mathbf{X}_{N-1} + \pi(N_1-1,0)\mathbf{Y}_{N-1})\mathbf{X}_{N-2} + \pi(N_1-1,0)\mathbf{Y}_{N-2} = \\ &= \pi_N\mathbf{X}_{N-1}\mathbf{X}_{N-2} + \pi(N_1-1,0)(\mathbf{Y}_{N-1}\mathbf{X}_{N-2} + \mathbf{Y}_{N-2}) = \\ &= \pi_N \prod_{i=0}^1 \mathbf{X}_{N-i-1} + \pi(N_1-1,0) \sum_{j=N-2}^{N-1} \mathbf{Y}_j \prod_{i=0}^{j-N+1} \mathbf{X}_{j-i-1}.\end{aligned}$$

Предположим, что соотношение верно для $n+1$

$$\pi_{n+1} = \pi_N \prod_{i=0}^{N-n-2} \mathbf{X}_{N-i-1} + \pi(N_1-1,0) \sum_{j=n+1}^{N-1} \mathbf{Y}_j \prod_{i=0}^{j-n-2} \mathbf{X}_{j-i-1},$$

тогда

$$\begin{aligned}\pi_n &= \pi_N \prod_{i=0}^{N-n-2} \mathbf{X}_{N-i-1} \mathbf{X}_n + \pi(N_1-1,0) \left(\sum_{j=n+1}^{N-1} \mathbf{Y}_j \prod_{i=0}^{j-n-2} \mathbf{X}_{j-i-1} \mathbf{X}_n + \mathbf{Y}_n \right) = \\ &= \pi_N \prod_{i=0}^{N-n-1} \mathbf{X}_{N-i-1} + \pi(N_1-1,0) \sum_{j=n}^{N-1} \mathbf{Y}_j \prod_{i=0}^{j-n-1} \mathbf{X}_{j-i-1}.\end{aligned}$$

Следовательно, сделанное предположение о зависимости π_n от π_N и $\pi(N_1-1,0)$ верно, таким образом

$$\pi_n = \pi_N \prod_{i=0}^{N-n-1} \mathbf{X}_{N-i-1} + \pi(N_1-1,0) \sum_{j=n}^{N-1} \mathbf{Y}_j \prod_{i=0}^{j-n-1} \mathbf{X}_{j-i-1}, \quad n = N_1, \dots, N-1.$$

л) Выражение π_n через π_N для $n > N$

Матрицы в матричных уравнениях СУР при $((n,1),(n,2))$ для $n > N$ не зависят от n , следовательно, следовательно, существует матрица $\mathbf{R}$ такая, что

$$\pi_n = \pi_{n-1} \mathbf{R}, \quad n = N+1, N+2, \dots$$

Выразим π_n через π_N , рекуррентно подставляя полученное соотношение

$$\pi_n = \pi_{n-1} \mathbf{R} = \pi_{n-2} \mathbf{R} \mathbf{R} = \dots = \pi_N \mathbf{R}^{n-N}, \quad n = N+1, N+2, \dots$$

Подставим данное соотношение в матричное уравнение СУР при $((n,1),(n,2))$

$$\pi_N \mathbf{R}^{n-1-N} \mathbf{A}_1 + \pi_N \mathbf{R}^{n-N} \mathbf{A}_{0N} + \pi_N \mathbf{R}^{n+1-N} \mathbf{A}_{2N} = \mathbf{0}, \quad n = N+1, N+2, \dots,$$

$$\pi_N \mathbf{R}^{n-1-N} (\mathbf{A}_1 + \mathbf{R} \mathbf{A}_{0N} + \mathbf{R}^2 \mathbf{A}_{2N}) = \mathbf{0}, \quad n = N+1, N+2, \dots$$

и получим квадратичное матричное уравнение относительно $\mathbf{R}$

$$\mathbf{A}_1 + \mathbf{R} \mathbf{A}_{0N} + \mathbf{R}^2 \mathbf{A}_{2N} = \mathbf{0}.$$

м) Выражение π_N через $\pi(N_1-1,0)$

Подставим полученные соотношения

$$\pi_{N-1} = \pi_N \mathbf{X}_{N-1} + \pi(N_1-1,0) \mathbf{Y}_{N-1},$$

$$\pi_{N+1} = \pi_N \mathbf{R}$$

в матричное уравнение СУР при $((N,1),(N,2))$

$$\pi_{N-1} \mathbf{A}_1 + \pi_N \mathbf{A}_{0,N+1} + \pi_{N+1} \mathbf{A}_{2,N+1} = \mathbf{0}$$

и получим

$$\pi_N = -\pi(N_1-1,0) \mathbf{Y}_{N-1} \mathbf{A}_1 (\mathbf{X}_{N-1} \mathbf{A}_1 + \mathbf{A}_{0,N+1} + \mathbf{R} \mathbf{A}_{2,N+1})^{-1}.$$

Обозначим $\mathbf{Y}_N = -\mathbf{Y}_{N-1} \mathbf{A}_1 (\mathbf{X}_{N-1} \mathbf{A}_1 + \mathbf{A}_{0,N+1} + \mathbf{R} \mathbf{A}_{2,N+1})^{-1}$, тогда

$$\pi_N = \pi(N_1-1,0) \mathbf{Y}_N.$$

н) Выражение π_n через $\pi(N_1-1,0)$ для $N_2 < n < N$

Подставим выражение π_N через $\pi(N_1-1,0)$ в π_n для $N_1 \leq n < N$

$$\begin{aligned}\pi_n &= \pi(N_1-1,0) \mathbf{Y}_N \prod_{i=0}^{N-n-1} \mathbf{X}_{N-i-1} + \pi(N_1-1,0) \sum_{j=n}^{N-1} \mathbf{Y}_j \prod_{i=0}^{j-n-1} \mathbf{X}_{j-i-1} = \\ &= \pi(N_1-1,0) \sum_{j=n}^N \mathbf{Y}_j \prod_{i=0}^{j-n-1} \mathbf{X}_{j-i-1}, \quad n = N_1, \dots, N-1,\end{aligned}$$

и объединим с соотношением для π_N

$$\pi_n = \pi(N_1-1,0) \sum_{j=n}^N \mathbf{Y}_j \prod_{i=0}^{j-n-1} \mathbf{X}_{j-i-1}, \quad n = N_1, \dots, N.$$

Подставим выражение π_{N_1} через $\pi(N_1-1,0)$ в π_n для $N_2 < n < N_1$

$$\begin{aligned}\pi_n &= \pi(N_1-1,0) \sum_{j=N_1}^N \mathbf{Y}_j \prod_{i=0}^{j-N_1-1} \mathbf{X}_{j-i-1} \prod_{i=0}^{N_1-n-1} \mathbf{X}_{N_1-i-1} = \\ &= \pi(N_1-1,0) \sum_{j=N_1}^N \mathbf{Y}_j \prod_{i=0}^{j-n-1} \mathbf{X}_{j-i-1}, \quad n = N_2 + 1, \dots, N_1 - 1.\end{aligned}$$

■

Замечание. Рекуррентные соотношения в теореме 5.4 выражают $\pi(n,u)$, $(n,u) \in \mathcal{X} \setminus \{(0,0)\}$ через $\pi(0,0)$. Вероятность $\pi(0,0)$ находится из условия нормировки $\pi \mathbf{1}^T = 1$.

5.4.3 Анализ показателей эффективности

Из леммы 5.4 и решения СУР $\pi \mathbf{Q} = \mathbf{0}$ с условием нормировки $\pi \mathbf{1}^T = 1$, вытекает следующее следствие.

Следствие 5.3. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика, следующие показатели эффективности рассчитываются по формулам:

вероятность того, что ни одна сессия потокового трафика не будет приостановлена

$$P_{\text{non-in-all}} = \sum_{m=1}^C \sum_{l=0}^{V-m} \frac{\alpha}{\lambda + (l+m)\mu + \alpha} \cdot \pi(l, m, 0, 1); \quad (5.66)$$

вероятность того, что сессия потокового трафика не будет приостановлена

$$P_{\text{non-in}} = \sum_{m=1}^C \sum_{l=0}^{V-m} \frac{\alpha}{\lambda + (l+m)\mu + \alpha} \cdot \pi(l, m, 0, 1) + \\ + \sum_{m=1}^C \sum_{l=V-m+1}^{V-1} \frac{V-l}{m} \cdot \frac{\alpha}{\lambda + (l+m)\mu + \alpha} \cdot \pi(l, m, 0, 1); \quad (5.67)$$

вероятность блокировки запроса на передачу потокового трафика

$$P_{\text{bl}} = \pi(V, 0, R, 0) + \pi(V, C, R - C, 1). \quad (5.68)$$

Проиллюстрируем поведение показателей эффективности обслуживания потокового трафика на примере сценария радиотелеметрии с самолетами. На рисунках 5.2, 5.3 и 5.4 представлены соответственно графики вероятности того, что ни одна сессия и конкретная сессия потокового трафика не будет приостановлена, и вероятность блокировки запроса на передачу потокового трафика в зависимости от предложенной нагрузки для разного среднего времени полной доступности диапазона радиочастот.

Графики на рисунках 5.2 и 5.3 показывают немонотонную зависимость вероятности «неприостановки» сессии от предложенной нагрузки. При низкой нагрузке индивидуальный диапазон радиочастот эффективно обрабатывает запросы, а LSA-диапазон используется редко, что снижает вероятность прерывания. С увеличением нагрузки LSA-диапазон активизируется, что повышает риск прерываний из-за своей ненадежности. Локальный экстремум графика отражает оптимальный баланс между диапазонами. При дальнейшем росте нагрузки вероятность «неприостановки» сессии падает, так как LSA-диапазон чаще отключается, и очередь переполняется. График также демонстрирует, что увеличение среднего времени между отказами улучшает качество обслуживания, смещая кривую вверх.

Из рисунка 5.4 видно, что вероятность блокировки монотонно растет с увеличением нагрузки. На низких уровнях нагрузки буфер успевает

освобождаться, и запросы редко теряются. С приближением к предельной емкости системы очередь заполняется, и вероятность блокировки резко возрастает. При высокой нагрузке даже небольшое увеличение интенсивности нагрузки вызывает значительный скачок блокировок.

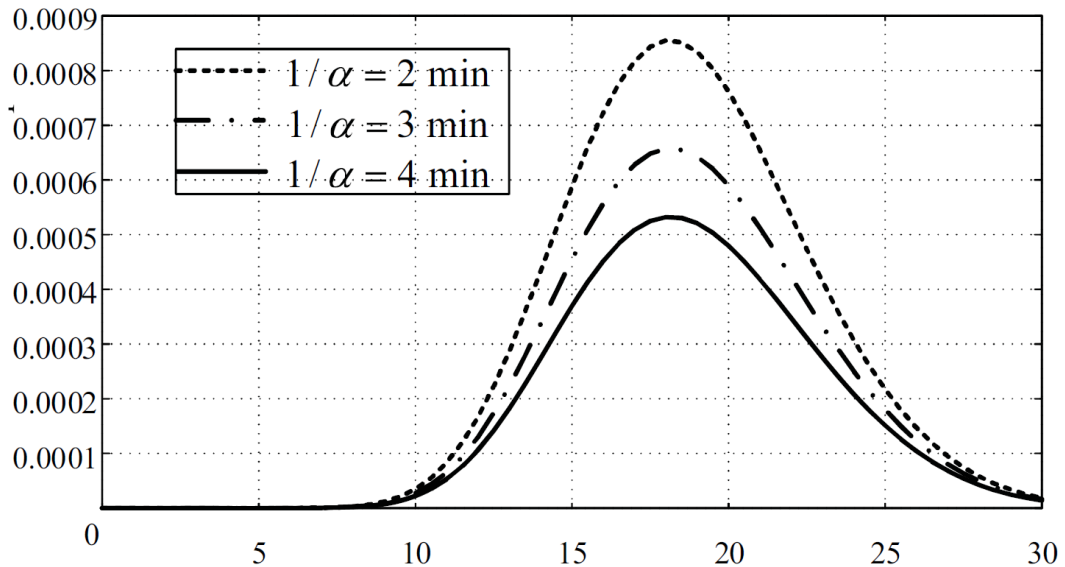


Рисунок 5.2 – Вероятность того, что ни одна сессия потокового трафика не будет приостановлена в зависимости от предложенной нагрузки (Эрл) для разного среднего времени полной доступности диапазона радиочастот (мин)

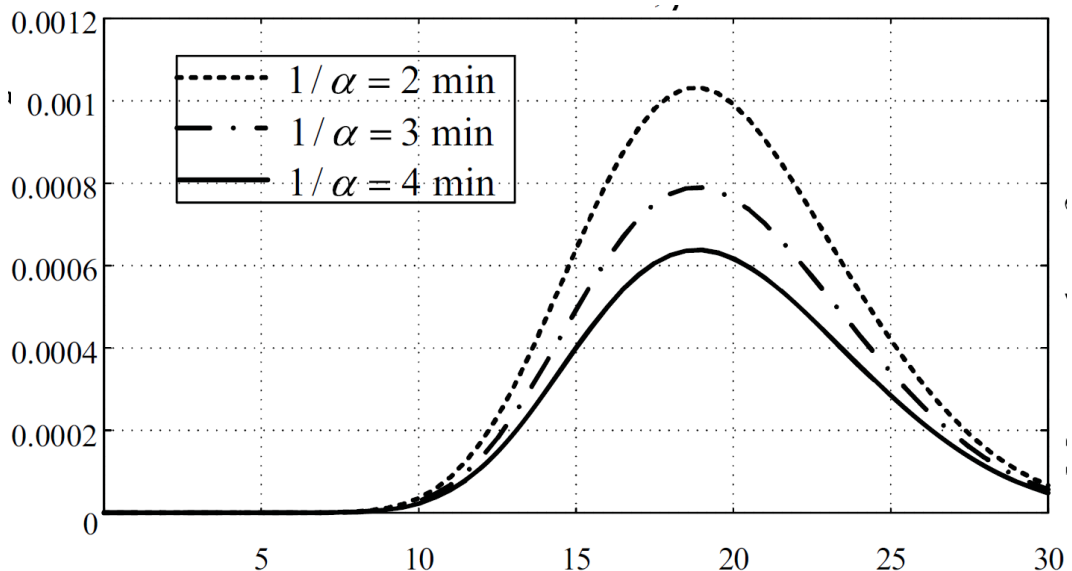


Рисунок 5.3 – Вероятность того, что сессия потокового трафика не будет приостановлена в зависимости от предложенной нагрузки (Эрл) для разного среднего времени полной доступности диапазона радиочастот (мин)

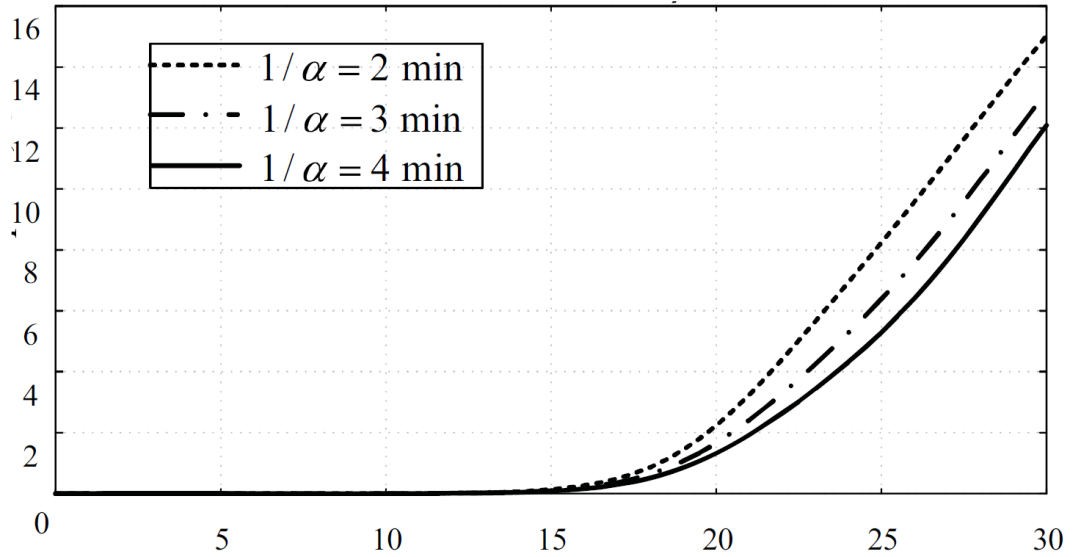


Рисунок 5.4 – Вероятность блокировки запроса на передачу потокового трафика в зависимости от предложенной нагрузки (Эрл) для разного среднего времени полной доступности диапазона радиочастот (мин)

5.5 УПРАВЛЯЕМАЯ СИСТЕМА ГИСТЕРЕЗИСНОГО ДОСТУПА К НЕНАДЕЖНОМУ РЕСУРСУ

5.5.1 Задача выбора порогов гистерезисного доступа как политики управления

Политикой назовем тройку параметров системы: объем ненадежного ресурса и пороги гистерезисного управления на выключение и включение ненадежного ресурса

$$f = \mathbf{a} = (C, N_1, N_2). \quad (5.69)$$

Тогда пространство состояний зависит от политики f

$$\mathcal{X}_f = \{(n, 0): 0 \leq n \leq N_1 - 1\} \cup \{(n, u): n \geq N_2 + 1, u = 1, 2\}, \quad (5.70)$$

а допустимое множество параметров системы выглядит следующим образом

$$\begin{aligned} \mathcal{A} = & \{(C, N_1, N_2): 0 \leq C \leq V - 1, 0 \leq N_2 < V - 1, N_1 \geq V\} \cup \\ & \cup \{(C, N_1, N_2): 0 \leq C \leq V - 1, N_2 \geq V - 1, N_1 \geq N_2 + 1\}. \end{aligned} \quad (5.71)$$

Ненулевые элементы матрицы интенсивностей переходов при политике f имеют вид, как записано в формуле (5.10). Функция стоимости в единицу времени за пребывание в состояниях системы определяется следующей леммой.

Лемма 5.6. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и гистерезисным доступом, функция стоимости в единицу времени за пребывание в состояниях системы имеет вид

$$\begin{aligned} g_f(\mathbf{x}) = g_f(n, u) = & \omega_1 \cdot n + \omega_2 \cdot (V + C \cdot 1\{u = 1, 2\}) + \\ & + \omega_3 \cdot \lambda \cdot 1\{(n, u) = (N_1 - 1, 0)\} + \\ & + \omega_4 \cdot \min(N_2 + 1, V + C) \mu \cdot 1\{(n, u) = (N_2 + 1, 1)\} + \\ & + \omega_4 \cdot \min(N_2 + 1, V) \mu \cdot 1\{(n, u) = (N_2 + 1, 2)\}, \quad \mathbf{x} = (n, u) \in \mathcal{X}_f. \end{aligned} \quad (5.72)$$

Доказательство.

Первое слагаемое в функции стоимости – это стоимость ω_1 за каждую из сессий трафика в системе, обслуживаемых на ненадежном и надежном ресурсах и в очереди. Второе слагаемое – стоимость ω_2 за использование надежного и ненадежного ресурса. Третье слагаемое – стоимость ω_3 за включение надежного ресурса. Последние два слагаемых – стоимость ω_4 за выключение надежного ресурса. ■

Тогда с учетом леммы 5.6, справедлива следующая теорема.

Теорема 5.5. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и гистерезисным доступом, средняя стоимость в единицу времени рассчитывается по формуле

$$\bar{g}_f = \sum_{\mathbf{x} \in \mathcal{X}_f} g_f(\mathbf{x}) \pi_f(\mathbf{x}), \quad (5.73)$$

где $\pi_f(\mathbf{x})$ стационарное распределение при политике f .

Доказательство.

Среднее стоимость в единицу времени будем использовать как целевую функцию в задаче оптимизации для нахождения оптимальной политики.

■

Оптимальной политикой назовем политику, которая минимизирует функцию стоимости

$$f^* = \arg \min_f \bar{g}_f. \quad (5.74)$$

Стационарное распределение при оптимальной политике f^* обозначим

$$\pi^*(x) = \pi_{f^*}(x), \quad x \in \mathcal{X}. \quad (5.75)$$

5.5.2 Применение генетического алгоритма для нахождения политики

Применим генетический алгоритм для расчета оптимальной политики f^* как решения задачи оптимизации (5.74). В терминологии генетического алгоритма, основные его шаги следующие:

- 1: Инициализация популяции – генерация начального набора случайных решений (особей), представляющих возможные варианты задачи.
- 2: Определение функции приспособленности – задание метрики, которая количественно оценивает качество каждой особи (например, минимизация стоимости или максимизация эффективности).
- 3: Оценка приспособленности – расчет значения функции приспособленности для каждой особи в текущей популяции.
- 4: Селекция родителей – выбор особей с высокой приспособленностью для участия в создании потомков (например, методом рулетки или турнира).
- 5: Скрещивание (кроссовер) – комбинирование генетической информации двух родителей для генерации новых решений - потомков.
- 6: Мутация – внесение случайных изменений в гены потомков для увеличения разнообразия популяции.

- 7: Формирование новой популяции – создание следующего поколения из потомков, лучших особей предыдущего поколения (элитизм) и, возможно, новых случайных решений.
- 8: Проверка условий останова – прекращение работы алгоритма при достижении критериев: максимальное число поколений, отсутствие улучшений или достаточная точность решения.

Представим политику

$$\begin{aligned}
 f = \mathbf{a} &= (a_1, a_2, a_3) = (C, N_1, N_2), \\
 a_i &\in \{A_i^{\min}, \dots, A_i^{\max}\}, \quad i = 1, 2, 3, \\
 A_1^{\min} &= 0, \quad A_1^{\max} = V - 1, \\
 (A_2^{\min} = V, \quad a_3 \leq V - 2) &\vee (A_2^{\min} = N_2 + 1, \quad a_3 \geq V - 1), \\
 A_3^{\min} &= 0
 \end{aligned} \tag{5.76}$$

в виде битовой строки, состоящей из трех бинарных сегментов

$$\begin{aligned}
 \mathbf{y} &= (y_1, y_2, y_3), \\
 y_i &= (y_{i, L_i-1}, y_{i, L_i-2}, \dots, y_{i1}, y_{i0}), \quad i = 1, 2, 3, \\
 y_{ij} &\in \{0, 1\}, \quad j = 0, \dots, L_i - 1, \quad i = 1, 2, 3,
 \end{aligned} \tag{5.77}$$

где L_i количество бит, выделенных для кодирования параметра a_i с точностью $0 < \varepsilon < 1$, выбирается исходя из соотношения

$$2^{L_i} < \frac{A_i^{\max} - A_i^{\min}}{\varepsilon} < 2^{L_i} - 1, \quad i = 1, 2, 3.$$

Отметим, что преобразование из битовой строки $\mathbf{y}$ в политику $f = \mathbf{a}$ выглядит следующим образом

$$a_i = A_i^{\min} + z_i \frac{A_i^{\max} - A_i^{\min}}{2^{L_i} - 1}, \quad i = 1, 2, 3, \tag{5.78}$$

где z_i десятичное число, полученное из бинарного сегмента y_i

$$z_i = \sum_{j=0}^{L_i-1} y_{ij} 2^j, \quad i = 1, 2, 3. \tag{5.79}$$

Таким образом, алгоритм расчета оптимальной политики можно представить следующим образом.

Алгоритм 5.1. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с приостановкой обслуживания трафика и гистерезисным доступом, расчет оптимальной политики по генетическому алгоритму записывается следующим образом:

1: Инициализация популяции – сгенерировать K битовых строк

$$\mathbf{y}^1 = (\mathbf{y}_1^1, \mathbf{y}_2^1, \mathbf{y}_3^1), \dots, \mathbf{y}^k = (\mathbf{y}_1^k, \mathbf{y}_2^k, \mathbf{y}_3^k), \dots, \mathbf{y}^K = (\mathbf{y}_1^K, \mathbf{y}_2^K, \mathbf{y}_3^K) \quad (5.80)$$

и преобразовать их в политики

$$\mathbf{f}_1 = \mathbf{a}_1 = (a_{11}, a_{12}, a_{13}), \dots, \mathbf{f}_k = \mathbf{a}_k = (a_{k1}, a_{k2}, a_{k3}), \dots, \mathbf{f}_K = \mathbf{a}_K = (a_{K1}, a_{K2}, a_{K3}). \quad (5.81)$$

2: Оценка приспособленности – рассчитать целевую функцию, преобразованную из задачи минимизации в максимизацию. Генетические алгоритмы ориентированы на максимизацию функции приспособленности, однако функция стоимости $\bar{g}_{f_k}$ минимизируется. Поэтому, чтобы получить задачу максимизации, ввести другую функцию

$$\tilde{g}_{f_k} = G - \bar{g}_{f_k} > 0, \quad k = 1, \dots, K, \quad (5.82)$$

где G константа, такая что $\tilde{g}_{f_k} > 0$.

3: Селекция родителей – рулеточная селекция. Чтобы исключить возможные отрицательные значения функции приспособленности и сместить ее в неотрицательную область, вычесть ее минимальное значение по популяции.

Рассчитать вероятность выбора особи-строки

$$p_k = \frac{\tilde{g}_{f_k} - \min_{k'=1, \dots, K} \tilde{g}_{f_{k'}}}{\sum_{k'=1}^K \left(\tilde{g}_{f_{k'}} - \min_{k'=1, \dots, K} \tilde{g}_{f_{k'}} \right)}, \quad k = 1, \dots, K. \quad (5.83)$$

Вычислить кумулятивные вероятности как накопленную сумму

$$q_k = \sum_{i=1}^k p_i, \quad k=1, \dots, K. \quad (5.84)$$

Сгенерировать случайное число $r \in [0,1]$, и выбрать в качестве родителя строку $k = 2, \dots, K$, если $r \in (q_{k-1}, q_k]$, или строку $k = 1$, если $r \in [0, q_1]$.

- 4: Скрещивание (кроссовер) – двухточечный кроссовер. Выбрать две случайные точки разрыва в битовой строке. Сделать обмен между родителями сегментов между точками.
- 5: Мутация – инверсия битов. Инвертировать каждый бит в особи независимо с вероятностью p_m .
- 6: Формирование новой популяции – генерационная замена с частичным сохранением элитных особей. Перенести лучшие особи в новое поколение без изменений. Заполнить потомками остальные места.
- 7: Проверка условий остановки – фиксированное число поколений. Остановить алгоритм после заданного числа итераций.

5.5.3 Анализ функции стоимости использования ненадежного ресурса

Проиллюстрируем поведение функции стоимости и ее влияние на выбор параметров оптимальной политики. На рисунках 5.5, 5.6 и 5.7 представлены соответственно графики функции стоимости в зависимости от порогов гистерезисного управления на включение и выключение ненадежного ресурса для разного объема ненадежного ресурса и в зависимости от объема ненадежного ресурса для разной интенсивности поступления запросов на передачу потокового трафика.

Из рисунка 5.5 видно, что для малого объема ненадежного ресурса минимальная стоимость достигается при относительно низких значениях порога его включения. При увеличении объема оптимальный порог смещается вправо. Это объясняется компромиссом между задержкой запросов и использованием

ненадежного ресурса. Большой объем ненадежного ресурса позволяет активировать его позже, что снижает частоту переключений и стоимость. Кроме того, при малом объеме кривая стоимости после оптимального значения растет медленнее, что указывает на меньшую чувствительность к превышению порога на включение ресурса. Ограниченный объем слабее влияет на общую нагрузку, снижая риски.

Рисунок 5.6 демонстрирует, что с ростом объема ненадежного ресурса оптимальный порог на выключение ресурса снижается. Это объясняется тем, что большой объем позволяет раньше его отключать, сокращая время его активной работы и связанную с этим стоимость. При малом объеме кривая стоимости становится более пологой, что указывает на меньшие риски преждевременного выключения ресурса. Даже при неоптимальном выборе порога на выключение система сохраняет приемлемую эффективность благодаря небольшому объему ненадежного ресурса.

Из рисунка 5.7 видно, что с уменьшением интенсивности поступления запросов на передачу потокового трафика, оптимальный объем ненадежного ресурса. Это связано с тем, что при меньшей нагрузке система может использовать больше надежный ресурс без риска переполнения очереди, так как надежный ресурс успешно обрабатывает запросы. В условиях интенсивного потока выгоднее минимизировать объем ненадежного ресурса, чтобы избежать излишней стоимости его обслуживания.

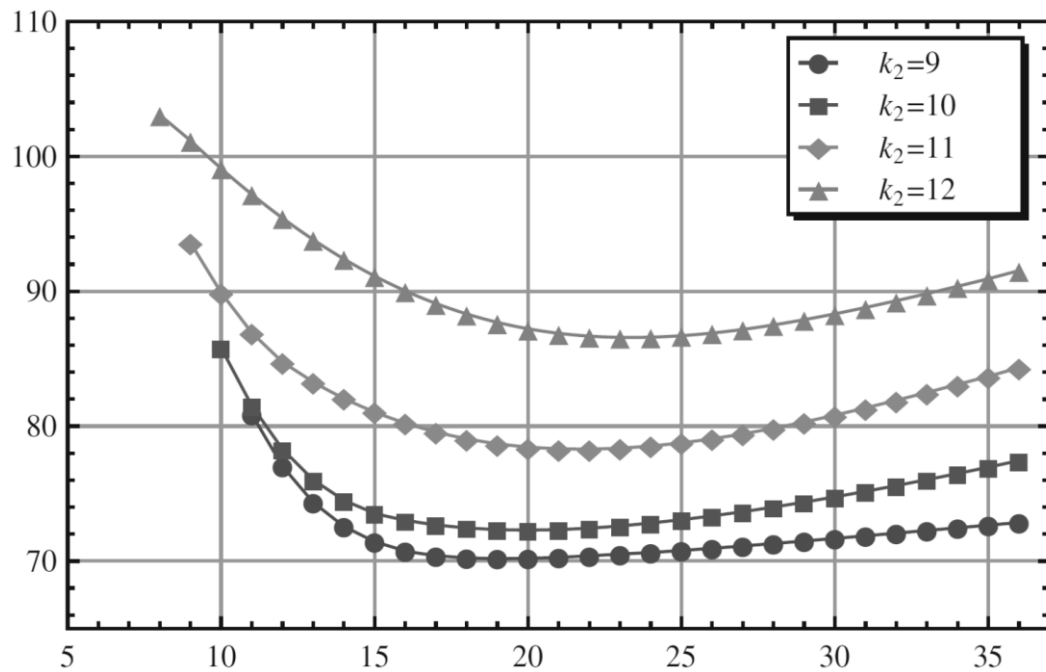


Рисунок 5.5 – Функция стоимости в зависимости от порога гистерезисного управления на включение ненадежного ресурса для разного объема ненадежного ресурса

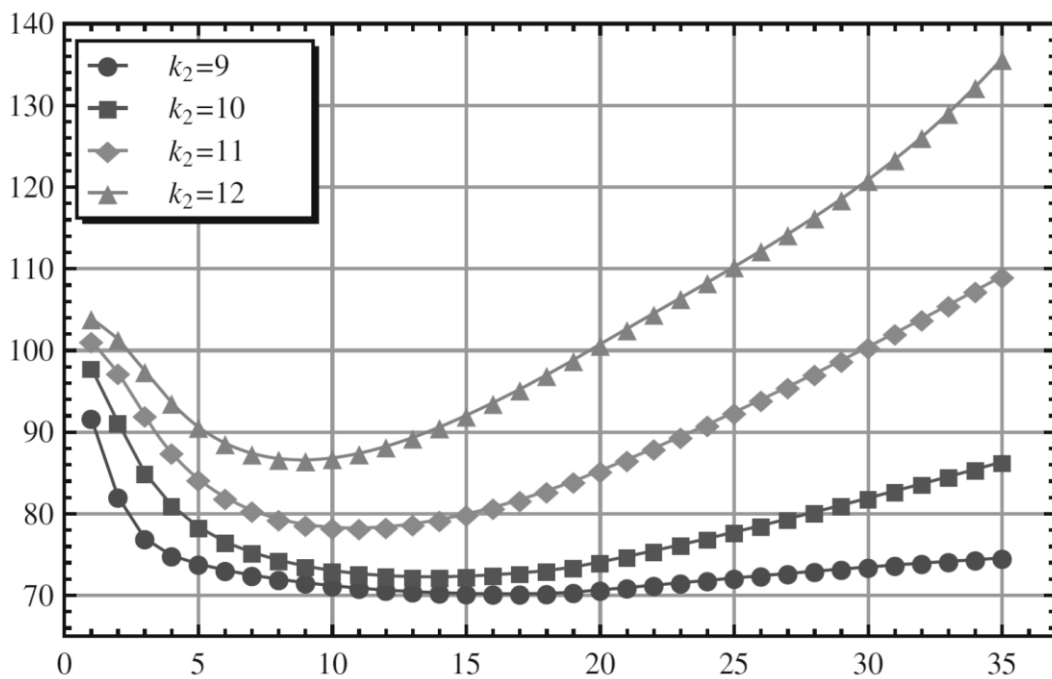


Рисунок 5.6 – Функция стоимости в зависимости от порога гистерезисного управления на выключение ненадежного ресурса для разного объема ненадежного ресурса

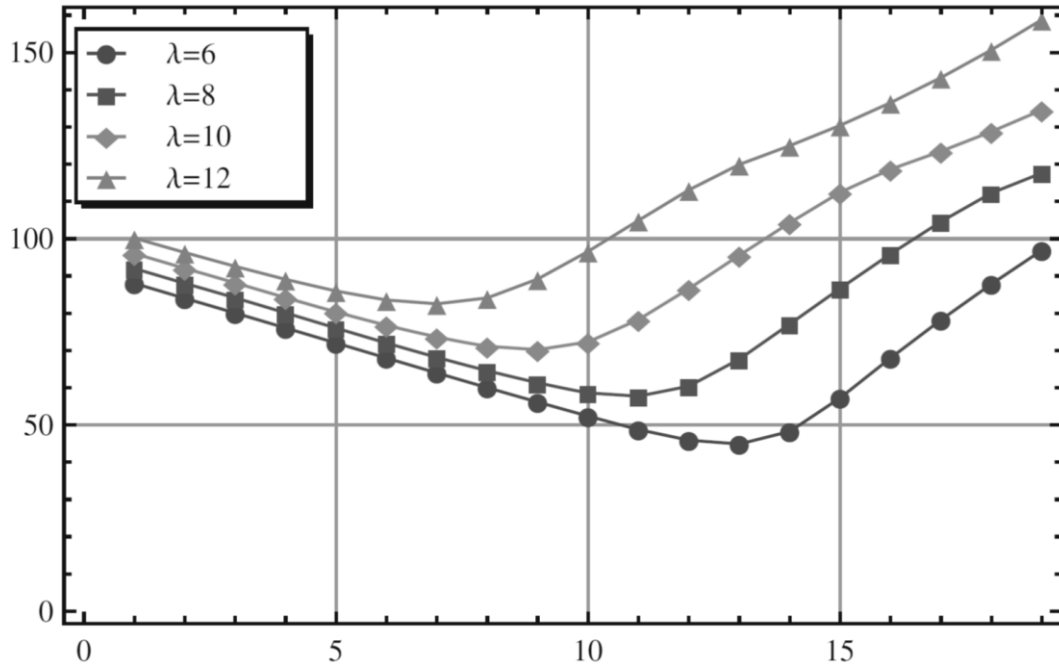


Рисунок 5.7 – Функция стоимости в зависимости от объема ненадежного ресурса для разной интенсивности поступления запросов на передачу потокового трафика

5.6 РЕСУРСНАЯ СИСТЕМА С ЗАВИСИМЫМ ЗАНЯТИЕМ РЕСУРСА

5.6.1 Модель требования к ресурсу

Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением мощности и скорости передачи трафика, требование к ресурсу для сессии потокового трафика в состоянии $(n,1)$ имеет вид

$$y(n,1) = \left(y_f = F, y_t = \frac{1}{n}, y_p = \frac{PL(N_0 + I_0)}{G_t G_r} (2^{bn/F} - 1) \right), \quad (n,1) \in \mathcal{X} \setminus \{(0,1)\}, \quad (5.85)$$

$$y(0,1) = (0,0,0).$$

Выведем вид требования к ресурсу для сессии потокового трафика в состоянии $(n,0)$. Для этого найдем максимальную мощность затухания сигнала, при которой для n пользователей достигается скорость b передачи трафика при мощности P_1 передающей антенны

$$b = \frac{1}{n} \cdot F \cdot \log_2 \left(1 + P_1 \cdot \frac{G_t G_r}{\text{PL}_{\max}(n)} \frac{1}{N_0 + I_0} \right), \quad n > 0,$$

откуда

$$\text{PL}_{\max}(n) = \frac{P_1 G_t G_r}{N_0 + I_0} (2^{bn/F} - 1), \quad n > 0.$$

Тогда скорость передачи трафика для n пользователей при мощности $\text{PL}_{\max}(n)$ затухания сигнала и мощности $P_0 < P_1$ передающей антенны имеет вид

$$\begin{aligned} b(n) &= \frac{1}{n} \cdot F \cdot \log_2 \left(1 + P_0 \cdot \frac{G_t G_r}{\text{PL}_{\max}(n)} \frac{1}{N_0 + I_0} \right) = \\ &= \frac{F}{n} \log_2 \left(1 + \frac{P_0 G_t G_r}{N_0 + I_0} \cdot \frac{N_0 + I_0}{P_1 G_t G_r} \frac{1}{2^{bn/F} - 1} \right) = \frac{F}{n} \log_2 \left(1 + \frac{P_0}{P_1} \frac{1}{2^{bn/F} - 1} \right), \quad n > 0. \end{aligned} \quad (5.86)$$

Поэтому требование к мощности передающей антенны равняется

$$\begin{aligned} y_p &= \frac{\text{PL}(N_0 + I_0)}{G_t G_r} \left[2^{\frac{b(n)n}{F}} - 1 \right] = \frac{\text{PL}(N_0 + I_0)}{G_t G_r} \left[2^{\frac{n}{F} \cdot F \cdot \log_2 \left(1 + \frac{P_0}{P_1} \frac{1}{2^{bn/F} - 1} \right)} - 1 \right] = \\ &= \frac{\text{PL}(N_0 + I_0)}{G_t G_r} \cdot \frac{P_0}{P_1} \frac{1}{2^{bn/F} - 1}. \end{aligned}$$

Следовательно, требование к ресурсу для сессии потокового трафика в состоянии $(n, 0)$ имеет вид

$$\begin{aligned} y(n, 0) &= \left(y_f = F, y_t = \frac{1}{n}, y_p = \frac{P_0}{P_1} \frac{\text{PL}(N_0 + I_0)}{G_t G_r} \frac{1}{2^{bn/F} - 1} \right), \quad (n, 0) \in \mathcal{X} \setminus \{(0, 0)\}, \\ y(0, 0) &= (0, 0, 0). \end{aligned} \quad (5.87)$$

Тогда множество вариантов занятия ресурса в состоянии (n, u) представимо в виде

$$\begin{aligned}
 \mathcal{Y}(n, 0) &= \left\{ (y_{01}, \dots, y_{0n}) = y_0[n]: y_{0i} = (y_{0i}^f, y_{0i}^t, y_{0i}^p), \right. \\
 &\quad y_{0i}^f = F, y_{0i}^t = \frac{1}{n}, y_{0i}^p = \frac{P_0}{P_1} \frac{PL_{0i}(N_0 + I_0)}{G_t G_r} \frac{1}{2^{bn/F} - 1}, \\
 &\quad PL_{0i} \geq 1, y_{0i}^p \leq P_0, i = 1, \dots, n \left. \right\}, \quad (n, 0) \in \mathcal{X} \setminus \{(0, 0)\}, \\
 \mathcal{Y}(n, 1) &= \left\{ (y_{11}, \dots, y_{1n}) = y_1[n]: y_{1i} = (y_{1i}^f, y_{1i}^t, y_{1i}^p), \right. \\
 &\quad y_{1i}^f = F, y_{1i}^t = \frac{1}{n}, y_{1i}^p = \frac{PL_{1i}(N_0 + I_0)}{G_t G_r} (2^{bn/F} - 1), \\
 &\quad PL_{1i} \geq 1, y_{1i}^p \leq P_1, i = 1, \dots, n \left. \right\}, \quad (n, 1) \in \mathcal{X} \setminus \{(0, 1)\}, \\
 \mathcal{Y}(0, 0) &= \mathcal{Y}(0, 1) = \{(0, 0, 0)\}, \\
 \mathcal{Y}(n, u) &= \emptyset, \quad (n, u) \notin \mathcal{X}.
 \end{aligned} \tag{5.88}$$

Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с постоянной мощностью и прерыванием обслуживания трафика, требование к ресурсу для сессии потокового трафика в состоянии (n, u) имеет вид

$$\begin{aligned}
 y(n, u) &= \left(y_f = F, y_t = b \left[F \log_2 \left(1 + \frac{P_u G_t G_r}{PL(N_0 + I_0)} \right) \right]^{-1}, y_p = P_u \right), \\
 (n, u) &\in \mathcal{X} \setminus \{(0, 0), (0, 1)\}, \\
 y(0, 0) &= y(0, 1) = (0, 0, 0).
 \end{aligned} \tag{5.89}$$

Тогда множество вариантов занятия ресурса в состоянии (n, u) представимо в виде

$$\begin{aligned}
 \mathcal{Y}(n, u) &= \left\{ (\mathbf{y}_{u1}, \dots, \mathbf{y}_{un}) = \mathbf{y}_u[n] : \mathbf{y}_{ui} = (y_{ui}^f, y_{ui}^t, y_{ui}^p), \right. \\
 &\quad y_{ui}^f = F, \quad y_{ui}^t = b \left[F \log_2 \left(1 + \frac{P_u G_t G_r}{\text{PL}_{ui} (N_0 + I_0)} \right) \right]^{-1}, \quad y_{ui}^p = P_u, \\
 &\quad \left. \text{PL}_{ui} \geq 1, \quad i = 1, \dots, n, \quad \sum_{i=1}^n y_{ui}^t \leq 1 \right\}, \quad (n, u) \in \mathcal{X} \setminus \{(0, 0), (0, 1)\}, \\
 \mathcal{Y}(0, 0) &= \mathcal{Y}(0, 1) = \{(0, 0, 0)\}, \\
 \mathcal{Y}(n, u) &= \emptyset, \quad (n, u) \notin \mathcal{X}.
 \end{aligned} \tag{5.90}$$

5.6.2 Вероятности занятия ресурса

Пользователи случайным образом располагаются в зоне покрытия базовой станции, поэтому мощность ξ_{PL} затухания сигнала с ФР $F_{\xi_{\text{PL}}}(x)$ и требование к ресурсу $\xi = (\xi_f, \xi_t, \xi_p)$ также являются СВ. Модель затухания сигнала и вид $F_{\xi_{\text{PL}}}(x)$ описаны в разделе 1.3.2.

Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением мощности и скорости передачи трафика, тогда справедлива следующая лемма.

Лемма 5.7. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением мощности и скорости передачи трафика, вероятности занятия ресурса рассчитываются по формулам

$$\Pr(\xi_0[n] \in \mathcal{Y}(n, 0)) = \left[F_{\xi_{\text{PL}}} \left(\frac{P_1 G_t G_r}{N_0 + I_0} (2^{bn/F} - 1) \right) \right]^n, \quad (n, 0) \in \mathcal{X}, \tag{5.91}$$

$$\Pr(\xi_1[n] \in \mathcal{Y}(n, 1)) = \left[F_{\xi_{\text{PL}}} \left(\frac{P_1 G_t G_r}{N_0 + I_0} \frac{1}{2^{bn/F} - 1} \right) \right]^n, \quad (n, 1) \in \mathcal{X}. \tag{5.92}$$

Доказательство.

Для состояний $(n,1)$ доказательство можно провести аналогично лемме 2.3.

Для состояний $(n,0)$ вывод следующий

$$\begin{aligned}
 \Pr(\xi_0[n] \in \mathcal{Y}(n,0)) &= \Pr(\xi_{0i}^p \leq P_0, i=1, \dots, n) = \prod_{i=1}^n \Pr(\xi_{0i}^p \leq P_0) = \\
 &= \prod_{i=1}^n \Pr\left(\frac{P_0}{P_1} \frac{\xi_{0i}^{\text{PL}} (N_0 + I_0)}{G_t G_r} \frac{1}{2^{bn/F} - 1} \leq P_0\right) = \\
 &= \prod_{i=1}^n \Pr\left(\xi_{0i}^{\text{PL}} \leq \frac{P_1 G_t G_r}{N_0 + I_0} (2^{bn/F} - 1)\right) = \\
 &= \prod_{i=1}^n F_{\xi_{0i}^{\text{PL}}} \left(\frac{P_1 G_t G_r}{N_0 + I_0} (2^{bn/F} - 1) \right) = \\
 &= \left[F_{\xi_{\text{PL}}} \left(\frac{P_1 G_t G_r}{N_0 + I_0} (2^{bn/F} - 1) \right) \right]^n, \quad (n,0) \in \mathcal{X}.
 \end{aligned}$$

■

Из леммы 5.7 вытекает следующее следствие.

Следствие 5.4. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением мощности и скорости передачи трафика, если все устройства пользователей расположены равномерно в зоне покрытия базовой станции и при любом своем местоположении находятся в прямой видимости, тогда вероятности занятия ресурса рассчитываются по формулам

$$\Pr(\xi_0[n] \in \mathcal{Y}(n,0)) = \left[\frac{1}{R^2} \left(\frac{P_1 G_t G_r}{N_0 + I_0} \frac{2^{bn/F} - 1}{\alpha} \right)^{2/\beta} - \left(\frac{h_{\text{BS}} - h_{\text{UT}}}{R} \right)^2 \right]^n, \quad (n,0) \in \mathcal{X}, \quad (5.93)$$

$$\Pr(\xi_1[n] \in \mathcal{Y}(n,1)) = \left[\frac{1}{R^2} \left(\frac{P_1 G_t G_r}{N_0 + I_0} \frac{1}{2^{bn/F} - 1} \frac{1}{\alpha} \right)^{2/\beta} - \left(\frac{h_{\text{BS}} - h_{\text{UT}}}{R} \right)^2 \right]^n, \quad (n,1) \in \mathcal{X}. \quad (5.94)$$

Доказательство.

Доказательство проводится с применением следствия 2.4.

■

Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с постоянной мощностью и прерыванием обслуживания трафика, справедлива следующая лемма.

Лемма 5.8. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с постоянной мощностью и прерыванием обслуживания трафика, вероятности занятия ресурса рассчитываются по формулам

$$\begin{aligned} \Pr(\xi_u[0] \in \mathcal{Y}(0, u)) &= 1, \quad u = 0, 1, \\ \Pr(\xi_u[n] \in \mathcal{Y}(n, u)) &\approx \Phi\left(\frac{1 - \mu(u) \cdot n}{\sigma(u) \cdot \sqrt{n}}\right), \quad (n, u) \in \mathcal{X} \setminus \{(0, 0), (0, 1)\}, \end{aligned} \quad (5.95)$$

где математическое ожидание $\mu(u)$ и дисперсия $\sigma^2(u)$ мощности затухания сигнала вычисляются соответственно по формулам

$$\mu(u) = \frac{b}{F} \int_{\text{PL}_{\min}}^{\text{PL}_{\max}} \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \frac{1}{x} \right) \right]^{-1} dF_{\xi_{\text{PL}}}(x), \quad u = 0, 1, \quad (5.96)$$

$$\sigma^2(u) = \left(\frac{b}{F} \right)^2 \int_{\text{PL}_{\min}}^{\text{PL}_{\max}} \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \frac{1}{x} \right) \right]^{-2} dF_{\xi_{\text{PL}}}(x) - [\mu(u)]^2, \quad u = 0, 1, \quad (5.97)$$

а $\Phi(x)$ ФР для стандартного нормального распределения.

Доказательство.

а) Выражение ФР ξ_{ui}^t через ФР ξ_{PL}

Из вида (5.90) множества $\mathcal{Y}(n, u)$ вариантов занятия ресурса, получим

$$\Pr(\xi_u[n] \in \mathcal{Y}(n, u)) = \Pr\left(\sum_{i=1}^n \xi_{ui}^t \leq 1\right), \quad n > 0, \quad (n, u) \in \mathcal{X}.$$

Выразим ФР СВ ξ_{ui}^t через ФР СВ ξ_{PL} . Поскольку $\xi_{\text{PL}} \in [\text{PL}_{\min}, \text{PL}_{\max}]$, поэтому

$\xi_{ui}^t \in [T_{\min}(u), T_{\max}(u)]$, где

$$T_{\min}(u) = \frac{b}{F} \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \frac{1}{\text{PL}_{\min}} \right) \right]^{-1},$$

$$T_{\max}(u) = \frac{b}{F} \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \frac{1}{\text{PL}_{\max}} \right) \right]^{-1}.$$

Следовательно,

$$\begin{aligned}
 F_{\xi_{ui}^t}(x) &= \Pr(\xi_{ui}^t \leq x) = \Pr\left(b \left[F \log_2 \left(1 + \frac{P_u G_t G_r}{\xi_{ui}^{\text{PL}} (N_0 + I_0)} \right) \right]^{-1} \leq x \right) = \\
 &= \Pr\left(\log_2 \left(1 + \frac{P_u G_t G_r}{\xi_{ui}^{\text{PL}} (N_0 + I_0)} \right) \geq \frac{b}{Fx} \right) = \Pr\left(\frac{P_u G_t G_r}{\xi_{ui}^{\text{PL}} (N_0 + I_0)} \geq 2^{b/Fx} - 1 \right) = \\
 &= \Pr\left(\xi_{ui}^{\text{PL}} \leq \frac{P_u G_t G_r}{N_0 + I_0} \frac{1}{2^{b/Fx} - 1} \right), \\
 x &\in [T_{\min}(u), T_{\max}(u)], \quad i=1, \dots, n, \quad n > 0, \quad (n, u) \in \mathcal{X}.
 \end{aligned}$$

Полученная вероятность представляет собой ФР СВ $\xi_{\bullet}^{\text{PL}}$, которые являются одинаково распределенными СВ с ФР $F_{\xi_{\text{PL}}}^{\text{PL}}(x)$, поэтому

$$\begin{aligned}
 F_{\xi_{ui}^t}(x) &= F_{\xi_{ui}^{\text{PL}}}^{\text{PL}}\left(\frac{P_u G_t G_r}{N_0 + I_0} \frac{1}{2^{b/Fx} - 1}\right) = F_{\xi_{\text{PL}}}^{\text{PL}}\left(\frac{P_u G_t G_r}{N_0 + I_0} \frac{1}{2^{b/Fx} - 1}\right) = F_{\xi_u^t}(x), \\
 x &\in [T_{\min}(u), T_{\max}(u)], \quad i=1, \dots, n, \quad n > 0, \quad (n, u) \in \mathcal{X}.
 \end{aligned}$$

б) *Выражение математического ожидания ξ_{ui}^t через ФР ξ_{PL}*

Выразим математическое ожидание СВ ξ_{ui}^t через ФР СВ ξ_{PL}

$$\begin{aligned}
 E[\xi_{ui}^t] &= \int_{T_{\min}(u)}^{T_{\max}(u)} x dF_{\xi_{ui}^t}(x) = \int_{T_{\min}(u)}^{T_{\max}(u)} x dF_{\xi_{\text{PL}}}^{\text{PL}}\left(\frac{P_u G_t G_r}{N_0 + I_0} \frac{1}{2^{b/Fx} - 1}\right), \\
 i &= 1, \dots, n, \quad n > 0, \quad (n, u) \in \mathcal{X}.
 \end{aligned}$$

Обозначим $y = \frac{P_u G_t G_r}{N_0 + I_0} \frac{1}{2^{b/Fx} - 1}$, откуда $x = \frac{b}{F} \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \frac{1}{y} \right) \right]^{-1}$ и

$$E[\xi_{ui}^t] = \int_{\text{PL}_{\min}}^{\text{PL}_{\max}} \frac{b}{F} \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \frac{1}{y} \right) \right]^{-1} dF_{\xi_{\text{PL}}}^{\text{PL}}(y), \quad i=1, \dots, n, \quad n > 0, \quad (n, u) \in \mathcal{X}.$$

в) *Выражение дисперсии ξ_u^t через ФР ξ_{PL}*

Выразим дисперсию СВ ξ_{ui}^t через ФР СВ ξ_{PL}

$$\text{Var}[\xi_{ui}^t] = E[(\xi_{ui}^t)^2] - (E[\xi_{ui}^t])^2, \quad i=1, \dots, n, \quad n > 0, \quad (n, u) \in \mathcal{X},$$

где

$$\mathbb{E}\left[\left(\xi_{ui}^t\right)^2\right]=\int_{\text{PL}_{\min}}^{\text{PL}_{\max}}\left(\frac{b}{F}\right)^2\left[\log_2\left(1+\frac{P_u G_t G_r}{N_0+I_0}\cdot\frac{1}{y}\right)\right]^{-2}dF_{\xi_{\text{PL}}}(y),$$

$$i=1,\dots,n, \quad n>0, \quad (n,u)\in\mathcal{X}.$$

г) Выражение ФР η_{nu} через ФР ξ_{PL}

Обозначим СВ $\eta_{nu}=\sum_{i=1}^n\xi_{ui}^t$, тогда по центральной предельной теореме,

$$\frac{\sum_{i=1}^n\xi_{ui}^t-n\cdot\mathbb{E}\left[\xi_{ui}^t\right]}{\sqrt{n\cdot\text{Var}\left[\xi_{ui}^t\right]}}\xrightarrow{n\rightarrow\infty}\mathcal{N}(0,1), \quad n>0, \quad (n,u)\in\mathcal{X},$$

то есть

$$\Pr\left(\frac{\sum_{i=1}^n\xi_{ui}^t-n\cdot\mathbb{E}\left[\xi_{ui}^t\right]}{\sqrt{n\cdot\text{Var}\left[\xi_{ui}^t\right]}}\leq x\right)=\Pr\left(\sum_{i=1}^n\xi_{ui}^t\leq x\cdot\sqrt{n\cdot\text{Var}\left[\xi_{ui}^t\right]}+n\cdot\mathbb{E}\left[\xi_{ui}^t\right]\right)\xrightarrow{n\rightarrow\infty}\Phi(x),$$

$$n>0, \quad (n,u)\in\mathcal{X}.$$

Обозначим $y=x\cdot\sqrt{n\cdot\text{Var}\left[\xi_{ui}^t\right]}+n\cdot\mathbb{E}\left[\xi_{ui}^t\right]$, откуда $x=\frac{y-n\cdot\mathbb{E}\left[\xi_{ui}^t\right]}{\sqrt{n\cdot\text{Var}\left[\xi_{ui}^t\right]}}$ и

$$\Pr\left(\sum_{i=1}^n\xi_{ui}^t\leq y\right)=F_{\eta_{nu}}(y)\xrightarrow{n\rightarrow\infty}\Phi\left(\frac{y-n\cdot\mathbb{E}\left[\xi_{ui}^t\right]}{\sqrt{n\cdot\text{Var}\left[\xi_{ui}^t\right]}}\right), \quad n>0, \quad (n,u)\in\mathcal{X}.$$

д) Получение выражения для $\Pr(\xi_u[n]\in\mathcal{Y}(n,u))$

Таким образом,

$$\Pr(\xi_u[n]\in\mathcal{Y}(n,u))=\Pr\left(\sum_{i=1}^n\xi_{ui}^t\leq 1\right)\xrightarrow{n\rightarrow\infty}\Phi\left(\frac{1-n\cdot\mathbb{E}\left[\xi_{ui}^t\right]}{\sqrt{n\cdot\text{Var}\left[\xi_{ui}^t\right]}}\right), \quad n>0, \quad (n,u)\in\mathcal{X}.$$

■

Из леммы 5.8 вытекает следующее следствие.

Следствие 5.5. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с постоянной мощностью и прерыванием обслуживания трафика, если все устройства пользователей расположены равномерно в зоне покрытия базовой станции и при любом своем местоположении находятся в прямой видимости, тогда математическое ожидание $\mu(u)$ и дисперсия $\sigma^2(u)$ мощности затухания сигнала в лемме 5.8 рассчитываются соответственно по формулам

$$\mu(u) = \frac{b}{F} \frac{2}{d_{\max}^2 \alpha \beta} \int_{d_{\min}}^{d_{\max}} \left[x^2 + (h_{\text{BS}} - h_{\text{UT}})^2 \right]^{(2-\beta)/2} \times$$

$$\times \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \frac{1}{\alpha} \left[x^2 + (h_{\text{BS}} - h_{\text{UT}})^2 \right]^{-\beta/2} \right) \right]^{-1} dx, \quad u = 0, 1, \quad (5.98)$$

$$\sigma^2(u) = \frac{b}{F} \frac{2}{d_{\max}^2 \alpha \beta} \int_{d_{\min}}^{d_{\max}} \left[x^2 + (h_{\text{BS}} - h_{\text{UT}})^2 \right]^{(2-\beta)/2} \times$$

$$\times \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \frac{1}{\alpha} \left[x^2 + (h_{\text{BS}} - h_{\text{UT}})^2 \right]^{-\beta/2} \right) \right]^{-2} dx, \quad u = 0, 1. \quad (5.99)$$

Доказательство.

Доказательство проводится с применением следствия 2.4. Плотность распределения СВ ξ_{PL} в этом случае имеет вид

$$f_{\xi_{\text{PL}}}(x) = \frac{dF_{\xi_{\text{PL}}}(x)}{dx} = \frac{d \left(\frac{[x/\alpha]^{2/\beta} - (h_{\text{BS}} - h_{\text{UT}})^2}{d_{\max}^2} \right)}{dx} =$$

$$= \frac{1}{d_{\max}^2} \cdot \frac{1}{\beta} \left(\frac{x}{\alpha} \right)^{2 \frac{1-\beta}{\beta}} \cdot 2 \frac{x}{\alpha} \cdot \frac{1}{\alpha} = \frac{2 \cdot x^{(2-\beta)/\beta}}{d_{\max}^2 \alpha^{2/\beta} \beta}.$$

Тогда математическое ожидание будет иметь вид

$$\mu(u) = \frac{b}{F} \frac{2}{d_{\max}^2 \alpha^{2/\beta} \beta} \int_{\text{PL}_{\min}}^{\text{PL}_{\max}} x^{(2-\beta)/\beta} \cdot \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \frac{1}{x} \right) \right]^{-1} dx, \quad u = 0, 1.$$

С учетом того, что

$$PL_{\min} = \alpha \left[d_{\min}^2 + (h_{BS} - h_{UT})^2 \right]^{\beta/2}, \quad PL_{\max} = \alpha \left[d_{\max}^2 + (h_{BS} - h_{UT})^2 \right]^{\beta/2},$$

получим

$$\mu(u) = \frac{b}{F} \frac{2}{d_{\max}^2 \alpha^{2/\beta} \beta} \int_{\alpha \left[d_{\min}^2 + (h_{BS} - h_{UT})^2 \right]^{\beta/2}}^{\alpha \left[d_{\max}^2 + (h_{BS} - h_{UT})^2 \right]^{\beta/2}} x^{(2-\beta)/\beta} \cdot \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \frac{1}{x} \right) \right]^{-1} dx,$$

$$u = 0, 1.$$

Обозначим $x = \alpha \left[y^2 + (h_{BS} - h_{UT})^2 \right]^{\beta/2}$, откуда $y = \sqrt{[x/\alpha]^{2/\beta} - (h_{BS} - h_{UT})^2}$ и

$$\begin{aligned} \mu(u) &= \frac{b}{F} \frac{2}{d_{\max}^2 \alpha^{2/\beta} \beta} \int_{d_{\min}}^{d_{\max}} \left(\alpha \left[y^2 + (h_{BS} - h_{UT})^2 \right]^{\beta/2} \right)^{(2-\beta)/\beta} \times \\ &\quad \times \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \left(\alpha \left[y^2 + (h_{BS} - h_{UT})^2 \right]^{\beta/2} \right)^{-1} \right) \right]^{-1} dy = \\ &= \frac{b}{F} \frac{2}{d_{\max}^2 \alpha \beta} \int_{d_{\min}}^{d_{\max}} \left[y^2 + (h_{BS} - h_{UT})^2 \right]^{(2-\beta)/2} \times \\ &\quad \times \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \frac{1}{\alpha \left[y^2 + (h_{BS} - h_{UT})^2 \right]^{\beta/2}} \right) \right]^{-1} dy, \quad u = 0, 1. \end{aligned}$$

А дисперсия будет иметь вид

$$\sigma^2(u) = \left(\frac{b}{F} \right)^2 \frac{2}{d_{\max}^2 \alpha^{2/\beta} \beta} \int_{PL_{\min}}^{PL_{\max}} x^{(2-\beta)/\beta} \cdot \left[\log_2 \left(1 + \frac{P_u G_t G_r}{N_0 + I_0} \cdot \frac{1}{x} \right) \right]^{-2} dx - [\mu(u)]^2,$$

$$u = 0, 1.$$

■

5.6.3 Условные вероятности выделения ресурса

Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением мощности и скорости передачи трафика, обозначим вероятности того, что еще одной сессии узкополосного и широкополосного трафика соответственно может быть выделен ресурс, при условии нахождения в состоянии (n, u)

$$p(n, u) = \Pr(\xi_u[n+1] \in \mathcal{Y}(n, u) \mid \xi_u[n] \in \mathcal{Y}(n, u)), \quad (n, u) \in \mathcal{X}. \quad (5.100)$$

Тогда справедлива следующая теорема.

Теорема 5.6. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением мощности и скорости передачи трафика, условные вероятности того, что еще одной сессии потокового трафика может быть выделен ресурс, рассчитываются по формулам

$$\begin{aligned} p(0, 0) &= F_{\xi_{\text{PL}}} \left(X \left(2^{b/F} - 1 \right) \right), \\ p(0, 1) &= F_{\xi_{\text{PL}}} \left(\frac{X}{2^{b/F} - 1} \right), \\ p(n, 0) &= \left[F_{\xi_{\text{PL}}} \left(X \left(2^{b(n+1)/F} - 1 \right) \right) \right]^{n+1} \left[F_{\xi_{\text{PL}}} \left(X \left(2^{b(n+1)/F} - 1 \right) \right) \right]^{-n}, \\ n > 0, (n, 0) &\in \mathcal{X}, \\ p(n, 1) &= \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{b(n+1)/F} - 1} \right) \right]^{n+1} \left[F_{\xi_{\text{PL}}} \left(\frac{X}{2^{b(n+1)/F} - 1} \right) \right]^{-n}, \\ n > 0, (n, 1) &\in \mathcal{X}. \end{aligned} \quad (5.101)$$

где $X = P_1 G_t G_r (N_0 + I_0)^{-1}$.

Доказательство.

Доказательство можно провести аналогично теореме 2.4 с применением леммы 5.7.

■

Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с постоянной мощностью и прерыванием обслуживания трафика, обозначим вероятности того, что еще одной сессии узкополосного и широкополосного трафика соответственно может быть выделен ресурс, при условии нахождения в состоянии (n, u)

$$p(n, u) = \Pr(\xi_u[n+1] \in \mathcal{Y}(n, u) \mid \xi_u[n] \in \mathcal{Y}(n, u)), \quad (n, u) \in \mathcal{X}. \quad (5.102)$$

Тогда справедливо следующее утверждение.

Утверждение 5.1. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с постоянной мощностью и прерыванием обслуживания трафика, условные вероятности того, что еще одной сессии потокового трафика может быть выделен ресурс, рассчитываются приближенно по формулам

$$\begin{aligned} p(0, u) &\approx \Phi\left(\frac{1 - \mu(u)}{\sigma(u)}\right), \quad u = 0, 1, \\ p(n, u) &\approx \left[\Phi\left(\frac{1 - \mu(u) \cdot (n+1)}{\sigma(u) \cdot \sqrt{n+1}}\right) \right]^{n+1} \left[\Phi\left(\frac{1 - \mu(u) \cdot n}{\sigma(u) \cdot \sqrt{n}}\right) \right]^{-n}, \\ (n, u) &\in \mathcal{X} \setminus \{(0, 0), (0, 1)\}. \end{aligned} \quad (5.103)$$

где математическое ожидание $\mu(u)$ и дисперсия $\sigma^2(u)$ мощности затухания сигнала вычисляются по формулам леммы 5.8.

Доказательство.

Доказательство можно провести аналогично теореме 2.4 с применением леммы 5.8. ■

Из теоремы 5.8 и решения СУР $\pi \mathbf{Q} = \mathbf{0}$ с условием нормировки $\pi \mathbf{1}^T = 1$, вытекает следующее следствие.

Следствие 5.6. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой со снижением мощности и скорости передачи трафика, средняя скорость передачи потокового трафика рассчитывается по формуле

$$\bar{b} = \frac{\sum_{n=1}^M b(n) \cdot \pi(n,0) + b \sum_{n=1}^M \pi(n,1)}{1 - \pi(0,0) + \pi(0,1)}, \quad (5.104)$$

а средняя скорость передачи потокового трафика, когда ненадежный ресурс недоступен вычисляется следующим образом

$$\bar{b}_0 = \frac{\sum_{n=1}^M b(n) \cdot \pi(n,0)}{\sum_{n=1}^M \pi(n,0)}. \quad (5.105)$$

Из утверждения 5.1 и решения СУР $\pi \mathbf{Q} = \mathbf{0}$ с условием нормировки $\pi \mathbf{1}^T = 1$, вытекает следующее следствие.

Следствие 5.7. Для модели с потоковым трафиком и частичным отказом изменяемого объема ресурса, схемой с постоянной мощностью и прерыванием обслуживания трафика, вероятность прерывания обслуживания сессии потокового трафика рассчитывается по формуле

$$P_{\text{in}} = \sum_{n=M_0+1}^{M_1-1} \frac{\alpha}{\lambda \cdot p(n,1) + m\mu_2 + \alpha} \cdot \binom{n-1}{n-M_0-1} \left[\binom{n}{n-M_0} \right]^{-1} \cdot \pi(n,1) + \\ + \frac{\alpha}{M_1\mu_2 + \alpha} \cdot \binom{M_1-1}{M_1-M_0-1} \left[\binom{M_1}{M_1-M_0} \right]^{-1} \cdot \pi(M_1,1). \quad (5.106)$$

Проиллюстрируем поведение показателей эффективности обслуживания потокового трафика на примере сценария умного города. Высокая плотность подключенных устройств, таких как сенсоры, IoT-системы и V2X-коммуникации, создает новые вызовы для традиционного управления спектром. В умных городах массово внедряются приложения, ориентированные на машины: автоматизированный транспорт, интеллектуальное энергоснабжение и системы общественной безопасности. Для их работы нужна гарантированная пропускная

способность и минимальные задержки. Однако фиксированное распределение спектра оказывается неэффективным из-за неравномерного и непредсказуемого трафика. Решение на основе LSA позволяет динамически перераспределять «простаивающие» частоты между владельцем частот и арендатором. Это обеспечивает гибкость и адаптивность системы. Особенно важно это для критически важных сервисов, таких как обмен видео в реальном времени между транспортными средствами или оперативное реагирование экстренных служб.

На рисунке 5.8 представлены графики средней скорости передачи потокового трафика в зависимости от радиуса зоны покрытия базовой станции для разной максимальной мощности передающей антенны. Увеличение радиуса соты не сильно влияет на среднюю скорость. При доступности полосы используется максимальную мощность, обеспечивая достаточный сигнал даже на границе соты. Однако при недоступности полосы мощность снижается, и пользователи на большем удалении от базовой станции сталкиваются с более сильным затуханием сигнала. Это приводит к снижению скорости.

На рисунках 5.9 и 5.10 представлены графики вероятности прерывания обслуживания сессии потокового трафика при максимальной мощности передающей антенны 0,2 Вт и 15,85 Вт соответственно в зависимости от радиуса зоны покрытия базовой станции для разного среднего времени полной доступности диапазона радиочастот. При малом радиусе соты высокая плотность устройств у базовой станции обеспечивает стабильное соединение, снижая риск недоступности сервиса. Однако с расширением зоны покрытия расстояние между устройствами и базовой станцией увеличивается, что ослабляет сигнал и ухудшает качество связи. Увеличение мощности передачи компенсирует негативное влияние большого радиуса соты. Это объясняется тем, что высокая мощность улучшает соотношение сигнал/шум, преодолевая затухание сигнала на больших расстояниях.

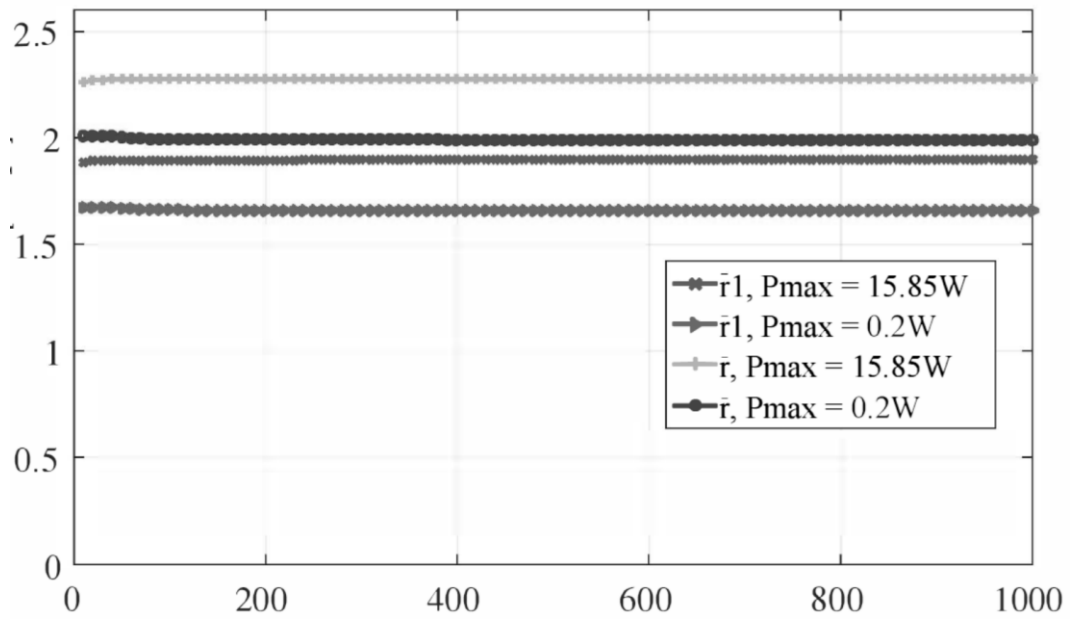


Рисунок 5.8 – Средняя скорость передачи потокового трафика (Мбит/с) в зависимости от радиуса зоны покрытия базовой станции (м) для разной максимальной мощности передающей антенны (Вт)

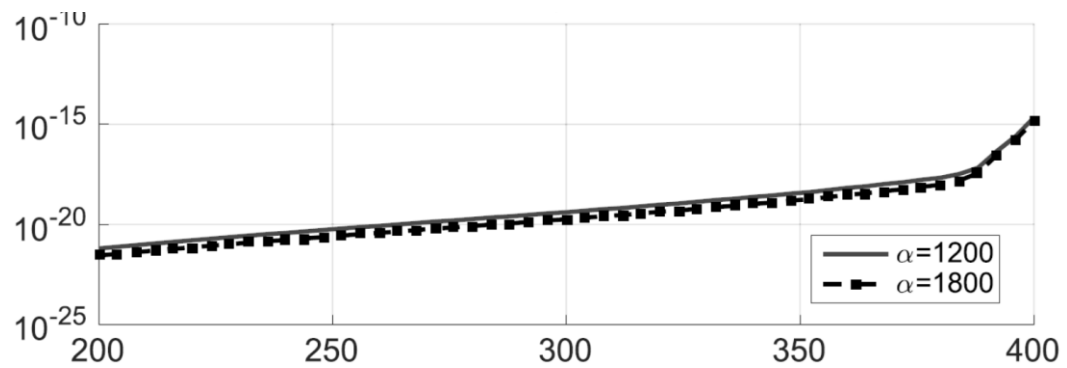


Рисунок 5.9 – Вероятность прерывания обслуживания сессии потокового трафика при максимальной мощности передающей антенны 0,2 Вт в зависимости от радиуса зоны покрытия базовой станции (м) для разного среднего времени полной доступности диапазона радиочастот (с)

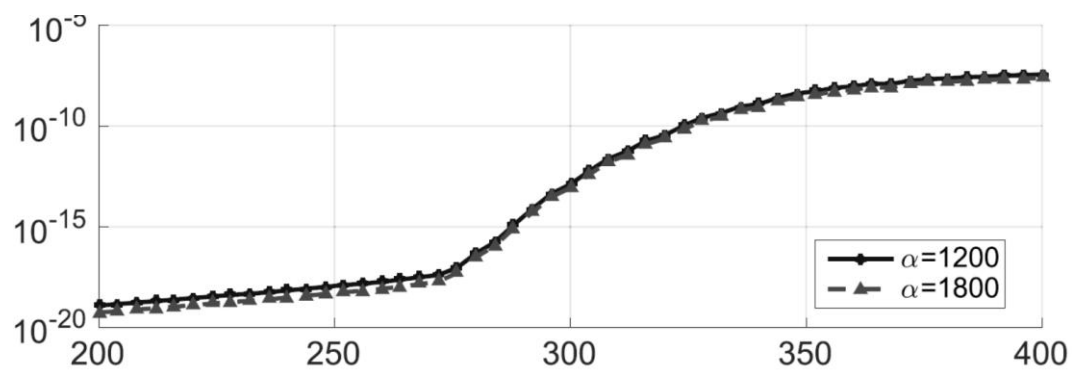


Рисунок 5.10 – Вероятность прерывания обслуживания сессии потокового трафика при максимальной мощности передающей антенны 15,85 Вт в зависимости от радиуса зоны покрытия базовой станции (м) для разного среднего времени полной доступности диапазона радиочастот (с)

ЗАКЛЮЧЕНИЕ

Основные научные результаты работы состоят в следующем.

1. Разработан комплекс моделей приоритетного доступа эластичного и потокового трафика в беспроводных сетях с применением механизмов прерывания и приостановки сессии, снижения скорости передачи и мощности сигнала. Модели применимы для анализа сценариев беспроводных сетей с несколькими классами пользователей, ненадежными и совместно используемыми ресурсами. Показателями эффективности приоритетного доступа являются характеристики случайных величин прерывания сессии, скорости передачи, объема используемого ресурса. Для классификации моделей комплекса предложена модификация обозначений Башарина – Кендалла.
2. Разработаны модели приоритетного доступа потокового трафика для услуг беспроводной сети. Для узкополосной сверхнадежной передачи с малой задержкой и широкополосной связи разработан матричный алгоритм расчета стационарного распределения для системы с двумя орбитами. Получены формулы для расчета вероятности прерывания и средней скорости широкополосного трафика для системы со случайным требованием к объему надежного ресурса, дискриминаторным разделением слота времени и снижением мощности сигнала. Для одноадресного видео по требованию и многоадресной видеоконференции предложен рекуррентный алгоритм расчета стационарного распределения. Формализована задача максимизации средней скорости видеоконференции по значениям уровней скорости и при ограничениях на вероятность блокировки и вероятность прерывания сессии.
3. Разработаны модели приоритетного доступа эластичного трафика с требованием к минимальной скорости передачи. Для обслуживания трафика в нелицензируемых радиочастотах построена модель системы массового обслуживания, допускающая мультипликативное распределение

вероятностей. Для нарезки радиоресурсов предложена модель управляемой системы массового обслуживания с перераспределением ресурса по сигналам по критерию максимума взвешенной суммы коэффициентов соответствия равному делению ресурса, результативности сигналов и использования ресурса. Формализована задача максимизации взвешенной суммы коэффициентов по параметру потока сигналов, управляющего нарезкой.

4. Разработаны модели обслуживания критически важных данных в миллиметровом диапазоне в виде систем с эластичным трафиком и ненадежным ресурсом. Время недоступности миллиметрового диапазона по причине его блокировки препятствием моделируется временем отказа ненадежных приборов. При отказе происходит приостановка или снижение скорости обслуживания трафика, а изменяемая активность пользователей моделируется марковским потоком МАР. Получены матричные алгоритмы расчета стационарного распределения и символьного вычисления преобразования Лапласа-Стилтьеса и моментов времени передачи эластичного трафика по вложенной цепи Маркова.
5. Разработаны модели совместного использования радиочастот с приоритетным доступом и случайным требованием потокового трафика к объему ненадежного ресурса. Получен рекуррентный алгоритм расчета среднего числа приостановленных сессий. Разработан матричный алгоритм символьного вычисления преобразования Лапласа-Стилтьеса периода занятости системы. Предложена модель управляемой системы массового обслуживания с гистерезисным занятием ненадежных приборов по критерию минимума средней стоимости использования ресурса и обслуживания сессий. Получен алгоритм расчета вероятности прерывания и средней скорости передачи трафика для двухресурсной системы массового обслуживания с зависимым занятием ресурса и снижением мощности сигнала.

СПИСОК ЛИТЕРАТУРЫ

1. Гольдштейн Б.С., Кучерявый А.Е. Сети связи пост-NGN. – СПб: Питер, 2013. – 160 с.
2. Гольдштейн Б.С., Соколов Н.А., Яновский Г.Г. Сети связи. – СПб.: БХВ-Петербург, 2010. – 400 с.
3. Деарт В.Ю. Мультисервисные сети связи. Транспортные сети и сети доступа. – М.: Брис-М, 2014. – 189 с.
4. Кучерявый Е.А. Управление трафиком и качество обслуживания в сети Интернет. – СПб.: Наука и Техника, 2004. – 336 с.
5. Молчанов Д.А., Бегишев В.О., Самуйлов К.Е., Кучерявый Е.А. Сети 5G/6G: архитектура, технологии, методы анализа и расчета. – М.: РУДН, 2022. – 516 с.
6. Пшеничников А.П., Росляков А.В. Будущие сети. – М.: Горячая линия-Телеком, 2022. – 256 с.
7. Росляков А.В. Сети фиксированной связи пятого поколения. – М.: Колос-с, 2024. – 232 с.
8. Самуйлов К.Е., Шалимов И.А., Васин Н.Н., Василевский В.В., Кулябов Д.С., Королькова А.В. Сети и телекоммуникации. – М.: Юрайт, 2025. – 464 с.
9. Смелянский Р.Л. Компьютерные сети: в 2 т. Т. 1. Системы передачи данных. – М.: Академия, 2011. – 304 с.
10. Смелянский Р.Л. Компьютерные сети: в 2 т. Т. 2. Сети ЭВМ. – М.: Академия, 2011. – 240 с.
11. Jiang W., Han B. Cellular communication networks and standards: The evolution from 1G to 6G. – Springer, 2024. – 246 p.
12. Пшеничников А.П. Этапы цифровизации сетей связи // Методические вопросы преподавания инфокоммуникаций в высшей школе. – 2019. – Т. 8, № 2. – С. 65-71.

13. Пшеничников А.П., Короткова В.И., Поскотин Л.С. Перспективные инфокоммуникационные технологии и сетевые услуги // Научные технологии в космических исследованиях Земли. – 2023. – Т. 15, № 3. – С. 57-64.
14. Shorgin S., Samouylov K., Gudkova I., Galinina O., Andreev S. On the benefits of 5G wireless technology for future mobile cloud computing // Proc. 1st International Science and Technology Conference on Modern Networking Technologies (MoNeTeC). – 2014. – Art. No. 6995601.
15. Кучерявый А.Е., Давуд А.А.А.А., Волков А.Н., Мутханна А.С.А., Парамонов А.И., Елагин В.С., Дунайцев Р.А., Владимиров С.С., Горбачева Л.С., Захаров М.В., Марочкина А.В., Анваржонов Б.Н.У., Паньков Б.О., Свечников Д.С. Сети связи шестого поколения: фрактальные архитектуры, голографические взаимодействия, телеприсутствие, сетевые роботы. – СПб: Питер, 2024. – 320 с.
16. Ateya A.A., Abd El-Latif A.A., Muthanna A., Volkov A., Koucheryavy A. Enabling metaverse and telepresence services in 6G networks. – River Publishers, 2024. – 254 p.
17. Росляков А.В. Сеть 2030: архитектура, технологии, услуги. – М.: Колос-с, 2022. – 278 с.
18. Iyer S., Kalla A., Lopez O.A., De Alwis C. Intelligent spectrum management: Towards 6G. – Wiley, 2024. – 282 p.
19. Prasad R., Mantri D.S., Pandey S.K., Mihovska A.D. 6G connectivity-systems, technologies, and applications: Digitalization of new technologies, 6G and evolution. – River Publishers, 2024. – 330 p.
20. Tripathi S.L., Mahmud M., Narmadha C., Alexander S.A. Development of 6G networks and technology. – Wiley, 2024. – 446 p.
21. Lin X. An overview of 5G Advanced evolution in 3GPP Release 18 // IEEE Communications Standards Magazine. – 2022. – Vol. 6, No. 3. – P. 77-83.
22. Ganesan K. 5G Advanced: Sidelink evolution // IEEE Communications Standards Magazine. – 2023. – Vol. 7, No. 1. – P. 58-63.

23. Liberg O., Hoymann C., Tidestav C., Larsson D.C., Rahman I., Blasco R., Falahati S., Blankenship Y. Introducing 5G Advanced // IEEE Communications Standards Magazine. – 2024. – Vol. 8, No. 1. – P. 52-57.
24. Yastrebova A., Kirichek R., Koucheryavy Y., Borodin A., Koucheryavy A. Future Networks 2030: architecture requirements // Proc. 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). – 2018. – Art. No. 8631208.
25. Makolkina M., Shypota N., Koucheryavy A. Development and investigation of model network IMT2020 with the use of MEC and voice assistant technologies // Lecture Notes in Computer Science. – 2020. – Vol. 12563. – P. 232-243.
26. Othman W.M., Ateya A.A., Nasr M.E., Muthanna A., ElAffendi M., Koucheryavy A., Hamdi A.A. Key enabling technologies for 6G: The role of UAVs, terahertz communication, and intelligent reconfigurable surfaces in shaping the future of wireless networks // Journal of Sensor and Actuator Networks. – 2025. – Vol. 14, No. 2. – Art. No. 30.
27. Вишневский В.М., Рыков В.В., Козырев Д.В., Иванова Н.М. Моделирование надежности привязных высотных беспилотных телекоммуникационных платформ. – М.: Техносфера, 2022. – 194 с.
28. Vishnevsky V.M., Selvamuthu D., Rykov V., Kozyrev D.V., Ivanova N., Krishnamoorthy A. Reliability assessment of tethered high-altitude unmanned telecommunication platforms: k-out-of-n reliability models and applications. – Springer, 2024. – 167 p.
29. Gaidamaka Y., Samouylov K. UVCS: Unit virtual coordinate system for UAV intra-swarm routing in GPS-denied environment // Mathematics. – 2023. – Vol. 11, No. 3. – Art. No. 694.
30. Gaydamaka A., Kabir M.T., Samuylov A., Moltchanov D., Tan B., Koucheryavy Y. System-level analysis of the directional radar coverage for UAV localization in dynamic swarms // IEEE Systems Journal. – 2025. – Vol. 19, No. 1. – P. 164-175.
31. Парамонов А.И., Маколкина М.А., Киричек Р.В., Выборнова А.И., Богданова Е.Г. Математические модели в сетях связи. – СПб.: СПбГУТ, 2018. – 229 с.

32. Ibrahimov B.G., Alieva A.A. Research and analysis indicators of the quality of service multimedia traffic using fuzzy logic // *Advances in Intelligent Systems and Computing*. – 2021. – Vol. 1306. – P. 773-780.
33. Ibrahimov B.G., Tagiyev A.D. Research of the performance multiservice telecommunication networks based on the architectural concept NGN and FN // *Lecture Notes in Mechanical Engineering*. – 2023. – P. 333-341.
34. Ibrahimov B.G., Aliyev A., Rafizade U.R. Research on the effectiveness of wireless networks cellular communications based on digital technologies // *Communications in Computer and Information Science*. – 2025. – Vol. 2225. – P. 55-65.
35. Tsitovich I.I., Chernushevich A.V. Numerical analysis of the effect of hysteresis in the access control of wireless broadband network on the functional efficiency // *Journal of Communications Technology and Electronics*. – 2012. – Vol. 57, No. 8. – P. 906-919.
36. Tsitovich I. Group polling method upon the independent activity of sensors in unsynchronized wireless monitoring networks // *Communications in Computer and Information Science*. – 2019. – Vol. 1141. – P. 436-448.
37. Tsitovich I. Two ways of group polling method application for sensors detecting in unsynchronized structured WSNs // *Communications in Computer and Information Science*. – 2020. – Vol. 1337. – P. 286-298.
38. Beschastnyi V., Ostrikova D., Moltchanov D., Gaidamaka Y., Koucheryavy Y., Samouylov K. Balancing latency and energy efficiency in mmWave 5G NR systems with multiconnectivity // *IEEE Communications Letters*. – 2022. – Vol. 26, No. 8. – P. 1952-1956.
39. Beschastnyi V., Ostrikova D., Moltchanov D., Gaidamaka Y., Koucheryavy Y., Samouylov K. Comparison of energy conservation strategies for 5G NR RedCap service in industrial environment // *Computer Networks*. – 2024. – Vol. 254. – Art. No. 110792.
40. Смелянский Р. Настоящее и будущее SDN&NFV // *Первая миля*. – 2016. – № 3(56). – С. 78-85.

41. Смелянский Р.Л., Васин В.В., Беззубцев С.О. Разработка отечественного коммутатора для программно-конфигурируемых сетей // Электронная техника. Серия 3: Микроэлектроника. – 2016. – № 1(161). – С. 9-17.
42. Смелянский Р.Л., Степанов Е.П. Применение методов с машинным обучением для управления сетевой вычислительной инфраструктурой // Доклады Российской академии наук. Математика, информатика, процессы управления. – 2024. – Т. 516, № 1. – С. 103-112.
43. Al-Ansi A., Al-Ansi A.M., Muthanna A., Elgendy I.A., Koucheryavy A. Survey on intelligence edge computing in 6G: Characteristics, challenges, potential use cases, and market drivers // Future Internet. – 2021. – Vol. 13, No. 5. – Art. No. 118.
44. Efrosinin D., Kochetkova I., Samouylov K., Stepanova N. Algorithmic analysis of a two-class multi-server heterogeneous queueing system with a controllable cross-connectivity // Lecture Notes in Computer Science. – 2020. – Vol. 12023. – P. 1-17.
45. Kochetkova I.A., Vlaskina A.S., Efrosinin D.V., Khakimov A.A., Burtseva S.A. To analysis of a two-buffer queueing system with cross-type service and additional penalties // Discrete and Continuous Models and Applied Computational Science. – 2021. – Т. 29, № 2. – С. 158-172.
46. Kushchazli A., Safargalieva A., Kochetkova I., Gorshenin A. Queueing model with customer class movement across server groups for analyzing virtual machine migration in cloud computing // Mathematics. – 2024. – Vol. 12, No. 3. – Art. No. 468.
47. Kushchazli A., Leonteva K., Kochetkova I., Khakimov A. Evaluating QoS in dynamic virtual machine migration: A multi-class queueing model for edge-cloud systems // Journal of Sensor and Actuator Networks. – 2025. – Vol. 14, No. 3. – Art. No. 47.
48. Vinel A., Ni Q., Staehle D., Turlikov A. Capacity analysis of reservation-based random access for broadband wireless access networks // IEEE Journal on Selected

- Areas in Communications. – 2009. – Vol. 27, No. 2. – P. 172-181, Art. No. 4769392.
49. Stepanov N., Turlikov A., Begishev V. Balancing the data transmission and random access phases in 6G mMTC radio technologies // IEEE Communications Letters. – 2023. – Vol. 27, No. 12. – P. 3419-3423.
 50. Cavallero S., Buratti C., Tsarev A., Cuzzo G., Khayrov E., Gaidamaka Y., Verdone R. Applying carrier sense multiple access to industrial IoT at terahertz frequencies // IEEE Internet of Things Journal. – 2024. – Vol. 11, No. 7. – P. 11986-11999.
 51. Samouylov K.E., Gaidamaka Y.V., Gudkova I.A., Zaripova E.R., Shorgin S.Y. Baseline analytical model for machine-type communications over 3GPP RACH in LTE-advanced networks // Communications in Computer and Information Science. – 2016. – Vol. 659. – P. 203-213.
 52. Medvedeva E., Zaripova E., Gudkova I., Semenova O., Vlaskina A., Gaidamaka Y. Discrete time Markov chain model for analyzing characteristics of RACH procedure under massive machine type communications // ACM International Conference Proceeding Series. – 2018. – Art. No. 59.
 53. Семенова О.В., Власкина А.С., Медведева Е.Г., Зарипова Э.Р., Гудкова И.А. Процедура установления соединения по радиоканалу случайного доступа с возможностью ретрансляции // Вестник Российского университета дружбы народов. Серия: Математика, информатика, физика. – 2018. – Т. 26, № 3. – С. 261-271.
 54. Adhikari M., Hazra A. 6G-enabled ultra-reliable low-latency communication in edge networks // IEEE Communications Standards Magazine. – 2022. – Vol. 6, No. 1. – P. 67-74.
 55. Gerami M., Singh B. Configured grant for ultra-reliable and low-latency communications: Standardization and beyond // IEEE Communications Standards Magazine. – 2022. – Vol. 6, No. 4. – P. 40-47.

56. Pourkabirian A., Kordafshari M.S., Jindal A., Anisi M.H. A vision of 6G URLLC: Physical-layer technologies and enablers // IEEE Communications Standards Magazine. – 2024. – Vol. 8, No. 2. – P. 20-27.
57. Saggese F., Moretti M., Popovski P. Power minimization of downlink spectrum slicing for eMBB and URLLC users // IEEE Transactions on Wireless Communications. – 2022. – Vol. 21, No. 12. – P. 11051-11065.
58. Cohen K.M., Park S., Simeone O., Popovski P., Shamai S. Guaranteed dynamic scheduling of ultra-reliable low-latency traffic via conformal prediction // IEEE Signal Processing Letters. – 2023. – Vol. 30. – P. 473-477.
59. Liu Y., Clerckx B., Popovski P. Network slicing for eMBB, URLLC, and mMTC: An uplink rate-splitting multiple access approach // IEEE Transactions on Wireless Communications. – 2024. – Vol. 23, No. 3. – P. 2140-2152.
60. Kavehmadavani F., Nguyen V.-D., Vu T.X., Chatzinotas S. Intelligent traffic steering in beyond 5G open RAN based on LSTM traffic prediction // IEEE Transactions on Wireless Communications. – 2023. – Vol. 22, No. 11. – P. 7727-7742.
61. Alsenwi M., Lagunas E., Chatzinotas S. Distributed learning framework for eMBB-URLLC multiplexing in open radio access networks // IEEE Transactions on Network and Service Management. – 2024. – Vol. 21, No. 5. – P. 5718-5732.
62. Tran D.-D., Ha V.N., Sharma S.K., Nguyen T.T., Chatzinotas S., Popovski P. Energy-efficient NOMA for 5G heterogeneous services: A joint optimization and deep reinforcement learning approach // IEEE Transactions on Communications. – 2025. – Vol. 73, No. 4. – P. 2448-2465.
63. Almekhlafi M., Chraiti M., Arfaoui M.A., Assi C., Ghrayeb A., Alloum A. A downlink puncturing scheme for simultaneous transmission of URLLC and eMBB traffic by exploiting data similarity // IEEE Transactions on Vehicular Technology. – 2021. – Vol. 70, No. 12. – P. 13087-13100.
64. Almekhlafi M., Arfaoui M.A., Elhattab M., Assi C., Ghrayeb A. Joint resource allocation and phase shift optimization for RIS-aided eMBB/URLLC traffic

- multiplexing // IEEE Transactions on Communications. – 2022. – Vol. 70, No. 2. – P. 1304-1319.
65. Almekhlafi M., Arfaoui M.A., Assi C., Ghrayeb A. Superposition-based URLLC traffic scheduling in 5G and beyond wireless networks // IEEE Transactions on Communications. – 2022. – Vol. 70, No. 9. – P. 6295-6309.
66. Koucheryavy A., Muthanna A., Volkov A. AI/machine learning for ultra-reliable low-latency communication // ITU News. – 2020. – No. 5. – P. 62-65.
67. Paramonov A., Peng J., Kashkarov D., Muthanna A., Elgendy I.A., Koucheryavy A., Maleh Y., Abd El-Latif A.A. Study and analysis of multiconnectivity for ultrareliable and low-latency features in networks and V2X communications // Wireless Communications and Mobile Computing. – 2021. – Vol. 2021. – Art. No. 1718099.
68. Gerasin I.S., Krasilov A.N., Khorov E.M. Dynamic multiplexing of URLLC traffic and eMBB traffic in an uplink using nonorthogonal multiple access // Journal of Communications Technology and Electronics. – 2020. – Vol. 65, No. 6. – P. 750-755.
69. Lebedeva I.V., Yusupov R.R., Krasilov A.N., Khorov E.M. Multiplexing of URLLC and eMBB traffic in a downlink channel with MU-MIMO // Journal of Communications Technology and Electronics. – 2022. – Vol. 67, No. 12. – P. 1506-1512.
70. Ivanova D., Adou Y., Markova E., Gaidamaka Y., Samouylov K. Mathematical framework for mixed reservation- and priority-based traffic coexistence in 5G NR systems // Mathematics. – 2023. – Vol. 11, No. 4. – Art. No. 1046.
71. Ivanova D., Zhibankova E., Markova E., Gaidamaka Y., Samouylov K. Performance modeling and comparison of URLLC and eMBB coexistence strategies in 5G new radio systems // Computer Networks. – 2024. – Vol. 255. – Art. No. 110904.
72. Li J., Nagalapur K.K., Stare E., Dwivedi S., Ashraf S.A., Eriksson P.-E., Engstrom U., Lee W.-H., Lohmar T. 5G New Radio for public safety mission critical

- communications // IEEE Communications Standards Magazine. – 2022. – Vol. 6, No. 4. – P. 48-55.
73. Municio E., Garcia-Aviles G., Garcia-Saavedra A., Costa-Perez X. O-RAN: Analysis of latency-critical interfaces and overview of time sensitive networking solutions // IEEE Communications Standards Magazine. – 2023. – Vol. 7, No. 3. – P. 82-89.
74. Hazra A., Munusamy A., Adhikari M., Awasthi L.K., Venu P. 6G-enabled ultra-reliable low latency communication for Industry 5.0: Challenges and future directions // IEEE Communications Standards Magazine. – 2024. – Vol. 8, No. 2. – P. 36-42.
75. Khorov E., Lyakhov A., Nasedkin I., Yusupov R., Famaey J., Akyildiz I.F. Fast and reliable alert delivery in mission-critical Wi-Fi HaLow sensor networks // IEEE Access. – 2020. – Vol. 8. – P. 14302-14313. – Art. No. 8957046.
76. Orsino A., Ometov A., Fodor G., Moltchanov D., Militano L., Andreev S., Yilmaz O.N.C., Tirronen T., Torsner J., Araniti G., Iera A., Dohler M., Koucheryavy Y. Effects of heterogeneous mobility on D2D-and drone-assisted mission-critical MTC in 5G // IEEE Communications Magazine. – 2017. – Vol. 55, No. 2. – P. 79-87, Art. No. 7842416.
77. Petrov V., Lema M.A., Gapeyenko M., Antonakoglou K., Moltchanov D., Sardis F., Samuylov A., Andreev S., Koucheryavy Y., Dohler M. Achieving end-to-end reliability of mission-critical traffic in softwarized 5G networks // IEEE Journal on Selected Areas in Communications. – 2018. – Vol. 36, No. 3. – P. 485-501.
78. Abdellah A.R., Muthanna A., Essai M.H., Koucheryavy A. Deep learning for predicting traffic in V2X networks // Applied Sciences (Switzerland). – 2022. – Vol. 12, No. 19. – Art. No. 10030.
79. Ateya A.A., Soliman N.F., Alkanhel R., Alhussan A.A., Muthanna A., Koucheryavy A. Lightweight deep learning-based model for traffic prediction in fog-enabled dense deployed IoT networks // Journal of Electrical Engineering and Technology. – 2023. – Vol. 18, No. 3. – P. 2275-2285.

80. Kochetkova I., Kushchazli A., Burtseva S., Gorshenin A. Short-term mobile network traffic forecasting using seasonal ARIMA and Holt-Winters models // *Future Internet*. – 2023. – Vol. 15, No. 9. – Art. No. 290.
81. Shrivastava V.K., Baek S., Baek Y. 5G Evolution for multicast and broadcast services in 3GPP Release 17 // *IEEE Communications Standards Magazine*. – 2022. – Vol. 6, No. 3. – P. 70-76.
82. Shrivastava R., Hegde S., Blume O. Sidelink evolution toward 5G-A/6G future considerations for standardization of group communications // *IEEE Communications Standards Magazine*. – 2023. – Vol. 7, No. 1. – P. 24-30.
83. Prasad A., Elmali B., Pauli V., Zheng N., Bhatooolaul D. Physical layer enhancements for high-reliability multicast and broadcast services in 5G Release-17 // *IEEE Communications Standards Magazine*. – 2024. – Vol. 8, No. 1. – P. 44-51.
84. Guzha E.A., Ivanov A.S., Kuznetsov E.S., Lyakhov A.I., Khorov E.M. Generalized mathematical model of reliable multicast transmission in modern Wi-Fi networks // *Journal of Communications Technology and Electronics*. – 2019. – Vol. 64, No. 8. – P. 870-879.
85. Krasilov A., Lebedeva I., Yusupov R., Khorov E. Resource-efficient multicast URLLC service in 5G systems // *Sensors*. – 2024. – Vol. 24, No. 8. – Art. No. 2536.
86. Chukhno N., Chukhno O., Pizzi S., Molinaro A., Iera A., Araniti G. Efficient management of multicast traffic in directional mmWave networks // *IEEE Transactions on Broadcasting*. – 2021. – Vol. 67, No. 3. – P. 593-605, Art. No. 9376272.
87. De La Fuente A., Interdonato G., Araniti G. User subgrouping and power control for multicast massive MIMO over spatially correlated channels // *IEEE Transactions on Broadcasting*. – 2022. – Vol. 68, No. 4. – P. 834-847.
88. Chukhno O., Chukhno N., Pizzi S., Orsino A., Maattanen H.-L., Torsner J., Molinaro A., Iera A., Araniti G. Mission-critical connectivity enhanced by IAB in beyond 5G: Interplay of sidelink, directional unicasting, and multicasting // *IEEE Open Journal of the Communications Society*. – 2023. – Vol. 4. – P. 1826-1838.

89. Bayat M., Wan K., Caire G. Coded caching over multicast routing networks // IEEE Transactions on Communications. – 2021. – Vol. 69, No. 6. – P. 3614-3627, Art. No. 9361706.
90. Amidzadeh M., Al-Tous H., Caire G., Tirkkonen O. Caching in cellular networks based on multipoint multicast transmissions // IEEE Transactions on Wireless Communications. – 2023. – Vol. 22, No. 4. – P. 2393-2408.
91. Amidzadeh M., Tirkkonen O., Caire G. Optimal multicast-cache-aided on-demand streaming in heterogeneous wireless networks via a path/surface following approach // IEEE Transactions on Wireless Communications. – 2024. – Vol. 23, No. 7. – P. 7833-7848.
92. Chukhno N., Chukhno O., Moltchanov D., Molinaro A., Gaidamaka Y., Samouylov K., Koucheryavy Y., Araniti G. Optimal multicasting in millimeter wave 5G NR with multi-beam directional antennas // IEEE Transactions on Mobile Computing. – 2023. – Vol. 22, No. 6. – P. 3572-3588.
93. Chukhno N., Chukhno O., Moltchanov D., Pizzi S., Gaydamaka A., Samuylov A., Molinaro A., Koucheryavy Y., Iera A., Araniti G. Models, methods, and solutions for multicasting in 5G/6G mmWave and sub-THz systems // IEEE Communications Surveys and Tutorials. – 2024. – Vol. 26, No. 1. – P. 119-159.
94. Бородакий В.Ю., Гудкова И.А., Острикова Д.Ю. Анализ предоставления услуг мультивещания в сетях LTE в виде системы с групповым обслуживанием заявок // Т-Comm: Телекоммуникации и транспорт. – 2013. – Т. 7, № 11. – С. 50-53.
95. Samouylov K.E., Gudkova I.A., Ostriкова D.Y. Modelling and performance analysis of multicast file repair in 3GPP LTE networks // Lecture Notes in Computer Science. – 2015. – Vol. 9247. – P. 383-392.
96. Бородакий В.Ю., Гудкова И.А., Острикова Д.Ю. Марковская модель восстановления файлов в сети LTE со случайным временем формирования группы // Т-Comm: Телекоммуникации и транспорт. – 2015. – Т. 9, № 9. – С. 5-10.

97. Orsino A., Araniti G., Scopelliti P., Gudkova I.A., Samouylov K.E., Iera A. Optimal subgroup configuration for multicast services over 5G-satellite systems // Proc. 12th IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB). – 2017. – Art. No. 7986134.
98. Ostrikova D., Rinaldi F., Beschastnyi V., Gudkova I., Militano L., Araniti G., Iera A., Samouylov K. Analytical model for multicast subgrouping in 5G-mobile eMBMS environment // Proc. 9th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). – 2017. – P. 13-19
99. Beschastnyi V.A., Ostrikova D.Yu., Zeifman A.I., Gudkova I.A. Optimal multicast subgrouping in mobility-aware 5G systems: challenges, modeling, and opportunities // CEUR Workshop Proceedings. – 2018. – Vol. 2332. – P. 13-22.
100. Beschastnyi V., Polovov M., Ostrikova D., Gudkova I., Araniti G., Shorgin S. Analytical model for performance analysis of video flow transmission to multicast subgroups in 5G wireless networks // AIP Conference Proceedings. – 2019. – Vol. 2116. – Art. No. 090004.
101. Beschastnyi V., Savich V., Ostrikova D., Gudkova I., Araniti G., Shorgin V. Analysis of machine-type communication data transmission by multicasting technology in 5G wireless networks // AIP Conference Proceedings. – 2019. – Vol. 2116. – Art. No. 090005.
102. Башарин Г.П. Лекции по математической теории телетрафика. – М.: РУДН, 2009. – 342 с.
103. Башарин Г.П., Бочаров П.П., Коган Я.А. Анализ очередей в вычислительных сетях: Теория и методы расчета. – М.: Наука, 1989. – 336 с.
104. Башарин Г.П., Толмачев А.Л. Теория сетей массового обслуживания и ее приложения к анализу информационно-вычислительных систем // Итоги науки и техники. Серия «Теория вероятностей. Математическая статистика. Теоретическая кибернетика». – 1983. – Т. 21. – С. 3-119.
105. Башарин Г.П., Харкевич А.Д., Шнепс М.А. Массовое обслуживание в телефонии. – М.: Наука, 1968. – 247 с.

106. Бенеш В.Э. Математические основы теории телефонных сообщений / пер. с англ. В.И. Неймана и А.Д. Харкевича; под ред. И.Н. Коваленко. – М.: Связь, 1986. – 291 с.
107. Вишневский В.М. Теоретические основы проектирования компьютерных сетей. – М.: Техносфера, 2003. – 512 с.
108. Вишневский В.М., Ляхов А.И., Портной С.Л., Шахнович И.В. Широкополосные беспроводные сети передачи информации. – М.: Техносфера, 2005. – 591 с.
109. Корнышев Ю.Н., Пшеничников А.П., Харкевич А.Д. Теория телетрафика. – М.: Радио и связь, 1996. – 272 с.
110. Крылов В.В., Самохвалова С.С. Теория телетрафика и ее приложения. – СПб.: БХВ-Петербург, 2005. – 288 с.
111. Меликов А.З., Пономаренко Л.А., Паладюк В.В. Телетрафик: модели, методы, оптимизация. – Киев: Политехника, 2007. – 256 с.
112. Пагано М., Рыков В.В., Хохлов Ю.С. Модели телетрафика. – М.: ИНФРА-М, 2019. – 178 с.
113. Пшеничников А.П. Теория телетрафика. – М.: Горячая линия-Телеком, 2017. – 212 с.
114. Степанов С.Н. Теория телетрафика: концепции, модели, приложения. – М.: Горячая линия-Телеком, 2015. – 868 с.
115. Шнепс М.А. Системы распределения информации: Методы расчета. – М.: Связь, 1979. – 344 с.
116. Штремер Х., Белендорф Э., Бининда Н. Теория телетрафика: Основы расчета систем проводной связи / пер. с нем. Г.П. Башарина; под ред. Г.П. Башарина. – М.: Связь, 1971. – 320 с.
117. Callegati F., Cerroni W., Raffaelli C. Traffic engineering: A practical approach. – Springer, 2023. – 222 p.
118. Giambene G. Queuing theory and telecommunications: Networks and applications. – Springer, 2021. – 413 p.
119. Iversen V.B. Teletraffic engineering handbook. – TUD, 2015. – 399 p.

120. Kelly F.P. Reversibility and stochastic networks. – Wiley, 2011. – 230 p.
121. Melikov A., Ponomarenko L. Multidimensional queueing models in telecommunication networks. – Springer, 2014. – 190 p.
122. Ponomarenko L. Kim C.S., Melikov A. Performance analysis and optimization of multi-traffic on communication networks. – Springer, 2010. – 194 p.
123. Ross K.W. Multiservice loss models for broadband telecommunication networks. – Springer, 1995. – 343 p.
124. Нейман В.И. Телетрафик и теория массового обслуживания // Автоматика и телемеханика. – 2009. – № 12. – С. 29-38.
125. Харкевич А.Д. 100 лет формуле А.К. Эрланга и процессу исследований по теории телетрафика // Автоматика и телемеханика. – 2009. – № 12. – С. 39-41.
126. Бочаров П.П., Печинкин А.В. Теория массового обслуживания. – М.: РУДН, 1995. – 529 с.
127. Вишневский В.М., Дудин А.Н., Клименок В.И. Стохастические системы с коррелированными потоками. Теория и применение в телекоммуникационных сетях. – М.: Техносфера, 2018. – 564 с.
128. Вишневский В.М., Ефросинин Д.В. Теория очередей и машинное обучение. – М.: ИНФРА-М, 2024. – 370 с.
129. Вишневский В.М., Семенова О.В. Системы поллинга: теория и применение в широкополосных беспроводных сетях. – М.: Техносфера, 2007. – 312 с.
130. Гарайшина И.Р., Моисеева С.П., Назаров А.А. Методы исследования коррелированных потоков и специальных систем массового обслуживания. – Томск: НТЛ, 2010. – 202 с.
131. Зейфман А.И., Бенинг В.Е., Соколов И.А. Марковские цепи и модели с непрерывным временем. – М.: Элекс-КМ, 2008. – 167 с.
132. Ивницкий В.А. Теория сетей массового обслуживания. – М.: Физматлит, 2004. – 772 с.
133. Карташевский В.Г. Основы теории массового обслуживания. – М.: Горячая линия-Телеком, 2013. – 130 с.

134. Кёниг Д., Рыков В.В., Шмидт Ф. Стационарные системы массового обслуживания с зависимостями // Итоги науки и техники. Серия «Теория вероятностей. Математическая статистика. Теоретическая кибернетика». – 1981. – Т. 18. – С. 95-186.
135. Клейнрок Л. Вычислительные системы с очередями / пер. с англ. Б.С. Цыбакова; под ред. Б.С. Цыбакова. – М.: Мир, 1979. – 600 с.
136. Клейнрок Л. Теория массового обслуживания / пер. с англ. И.И. Грушко; под ред. В.И. Неймана. – М.: Машиностроение, 1979. – 432 с.
137. Климов Г.П. Теория массового обслуживания. – М.: МГУ, 2011. – 312 с.
138. Кокс Д.Р., Смит У.Л. Теория очередей / пер. с англ. В.В. Рыкова и Ю.Б. Рождественского; под ред. А.Д. Соловьева. – М.: Мир, 1966. – 218 с.
139. Кузнецов Д.Ю., Назаров А.А. Адаптивные сети случайного доступа. – Томск: Дельтаплан, 2002. – 254 с.
140. Лукашенко О.В., Морозов Е.В. Введение в гауссовские системы обслуживания. – Петрозаводск: ПГУ, 2013. – 80 с.
141. Моисеев А.Н., Назаров А.А. Бесконечнолинейные системы и сети массового обслуживания. – Томск: НТЛ, 2015. – 240 с.
142. Морозов Е.В., Некрасова Р.С. Регенеративный метод и его применение в анализе систем обслуживания. – Петрозаводск: ПГУ, 2013. – 75 с.
143. Назаров А.А. Асимптотический анализ марковизируемых систем. – Томск: ТГУ, 1991. – 158 с.
144. Назаров А.А., Моисеева С.П. Метод асимптотического анализа в теории массового обслуживания. – Томск: НТЛ, 2006. – 112 с.
145. Назаров А.А., Терпугов А.Ф. Теория массового обслуживания. – Томск: НТЛ, 2010. – 228 с.
146. Печинкин А.В., Разумчик Р.В. Системы массового обслуживания в дискретном времени. – М.: Физматлит, 2018. – 432 с.
147. Рыков В.В., Козырев Д.В. Основы теории массового обслуживания (Основной курс: марковские модели, методы марковизации). – М.: ИНФРА-М, 2016. – 223 с.

148. Соколов А.Н., Соколов Н.А., Зайцев В.С. Однолинейные системы телетрафика. – СПб.: СПбГУТ, 2019. – 118 с.
149. Хинчин А.Я. Работы по математической теории массового обслуживания. – М.: Физматгиз, 1963. – 236 с.
150. Цициашвили Г.Ш., Осипова М.А. Актуальные задачи моделирования и анализа сетей массового обслуживания. – Казань: Бук, 2020. – 88 с.
151. Bocharov P.P., D'Apice C., Pechinkin A.V. Queueing theory. – De Gruyter, 2003.
152. Dudin A.N., Klimenok V.I., Vishnevsky V.M. The theory of queueing systems with correlated flows. – Springer, 2020. – 410 p.
153. Morozov E., Steyaert B. Stability analysis of regenerative queueing models: Mathematical methods and applications. – Springer, 2021. – 185.
154. Tsitsiashvili G.S., Osipova M.A., Losev A.S. Analysis of stochastic models: analytical results and computational algorithms. – Казань: Бук, 2023. – 80 с.
155. Рыков В.В. Сети обслуживания «прозрачных» требований // Автоматика и телемеханика. – 2001. – № 5. – С. 147-158.
156. Наумов В.А., Самуйлов К.Е., Гайдамака Ю.В. Мультипликативные решения конечных цепей Маркова. – М.: РУДН, 2015. – 159 с.
157. Наумов В.А., Самуйлов К.Е., Яркина Н.В. Теория телетрафика мультисервисных сетей. – М.: РУДН, 2007. – 191 с.
158. Naumov V., Gaidamaka Y., Yarkina N., Samouylov K. Matrix and analytical methods for performance analysis of telecommunication systems. – Springer, 2022. – 308 p.
159. Samuylov A., Moltchanov D., Kovalchukov R., Pirmagomedov R., Gaidamaka Y., Andreev S., Koucheryavy Y., Samouylov K. Characterizing resource allocation trade-offs in 5G NR serving multicast and unicast traffic // IEEE Transactions on Wireless Communications. – 2020. – Vol. 19, No. 5. – P. 3421-3434, Art. No. 9003488.
160. Yarkina N., Moltchanov D., Gaydamaka A., Koucheryavy Y. Coexistence of multicast and unicast services in mmWave/sub-THz self-backhauled systems: User

- associations and performance gains // IEEE Transactions on Vehicular Technology. – 2025. – Vol. 74, No. 3. – P. 4608-4624.
161. Gudkova I., Plaksina O. Performance measures computation for a single link loss network with unicast and multicast traffics // Lecture Notes in Computer Science. – 2010. – Vol. 6294. – P. 256-265.
162. Башарин Г.П., Клапоущак С.Н., Русина Н.В. Мультисервисная модель Эрланга с потоковым трафиком и пороговой стратегией доступа для эластичного трафика // Т-Comm: Телекоммуникации и транспорт. – 2012. – Т. 6, № 7. – С. 19-22.
163. Башарин Г.П., Клапоущак С.Н., Митькина Н.В. Математическая модель адаптивной многоскоростной системы с эластичным трафиком // Вестник Российского университета дружбы народов. Серия: Математика, информатика, физика. – 2008. – № 3. – С. 31-39.
164. Dudin A.N., Dudina O.S., Dudin S.A., Kostyukova O.I. Optimization of road design via the use of a queueing model with transit and local users and processor sharing discipline // Optimization. – 2022. – Vol. 71, No. 11. – P. 3147-3164.
165. D'Apice C., Dudin A., Dudin S., Manzo R. Priority queueing system with many types of requests and restricted processor sharing // Journal of Ambient Intelligence and Humanized Computing. – 2023. – Vol. 14, No. 9. – P. 12651-12662.
166. Dudin A., Dudin S., Manzo R., Rarità L. Queueing system with batch arrival of heterogeneous orders, flexible limited processor sharing and dynamical change of priorities // AIMS Mathematics. – 2024. – Vol. 9, No. 5. – P. 12144-12169.
167. Зорин А.В., Федоткин М.А. Анализ и оптимизация процессов с разделением времени, функционирующих в случайной среде // Вестник Нижегородского университета им. Н.И. Лобачевского. Серия: Математика. – 2004. – № 1. – С. 92-103.
168. Зорин А.В., Федоткин М.А. Анализ процессов с разделением времени // Вестник Нижегородского университета им. Н.И. Лобачевского. Серия: Математика. – 2003. – № 1. – С. 18-28.

169. Яшков С.Ф., Яшкова А.С. Разделение процессора: обзор математической теории // Информационные процессы. – 2007. – Т. 7, № 3. – С. 248-322.
170. Яшков С.Ф. О нестационарном распределении числа требований в системе M/G/1–EPS / С. Ф. Яшков // Информационные процессы. – 2009. – Т. 9, № 4. – С. 364-367.
171. Яшков С.Ф., Яшкова А.С. О распределениях периода занятости и времени пребывания в системе M/G/1EPS с катастрофами // Автоматика и телемеханика. – 2009. – № 12. – С. 134-146.
172. Raaijmakers Y., Borst S., Boxma O. Stability and tail behavior of redundancy systems with processor sharing // Performance Evaluation. – 2021. – Vol. 147. – Art. No. 102195.
173. Kapodistria S., Saxena M., Boxma O.J., Kella O. Workload analysis of a two-queue fluid polling model // Journal of Applied Probability. – 2023. – Vol. 60, No. 3. – P. 1003-1030.
174. Raaijmakers Y., Borst S., Boxma O. Fork–join and redundancy systems with heavy-tailed job sizes // Queueing Systems. – 2023. – Vol. 103, No. 1-2. – P. 131-159.
175. Devos A., Walraevens J., Fiems D., Bruneel H. Heavy-traffic comparison of a discrete-time generalized processor sharing queue and a pure randomly alternating service queue // Mathematics. – 2021. – Vol. 9, No. 21. – Art. No. 2723.
176. Fiems D. Strategic revenue management for discriminatory processor sharing queues // Lecture Notes in Computer Science. – 2023. – Vol. 14231. – P. 3-17.
177. Fiems D., Prabhu B.J. Strategic routing in heterogeneous discriminatory processor sharing queues // Lecture Notes in Computer Science. – 2025. – Vol. 15185. – P. 14-23.
178. Гудкова И.А. Методы анализа вероятностно-временных характеристик модели мультисервисной сети с потоковым и эластичным трафиком: диссертация на соискание ученой степени кандидата физико-математических наук. – Москва, 2011. – 130 с.

179. Гудкова И.А., Маркова Е.В. Модель управления доступом к радиоресурсам с индивидуальными потолками на скорость передачи данных в сети LTE // T-Comm: Телекоммуникации и транспорт. – 2014. – Т. 8, № 8. – С. 28-31.
180. Shorgin S.Y., Samouylov K.E., Gudkova I.A., Markova E.V., Sopin E.S. Approximating performance measures of radio admission control model for non real-time services with maximum bit rates in LTE // AIP Conference Proceedings. – 2015. – Vol. 1648. – Art. No. 250004.
181. Gudkova I.A., Markova E.V., Abaev P.O., Antonova V.M. Analytical modelling and simulation of admission control scheme for non-real time services in LTE networks // Proc. 29th European Conference on Modelling and Simulation (ECMS). – 2015. – P. 689-695.
182. Степанов С.Н., Степанов М.С. Планирование ресурса передачи при совместном обслуживании мультисервисного трафика реального времени и эластичного трафика данных // Автоматика и телемеханика. – 2017. – № 11. – С. 79-93.
183. Степанов С.Н., Степанов М.С. Приближенный метод оценки характеристик совместного обслуживания трафика реального времени и эластичного трафика данных в мультисервисных узлах доступа // Автоматика и телемеханика. – 2023. – № 11. – С. 93-114.
184. Borodakiy V.Y., Buturlin I.A., Gudkova I.A., Samouylov K.E. Modelling and analysing a dynamic resource allocation scheme for M2M traffic in LTE networks // Lecture Notes in Computer Science. – 2013. – Vol. 8121. – P. 420-426.
185. Gudkova I., Samouylov K., Buturlin I., Borodakiy V., Gerasimenko M., Galinina O., Andreev S. Analyzing impacts of coexistence between M2M and H2H Communication on 3GPP LTE System // Lecture Notes in Computer Science. – 2014. – Vol. 8458. – P. 162-174.
186. Бутурлин И.А., Гудкова И.А., Чукарин А.В. Модель распределения радиоресурсов с фиксированным диапазоном для трафика межмашинного взаимодействия в сети LTE // T-Comm: Телекоммуникации и транспорт. – 2014. – Т. 8, № 8. – С. 14-18.

187. Башарин Г.П., Самуйлов К.Е., Яркина Н.В., Гудкова И.А. Новый этап развития математической теории телетрафика // Автоматика и телемеханика. – 2009. – № 12. – С. 16-28.
188. Самуйлов К.Е., Абаев П.О., Гайдамака Ю.В., Гудкова И.А., Королькова А.В., Кулябов Д.С., Щукина О.Н. Мультисервисные сети связи. – М.: РУДН, 2013. – 363 с.
189. Basharin G.P., Gaidamaka Yu.V., Samouylov K.E. Mathematical theory of teletraffic and its application to the analysis of multiservice communication of next generation networks // Automatic Control and Computer Sciences. – 2013. – Vol. 47, No. 2. – P. 62-69.
190. Samuylov A., Moltchanov D., Gaidamaka Y., Andreev S., Koucheryavy Y. Random triangle: A baseline model for interference analysis in heterogeneous networks // IEEE Transactions on Vehicular Technology. – 2016. – Vol. 65, No. 8. – P. 6778-6782, Art. No. 7275184.
191. Kovalchukov R., Moltchanov D., Samuylov A., Ometov A., Andreev S., Koucheryavy Y., Samouylov K. Analyzing effects of directionality and random heights in drone-based mmWave communication // IEEE Transactions on Vehicular Technology. – 2018. – Vol. 67, No. 10. – P. 10064-10069, Art. No. 8412525.
192. Shurakov A., Rozhkova P., Khakimov A., Mokrov E., Prihodko A., Begishev V., Koucheryavy Y., Komarov M., Gol'Tsman G. Dynamic blockage in indoor reflection-aided sub-terahertz wireless communications // IEEE Access. – 2023. – Vol. 11. – P. 134677-134689.
193. Shurakov A., Moltchanov D., Prihodko A., Khakimov A., Mokrov E., Begishev V., Belikov I., Koucheryavy Y., Gol'tsman G. Empirical blockage characterization and detection in indoor sub-THz communications // Computer Communications. – 2023. – Vol. 201. – P. 48-58.
194. Orlov Y., Kirina-Lilinskaya E., Samuylov A., Ometov A., Moltchanov D., Gaimamaka Y., Andreev S., Samouylov K. Time-dependent SIR analysis in

- shopping malls using fractal-based mobility models // Lecture Notes in Computer Science. – 2017. – Vol. 10372. – P. 16-25.
195. Orlov Y.N., Kislitsyn A.A. Chernoff approximations for nonstationary random walk modeling // Lobachevskii Journal of Mathematics. – 2019. – Vol. 40, No. 12. – P. 2095-2102.
196. Poluektov D., Polovov M., Kharin P., Stusek M., Zeman K., Masek P., Gudkova I., Hosek J., Samouylov K. On the performance of LoRaWAN in smart city: end-device design and communication coverage // Lecture Notes in Computer Science. – 2019. – Vol. 11965. – P. 15-29.
197. Макеева Е.Д., Кочеткова И.А., Шоргин С.Я. Вероятностная модель затухания мощности сигнала в сценариях 3GPP TR 38.901 разворачивания сети 5G // Информатика и ее применения. – 2024. – Т. 18, № 2. – С. 25-31.
198. Tikhonenko O., Ziółkowski M. Queueing systems with random volume customers and their performance characteristics // Journal of Information and Organizational Sciences. – 2021. – Vol. 45, No. 1. – P. 21-38.
199. Tikhonenko O., Ziółkowski M., Kempa W.M. Queueing systems with random volume customers and a sectorized unlimited memory buffer // International Journal of Applied Mathematics and Computer Science. – 2021. – Vol. 31, No. 3. – P. 471-486.
200. Tikhonenko O., Ziółkowski M. MX/G/1/ ∞ single-server queueing system with random volume customers and multiple vacations // Bulletin of the Polish Academy of Sciences: Technical Sciences. – 2024. – Vol. 72, No. 5. – Art. No. e150813.
201. Naumov V.A., Samuilov K.E. Analysis of networks of the resource queueing systems // Automation and Remote Control. – 2018. – Vol. 79, No. 5. – P. 822-829.
202. Naumov V., Samouylov K. Product-form markovian queueing systems with multiple resources // Probability in the Engineering and Informational Sciences. – 2021. – Vol. 35, No. 1. – P. 180-188.
203. Naumov V., Samouylov K. Resource system with losses in a random environment // Mathematics. – 2021. – Vol. 9, No. 21. – Art. No. 2685.

204. Samouylov K., Gaidamaka Y. Analysis of loss systems with overlapping resource requirements // Statistical Papers. – 2018. – Vol. 59, No. 4. – P. 1463-1470.
205. Naumov V.A., Gaidamaka Y.V., Samouylov K.E. Computing the stationary distribution of queueing systems with random resource requirements via fast fourier transform // Mathematics. – 2020. – Vol. 8, No. 5. – Art. No. 772.
206. Moltchanov D., Sopin E., Begishev V., Samuylov A., Koucheryavy Y., Samouylov K. A Tutorial on mathematical modeling of 5G/6G millimeter wave and terahertz cellular systems // IEEE Communications Surveys and Tutorials. – 2022. – Vol. 24, No. 2. – P. 1072-1116.
207. Maslov A., Sopin E., Samouylov K. Accurate approximation for resource queueing systems with losses and signals // Mathematics. – 2025. – Vol. 13, No. 4. – Art. No. 619.
208. Lisovskaya E., Fedorova E., Salimzyanov R., Moiseeva S. Resource retrial queue with two orbits and negative customers // Mathematics. – 2022. – Vol. 10, No. 3. – Art. No. 321.
209. Pankratova E., Moiseeva S., Farkhadov M. Infinite-server resource queueing systems with different types of Markov-modulated poisson process and renewal arrivals // Mathematics. – 2022. – Vol. 10, No. 16. – Art. No. 2962.
210. Bushkova T., Moiseeva S., Moiseev A., Sztrik J., Lisovskaya E., Pankratova E. Using infinite-server resource queue with splitting of requests for modeling two-channel data transmission // Methodology and Computing in Applied Probability. – 2022. – Vol. 24, No. 3. – P. 1753-1772.
211. Pankratova E.V., Moiseeva S.P., Farhadov M.P., Moiseev A. Heterogeneous system MMPP/GI(2)/ ∞ with random customers capacities // Journal of Siberian Federal University - Mathematics and Physics. – 2019. – Vol. 12, No. 2. – P. 231-239.
212. Pankratova E., Moiseeva S., Farkhadov M. Infinite-server resource queueing systems with different types of Markov-modulated poisson process and renewal arrivals // Mathematics. – 2022. – Vol. 10, No. 16. – Art. No. 2962.

213. Моисеева С.П., Бушкова Т.В., Панкратова Е.В., Фархадов М.П., Имомов А.А. Асимптотический анализ ресурсной гетерогенной СМО (MMPP + 2M)(2,ν)/GI(2)/∞ при условии эквивалентно растущего времени обслуживания // Автоматика и телемеханика. – 2022. – № 8. – С. 81-99.
214. Рыков В.В. Маркированные марковские процессы и их применение // XIV Всероссийское совещание по проблемам управления: сборник научных трудов. – М.: ИПУ РАН, 2024. – С. 2424-2428.
215. Samouylov K., Gudkova I., Markova E., Dzantiev I. On analyzing the blocking probability of M2M transmissions for a CQI-based RRM scheme model in 3GPP LTE // Communications in Computer and Information Science. – 2016. – Vol. 638. – P. 327-340.
216. Маркова Е.В., Гольская А.А., Дзантиев И.Л., Гудкова И.А., Шоргин С.Я. Сравнительный анализ показателей эффективности модели беспроводной сети межмашинного взаимодействия, работающей в рамках двух политик разделения радиоресурсов // Информатика и ее применения. – 2019. – Т. 13, № 1. – С. 108-116.
217. Ahmadian A., Galinina O., Gudkova I.A., Andreev S., Shorgin S., Samouylov K. On capturing spatial diversity of joint M2M/H2H dynamic uplink transmissions in 3GPP LTE cellular system // Lecture Notes in Computer Science. – 2015. – Vol. 9247. – P. 407-421.
218. Pesando L., Fischer J.K., Shariati B., Freund R., Cananao J., Li H., Lin Y., Ferveur O., Jiang M., Jin J., Hillerkuss D., Brunner M., Zhou J., Del Junco J., Mkinsi H., Liu X. Standardization of the 5th generation fixed network for enabling end-to-end network slicing and quality-assured services // IEEE Communications Standards Magazine. – 2022. – Vol. 6, No. 4. – P. 96-103.
219. Dhanasekaran R.M., Ping J., Gomez G.P. End-to-end network slicing security across standards organizations // IEEE Communications Standards Magazine. – 2023. – Vol. 7, No. 1. – P. 40-47.

220. Lee H.-S., Moon S., Kim D.-Y., Lee J.-W. Packet-based fronthauling in 5G networks: Network slicing-aware packetization // IEEE Communications Standards Magazine. – 2023. – Vol. 7, No. 2. – P. 56-63.
221. Maule M., Vardakas J., Verikoukis C. 5G RAN slicing: Dynamic single tenant radio resource orchestration for eMBB traffic within a multi-slice scenario // IEEE Communications Magazine. – 2021. – Vol. 59, No. 3. – P. 110-116, Art. No. 9422343.
222. Ojaghi B., Adelantado F., Antonopoulos A., Verikoukis C. SlicedRAN: Service-aware network slicing framework for 5G radio access networks // IEEE Systems Journal. – 2022. – Vol. 16, No. 2. – P. 2556-2567.
223. Rezazadeh F., Zanzi L., Devoti F., Chergui H., Costa-Perez X., Verikoukis C. On the specialization of FDRL agents for scalable and distributed 6G RAN slicing orchestration // IEEE Transactions on Vehicular Technology. – 2023. – Vol. 72, No. 3. – P. 3473-3487.
224. Addad R.A., Dutra D.L.C., Taleb T., Flinck H. Toward using reinforcement learning for trigger selection in network slice mobility // IEEE Journal on Selected Areas in Communications. – 2021. – Vol. 39, No. 7. – P. 2241-2253. – Art. No. 9439923.
225. Benzaid C., Taleb T., Song J. AI-based autonomic and scalable security management architecture for secure network slicing in B5G // IEEE Network. – 2022. – Vol. 36, No. 6. – P. 165-174.
226. Addad R.A., Dutra D.L.C., Taleb T., Flinck H. AI-based network-aware service function chain migration in 5G and beyond networks // IEEE Transactions on Network and Service Management. – 2022. – Vol. 19, No. 1. – P. 472-484.
227. Bakri S., Brik B., Ksentini A. On using reinforcement learning for network slice admission control in 5G: Offline vs. online // International Journal of Communication Systems. – 2021. – Vol. 34, No. 7. – Art. No. e4757.
228. Bakri S., Frangoudis P.A., Ksentini A., Bouaziz M. Data-driven RAN slicing mechanisms for 5G and beyond // IEEE Transactions on Network and Service Management. – 2021. – Vol. 18, No. 4. – P. 4654-4668.

229. Boutiba K., Ksentini A., Brik B., Challal Y., Balla A. NRflex: Enforcing network slicing in 5G New Radio // *Computer Communications*. – 2022. – Vol. 181. – P. 284-292.
230. Zhirnov N.S., Lyakhov A.I., Khorov E.M. Mathematical Model of a network slicing approach for video and web traffic // *Journal of Communications Technology and Electronics*. – 2019. – Vol. 64, No. 8. – P. 890-899.
231. Liubogoshchev M., Zudin D., Krasilov A., Krotov A., Khorov E. DeSlice: An architecture for QoE-Aware and Isolated RAN Slicing // *Sensors*. – 2023. – Vol. 23, No. 9. – Art. No. 4351.
232. Yarkina N., Gaidamaka Y., Correia L.M., Samouylov K. An analytical model for 5G network resource sharing with flexible SLA-oriented slice isolation // *Mathematics*. – 2020. – Vol. 8, No. 7. – Art. No. 1177.
233. Yarkina N., Correia L.M., Moltchanov D., Gaidamaka Y., Samouylov K. Multi-tenant resource sharing with equitable-priority-based performance isolation of slices for 5G cellular systems // *Computer Communications*. – 2022. – Vol. 188. – P. 39-51.
234. Yarkina N., Gaydamaka A., Moltchanov D., Koucheryavy Y. Performance assessment of an ITU-T compliant machine learning enhancements for 5G RAN network slicing // *IEEE Transactions on Mobile Computing*. – 2024. – Vol. 23, No. 1. – P. 719-736.
235. Adou Y., Markova E., Gaidamaka Y. Modeling and analyzing preemption-based service prioritization in 5G networks slicing framework // *Future Internet*. – 2022. – Vol. 14, No. 10. – Art. No. 299.
236. Adou Y., Markova E., Gaidamaka Y. Streaming and elastic traffic service in 5G-sliced wireless networks and mutual utilization of guaranteed resource units // *Future Internet*. – 2024. – Vol. 16, No. 11. – Art. No. 397.
237. Ageev K., Garibyan A., Golskaya A., Gaidamaka Y., Sopin E., Samouylov K., Correia L.M. Modelling of virtual radio resources slicing in 5G networks // *Communications in Computer and Information Science*. – 2019. – Vol. 1109. – P. 150-161.

238. Ageev K., Sopin E., Samouylov K. Resource sharing model with minimum allocation for the performance analysis of network slicing // Communications in Computer and Information Science. – 2021. – Vol. 1391. – P. 378-389.
239. Vlaskina A., Polyakov N., Gudkova I. Modeling and performance analysis of elastic traffic with minimum rate guarantee transmission under network slicing // Lecture Notes in Computer Science. – 2019. – Vol. 11660. – P. 621-634.
240. Власкина А.С., Поляков Н.А., Гудкова И.А., Гайдамака Ю.В. Анализ вероятностно-временных характеристик обслуживания эластичного трафика с минимальной скоростью в сегменте беспроводной сети с нарезкой радиоресурсов // Известия Саратовского университета. Новая серия. Серия: Математика. Механика. Информатика. – 2020. – Т. 20, № 3. – С. 378-387.
241. Кочеткова И.А., Власкина А.С., Эсенг Манге П.Л.Э., Шоргин В.С. Задача разделения ресурса беспроводной сети по задержке передачи эластичного трафика // Системы и средства информатики. – 2024. – Т. 34, № 1. – С. 23-32.
242. Kochetkova I., Satin Y., Kovalev I., Makeeva E., Chursin A., Zeifman A. Convergence bounds for limited processor sharing queue with impatience for analyzing non-stationary file transfer in wireless network // Mathematics. – 2022. – Vol. 10, No. 1. – Art. No. 30.
243. Зейфман А.И. О нестационарной модели Эрланга // Автоматика и телемеханика. – 2009. – № 12. – С. 71-80.
244. Zeifman A., Satin Y., Kovalev I., Ammar S.I. Ergodicity and perturbation bounds for $Mt/Mt/1$ queue with balking, catastrophes, server failures and repairs // RAIRO - Operations Research. – 2021. – Vol. 55, No. 4. – P. 2223-2240.
245. Zeifman A., Satin Y., Sipin A. Bounds on the rate of convergence for $Mxt/MXt/1$ queueing models // Mathematics. – 2021. – Vol. 9, No. 15. – Art. No. 1752.
246. Usov I., Satin Y., Zeifman A. Estimating the rate of convergence of the $PH/M/1$ model by reducing to quasi-birth-death processes // Mathematics. – 2023. – Vol. 11, No. 6. – Art. No. 1494.
247. Вытовтов К.А., Барабанова Е.А. Аналитический метод анализа неоднородных непрерывных марковских процессов с кусочно-постоянными

- интенсивностями перехода // Автоматика и телемеханика. – 2021. – № 12. – С. 90-104.
248. Vishnevsky V., Vytovtov K., Barabanova E., Semenova O. Analysis of a MAP/M/1/N Queue with Periodic and Non-Periodic Piecewise Constant Input Rate // Mathematics. – 2022. – Vol. 10, No. 10. – Art. No. 1684.
249. Fendick K., Whitt W. Queues with path-dependent arrival processes // Journal of Applied Probability. – 2021. – Vol. 58, No. 2. – P. 484-504.
250. Fendick K., Whitt W. Heavy traffic limits for queues with non-stationary path-dependent arrival processes // Queueing Systems. – 2022. – Vol. 101, No. 1-2. – P. 113-135.
251. Fendick K., Whitt W. Queueing networks with path-dependent arrival processes // Queueing Systems. – 2023. – Vol. 105, No. 1-2. – P. 17-46.
252. Димитров Б.Н., Рыков В.В., Круглый З.Л. Периодические пуассоновские процессы и распределения с почти отсутствующей памятью // Автоматика и телемеханика. – 2004. – № 10. – С. 85-100.
253. Ховард Р.А. Динамическое программирование и марковские процессы / пер. с англ. В.В. Рыкова; под ред. Н.П. Бусленко. – М.: Советское радио, 1964. – 189 с.
254. Рыков В.В. Управляемые системы массового обслуживания // Итоги науки и техники. Серия «Теория вероятностей. Математическая статистика. Теоретическая кибернетика». – 1975. – Т. 12. – С. 43-153.
255. Горцев А.М., Назаров А.А., Терпугов А.Ф. Управление и адаптация в системах массового обслуживания. – Томск: ТГУ, 1978. – 208 с.
256. Назаров А.А. Управляемые системы массового обслуживания. – Томск: ТГУ, 1984. – 234 с.
257. Kitaev M.Yu., Rykov V.V. Controlled queueing systems. – CRC Press, 1995. – 304 p.
258. Пройдакова Е.В., Федоткин М.А. Построение и исследование математических моделей реальных процессов в виде управляющих систем. – М.: Русайнс, 2024. – 334 с.

259. Рыков В.В. Управляемые марковские процессы с конечным числом состояний и управлений // Теория вероятностей и ее применения. – 1966. – Т. 11, № 2. – Р. 343-351.
260. Рыков В.В. Об условиях монотонности оптимальных политик управления системами массового обслуживания // Автоматика и телемеханика. – 1999. – № 9. – С. 92-106.
261. Efrosinin D., Stepanova N. Estimation of the optimal threshold policy in a queue with heterogeneous servers using a heuristic solution and artificial neural networks // Mathematics. – 2021. – Vol. 9, No. 11. – Art. No. 1267.
262. Efrosinin D., Stepanova N. Optimal Control of Degrading Units through Threshold-Based Control Policies // Mathematics. – 2022. – Vol. 10, No. 21. – Art. No. 4098.
263. Efrosinin D., Vishnevsky V., Stepanova N. Optimal scheduling in general multi-queue system by combining simulation and neural network techniques // Sensors. – 2023. – Vol. 23, No. 12. – Art. No. 5479.
264. Fedotkin M.A., Zorine A.V. Optimization of control of conflict flows with repeated service // Journal of Mathematical Sciences (United States). – 2013. – Vol. 191, No. 4. – P. 492-505.
265. Rachinskaya M., Fedotkin M. Research of a multidimensional Markov Chain as a model for the class of queueing systems controlled by a threshold priority algorithm // Reliability: Theory and Applications. – 2018. – Vol. 13, No. 1. – P. 47-58.
266. Gomes A., Kibilda J., Marchetti N., Dasilva L.A. Dimensioning spectrum to support ultra-reliable low-latency communication // IEEE Communications Standards Magazine. – 2023. – Vol. 7, No. 1. – P. 88-93.
267. Wang S., Sun C. Blockchain empowered dynamic spectrum sharing: Standards, state of research and road ahead // IEEE Communications Standards Magazine. – 2023. – Vol. 7, No. 3. – P. 72-80.
268. Han X., Liang C., Liu R., Wei X., Chen M., Li Y.-N.R., Jin S. Flexible spectrum orchestration of carrier aggregation ForSg-advanced // IEEE Communications Standards Magazine. – 2023. – Vol. 7, No. 4. – P. 68-74.

269. Loginov V.A., Lyakhov A.I., Khorov E.M. Coexistence of Wi-Fi and LTE-LAA networks: Open issues // Journal of Communications Technology and Electronics. – 2018. – Vol. 63, No. 12. – P. 1530-1537.
270. Loginov V., Troegubov A., Lyakhov A., Khorov E. Coexistence of NR-U operators in multichannel scenarios: Fair cooperation or endless struggle for channel resources // Ad Hoc Networks. – 2025. – Vol. 173. – Art. No. 103798.
271. Andreev S., Galinina O., Pyattaev A., Johnsson K., Koucheryavy Y. Analyzing assisted offloading of cellular user sessions onto D2D links in unlicensed bands // IEEE Journal on Selected Areas in Communications. – 2015. – Vol. 33, No. 1. – P. 67-80, Art. No. 6952957.
272. Maule M., Moltchanov D., Kustarev P., Komarov M., Andreev S., Koucheryavy Y. Delivering fairness and QoS guarantees for LTE/Wi-Fi coexistence under LAA operation // IEEE Access. – 2018. – Vol. 6. – P. 7359-7373.
273. Lu X., Petrov V., Moltchanov D., Andreev S., Mahmoodi T., Dohler M. 5G-U: conceptualizing integrated utilization of licensed and unlicensed spectrum for future IoT // IEEE Communications Magazine. – 2019. – Vol. 57, No. 7. – P. 92-98, Art. No. 8722595.
274. Lu X., Sopin E., Petrov V., Galinina O., Moltchanov D., Ageev K., Andreev S., Koucheryavy Y., Samouylov K., Dohler M. Integrated use of licensed-and unlicensed-band mmWave radio technology in 5G and beyond // IEEE Access. – 2019. – Vol. 7. – P. 24376-24391, Art. No. 8643932.
275. Daraseliya A., Korshykov M., Sopin E., Moltchanov D., Andreev S., Samouylov K. Coexistence analysis of 5G NR unlicensed and WiGig in millimeter-wave spectrum // IEEE Transactions on Vehicular Technology. – 2021. – Vol. 70, No. 11. – P. 11721-11735.
276. Ponomarenko-Timofeev A., Pyattaev A., Andreev S., Koucheryavy Y., Mueck M., Karls I. Highly dynamic spectrum management within licensed shared access regulatory framework // IEEE Communications Magazine. – 2016. – Vol. 54, No. 3. – P. 100-109, Art. No. 7432155.

277. Masek P., Mokrov E., Zeman K., Ponomarenko-Timofeev A., Pyattaev A., Nesterov S., Andreev S., Hosek J., Samouylov K., Koucheryavy Y. A practical perspective on 5G-ready highly dynamic spectrum management with LSA // Wireless Communications and Mobile Computing. – 2018. – Vol. 2018. – Art. No. 2103868.
278. Morozov E., Rogozin S., Nguyen H.Q., Phung-Duc T. Modified erlang loss system for cognitive wireless networks // Mathematics. – 2022. – Vol. 10, No. 12. – Art. No. 2101.
279. Sadreddini Z., Masek P., Cavdar T., Ometov A., Hosek J., Gudkova I., Andreev S. Dynamic resource sharing in 5G with LSA: criteria-based management framework // Wireless Communications and Mobile Computing. – 2018. – Vol. 2018. – Art. No. 7302025.
280. Гайдамака Ю.В., Сопин Э.С., Гудкова И.А., Андреев С.Д., Самуйлов К.Е., Шоргин С.Я. Модели и методы анализа и расчета показателей эффективности беспроводных гетерогенных сетей. – М.: ФИЦ ИУ РАН, 2018. – 71 с.
281. Нетес В.А. Основы теории надежности. – М.: Горячая линия-Телеком, 2024. – 102 с.
282. Sindhu S., Krishnamoorthy A. MAP /Ek /1 queue with working vacation providing main service only in normal mode of service // Queueing Models and Service Management. – 2023. – Vol. 6, No. 2. – P. 1-7.
283. Sindhu S., Krishnamoorthy A., Kozyrev D. On queues with working vacation and interdependence in arrival and service processes // Mathematics. – 2023. – Vol. 11, No. 10. – Art. No. 2280.
284. Dudin A., Krishnamoorthy A., Dudin S., Dudina O. Queueing system with control by admission of retrial requests depending on the number of busy servers and state of the underlying process of Markov arrival process of primary requests // Annals of Operations Research. – 2024. – Vol. 335, No. 1. – P. 135-150.
285. Jain M., Dhibar S., Sanga S.S. Markovian working vacation queue with imperfect service, balking and retrial // Journal of Ambient Intelligence and Humanized Computing. – 2022. – Vol. 13, No. 4. – P. 1907-1923.

286. Kumar A., Jain M. Cost optimization of an unreliable server queue with two stage service process under hybrid vacation policy // *Mathematics and Computers in Simulation*. – 2023. – Vol. 204. – P. 259-281.
287. Rani S., Jain M., Meena R.K. Queueing modeling and optimization of a fault-tolerant system with reboot, recovery, and vacationing server operating under admission control policy // *Mathematics and Computers in Simulation*. – 2023. – Vol. 209. – P. 408-425.
288. Panda G., Goswami V. Equilibrium joining strategies of positive customers in a markovian queue with negative arrivals and working vacations // *Methodology and Computing in Applied Probability*. – 2022. – Vol. 24, No. 3. – P. 1439-1466.
289. Goswami V., Panda G. Synchronized abandonment in discrete-time renewal input queues with vacations // *Journal of Industrial and Management Optimization*. – 2022. – Vol. 18, No. 6. – P. 4373-4391.
290. Goswami V., Panda G. Analysis of renewal batch arrival queues with multiple vacations and geometric abandonment // *Methodology and Computing in Applied Probability*. – 2024. – Vol. 26, No. 2. – Art. No. 20.
291. Divya V., Krishnamoorthy A., Vishnevsky V.M., Kozyrev D.V. On a queueing system with processing of service items under vacation and N-policy with impatient customers // *Queueing Models and Service Management*. – 2020. – Vol. 3, No. 2. – P. 167-201.
292. Вишневский В.М., Семенова О.В. Об одной модели оценки производительности широкополосного гибридного канала связи на основе лазерной и радиотехнологий // *Проблемы информатики*. – 2010. – № 2(6). – С. 43-58.
293. Вишневский В.М., Андронов А.М. Оценка производительности беспроводных гибридных систем, функционирующих в полумарковской случайной среде // *Автоматика и телемеханика*. – 2017. – № 12. – С. 70-83.
294. Danilyuk E.Y., Moiseeva S.P., Sztrik J. Asymptotic analysis of retrial queueing system m/m/1 with impatient customers, collisions and unreliable server // *Journal*

- of Siberian Federal University - Mathematics and Physics. – 2020. – Vol. 13, No. 2. – P. 218-230.
295. Nazarov A., Paul S., Lizyura O. Asymptotic analysis of markovian retrial queue with unreliable server and multiple types of outgoing calls // *Global and Stochastic Analysis*. – 2021. – Vol. 8, No. 3. – P. 151-158.
296. Danilyuk E., Plekhanov A., Moiseeva S., Sztrik J. Asymptotic diffusion analysis of retrial queueing system M/M/1 with impatient customers, collisions and unreliable servers // *Axioms*. – 2022. – Vol. 11, No. 12. – Art. No. 699.
297. Do N.H., Van Do T., Melikov A. Equilibrium customer behavior in the M/M/1 retrial queue with working vacations and a constant retrial rate // *Operational Research*. – 2020. – Vol. 20, No. 2. – P. 627-646.
298. Melikov A., Aliyeva S., Sztrik J. Retrial queues with unreliable servers and delayed feedback // *Mathematics*. – 2021. – Vol. 9, No. 19. – Art. No. 2415.
299. Masek P., Mokrov E., Pyattaev A., Zeman K., Ponomarenko-Timofeev A., Samuylov A., Sopin E., Hosek J., Gudkova I.A., Andreev S., Novotny V., Koucheryavy Y., Samouylov K. Experimental evaluation of dynamic licensed shared access operation in live 3GPP LTE system // *Proc. 59th IEEE Global Communications Conference (GLOBECOM)*. – 2016. – Art. No. 7841826.
300. Mokrov E., Gudkova I. Performance evaluation of dynamic LSA operation through a model of a stand-alone cell // *CEUR Workshop Proceedings*. – 2016. – Vol. 1763. – P. 35-41.
301. Mokrov E., Ponomarenko-Timofeev A., Gudkova I., Masek P., Hosek J., Andreev S., Koucheryavy Y., Gaidamaka Y. Modeling transmit power reduction for a typical cell with licensed shared access capabilities // *IEEE Transactions on Vehicular Technology*. – 2018. – Vol. 67, No. 6. – P. 5505-5509.
302. Gudkova I., Korotysheva A., Zeifman A., Shilova G., Korolev V., Shorgin S., Razumchik R. Modeling and analyzing licensed shared access operation for 5G network as an inhomogeneous queue with catastrophes // *Proc. 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. – 2016. – P. 282-287, Art. No. 7765372.

- 303. Markova E., Satin Y., Kochetkova I., Zeifman A., Sinitcina A. Queuing system with unreliable servers and inhomogeneous intensities for analyzing the impact of non-stationarity to performance measures of wireless network under licensed shared access // Mathematics. – 2020. – Vol. 8, No. 5. – Art. No. 800.
- 304. Naumov V. A Matrix-multiplicative solution for multi-dimensional QBD processes // Mathematics. – 2024. – Vol. 12, No. 3. – Art. No. 444.
- 305. Naumov V., Samouylov K. Multi-dimensional Markov chains of M/G/1 type // Mathematics. – 2025. – Vol. 13, No. 2. – Art. No. 209.
- 306. Naumov V., Samouylov K. Computing the matrix G of multi-dimensional Markov chains of M/G/1 type // Mathematics. – 2025. – Vol. 13, No. 8. – Art. No. 1223.
- 307. Efrosinin D., Vishnevsky V., Stepanova N., Sztrik J. Use cases of machine learning in queueing theory based on a GI/G/K system // Mathematics. – 2025. – Vol. 13, No. 5. – Art. No. 776.
- 308. Tsitsiashvili G., Radchenkova T. Queuing system without queue and deterministic service time // Reliability: Theory and Applications. – 2022. – Vol. 17, No. 3. – P. 301-305.
- 309. Tsitsiashvili G., Osipova M. Invariance of stationary distributions of exponential networks with prohibitions and determination of maximum prohibitions // Computation. – 2025. – Vol. 13, No. 4. – Art. No. 89.
- 310. Tsitsiashvili G. Construction and analysis of queueing and reliability models using random graphs // Mathematics. – 2021. – Vol. 9, No. 19. – Art. No. 2511.
- 311. Джейсуол Н. Очереди с приоритетами / пер. с англ. И.С. Нефедовой и В.С. Манускевича; под ред. В.В. Калашникова. – М.: Мир, 1973. – 280 с.
- 312. Гнеденко Б.В., Даниелян Э.А., Димитров Б.Н., Климов Г.П., Матвеев В.Ф. Приоритетные системы обслуживания. – М.: МГУ, 1973. – 448 с.
- 313. Бронштейн О.И., Духовный И.М. Модели приоритетного обслуживания в информационно-вычислительных сетях. – М.: Наука, 1976. – 220 с.
- 314. Климов Г.П., Мишкой Г.К. Приоритетные системы обслуживания с ориентацией. – М.: МГУ, 1979. – 223 с.

- 315. Bruneel H., Devos A. Explicit solutions for coupled parallel queues // Mathematics. – 2024. – Vol. 12, No. 15. – Art. No. 2345.
- 316. Bruneel H., Devos A. Coupled queues with server interruptions: Some solutions // Performance Evaluation. – 2025. – Vol. 167. – Art. No. 102466.
- 317. Bruneel H. Analysis of a threshold-based priority queue // Queueing Systems. – 2025. – Vol. 109, No. 1. – Art. No. 8.
- 318. De Clercq S., Walraevens J. Delay analysis of a two-class priority queue with external arrivals and correlated arrivals from another node // Annals of Operations Research. – 2020. – Vol. 293, No. 1. – P. 57-72.
- 319. Walraevens J., Van Giel T., De Vuyst S., Wittevrongel S. Asymptotics of waiting time distributions in the accumulating priority queue // Queueing Systems. – 2022. – Vol. 101, No. 3-4. – P. 221-244.
- 320. Fiems D., Walraevens J. Recursive calculation of moments in priority queues // Operations Research Letters. – 2022. – Vol. 50, No. 5. – P. 524-528.
- 321. Ghosh S., Banik A.D., Chaudhry M.L. Analysis of BMAP/R/1 queues under gated-limited service with the server's single vacation policy // Infosys Science Foundation Series in Mathematical Sciences. – 2020. – P. 103-128.
- 322. Kim N.K., Yang W.S., Kim K., Chaudhry M.L. On the largest-group-first priority rule in queueing // Industrial Engineering and Management Systems. – 2022. – Vol. 21, No. 2. – P. 267-276.
- 323. Бронштейн О.И., Рыков В.В. Об оптимальных приоритетах в системах массового обслуживания // Известия Академии наук СССР. Техническая кибернетика. – 1965. – № 6. – С. 28-37.
- 324. Рыков В.В. Регенерирующие процессы с вложенными периодами регенерации и их применение при исследовании приоритетных систем массового обслуживания // Кибернетика. – 1975. – № 6. – С. 105-111.
- 325. Мишкой Г.К., Рыков В.В., Джиордано С., Бежан А.Ю. Многомерные аналоги уравнения Кендалла для приоритетных систем: вычислительные аспекты // Автоматика и телемеханика. – 2008. – № 6. – С. 82-95.

- 326. Nair D.V., Krishnamoorthy A., Melikov A., Aliyeva S. MMAP/(PH,PH)/1 queue with priority loss through feedback // Mathematics. – 2021. – Vol. 9, No. 15. – Art. No. 1797.
- 327. Dudin A., Dudina O., Dudin S., Melikov A. Analysis of a queueing model with flexible priority, batch arrival, and impatient customers // Computation. – 2025. – Vol. 13, No. 3. – Art. No. 77.
- 328. Башарин Г.П. Об обслуживании двух потоков с относительным приоритетом на полностью доступной системе с ограниченным числом мест для ожидания // Известия Академии наук СССР. Техническая кибернетика. 1967. – № 2. – С. 72-86.
- 329. Башарин Г.П. Обслуживание двух потоков на однолинейной системе с ограниченным числом мест для ожидания и абсолютным приоритетом // Известия Академии наук СССР. Техническая кибернетика. 1967. – № 5. – С. 106-116.
- 330. Башарин Г.П. О пуассоновских обслуживающих системах с абсолютным приоритетом и обратной связью // Сборник «Массовое обслуживание в системах передачи информации». – М.: Наука, 1969. – С. 3-20.
- 331. Степанов С.Н. Численные методы расчета систем с повторными вызовами. – М.: Наука, 1983. – 230 с.
- 332. Stepanov S.N. Generalized model with repeated calls in case of extreme load // Queueing Systems. – 1997. – Vol. 27, No. 1-2. – P. 131-151.
- 333. Stepanov S.N. Markov models with retrials: The calculation of stationary performance measures based on the concept of truncation // Mathematical and Computer Modelling. – 1999. – Vol. 30, No. 3-4. – P. 207-228.
- 334. Stepanov S., Stepanov M. Estimation of the performance measures of a group of servers taking into account blocking and call repetition before and after server occupation // Mathematics. – 2021. – Vol. 9, No. 21. – Art. No. 2811.
- 335. Nazarov A., Moiseev A., Phung-Duc T., Paul S. Diffusion limit of multi-server retrial queue with setup time // Mathematics. – 2020. – Vol. 8, No. 12. – Art. No. 2232.

- 336. Nazarov A., Phung-Duc T., Paul S., Lizyura O. Diffusion limit for single-server retrial queues with renewal input and outgoing calls // *Mathematics*. – 2022. – Vol. 10, No. 6. – Art. No. 948.
- 337. Nazarov A., Fedorova E., Lizyura O., Salimzyanov R. Asymptotic diffusion method for retrial queues with state-dependent service rate // *Mathematics*. – 2023. – Vol. 11, No. 14. – Art. No. 3140.
- 338. Sztrik J., Tóth Á., Pintér Á., Bács Z. The effect of operation time of the server on the performance of finite-source retrial queues with two-way communications to the orbit // *Journal of Mathematical Sciences (United States)*. – 2022. – Vol. 267, No. 2. – P. 196-204.
- 339. Tóth Á., Sztrik J. Analysis of retrial queueing systems with two-way communication and impatient customers using simulation // *Annales Mathematicae et Informaticae*. – 2023. – Vol. 58. – P. 160-169.
- 340. Dragieva V.I., Phung-Duc T. A finite-source M/G/1 retrial queue with outgoing calls // *Annals of Operations Research*. – 2020. – Vol. 293, No. 1. – P. 101-121.
- 341. Devos A., Walraevens J., Phung-Duc T., Bruneel H. Analysis of the queue lengths in a priority retrial queue with constant retrial policy // *Journal of Industrial and Management Optimization*. – 2020. – Vol. 16, No. 6. – P. 2813-2842.
- 342. Phung-Duc T., Akutsu K., Kawanishi K., Salameh O., Wittevrongel S. Queueing models for cognitive wireless networks with sensing time of secondary users // *Annals of Operations Research*. – 2022. – Vol. 310, No. 2. – P. 641-660.
- 343. Li K., Wang J. Equilibrium balking strategies in the single-server retrial queue with constant retrial rate and catastrophes // *Quality Technology and Quantitative Management*. – 2021. – Vol. 18, No. 2. – P. 156-178.
- 344. Wang Z., Wang J. Strategic priority-purchasing and joining rules in a retrial queue // *IMA Journal of Management Mathematics*. – 2021. – Vol. 32, No. 2. – P. 161-194.
- 345. Zhang Y., Wang J. Managing retrial queueing systems with boundedly rational customers // *Journal of the Operational Research Society*. – 2023. – Vol. 74, No. 3. – P. 748-761.

- 346. Klimenok V.I., Dudin A.N., Vishnevsky V.M., Semenova O.V. Retrial BMAP/PH/N queueing system with a threshold-dependent inter-retrial time distribution // Mathematics. – 2022. – Vol. 10, No. 2. – Art. No. 269.
- 347. Melikov A., Aliyeva S., Sztrik J. Analysis of instantaneous feedback queue with heterogeneous servers // Mathematics. – 2020. – Vol. 8, No. 12. – Art. No. 2186.
- 348. Melikov A., Chakravarthy S.R., Aliyeva S. A retrial queueing model with feedback // Queueing Models and Service Management. – 2023. – Vol. 6, No. 1. – P. 63-95.
- 349. Morozov E.V., Peshkova I.V., Rumyantsev A.S. Bounds and maxima for the workload in a multiclass orbit queue // Mathematics. – 2023. – Vol. 11, No. 3. – Art. No. 564.
- 350. Morozov E., Rumyantsev A., Dey S., Deepak T.G. Performance analysis and stability of multiclass orbit queue with constant retrial rates and balking // Performance Evaluation. – 2019. – Vol. 134. – Art. No. 102005.
- 351. Алиев Т.И., Махарева Э. Дисциплины обслуживания на основе матрицы приоритетов // Научно-технический вестник информационных технологий, механики и оптики. – 2014. – Т. 6, № 94. – С. 91-97.
- 352. Samouylov K., Gudkova I., Markova E. Formalizing set of multiservice models for analyzing pre-emption mechanisms in wireless 3GPP networks // Communications in Computer and Information Science. – 2016. – Vol. 601. – P. 61-71.
- 353. Самуйлов К.Е., Гудкова И.А., Маркова Е.В. К разработке исходных данных для сценария приоритетного управления доступом в мультисервисной сети LTE // Т-Comm: Телекоммуникации и транспорт. – 2015. – Т. 9, № 10. – С. 22-27.
- 354. Гудкова И.А., Маркова Е.В., Матвейчук И.В. Анализ одной схемы управления доступом к радиоресурсам сети LTE // Т-Comm: Телекоммуникации и транспорт. – 2012. – Т. 6, № 7. – С. 67-70.
- 355. Gudkova I.A., Samouylov K.E. Modelling a radio admission control scheme for video telephony service in wireless networks // Lecture Notes in Computer Science. – 2012. – Vol. 7469. – P. 208-215.

356. Кочеткова И.А., Самуйлов К.Е. Классификация моделей схем приоритетного доступа в беспроводных сетях // XIV Всероссийское совещание по проблемам управления: сборник научных трудов. – М.: ИПУ РАН, 2024. – С. 2350-2354.
357. Кочеткова И.А., Самуйлов К.Е. Мультисервисные системы с приоритетным обслуживанием трафика // М.: Техносфера, 2025. – 160 с.
358. Makeeva E., Polyakov N., Kharin P., Gudkova I. Probability model for performance analysis of joint URLLC and eMBB transmission in 5G networks // Lecture Notes in Computer Science. – 2019. – Vol. 11660. – P. 635-648.
359. Харин П.А., Макеева Е.Д., Кочеткова И.А., Ефросинин Д.В., Шоргин С.Я. Система массового обслуживания с орбитами для анализа совместного обслуживания трафика с малыми задержками URLLC и широкополосного доступа eMBB в беспроводных сетях пятого поколения // Информатика и ее применения. – 2020. – Т. 14, № 4. – С. 17-24.
360. Makeeva E., Kochetkova I., Alkanhel R. Retrial queueing system for analyzing impact of priority ultra-reliable low-latency communication transmission on enhanced mobile broadband quality of service degradation in 5G networks // Mathematics. – 2023. – Vol. 11, No. 18. – Art. No. 3925.
361. Кочеткова И.А., Кущазли А.И., Харин П.А., Шоргин С.Я. Модель схемы приоритетного доступа трафика URLLC и eMBB в сети пятого поколения в виде ресурсной системы массового обслуживания // Информатика и ее применения. – 2021. – Т. 15, № 4. – С. 87-92.
362. Kochetkova I., Makeeva E., Ageeva A., Gorshenin A. Model for analyzing impact of path loss on eMBB bit rate degradation under priority URLLC transmission in 5G network // Lecture Notes in Computer Science. – 2022. – Vol. 13766. – P. 176-189.
363. Агеева А.С., Макеева Е.Д., Кочеткова И.А. Расчет характеристик модели приоритетного доступа трафика URLLC с адаптивным изменением скорости трафика eMBB в условиях затухания сигнала сети 5G // Свидетельство о регистрации программы для ЭВМ №2023663827, Опубл.28.06.2023 Бюл. №7.

364. Kushchazli A., Ageeva A., Kochetkova I., Kharin P., Chursin A., Shorgin S. Model of radio admission control for URLLC and adaptive bit rate eMBB in 5G network // CEUR Workshop Proceedings. – 2021. – Vol. 2946. – P. 74-84.
365. Кочеткова И.А., Кушазли А.И., Харин П.А., Шоргин С.Я. Модель для анализа приоритетного доступа трафика URLLC при прерывании обслуживания и снижении скорости передачи сессий eMBB в сети 5G // Системы и средства информатики. – 2021. – Т. 31, № 3. – С. 123-134.
366. Макеева Е.Д., Кочеткова И.А., Шоргин В.С. Модель для выбора уровней скорости широкополосного трафика eMBB в условиях приоритетной передачи трафика URLLC в сети 5G // Системы и средства информатики. – 2023. – Т. 33, № 4. – С. 60-68.
367. Агеева А.С., Кушазли А.И., Харин П.А., Кочеткова И.А. Расчет вероятностно-временных характеристик для модели приоритетного доступа трафика URLLC с адаптивным изменением скорости трафика eMBB в сети 5G // Свидетельство о регистрации программы для ЭВМ №2021661640 Оpubл.14.07.2021 Бюл. №7.
368. Бородакий В.Ю., Гудкова И.А., Маркова Е.В. Рекуррентный алгоритм для расчета характеристик модели приоритетного управления доступом в сети LTE // Т-Comm: Телекоммуникации и транспорт. – 2013. – Т. 7, № 11. – С. 45-49.
369. Borodakiy V.Y., Gudkova I.A., Markova E.V., Samouylov K.E. Modelling and performance analysis of pre-emption based radio admission control scheme for video conferencing over LTE // Proc. 6th ITU Kaleidoscope Academic Conference (K). – 2014. – P. 53-59, Art. No. 6858480.
370. Гудкова И.А., Маркова Е.В. Рекуррентный алгоритм расчета стационарного распределения вероятностей состояний модели с прерыванием одноадресных соединений трафиком мультимедиа // Вестник Российского университета дружбы народов. Серия: Математика, информатика, физика. – 2015. – № 3. – С. 10-17.

371. Borodakiy V.Y., Samouylov K.E., Gudkova I.A., Markova E.V. Analyzing mean bit rate of multicast video conference in LTE network with adaptive radio admission control scheme // *Journal of Mathematical Sciences (United States)*. – 2016. – Vol. 218, No. 3. – P. 257-268.
372. Markova E., Moltchanov D., Gudkova I., Samouylov K., Koucharyavy Y. Performance assessment of QoS-aware LTE sessions offloading onto LAA/WiFi systems // *IEEE Access*. – 2019. – Vol. 7. – P. 36300-36311, Art. No. 8672559.
373. Markova E., Moltchanov D., Sinitsyna A., Ivanova D., Filipova V., Gudkova I., Samouylov K. Analytical models for schedule-based license assisted access (LAA) LTE systems // *Lecture Notes in Computer Science*. – 2018. – Vol. 11118. – P. 210-223.
374. Кочеткова И.А., Власкина А.С., Ву Н.Н., Шоргин В.С. Система массового обслуживания с управляемым по сигналам перераспределением приборов для анализа нарезки ресурсов сети 5G // *Информатика и ее применения*. – 2021. – Т. 15, № 3. – С. 91-97.
375. Филиппова В.С., Хусайнова Ф.Д., Власкина А.С., Кочеткова И.А., Бурцева С.А. Расчет показателей эффективности модели управления нарезкой радиоресурсов беспроводной сети между двумя виртуальными операторами по сигналам контроллера // *Свидетельство о регистрации программы для ЭВМ №2022660727 Оpubл.08.06.2022 Бюл. №6*.
376. Власкина А.С., Бурцева С.А., Кочеткова И.А., Шоргин С.Я. Управляемая система массового обслуживания с эластичным трафиком и сигналами для анализа нарезки ресурсов в сети радиодоступа // *Информатика и ее применения*. – 2022. – Т. 16, № 3. – С. 90-96.
377. Kochetkova I., Leonteva K., Ghebrial I., Vlaskina A., Burtseva S., Kushchazli A., Samouylov K. Controllable queuing system with elastic traffic and signals for resource capacity planning in 5G network slicing // *Future Internet*. – 2024. – Vol. 16, No. 1. – Art. No. 18.
378. Ghebrial I., Leonteva K., Kochetkova I., Shorgin S. Markov decision process and artificial neural network for resource capacity planning in 5G network slicing //

- Communications in Computer and Information Science. – 2024. – Vol. 2163. – P. 33-47.
379. Леонтьева К.А., Гебриал И.Е.З., Бурдина К.П., Бурцева С.А., Власкина А.С., Кочеткова И.А. Расчет стационарного распределения вероятностей состояний модели нарезки радиоресурсов для заданной политики перераспределения ресурса между двумя виртуальными операторами // Свидетельство о регистрации программы для ЭВМ №2023662074 Оpubл.06.06.2023 Бюл. №6.
380. Власкина А.С., Кочеткова И.А. Расчет оптимальной политики управляемого по внешним событиям динамического разделения ресурса для обслуживания двух потоков заявок на выполнение работы заданного объема // Свидетельство о регистрации программы для ЭВМ №2023667476 Оpubл. 15.08.2023. Бюл. №8.
381. Гебриал И.Е.З., Леонтьева К.А., Кочеткова И.А. Расчет оптимальной политики управляемого контроллером перераспределения ресурса между двумя виртуальными операторами с использованием нейронной сети // Свидетельство о регистрации программы для ЭВМ №2024662703 Оpubл.30.05.2024. Бюл. №6.
382. Kochetkova I., Vlaskina A., Burtseva S., Savich V., Hosek J. Analyzing the effectiveness of dynamic network slicing procedure in 5G network by queuing and simulation models // Lecture Notes in Computer Science. – 2020. – Vol. 12525. – P. 71-85.
383. Ву Н.Н., Власкина А.С., Кочеткова И.А., Бурцева С.А. Расчет показателей эффективности модели динамической нарезки между двумя сегментами радиоресурсов с управлением по внешнему событию // Свидетельство о регистрации программы для ЭВМ RU №2021661716, Оpubл.14.07.2021 Бюл. № 7.
384. Samouylov K., Naumov V., Sopin E., Gudkova I., Shorgin S. Sojourn time analysis for processor sharing loss system with unreliable server // Lecture Notes in Computer Science. – 2016. – Vol. 9845. – P. 284-297.

- 385. Samouylov K., Sopin E., Gudkova I. Sojourn time analysis for processor sharing loss queueing system with service interruptions and MAP arrivals // Communications in Computer and Information Science. – 2016. – Vol. 678. – P. 406-417.
- 386. Mokrov E., Sopin E., Markova E., Poluektov D., Gudkova I., Masek P., Hosek J. Modeling and response time analysis for web browsing under interruptions in LTE network // Proc. 31st European Conference on Modelling and Simulation (ECMS), 2017. – P. 706-712.
- 387. Ometov A., Sopin E., Gudkova I., Andreev S., Gaidamaka Y.V., Koucheryavy Y. Modeling unreliable operation of mmWave-based data sessions in mission-critical PPDR services // IEEE Access. – 2017. – Vol. 5. – P. 20536-20544, Art. No. 8055435.
- 388. Borodakiy V.Y., Samouylov K.E., Gudkova I.A., Ostrikova D.Y., Ponomarenko-Timofeev A.A., Turlikov A.M., Andreev S.D. Modeling unreliable LSA operation in 3GPP LTE cellular networks // Proc. 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). – 2014. – P. 390-396, Art. No. 7002133.
- 389. Efrosinin D., Samouylov K., Gudkova I. Busy period analysis of a queueing system with breakdowns and its application to wireless network under licensed shared access regime // Lecture Notes in Computer Science. – 2016. – Vol. 9870. – P. 426-439.
- 390. Gudkova I., Samouylov K., Ostrikova D., Mokrov E., Ponomarenko-Timofeev A., Andreev S., Koucheryavy Y. Service failure and interruption probability analysis for licensed shared access regulatory framework // Proc. 7th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). – 2015. – P. 123-131, Art. No. 7382416.
- 391. Efrosinin D., Gudkova I., Stepanova N. Performance analysis and optimal control for queueing system with a reserve unreliable server pool // Communications in Computer and Information Science. – 2019. – Vol. 1109. – P. 109-120.

- 392. Gudkova I., Markova E., Masek P., Andreev S., Hosek J., Yarkina N., Samouylov K., Koucheryavy Y. Modeling the utilization of a multi-tenant band in 3GPP LTE system with licensed shared access // Proc. 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). – 2016. – P. 119-123, Art. No. 7765343.
- 393. Golskaia A.A., Markova E.V., Dzantiev I.L., Gudkova I.A. Analysis of performance measures for scheme model with limit power policy within LSA framework // CEUR Workshop Proceedings. – 2017. – Vol. 1995. – P. 42-52.
- 394. Markova E., Dzantiev I., Gudkova I., Shorgin S. Analyzing impact of path loss models on probability characteristics of wireless network with licensed shared access framework // Proc. 9th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). – 2017. – P. 20-25.
- 395. Гудкова И.А., Шоргин С.Я. Вероятностная модель совместного использования ресурсов беспроводной сети с адаптивным управлением мощностью // Информатика и ее применения. – 2017. – Т. 11, № 3. – С. 90-98.
- 396. Markova E., Gudkova I., Ometov A., Dzantiev I., Andreev S., Koucheryavy Y., Samouylov K. Flexible spectrum management in a smart city within licensed shared access framework // IEEE Access. – 2017. – Vol. 5. – P. 22252-22261, Art. No. 8055552.
- 397. Markova E., Poluektov D., Ostrikova D., Gudkova I., Dzantiev I., Samouylov K., Shorgin V. Analysis of admission control schemes models for wireless network under licensed shared access framework // Lecture Notes in Computer Science. – 2017. – Vol. 10531. – P. 536-549.
- 398. Markova E.V., Poluektov D.S., Ostrikova D.U., Gudkova I.A. On comparative analysis of shared access scheme models with service interruption in 3GPP LTE network with LSA framework // CEUR Workshop Proceedings. – 2017. – Vol. 2064. – P. 112-120.

399. Samouylov K., Gudkova I., Markova E., Yarkina N. Queuing model with unreliable servers for limit power policy within licensed shared access framework // Lecture Notes in Computer Science. – 2016. – Vol. 9870. – P. 404-413.
400. Ivanova D., Karnauhova E., Markova E., Gudkova I. Analyzing of licensed shared access scheme model with service bit rate degradation in 3GPP network // Communications in Computer and Information Science. – 2017. – Vol. 800. – P. 231-242.