



Молчанов Дмитрий Александрович — доктор технических наук, профессор кафедры прикладной информатики и теории вероятностей РУДН. Автор более 200 публикаций по сетям связи, гетерогенным беспроводным сетям, Интернету вещей, прикладной теории массового обслуживания. За свою карьеру прочитал более 70 курсов по беспроводным и проводным сетевым технологиям на русском и английском языках. В сферу научных интересов входят исследования и разработка систем 5G+/6G.



Бегисhev Вячеслав Олегович — кандидат физико-математических наук, доцент кафедры прикладной информатики и теории вероятностей РУДН. Автор более 50 научных публикаций по темам, связанным с областями изучения сетей беспроводной связи 5G+/6G, анализом качества обслуживания, оценкой производительности беспроводных сетей в mmWave/ТГц диапазонах частот.



Самуйлов Константин Евгеньевич — доктор технических наук, профессор, заведующий кафедрой прикладной информатики и теории вероятностей РУДН. Член редколлегии и рецензент многих научных журналов, сопредседатель и член программных комитетов нескольких международных конференций. Его исследовательские интересы включают прикладные задачи теории вероятностей, теории массового обслуживания и математической теории телетрафика, модели и методы анализа производительности сетей 5G+/6G, алгоритмы распределения ресурсов в гетерогенных беспроводных сетях и др. Автор более 250 научных статей и семи книг.



Кучерявый Евгений Андреевич — доктор технических наук, профессор Департамента электронной инженерии МИЭМ НИУ ВШЭ. Автор более 350 публикаций в области передовых проводных и беспроводных сетей и коммуникаций. Его текущие исследовательские интересы включают различные аспекты гетерогенных сетей и систем беспроводной связи 5G/5G+/6G, а также Интернет вещей. Его h-index в Scopus на момент публикации книги составлял 42.



9 785209 110897

АРХИТЕКТУРА, ТЕХНОЛОГИИ,
МЕТОДЫ АНАЛИЗА И РАСЧЕТА

СЕТИ 5G/6G:

Д.А. Молчанов, В.О. Бегисhev,
К.Е. Самуйлов, Е.А. Кучерявый

Д.А. Молчанов, В.О. Бегисhev,
К.Е. Самуйлов, Е.А. Кучерявый

СЕТИ 5G/6G:

АРХИТЕКТУРА, ТЕХНОЛОГИИ,
МЕТОДЫ АНАЛИЗА
И РАСЧЕТА



Москва
Российский университет дружбы народов
2022

УДК 004(035.3)
ББК 16.2
С33

Утверждено
РИС Ученого совета
Российского университета
дружбы народов



Издание осуществлено при финансовой поддержке Российского фонда фундаментальных исследований по проекту № 21-17-00012, не подлежит продаже

Рецензенты:

заведующий кафедрой сетей связи и систем коммутации МТУСИ
доктор технических наук, профессор *С.Н. Степанов*;

профессор кафедры прикладной информатики и теории вероятностей РУДН
доктор физико-математических наук, профессор *Ю.В. Гайдамака*

Авторы:

Д.А. Молчанов, В.О. Бегишев, К.Е. Самуйлов, Е.А. Кучерявый

С33 Сети 5G/6G: архитектура, технологии, методы анализа и расчета : монография / Д. А. Молчанов, В. О. Бегишев, К. Е. Самуйлов, Е. А. Кучерявый. – Москва : РУДН, 2022. – 516 с. : ил.

Монография представляет собой, с одной стороны, доступное руководство по основным технологиям сетей 5G/6G, а с другой – последовательное изложение методов анализа и расчета показателей эффективности их характеристик в рамках конкретных сценариев и системных моделей. Показаны основные приемы и принципы применения теории стохастического моделирования беспроводных сетей. Рассмотрены общие вопросы архитектуры, основных услуг и понятия сетей 5G/6G, а также исследованы модели, характеристики, методы анализа радиоканалов и эффективности обслуживания абонентов сетей 5G/6G.

Издание предназначено для научных работников, специалистов в области анализа показателей эффективности функционирования систем и сетей телекоммуникаций, преподавателей профильных кафедр, аспирантов и студентов старших курсов естественно-научных и инженерно-технических специальностей.

This book serves two purposes. At one hand, it provides an easy-to-follow guide on the critical technologies comprising 5G/6G cellular systems, while, on the other hand, it presents detailed methods for performance evaluation and optimization of user- and system-centric metrics of interest for an extensive set of 5G/6G services, deployment scenarios, and use-cases. Specifically, we consider architectural aspects of 5G/6G systems, their future applications and services, and then proceed illustrating models, characteristics, methods for analyzing the efficiency of radio channels and service performance in these systems.

This book can be utilized by undergraduate and graduate students as well as professionals in the field of performance evaluation and optimization of wireless networking technologies.

ISBN 978-5-209-11089-7 © Молчанов Д.А., Бегишев В.О., Самуйлов К.Е., Кучерявый Е.А., 2022
© Оформление. Российский университет дружбы народов, 2022

ОГЛАВЛЕНИЕ

Предисловие	10
Введение	12
Глава 1. Архитектура, услуги и основные понятия сетей связи 5G/6G	19
1.1. Сотовые сети 5G	19
1.1.1. Общая характеристика сетей 5G	19
1.1.2. Требования к характеристикам обслуживания	20
1.1.3. Процесс стандартизации сетей радиодоступа	21
1.1.4. Базовые услуги сетей 5G.....	24
1.1.5. Спектр частот для сетей радиодоступа.....	25
1.2. Технология радиодоступа 5G NR.....	26
1.2.1. Организация беспроводного интерфейса NR.....	27
1.2.2. Дуплексная передача.....	29
1.2.3. Канальное кодирование.....	30
1.2.4. Планирование передач и каналы управления.....	30
1.2.5. Формирование луча.....	31
1.2.6. Доступ к сети.....	32
1.3. Сети доступа 5G NR в миллиметровом диапазоне частот.....	33
1.3.1. Потери распространения.....	33
1.3.2. Влияние погодных условий.....	34
1.3.3. Динамическая блокировка.....	35
1.4. Программируемость и виртуализация сетевых функций.....	36
1.4.1. Программируемая архитектура 3GPP 5G.....	38
1.4.2. Модульность и принципы использования.....	40
1.5. Сетевой слайсинг.....	42
1.6. Интернет вещей.....	45
1.7. Технологии прямого взаимодействия устройств.....	46
1.8. Гетерогенные сети доступа.....	47
1.9. Терагерцовый интерфейс сетей 6G.....	49
1.9.1. Приложения и сценарии использования.....	51

1.9.2. Физический уровень IEEE 802.15 3D.....	54
1.9.3. Канальный уровень IEEE 802.15 3D.....	57

**ЧАСТЬ I. МОДЕЛИ, ХАРАКТЕРИСТИКИ
И МЕТОДЫ АНАЛИЗА РАДИОКАНАЛА СЕТЕЙ 5G/6G**

ГЛАВА 2. МОДЕЛИ КОМПОНЕНТОВ СЕТЕЙ 5G/6G.....	60
2.1. Модели распространения.....	60
2.1.1. Моделирование лучей.....	61
2.1.2. Эмпирические модели распространения.....	62
2.1.3. Кластерная модель распространения.....	63
2.2. Модели фазированных антенных решеток.....	64
2.2.1. Трехмерная модель.....	64
2.2.2. Упрощенная двухмерная модель.....	66
2.2.3. Связь с параметрами антенных решеток.....	67
2.2.4. Применение моделей антенн.....	69
2.3. Модели блокировки.....	70
2.3.1. Статический случай.....	71
2.3.2. Подвижность пользователя.....	72
2.3.3. Движущиеся блокаторы.....	75
2.4. Модели параметров канала связи.....	77
2.4.1. Блокировки в моделях распространения.....	78
2.4.2. Область покрытия точкой доступа.....	79
2.4.3. Распределение требований к ресурсам.....	80
2.4.4. Интенсивность прерываний.....	82
2.4.5. Упрощенная модель прерываний.....	84
ГЛАВА 3. ВЕРОЯТНОСТНЫЕ МОДЕЛИ ДЛЯ АНАЛИЗА ХАРАКТЕРИСТИК СЕТЕЙ РАДИОДОСТУПА 5G/6G.....	85
3.1. Показатели эффективности сетей радиодоступа.....	85
3.2. Стохастическая геометрия и ее применение к анализу сетей радиодоступа 5G/6G.....	86
3.2.1. Базовая модель взаимодействия устройств.....	87
3.2.2. Применение теоремы Кэмпбелла.....	93
3.2.3. Оценка распределения скорости передачи.....	95

3.3. Ресурсные системы массового обслуживания и их применение к анализу сетей 5G/6G.....	96
3.3.1. Построение моделей ресурсных систем массового обслуживания.....	96
3.3.2. Базовая ресурсная система массового обслуживания.....	98
3.4. Пример построения модели стохастической геометрии прямого взаимодействия устройств.....	101
ГЛАВА 4. ПОМЕХА И ОТНОШЕНИЕ СИГНАЛ–ШУМ В СЕТЯХ 5G/6G С НАПРАВЛЕННЫМИ АНТЕННАМИ И БЛОКИРОВКОЙ	106
4.1. Помехи в сетях с направленными антенными решетками	106
4.1.1. Двухмерный сценарий развертывания.....	106
4.1.2. Трехмерный сценарий развертывания.....	112
4.1.3. Численные примеры.....	118
4.2. Отношение сигнал–шум в сетях с направленными антенными решетками	121
4.2.1. Двухмерный сценарий развертывания.....	122
4.2.2. Трехмерный сценарий развертывания.....	125
4.2.3. Численные примеры.....	127
 ЧАСТЬ II. МОДЕЛИ И МЕТОДЫ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ОБСЛУЖИВАНИЯ АБОНЕНТОВ В СЕТЯХ 5G/6G 	
ГЛАВА 5. МЕХАНИЗМЫ ПОВЫШЕНИЯ КАЧЕСТВА ОБСЛУЖИВАНИЯ АБОНЕНТОВ СЕТЕЙ 5G/6G – МУЛЬТИСВЯЗНОСТЬ И РЕЗЕРВИРОВАНИЕ РЕСУРСОВ.....	132
5.1. Базовая ресурсная модель.....	133
5.1.1. Агрегирование потоков.....	134
5.1.2. Рекуррентный алгоритм.....	136
5.2. Ресурсная модель с сигналами.....	136
5.2.1. Стационарное распределение.....	137
5.2.2. Показатели эффективности	140
5.3. Резервирование ресурсов.....	141
5.3.1. Построение модели.....	141
5.3.2. Система уравнений равновесия.....	142

5.4. Мультиязычность и резервирование ресурсов.....	143
5.4.1. Построение модели мультиязычности.....	144
5.4.2. Модель отдельного узла сети.....	146
5.4.3. Итерационный алгоритм.....	151
5.5. Численный анализ.....	153
5.5.1. Модель резервирования ресурсов.....	153
5.5.2. Модель мультиязычности.....	157
5.5.3. Совместное применение резервирования и мультиязычности.....	159
 ГЛАВА 6. МНОГОАДРЕСНЫЕ СОЕДИНЕНИЯ В СЕТЯХ ДОСТУПА 5G/6G.....	 162
6.1. Проблемы применения многоадресных соединений в сетях 5G/6G.....	162
6.2. Системная модель.....	165
6.3. Мультисервисная модель с одноадресными и многоадресными соединениями.....	168
6.4. Численный анализ.....	171
6.4.1. Оценка точности модели.....	173
6.4.2. Эффекты обслуживания многоадресных и одноадресных сессий.....	174
6.4.3. Оптимальное расстояние между базовыми станциями.....	178
6.4.4. Использование одноадресных сессий для многоадресной передачи.....	183
 ГЛАВА 7. УПРАВЛЕНИЕ ПЕРЕГРУЗКАМИ В СЕТЯХ 5G/6G СРЕДСТВАМИ БПЛА.....	 187
7.1. Причины возникновения перегрузок.....	187
7.2. Стандартизация применения БПЛА.....	189
7.2.1. Интеграция БПЛА в беспроводную сеть.....	189
7.2.2. Архитектурное решение IAB.....	190
7.2.3. Ретрансляция для поддержки IAB.....	192
7.3. Оценка производительности сетей 5G/6G на основе БПЛА.....	192
7.3.1. Системная модель.....	193
7.3.2. Численный анализ.....	194

7.4. Совместная оптимизация участков доступа и транзита.....	198
7.4.1. Системная модель.....	198
7.4.2. Метод совместной оптимизации доступа и транзита.....	201
7.4.3. Численный анализ	201
7.5. Обслуживание на абонентском участке.....	206
7.5.1. Системная модель.....	206
7.5.2. Анализ модели.....	208
7.5.3. Численный анализ.....	214
7.6. Использование многошаговой ретрансляции.....	218
7.6.1. Концепция БПЛА-мостов.....	218
7.6.2. Системная модель.....	219
7.6.3. Анализ модели.....	223
7.6.4. Численный анализ.....	230

ЧАСТЬ III. УСЛУГИ И ФУНКЦИОНАЛЬНЫЕ ВОЗМОЖНОСТИ СЕТЕЙ 5G/6G

ГЛАВА 8. УСЛУГИ СВЕРХНАДЕЖНОЙ ДОСТАВКИ С УЛЬТРАМАЛОЙ ЗАДЕРЖКОЙ.....	238
8.1. Концепция Индустрии 4.0.....	238
8.1.1. Сценарии использования беспроводной связи на производстве.....	240
8.1.2. Сети 5G/6G для индустрии будущего.....	242
8.2. Поддержка услуг URLLC.....	243
8.2.1. Системная модель.....	244
8.2.2. Анализ системы.....	246
8.2.3. Численный анализ.....	249
8.3. Системы с кэшированием.....	254
8.3.1. Прогнозируемый выбор режима с учетом мобильности	257
ГЛАВА 9. МЕТОДЫ УПРАВЛЕНИЯ РЕСУРСАМИ В СЕТЯХ NB-IoT....	262
9.1. Технологии сотовой связи для IoT.....	262
9.2. Методы анализа NB-IoT.....	264
9.2.1. Модель соты LTE.....	266
9.2.2. Стратегии разделения ресурсов	266

9.3. Модель разделения ресурсов.....	268
9.3.1. Характеристики обслуживания.....	270
9.3.2. Численный анализ.....	271
9.3.3. Оценка производительности NB-IoT.....	274
9.4. Использование ретрансляции в NB-IoT.....	276
9.4.1. Системная модель.....	276
9.5. Аналитическая модель.....	280
9.5.1. Базовая модель.....	280
9.5.2. Расширенная система с ретрансляцией.....	287
9.6. Численный анализ.....	290
9.6.1. Базовый сценарий.....	291
9.6.2. Расширенная система с ретрансляцией.....	294
ГЛАВА 10. РАЗДЕЛЕНИЕ РЕСУРСОВ ПРИ ИСПОЛЬЗОВАНИИ ТЕХНОЛОГИИ D2D.....	299
10.1. Возможности D2D в сотовых сетях.....	299
10.2. Архитектурные аспекты построения D2D.....	302
10.3. Оценка возможностей архитектуры D2D.....	310
10.3.1. Методология оценки.....	310
10.3.2. Имитационное моделирование D2D.....	311
10.3.3. Сценарий применения гетерогенных сетей.....	315
10.4. Результаты имитационного моделирования.....	317
10.5. Аспекты практической реализации.....	322
10.5.1. Программная платформа.....	322
10.5.2. Результаты экспериментов.....	329
10.5.3. Перспективы внедрения.....	334
10.6. Дальнейшее развитие D2D.....	335
10.7. Mesh-сети с использованием технологии блокчейн... 336	336
10.7.1. Использование блокчейн в Mesh-сетях.....	337
10.7.2. Пример сценария стимулирования пользователей.....	339
10.7.3. Открытые исследовательские задачи.....	347
ГЛАВА 11. ЗАДАЧА СПРАВЕДЛИВОГО РАЗДЕЛЕНИЯ РЕСУРСОВ ГЕТЕРОГЕННОЙ СЕТИ РАДИОДОСТУПА.....	349
11.1. Архитектура гетерогенной сети доступа.....	349
11.2. Выбор критерия разделения ресурсов.....	353

11.3. Критерий для гетерогенной сети радиодоступа.....	357
11.4. Численный анализ.....	364
11.5. Гетерогенная облачная сеть радиодоступа H-CRAN..	372
11.6. Разделение ресурсов в H-CRAN.....	376
11.7. Численный анализ решения H-CRAN.....	379
11.8. Аспекты практической реализации	382

ЧАСТЬ IV. ТЕРАГЕРЦОВЫЕ СЕТИ ДОСТУПА СИСТЕМ СВЯЗИ 6G

ГЛАВА 12. ХАРАКТЕРИСТИКИ КАНАЛОВ СВЯЗИ	385
12.1. Особенности канала связи в терагерцовом диапазоне.....	386
12.2. Окна прозрачности и характеристики канала связи.....	390
12.3. Особенности устройств связи в терагерцовом диапазоне частот.....	400
ГЛАВА 13. ПРОБЛЕМЫ И РЕШЕНИЯ ДЛЯ СЕТЕЙ СВЯЗИ 6G В ТЕРАГЕРЦОВОМ ДИАПАЗОНЕ ЧАСТОТ	405
13.1. Эффект микромобильности.....	405
13.2. Методы защиты от прослушивания на физическом уровне	421
13.3. Доступ на последнем метре.....	436
ГЛАВА 14. ПРИЛОЖЕНИЯ ТЕРАГЕРЦОВЫХ СИСТЕМ СВЯЗИ.....	447
14.1. Обзор приложений терагерцовых систем.....	447
14.2. Связь между компонентами вычислительных систем.....	453
14.3. Связь внутри вычислительных элементов.....	465
ЗАКЛЮЧЕНИЕ	496
БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....	497
СПИСОК СОКРАЩЕНИЙ.....	510

ПРЕДИСЛОВИЕ

Мысленно держу эту книгу в руках и не верю, что возможно было ее написать. Сети 5G/6G: архитектура, технологии, методы анализа и расчета! Практически все, что надо для научных исследований, обучения, проектной и эксплуатационной деятельности. Конечно же, это могло бы быть и несколько книг, таких же больших, как эта. Однако в том и мастерство авторов, широко и хорошо известных ученых, что им удалось в одной книге изложить практически все, что известно на сегодняшний день по сетям связи пятого и шестого поколений.

На самом деле, читая эту книгу, понимаешь главное – до сетей четвертого поколения были в основном гомогенные сети, что вполне соответствовало уровню развития общества, а сети связи пятого и последующих поколений – это гетерогенные сети эпохи становления и развития цифровой экономики. Вспомним, как все это начиналось. Концепция Интернета вещей привнесла в сети связи совершенно иное представление о числе устройств, которые можно и нужно будет подключить к сети, рекомендации международных организаций только закрепили это в своих стандартах – 1 млн устройств на 1 кв. км. Концепция Тактильного интернета, возникшая от неуместного желания ученых передать по сети связи тактильные ощущения, потребовала от сети связи обеспечить круговую задержку величиной в 1 мс. А это в сто раз меньше, чем в существующих сетях. Здесь и сработал знаменитый закон Г.В.Ф. Гегеля о переходе количества в качество. Да, сети связи пятого и последующих поколений – совершенно новый этап развития сетей связи и общества в целом.

Миллион устройств на квадратный километр приводит нас к гетерогенным сетям, ибо, только интегрируя все возможные ресурсы любых сетевых структур, мы можем обеспечить обслуживание трафика с заданным уровнем качества такого невиданного числа пользователей. Вполне естественным теперь кажется, что Интернет вещей имеет первостепенное значение для ускорения развития многих отраслей нашей жизнедеятельности. Тактильный интернет получил свое развитие в сверхнадежных сетях с ультрамалыми задержками, и здесь научные исследования в области сетей связи столкнулись с проблемами фундаментальных ограничений по скорости распространения света. Отсюда – децентрализация сетей связи и предпосылки для децентрализации экономики. Не погрешив против истины, можно с уверенностью сказать, что сети 5G/6G – основа

построения эффективной цифровой экономики и преодоления цифрового разрыва между регионами Российской Федерации.

Книга состоит из четырех частей, которые делятся на четырнадцать глав. В последней главе мы видим прелюдию к еще более масштабным научным исследованиям по сетям связи шестого поколения. Как в свое время было отмечено в известном философском труде «Материализм и эмпириокритицизм»: электрон так же неисчерпаем, как атом. Все это в полной мере относится и к паре 6G и 5G. Науку в области сетей связи ожидают прекрасные времена новых открытий и познания сути сетей последующих поколений. Дверь в такое будущее немного приоткрывает эта книга.

*Доктор технических наук, профессор,
Почетный член ИТОРЭС им. А.С. Попова
А.Е. Кучерявый*

ВВЕДЕНИЕ

Мировая телекоммуникационная сеть представляет собой самый большой технический объект, который когда-либо создавало человеческое общество для обеспечения своей жизнедеятельности. Эта сеть требует постоянного развития и совершенствования, следуя за потребностями человека в передаче, хранении и обработке огромных, постоянно растущих объемов передаваемой по сети информации. От производительности телекоммуникационной сети напрямую зависит качество обслуживания (QoS, Quality of Service) потоков трафика и качество восприятия (QoE, Quality of Experience) предоставляемых услуг пользователем. Развитие телекоммуникационных технологий и разработка телекоммуникационных систем опережает процессы проектирования, создания и эксплуатации сетей, а именно на этих этапах обнаруживаются проблемы с качеством предоставления услуг пользователям. Поэтому практически параллельно с созданием технологий ведутся исследования, разрабатываются системные модели, оцениваются возможные сетевые сценарии применения технологий, создаются новые, развиваются и модифицируются существующие методы и даже теории, направленные на моделирование и анализ производительности новых сетей и услуг связи.

На рубеже 2020-х годов международные стандартизирующие организации завершили разработку стандартов сетей 5-го поколения (5G, 5th Generation), производители выпустили на рынок оборудование и программное обеспечение, а операторы приступили к коммерческой эксплуатации сетей. За предшествующее десятилетие научное сообщество интенсивно проводило исследования и развивало модели и методы для анализа производительности новых телекоммуникационных систем.

Технологический цикл поколений мобильной связи составляет около 10 лет. Для того чтобы коммерческие сети 5G начали массово запускаться в 2020–2021 годах, примерно с 2011 года по всему миру создавались многочисленные организации и консорциумы для разработки беспроводных технологий 5G, а с 2021 года такой же процесс разворачивается с разработкой мобильных технологий 6-го поколения (6G). К моменту издания этой книги уже определились основные проблемы и особенности построения сетей 6G. Одному из существенных аспектов – радиодоступу в терагерцовом диапазоне частот сетей 6G – посвящена заключительная часть книги.

Книга, которую можно в определенной мере считать и учебником, написана по результатам примерно пятилетних исследований в области беспроводных сетей 5G/6G, проведенных коллективом ученых и специалистов – инженеров-связистов из Технологического университета г. Тампере, Финляндия (ныне Университет г. Тампере, англ. Tampere University, TAU) и математиков-прикладников из Российского университета дружбы народов (РУДН). За это время по тематике исследований участниками коллектива были защищены четыре докторские и более десяти диссертаций на соискание ученой степени кандидата наук и ученой степени Doctor of Technology (PhD). Опубликовано более ста статей (в список литературы мы включили ровно сто из них) в научных журналах первой, высшей четверти (Q1) по импакт-фактору JCR Science Edition, реализовано более двадцати научно-исследовательских проектов в области сетей 4G/5G/6G. Книга покрывает лекционную часть нескольких спецкурсов по беспроводным технологиям и сетям 5G, читаемых авторами в РУДН, ВШЭ (Высшая школа экономики) и в TAU (Tampere University). Издание не содержит методических материалов и руководств по практическим занятиям и лабораторным работам, и только поэтому авторы не позиционируют его как полноценный учебник.

Показанный в работе процесс исследований показателей эффективности беспроводных сетей является заведомо итеративным и представляется следующим образом.

Во-первых, для анализа беспроводной технологии предлагается исследовать конкретный, но одновременно и достаточно общий сценарий ее применения. Сценарий, с одной стороны, должен отражать будущие потребности пользователей в сетевых ресурсах, иначе он не будет актуальным, а с другой стороны, должен максимально включать метрики сетей 5G, изучение которых до сих пор не проводилось или почти не проводилось, т. е. решаемые в исследованиях задачи должны быть новыми, или, как говорят, ультрановыми, если ранее их постановки вообще отсутствовали.

Определившись со сценарием, во-вторых, строится системная модель, которая включает соответствующие сценарию компоненты сетей 5G, такие, как модель распространения сигналов, модель фазированных антенных решеток, модель блокировки прямой видимости, модель параметров канала связи, модель доступа, модель обслуживания абонентов и, может быть, другие модели. Если возможно, в основу кладутся готовые модели, но, что важно в опережающих исследова-

дованиях, разрабатывается системная модель, зачастую представляющая самостоятельную ценность. Для анализа показателей эффективности системной модели применяются, в зависимости от необходимости, три основных метода моделирования: аналитическое или статистическое (метод Монте-Карло) моделирование; имитационное моделирование и, наконец, измерения на лабораторном оборудовании. В книге мы не рассматриваем натуральный эксперимент в реальных условиях, редко обращаемся к измерениям и в основном рассматриваем аналитические, статистические и имитационные модели.

Наконец, в-третьих, наступает этап построения моделей и разработки методов их анализа и расчета интересующих характеристик в соответствии с исследуемым сценарием применения беспроводной технологии и разработанной системной моделью. Из-за случайного характера потоков запросов пользователей, случайных длительностей пользовательских сессий и случайных объемов занимаемых сетевых ресурсов, в большинстве случаев моделирование основывается на аппарате различных разделов теории вероятностей, в первую очередь на теории массового обслуживания, теории марковских случайных процессов, точечных случайных процессов и методов стохастической геометрии.

Методы анализа моделей, в свою очередь, используют другие математические дисциплины, в том числе матричные методы линейной алгебры. Исследования показали, что в части теории массового обслуживания востребованными оказались модели, получившие название ресурсных систем массового обслуживания. Эти системы позволяют наиболее адекватно описать процессы предоставления и разделения радиоресурсов, например, таких, как в сетях стандарта 5G New Radio (5G NR). Прогнозы развития сетевых технологий свидетельствуют о том, что модели будут востребованы и в дальнейшем, и для этого специалистами по математической теории телетрафика уже созданы основы теории ресурсных систем массового обслуживания. К этим моделям добавляются методы точечных случайных процессов и стохастической геометрии, которые позволяют учитывать расположение пользователей и устройств в пространстве, случайные расстояния между ними, конфигурацию антенн и помещений, в случае высокочастотных технологий конфигурацию блокирующих передачу препятствий (так называемых блокираторов) и другие параметры моделирования. Эти параметры, как показано в этой книге, так же как и в других источниках, могут быть описаны

геометрически, а с учетом случайного расположения объектов – в терминах стохастической геометрии. Таким образом, и возникает новая дисциплина, которую мы назвали теорией стохастического анализа беспроводных сетей и для которой в книге показываем различные аспекты ее применения к сетям 5G/6G.

Таким образом, книга представляет собой, с одной стороны, доступное руководство по основным технологиям сетей 5G/6G, а с другой стороны, последовательное изложение методов для анализа и расчета показателей эффективности их характеристик в рамках конкретных сценариев и системных моделей. При этом показаны основные приемы и принципы применения теории стохастического моделирования беспроводных сетей.

Книга состоит из четырнадцати глав, причем первая глава, посвященная общим вопросам архитектуры, основных услуг и определения беспроводных сетей 5G/6G, является вводной, а остальные главы разбиты на четыре части.

В части I исследованы модели, характеристики и методы анализа радиоканалов сетей 5G/6G. Здесь, в главе 2, показаны модели основных компонентов этих сетей, включая модели распространения радиосигналов, модели фазированных антенных решеток, модели блокировки пути прямой видимости и модели параметров канала связи. В главе 3 кратко изложены теоретические основы стохастического анализа показателей эффективности сетей доступа 5G/6G, включая основные характеристики сетей радиодоступа, используемые в книге методы стохастической геометрии, методы теории массового обслуживания и модели прямого взаимодействия устройств. В главе 4 дано определение помехи и отношения *сигнал–шум* в системах с направленными антеннами и блокировкой и рассмотрены двух- и трехмерные сценарии развертывания.

Часть II посвящена построению моделей и изложению методов повышения эффективности обслуживания абонентов в сетях 5G/6G. В главе 5 рассматриваются методы повышения уровня характеристик обслуживания абонентов в сетях путем использования механизмов мультисвязности и резервирования ресурсов. Предложены базовая ресурсная модель массового обслуживания и ресурсная модель с сигналами. Методы обслуживания многоадресных соединений в сетях доступа 5G/6G изложены в главе 6, где построена модель системы, предложен метод ее анализа и приведен численный эксперимент. Глава 7 посвящена методам борьбы с перегрузками в сетях

доступа 5G/6G с использованием беспилотных летательных аппаратов (БПЛА). В ней обсуждаются причины возникновения перегрузок, текущее состояние стандартизации. Показан подход к оценке параметров производительности сетей доступа 5G NR на основе БПЛА, решается задача совместной оптимизации участков доступа и транзита, анализируется организация обслуживания на абонентском участке и метод использования многошаговой ретрансляции.

В части III исследованы услуги и функциональные возможности сетей 5G/6G. В главе 8 изложены методы поддержки услуги ультранадежной доставки с низкой задержкой, в необходимом для контекста книги объеме определена концепция Индустрии 4.0. Описаны методы поддержки услуги URLLC (5G ultra reliable low latency communications) в 5G NR и охарактеризованы системы с кэшированием.

Глава 9 включает методы распределения ресурсов в сетях узкополосного Интернета вещей NB-IoT (Narrowband Internet of Things), в том числе сотовые технологии для Интернета вещей, методы анализа технологии NB-IoT, модель оценки показателей производительности, метод использования методов ретрансляции в NB-IoT, построение и анализ аналитической модели.

В главе 10 предложены методы распределения ресурсов при использовании технологии прямого взаимодействия устройств D2D (device-to-device communications), показаны ее возможности в сотовых сетях, дана оценка предложенной архитектуры, отражены результаты имитационного моделирования, обсуждены вопросы практической реализации и дальнейшего развития D2D, а также Mesh-сети с использованием технологии блокчейн.

Глава 11 содержит методы распределения ресурсов в гетерогенных сетях доступа. Здесь представлена архитектура гетерогенной сети доступа, показан выбор критерия для справедливого распределения ресурсов, даны определение объединенного критерия и некоторые численные результаты. Исследована гетерогенная облачная сеть радиодоступа H-CRAN (Heterogeneous Cloud Radio Access Networks), метод распределения ее ресурсов, рассматриваются численные результаты анализа H-CRAN и некоторые аспекты практической реализации.

Часть IV книги состоит из трех глав и посвящена терагерцовым сетям доступа систем связи 6G. На момент написания книги аспекты построения терагерцовых сетей являлись наименее исследованной

областью. В главе 12 изложены особенности каналов связи в терагерцовом диапазоне, определены окна прозрачности и характеристики канала связи, обсуждены некоторые инженерные компромиссные решения. Глава 13 посвящена проблемам и решениям сетей 6G в терагерцовом диапазоне частот, в том числе эффектам микромобильности, проблемам безопасности физического уровня и проблемам доступа на последнем метре. В главе 14 содержится обзор специальных приложений терагерцовых сетей, обсуждаются вопросы связи внутри вычислительных элементов.

Авторы и их научная команда непрерывно ведут исследования и продолжают публикацию полученных результатов. Уже при написании книги было ясно, что на момент ее выхода в печать появится новый актуальный материал для продолжения книги, интересный для всех категорий читателей от магистра и аспиранта до научного сотрудника и профессора. Понимая это, мы планировали, что при ее появлении будут узнаваемы не только его авторы, но и большой уникальный коллектив ученых и специалистов, которым мы выражаем большую благодарность за сотрудничество и после этого переходим к изложению глав книги. Перечисляем с глубокой признательностью и с комплиментами наших коллег и в большинстве случаев соавторов по научным статьям в алфавитном порядке: Андреев С.Д., Бесчастный В.А., Вихрова О.Г., Вишневский В.М., Гапеенко М.А., Гайдамака Ю.В., Галинина О.С., Киричек Р.В., Комаров М.М., Кочеткова (Гудкова) И.А., Кучерявый А.Е., Ковальчуков Р.Н., Маркова Е.В., Мокров Е.В., Мухтанна А., Наумов В.А., Ометов А.Я., Орлов Ю.Н., Острикова Д.Ю., Парамонов А.И., Петров В.И., Пяттаев А.В., Самуйлов А.К., Соломитский Д.Н., Сопин Э.С., Тюрликов А.М., Чухно Н.В., Чухно О.В., Хакимов А.А., Яркина Н.В.

Мы благодарим также наших коллег из зарубежных университетов, научных центров и промышленных компаний, в кооперации с которыми был получен и опубликован ряд важных результатов, нашедших отражение в этой книге. Имена и фамилии зарубежных коллег мы пишем в англоязычной транскрипции, как они их указывают в списке авторов научных статей, опубликованных в соавторстве с нами. Аффiliation наших зарубежных коллег, география мест работы которых, наверное, не покрывает лишь южный и северный полюса планеты – и для краткости здесь не приводится, можно найти в наших с ними совместных статьях. Они – наши друзья и коллеги Aaltonen T., Abuarqoub A., Akdeniz M.R., Algarni A.D., Antonakoglou K., Araniti G., Aryafar E., Ateya A.A., Balasubramaniam S.,

Barros M.T., Burget R., Channa A., Correia L.M., Costa M., Curado M., Destino G., Dohler M., Eckhardt J.M., Fodor G., Frolka J., Gerla M., Gupta R., Göktepe B., Haneda K., Hellge C., Himayat N., Hosek J., Hossain E., Huang K., Hussain O., Iera A., Johnsson K., Jornet J., Juntti M., Kannisto J., Karls I., Karttunen A., Kassler A., Kavuri S., Koivisto M., Kokkonieniemi J., Kröpfl F., Kurner T., Lehtomäki J., Lema M.A., Leppänen K., Lohan E.S., Lu X., Ma K., Maattanen H.-L., Mahmoodi T., Mao W., Masek P., Maule M., Melman A., Mikkonen T., Militano L., Mueck M., Muhammad K., Mäkitalo N., Naderpour R., Nguyen S.L.H., Nikopour H., Olsson T., Orhan O., Orsino A., Pierobon M., Pizzi S., Pokorny J., Popescu N., Pouttu A., Qian H., Raatikainen M., Rakheja S., Rangan S., Rinaldi F., Roubicek M., Räisänen A., Saafi S., Sardis F., Schierl T., Singh S., Skibinska J., Slezak C., Stachowiak M., Stusek M., Szyszkowicz S.S., Tabassum H., Talwar S., Tirronen T., Torsner J., Töllli A., Valkama M., Villa J., Wang H., Wiedermann W., Wirdatmadja S., Yang M., Yanikomeroğlu H., Yeh S.-P., Yilmaz O.N.C., Zeman O., Zhu H., Zhu J.

Мы перечисляем наших соавторов не только для того, чтобы выразить им нашу признательность, но и для того, чтобы показать читателю, что практически каждая статья из списка литературы является результатом отдельного исследования либо важным этапом этих исследований. Таким образом, мы подчеркиваем, что исследования такого уровня высокоэффективны в рамках международной стабильной коллаборации, объединяющей как известных ученых с мировым именем, так и талантливую, желающую понять суть современных проблем беспроводной связи молодежь.

Наконец, мы признательны нашим семьям за терпение и поддержку в наших исследованиях и нашей очень сложной, но крайне интересной и богатой событиями научной жизни, результаты которой частично описаны нами в этой книге.

Д.А. Молчанов, В.О. Бегушев, К.Е. Самуйлов, Е.А. Кучерявый

ГЛАВА 1

АРХИТЕКТУРА, УСЛУГИ И ОСНОВНЫЕ ПОНЯТИЯ СЕТЕЙ СВЯЗИ 5G/6G

Цель главы – познакомить читателя с общими понятиями беспроводных сетей пятого и шестого поколений. В параграфе 1.1 приводится общая характеристика сетей 5G, определяется набор услуг, которые эти сети должны поддерживать, а также доступный для использования спектр частот. В параграфе 1.2 рассматриваются основы сетей 5G – технология радиодоступа «Новое Радио» (NR, New Radio). Специфика использования 5G в миллиметровом диапазоне частот обсуждается в 1.3. Специальные функциональные возможности архитектуры сетей 5G – сегментация (далее также используется термин «слайсинг», от англ. slicing), программируемость и виртуализация сетевых функций – раскрываются в параграфах 1.4 и 1.5. Интернет вещей и беспроводные технологии для его поддержки представлены в 1.6. Параграф 1.7 посвящен технологии прямого взаимодействия устройств, которая, как ожидается, будет играть существенную роль в сетях 5G/6G. В 1.8 рассматриваются гетерогенные сети радиодоступа. Наконец, в 1.9 для сетей 6G приводятся варианты использования терагерцового радиодоступа.

1.1. Сотовые сети 5G

1.1.1. Общая характеристика сетей 5G

Задачей беспроводных сетей пятого поколения (5G, 5th Generation) является удовлетворение возрастающих потребностей в мобильной связи государств, предприятий, а также отдельных граждан. Предполагается, что технологии 5G будут играть ключевую роль в превращении городов в «умные» города, что позволит гражданам и обществу в целом получить социально-экономические преимущества, которые дает цифровая экономика.

Технология 5G направлена на повышение качества обслуживания конечных пользователей, предлагая новые приложения и услуги с гигабитной скоростью передачи данных. Сети 5G будут опираться на достижения мобильных сетей предыдущих поколений, которые изменили общество, предоставив новые услуги и новые бизнес-модели. Для операторов беспроводной связи появляется возможность

не только предложить новые услуги, но и разрабатывать собственные новые решения и услуги для потребителей и промышленности в различных секторах экономики, и все это по доступной цене. Сети 5G – это реализация проводных и беспроводных конвергентных сетей, которая предлагает в том числе возможности интеграции систем управления сетями.

Коммерческие сети 5G начали эксплуатироваться после 2020 года, когда завершился процесс их стандартизации на международном уровне. Ассоциация GSM (Global System for Mobile Communications) прогнозирует, что число соединений 5G достигнет 1,1 миллиарда, что составит около 12 % от общего количества мобильных подключений, а общие доходы операторов вырастут в среднем на 2,5 %, достигнув 1,3 триллиона долларов США к 2025 году.

Прогнозируется, что технология 5G значительно увеличит скорость передачи данных и уменьшит задержку (менее 1 мс на беспроводном участке доступа) по сравнению с сетями предыдущих поколений. Такое требование необходимо для критически важных служб, в которых данные чувствительны ко времени передачи. Сети 5G могут предоставлять широкий спектр высокоскоростных широкополосных услуг и предлагать альтернативу сетям радиодоступа последней мили.

1.1.2. Требования к характеристикам обслуживания

В соответствии с рекомендацией МСЭ-Р М.2083 (рис. 1.1) пиковая скорость передачи данных в сетях радиодоступа 5G должна достигать 10 Гбит/с. Однако при определенных условиях требуется поддержка пиковой скорости до 20 Гбит/с. Для случаев эксплуатации сетей на больших площадях, например в городских и пригородных районах, скорость передачи данных для пользователя должна составлять до 100 Мбит/с. В зонах с экстремальными требованиями ожидаемая пользователем скорость передачи данных может достигать более высоких значений, например, до 1 Гбит/с внутри помещений.

Ожидается, что эффективность использования спектра частот будет в три раза выше по сравнению с сетями 4G. Предполагается, что сети 5G будут предоставлять пространственную емкость, достигающую 10 Мбит/с/м² в зонах с высокими требованиями, в частности на городских площадях, стадионах, в торговых центрах и т.д. Потребление энергии сетями стандарта подвижной связи 2020

(IMT-2020) не должно быть выше, чем у сетей четвертого поколения 4G. Следовательно, на этапе разработки эффективность энергопотребления как базовых станций (БС), так и абонентских устройств (АУ) должна быть повышена, по крайней мере, в такой же степени, как и в 4G.

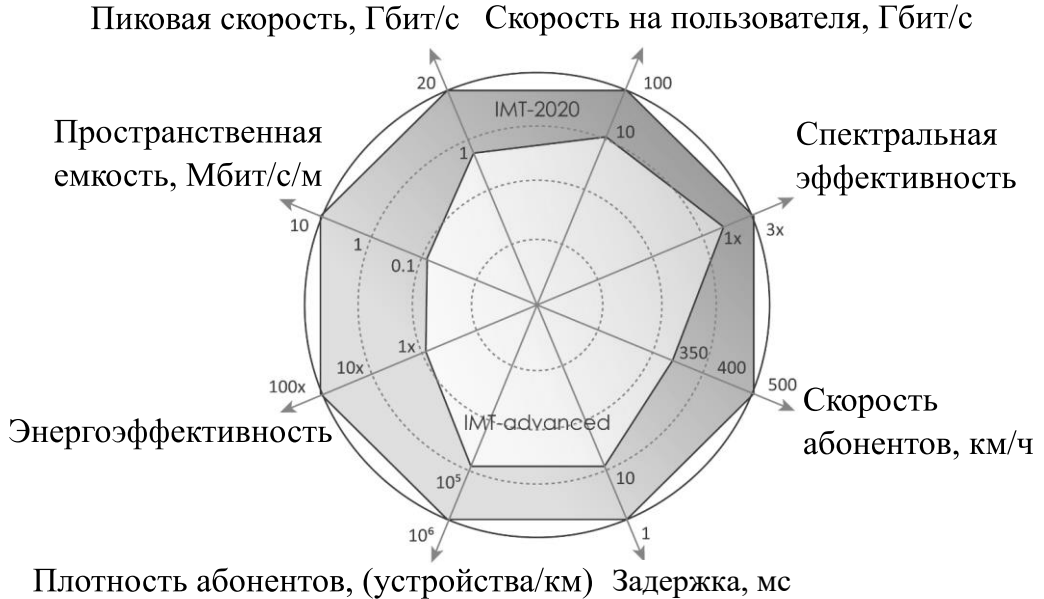


Рис. 1.1. Требования рекомендации МСЭ-Р М.2083 к сетям радиодоступа 5G

Необходимо также обеспечить поддержку абонентов с высокой мобильностью, имеющую скорость движения до 500 км/ч, при этом требуемые параметры качества обслуживания не должны ухудшаться. Такие услуги предусмотрены, в частности, для скоростных поездов. Наконец, сети радиодоступа 5G должны будут поддерживать плотность абонентов до 10⁶/км², например в сверхплотных сценариях межмашинного взаимодействия.

1.1.3. Процесс стандартизации сетей радиодоступа

Телекоммуникационное сообщество многое сделало для улучшения качества нашей жизни. Со времен разработки телеграфа средства связи непрерывно улучшали свои возможности и открывали новые горизонты для человека. Эту пользу оценили во всем мире и

практически в каждой стране мира существует своя телекоммуникационная инфраструктура, операторы связи и производители оборудования. В связи с этим возникла необходимость разрабатывать единые правила для систем связи, так как отдельные телекоммуникационные сети так бы и могли остаться разрозненными или же их объединение превратилось бы в очень сложный и дорогостоящий процесс. Поэтому в середине XX века начали появляться стандартизирующие организации, занимающиеся разработкой единых норм и правил организации и взаимодействия средств связи.

Разработкой стандартов сетей радиодоступа 5G занимается нескольких международных организаций. Среди них есть как официальные организации, занимающиеся подготовкой спецификаций системы в целом, так и специализированные отраслевые альянсы с узкой специализацией в той или иной области.

В настоящее время организации в международной системе стандартизации можно распределить по следующим группам.

Официальные международные организации стандартизации:

– *Международная организация по стандартизации (International Organization for Standardization, ISO)*. Создана в 1946 г., включает в себя национальные организации стандартизации из 157 стран мира, в том числе ANSI (США), Федеральное агентство по техническому регулированию и метрологии (Россия), BSI (Великобритания), AFNOR (Франция) и др., обладает полномочиями для координирования на международном уровне разработки различных промышленных стандартов и принятия их в качестве международных стандартов.

– *Международный союз электросвязи, МСЭ (International Telecommunication Union, ITU)*. Занимается стандартизацией международных средств связи и состоит из трех секторов:

а) стандартизации телекоммуникаций (ITU-T) – занимается вопросами, связанными с телефонными системами и системами передачи данных;

б) радиосвязи (ITU-R) – распределяет радиочастоты между конкурирующими компаниями, решает спорные вопросы в данной области;

в) развития (ITU-D) – занимается вопросами стратегии и политики развития систем электросвязи.

Региональные организации стандартизации:

– *Европейский институт стандартизации в области телекоммуникаций (ETSI, European Telecommunications Standards*

Institute). Создан в 1988 г., отвечает за стандартизацию информационных и телекоммуникационных технологий в пределах Европы;

– *Центр сетевых информационных технологий Азиатско-Тихоокеанского региона (APNIC, Asia Pacific Network Information Centre)*, отвечает за распределение сетевых ресурсов в Азиатско-Тихоокеанском регионе.

Национальные организации стандартизации:

– *Федеральное агентство по техническому регулированию и метрологии (Россия)*;

– *Американский институт национальных стандартов (ANSI, American National Standards Institute)*.

Некоммерческие промышленные консорциумы:

– *Сообщество инженеров по электротехнике и электронике (IEEE, Institute of Electrical and Electronic Engineers)*. Целью организации является продвижение теоретических и прикладных достижений электротехнической и электронной индустрии;

– *Рабочая группа по проектированию интернет-технологий (IETF, Internet Engineering Task Force)*. Является сообществом разработчиков, операторов, изготовителей и исследователей в области сетевых технологий.

Стандартизация сетей радиодоступа 5G/6G проводится консорциумом 3GPP (3rd Generation Partnership Project), который отвечает за разработку и поддержку стандартов текущего и предшествующих поколений технологий мобильной связи. Работа по стандартизации в 3GPP организована в группах по техническим спецификациям, а именно: сети радиодоступа, обслуживание, базовая сеть и абонентское устройство (АУ). 3GPP использует поэтапный подход для определения стандартов. Выделяют три стадии стандартизации, ориентированные на определение услуг, архитектуру и подробные спецификации интерфейсов.

На сегодняшний день стандартизирующие организации отвечают не только за выпуск стандартов и регламентов. В их фокус внимания входят задачи поддержания открытого диалога между всеми заинтересованными участниками: производителями оборудования, государственными регулирующими органами, операторами сетей, поставщиками услуг и др. При этом необходимо учесть все постановления и требования законодательства в каждом государстве, а также проводить регулярные встречи, решать спорные вопросы. Задачей некоторых организаций является популяризация какой-либо технологии, т.е. целью является донесение до потребителя преимуществ новых разработок и выгоды от их использования.

1.1.4. Базовые услуги сетей 5G

Заинтересованные стороны отрасли «Связь» совместно с организациями, занимающимися стандартизацией сетей радиодоступа 5G, определили несколько потенциальных вариантов их использования. На основе проведенного анализа предложена поддержка трех принципиально разных категорий услуг:

- расширенная услуга широкополосного доступа (eMBB, Enhanced Mobile Broadband);
- услуга сверхнадежной связи с ультрамалой задержкой (URLLC, Ultra-Reliable Low-Latency Communications);
- услуга сверхплотного межмашинного взаимодействия (mMTC, Massive Machine-Type Communications).

Услуга eMBB направлена на удовлетворение потребностей пользователей в сверхвысокой скорости на абонентском участке доступа. К приложениям, требующим таких скоростей, можно отнести просмотр видео высокого разрешения, очки и шлемы дополненной и виртуальной реальности, скачивание большого объема данных. Услуга перенесет высокоскоростную мобильную широкополосную связь в зоны с высокой плотностью населения, позволит потребителям использовать высокоскоростную потоковую передачу для стационарных и мобильных устройств и подготовит вектор развития корпоративным услугам связи для организации совместной работы. Многие операторы связи также рассматривают eMBB в качестве решения последней мили в тех районах, где отсутствуют проводные решения.

Критически важные показатели, такие как задержка и обеспечение безопасности, в сетях 5G будут играть важную роль в развитии интеллектуальных транспортных систем будущего, позволив техническим средствам передвижения связываться друг с другом, создавая новые возможности для внедрения беспилотных легковых и грузовых автомобилей. Например, беспилотное транспортное средство, управляемое через облачную автономную систему вождения, должно быть способно останавливаться, ускоряться или поворачивать согласно получаемым инструкциям. Любая задержка передачи информации в сети или потеря сигнала от базовой станции может привести к катастрофическим последствиям. Ультрамалая задержка предоставляет широкие возможности для удаленной хирургии, автоматизации производства и контроля процессов в реальном времени.

Ожидается, что сети радиодоступа 5G также будут способствовать развитию «умных» городов и Интернета вещей (IoT, Internet of

Things) путем развертывания сенсорных сетей в городах и сельской местности. Безопасность и надежность, встроенные в сети радиодоступа 5G, сделают их пригодными для обеспечения общественной безопасности, а также для использования в критически важных службах: полиции, службах безопасности, энергетике, водоснабжении и здравоохранении.

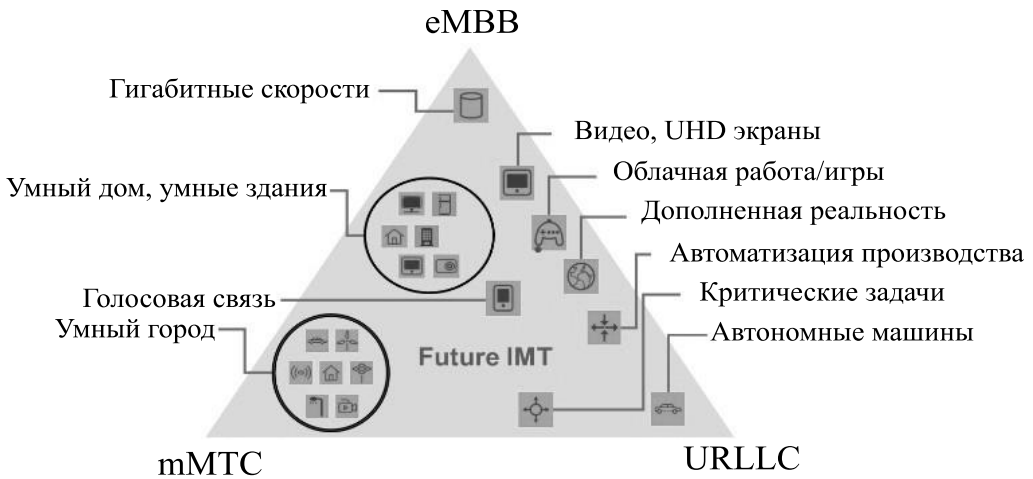


Рис. 1.2. Диаграмма услуг и приложений сетей радиодоступа 5G

Спектр услуг, предоставляемый сетями 5G, не ограничивается eMBB, URLLC и mMTC. Скорее они определяют типичные приложения с конкретными требованиями к характеристикам обслуживания. Такие виды услуг, которые будут поддерживаться в сетях радиодоступа 5G, удобно представить в виде треугольной диаграммы, как показано на рис. 1.2. Грани треугольника определяют фиксированные требования к скорости доступа, надежности передачи и задержке. Как внутри треугольника, так и на его гранях располагаются приложения, обладающие теми или иными наборами требований.

1.1.5. Спектр частот для сетей радиодоступа

Для удовлетворения требований по емкости сетей радиодоступа 5G (см. рис. 1.1) потребовался более широкий спектр частот по сравнению с сетями 4G. Были предприняты согласованные усилия по гармонизации спектра 5G. Основная часть задачи определения используемого спектра для сетей радиодоступа 5G решается МСЭ-Р. Принимая во внимание, что реализация узконаправленных антенн становится более доступной на более высоких частотах, МСЭ-Р

исследует техническую реализацию будущего спектра 5G на частотах от 24 и до 72 ГГц. Полосы ниже и выше 6 ГГц могут использоваться в качестве дополнительного спектра.

Варианты использования 5G потенциально могут быть удовлетворены различными частотами спектра. Например, приложения с малой задержкой и малой дальностью (подходящие для густонаселенных городских районов), вероятно, будут использовать миллиметровый диапазон частот (выше 24 ГГц). Приложения большой дальности с низкой пропускной способностью (более подходящие для сельской местности) будут использовать сети доступа, работающие на частотах ниже 1 ГГц. В то время как более низкие частоты имеют лучшие характеристики распространения сигнала и, как следствие, характеризуются лучшим покрытием, более высокие частоты предоставляют большую емкость.

Многослойный подход является одним из возможных вариантов развертывания сетей радиодоступа 5G. В этом подходе различают три зоны: зона покрытия, зона покрытия и емкости и зона емкости сети. Первая зона использует спектр ниже 2 ГГц (например, 700 МГц), обеспечивая большой охват покрытия и проникновения сигнала связи внутрь помещений. Вторая зона реализуется в диапазоне частот 2–6 ГГц (например, 3,5 ГГц) для обеспечения наилучшего баланса между пропускной способностью и покрытием сети. Наконец, третья зона использует спектр выше 6 ГГц (например, 28 ГГц) для обеспечения чрезвычайно высоких скоростей передачи данных.

1.2. Технология радиодоступа 5G NR

В параграфе представлен краткий обзор технических компонентов и возможностей радиоинтерфейса «Новое радио» (англ. New Radio – NR), который разрабатывается консорциумом 3GPP. Предполагается, что именно NR обеспечит новые варианты использования, требующие дальнейшего повышения скорости передачи данных, уменьшения задержки, увеличения сетевого покрытия и повышения надежности обслуживания. Ключевые компоненты для достижения этих целей включают гибкую структуру кадра физического уровня, структуру кадров с оптимизированной задержкой, использование антенных решеток, а также гибкое взаимодействие между сетями доступа, работающими в высокочастотном и низкочастотном спектре.

1.2.1. Организация беспроводного интерфейса NR

Аналогично LTE (Long Term Evolution) сеть радиодоступа NR основана на методе ортогонального частотного разделения каналов с мультиплексированием (OFDM, Orthogonal Frequency-Division Multiplexing) с возможностью предварительного кодирования дискретным преобразованием Фурье (DFT, Discrete Fourier Transform) Технология NR поддерживает гибкую структуру беспроводного интерфейса с разнесением поднесущих в диапазоне от 15 до 240 кГц с пропорциональным изменением длительности циклического префикса. Малое разнесение поднесущих позволяет использовать относительно длинный циклический префикс при разумных накладных расходах.

Хотя спецификация физического уровня не зависит от полосы частот, не все поддерживаемые структуры беспроводного интерфейса актуальны для всех полос частот NR. Для полосы частот 0,45–6 ГГц доступное разделение поднесущих составляет 15, 30 и 60 МГц, тогда как максимальный размер канала, соответствующий этим поднесущим, равен 50, 100 и 200 МГц. Минимальный и максимальный размеры каналов для полосы частот 24–72 ГГц составляет 100, 200 и 400 МГц, что позволяет использовать разнесение несущих в 60 кГц и 120 кГц для реализации 3300 поднесущих. В тех случаях, когда необходимо поддерживать еще большую полосу пропускания, можно использовать функцию агрегации несущих. Для NR еще не определен спектр частот между 6 ГГц и 24 ГГц, но требования могут быть представлены на более поздней стадии разработки стандарта, если такой спектр станет доступным.

Не все устройства NR должны поддерживать полную полосу пропускания, что значительно влияет на организацию, например, каналов управления. Кроме того, NR позволяет адаптировать ширину полосы канала на стороне АУ для снижения энергопотребления устройства. Следовательно, NR определяет полосы частот (BR, bandwidth part) для АУ. Если АУ способно к одновременному приему нескольких BR, то существует возможность на одной несущей смешивать передачи разных структур кадра. Отметим, что в версии 15 стандарта 3GPP (Release 15) определяется поддержка только одной BR на приемнике.

Структура кадра NR показана на рис. 1.3. Радиокадр длиной 10 мс делится на десять подкадров, длительность каждого из которых 1 мс. Подкадр, в свою очередь, делится на интервалы, состоящие

из 14 символов OFDM. Таким образом, длительность интервала в миллисекундах зависит от выбранной структуры кадра. Для разнесения поднесущих 15 кГц интервал NR имеет ту же структуру, что и подкадр LTE. Поскольку временной интервал определяется как фиксированное количество символов OFDM, большие интервалы разделения поднесущих приводят к более короткому интервалу. Такое свойство могло бы быть использовано для реализации передач с уменьшенной задержкой. Однако, поскольку циклический префикс также уменьшается при увеличении разнесения поднесущих, это является невозможным. В результате NR использует более эффективный подход к уменьшению задержки передачи, вводя дополнительную гранулярность, называемую «мини-интервалом» передачи. Такие передачи могут прерывать уже запланированные соединения с другими устройствами, если возникнет срочная необходимость отправить данные, требующие очень низкой задержки. Данная функциональная возможность используется для поддержки услуги URLLC.

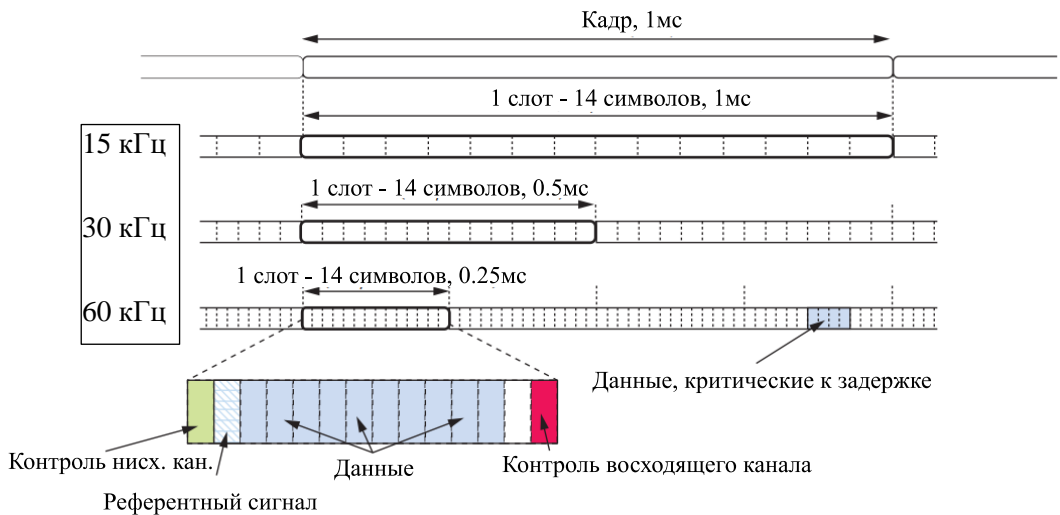


Рис. 1.3. Структура кадра технологии радиодоступа NR

Возможность инициировать передачу данных полезна не только на границах интервалов, но и при работе в нелицензированном спектре. В этих условиях, прежде чем начать отправку данных, необходимо удостовериться, что радиоканал доступен. Это реализуется с помощью процедуры сканирования эфира перед передачей (LBT, Listen Before Talk). Как только канал окажется доступным, необходимо немедленно начать процедуру отправки данных, а не ждать

начала интервала, чтобы избежать ситуации, когда другой передатчик инициирует соединение.

Работа NR в области миллиметрового диапазона частот является еще одним примером эффективного использования мини-слотов, поскольку доступная ширина полосы очень велика, и даже нескольких символов OFDM может быть достаточно для завершения передачи. Это особенно полезно в сочетании с процедурой аналогового формирования луча, когда передачи на множество устройств могут быть мультиплексированы только во времени.

Поддержка ультранизкой задержки на беспроводном интерфейсе является важной частью NR, которая влияет на многие решения. Одним из примеров являются передачи с фронтальной загрузкой. Имея опорные сигналы и управляющую сигнализацию нисходящего канала, в котором содержится информация о планировании и начале передачи, АУ может немедленно начать обработку принятых данных без предварительной буферизации. Это позволяет минимизировать задержку декодирования.

1.2.2. Дуплексная передача

Дуплексная схема связи обычно определяется доступным для оператора спектром частот. Для низкочастотных диапазонов используется частотное разделение исходящего и восходящего каналов (FDD, Frequency-Division Duplexing). На высоких частотах (например, в миллиметровом диапазоне) рекомендуется использовать непарные значения спектра, что требует временного разделения исходящего и восходящего каналов (TDD, Time Division Duplex).

NR, подобно LTE, будет поддерживать оба метода дуплексной передачи. Однако NR поддерживает динамическое выделение временных интервалов, в отличие от LTE, где назначение временных интервалов в схеме TDD не изменяется со временем. В динамическом TDD временной интервал (или его часть) может быть выделена либо для восходящего канала связи, либо для нисходящего в соответствии с решением планировщика. Это позволяет эффективно обрабатывать быстрые изменения трафика, которые особенно заметны при плотном развертывании с относительно небольшим числом пользователей на БС. Предусматривается также возможность статического конфигурирования направления передачи некоторых из слотов. Эта возможность может позволить снизить потребление энергии АУ,

поскольку нет необходимости отслеживать управление нисходящего канала в слотах, которые заранее зарезервированы для передачи.

1.2.3. Канальное кодирование

Канальное кодирование для передачи данных NR основано на кодах проверки четности с малой плотностью (LDPC, Low Density Parity Check). Коды LDPC привлекательны с точки зрения реализации, особенно при скоростях передачи данных несколько гигабит в секунду. Также применяется система гибридного автоматического переспроса (HARQ, Hybrid Automatic Repeat Request) с использованием наращивания избыточности, где БС повторно передает ошибочно принятые данные и устройство объединяет полезную информацию из нескольких попыток передачи.

Для каналов управления физического уровня HARQ не используется. Для таких передач были выбраны полярные коды. Дополнительно для контрольных каналов могут использоваться коды Рида–Мюллера.

1.2.4. Планирование передач и каналы управления

Процесс планирования передач в NR очень схож с тем, который реализован в LTE. Каждое АУ контролирует несколько физических каналов управления нисходящего канала связи (PDCCH, Physical Downlink Control Channel), как правило, один раз в слот, хотя можно настроить и более частый мониторинг для поддержки трафика, требующего очень низкой задержки. После обнаружения PDCCH АУ подчиняется решению планировщика и принимает (или передает) данные, известные как транспортный блок. PDCCH передаются в одном или нескольких наборах ресурсов управления, каждый из которых имеет длину от одного до трех символов OFDM.

После приема данных по нисходящей линии связи устройство посылает подтверждение передачи, используя протокол HARQ, указывающий на успешность выполнения процедуры. Учитывая крайне высокие скорости передачи данных, поддерживаемые NR, размер транспортного блока может быть очень большим. Повторная передача всего транспортного блока в этом случае может стать неэффективной. Поэтому NR поддерживает повторные передачи с более высокой степенью детализации, известной как группа кодовых блоков (CBG, Code Block Group). Критически важная передача может

использовать только один или несколько символов OFDM и, следовательно, создавать сильные помехи другим устройствам. В этом случае достаточно повторной передачи только той части данных, которые были получены в пределах CBG, и нет необходимости пересылать весь блок данных.

Подтверждение передачи по протоколу HARQ, а также информация о состоянии канала для работы с антенными решетками и запрос на планирование передачи осуществляются с использованием восходящего физического канала управления PUSCH (Physical Uplink Control Channel). Существует несколько различных форматов PUSCH в зависимости от объема информации и продолжительности передачи. Короткий PUSCH передается в одном или двух символах временного интервала и может моментально предоставить обратную связь. Для сравнения, в LTE задержка составляет 3 мс. Для ситуаций, когда длительность короткого PUSCH слишком мала, чтобы обеспечить достаточное покрытие соты, существует также возможность использования более длинных PUSCH.

1.2.5. Формирование луча

Поддержка большого числа активных антенных элементов как для передачи, так и для приема является ключевой особенностью NR. На высоких частотах антенные решетки в основном используются для формирования луча и, как следствие, увеличения покрытия соты, в то время как на более низких частотах они обеспечивают функции пространственного разнесения передачи для снижения уровня помехи, а также могут быть использованы для осуществления одновременного обслуживания нескольких АУ. Каналы NR, включая те, которые используются для управления и синхронизации, разработаны с учетом возможности использования направленных передач. Информация о состоянии канала (CSI, channel state information) для конфигурации антенных решеток может быть получена на основе передачи опорных сигналов CSI в нисходящем канале связи.

NR поддерживает аналоговое формирование луча в дополнение к цифровому методу. В соответствии с этим луч может формироваться после цифро-аналогового преобразования, что позволит упростить реализацию, по крайней мере, на начальном этапе. Аналоговое формирование луча приводит к тому, что приемный или передающий

луч в каждый момент времени может быть сформирован только в одном направлении и требует дополнительной развертки, когда один и тот же сигнал повторяется в нескольких символах OFDM, но в разных передающих лучах. Например, такое возможно для услуги мультимедиа.

В стандарте технологии NR определены процедуры управления лучом и протоколы сигнализации. Каждый передающий луч идентифицирован таким образом, чтобы помочь АУ в выборе приемного луча (в случае аналогового формирования приемного луча), который будет использоваться для управления и приема данных. Для большого числа антенных элементов в фазированных антенных решетках лучи являются узкими, и они могут быть легко «потеряны» при изменении местоположения АУ. В стандартах NR также определены процедуры восстановления луча, инициированные АУ. Более того, так как каждая сота может иметь несколько БС (каждая со своими лучами), процедуры управления обеспечивают прозрачную для АУ мобильность для беспроводной передачи вызова между лучами разных БС.

1.2.6. Доступ к сети

Метод доступа к сети NR в целом аналогичен LTE и предусматривает четырехэтапную процедуру случайного доступа. Однако существуют указанные далее принципиальные отличия.

В LTE логические сигнальные каналы случайного доступа расположены в центре несущей и повторяются каждые 5 мс. Остановившись на каждой несущей частоте в течение по меньшей мере 5 мс, устройство гарантированно услышит логический канал случайного доступа. Таким образом, АУ должно прослушивать все возможные несущие частоты. В NR блок синхронизации передается один раз каждые 20 мс. Из-за более длительного периода между последовательными блоками синхронизации АУ, прослушивающее несущие NR, должно задерживаться на каждой возможной частоте в течение более длительного времени. Чтобы уменьшить общее время поиска, NR использует случайное положение блока синхронизации. Последнее позволяет значительно снизить задержку доступа.

1.3. Сети доступа 5G NR в миллиметровом диапазоне частот

До настоящего времени спектр миллиметрового диапазона частот не использовался системами сотовой связи. Причина заключается в специфических особенностях распространения радиоволн, включая высокие потери распространения, атмосферного и дождевого поглощения, низкой дифракции вокруг препятствий, более высокого рассеяния из-за шероховатости материалов, больших потерь при проникновении через объекты и, кроме того, из-за сильного фазового шума.

1.3.1. Потери распространения

Согласно стандартной модели распространения Фрииса увеличение частоты ведет к значительному росту потерь распространения. Однако при одинаковом размере физической апертуры передающая и приемная антенны на более высоких частотах излучают и принимают больше энергии, используя более узкие диаграммы направленности. Например, рост рабочей частоты значительно увеличивает потери распространения. Однако если апертура антенны поддерживается постоянной на приемнике или передатчике по мере увеличения частоты, тогда потери в свободном пространстве остаются неизменными. На практике можно поддерживать одинаковую эффективность апертуры антенны с помощью антенных решеток как на передающей, так и на приемной стороне.

Влияние числа антенных элементов БС на отношение сигнала-к-шуму (ОСШ), как функции от расстояния до БС, проиллюстрировано на рис. 1.4 для излучаемой мощности $P_T = 23$ дБм, где N_V и N_H – число антенных элементов в горизонтальной и вертикальной плоскостях. Отметим, что использование антенных решеток позволяет значительно увеличить потенциальную зону обслуживания. На сегодняшний день существуют практические реализации антенных решеток, в которых количество элементов превышает 512.

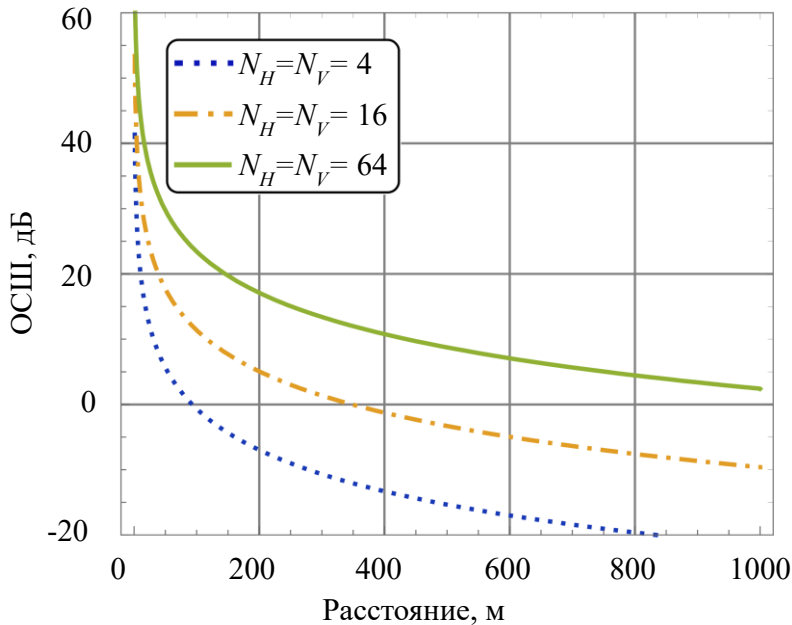


Рис. 1.4. Зависимость расстояния до БС и от числа антенных элементов

1.3.2. Влияние погодных условий

Обратим внимание, что наиболее существенное влияние на качество сигнала в миллиметровом диапазоне частот оказывает листва, причем величина падения сигнала достигает 2 дБ/м. Нарушения, вызванные сильным снегом, туманом и облаками, незначительны (менее 1 дБ/км) для реальных расстояний взаимодействия с БС. Влияние дождя обычно характеризуется дополнительным затуханием около 10 дБ/км, что может серьезно отразиться на характеристике канала связи.

Атмосферное поглощение

Атмосферное поглощение вносит дополнительные потери при распространении миллиметровых радиоволн. Основными компонентами, относящимися к поглощению в рассматриваемом диапазоне частот, являются кислород и пары воды. График поглощения показан на рис. 1.5.

Особо следует отметить поглощение кислородом, которое достигает максимальных значений на частоте 60 ГГц. Однако данный диапазон является нелицензируемым и не будет использован для коммерческих сотовых систем связи.

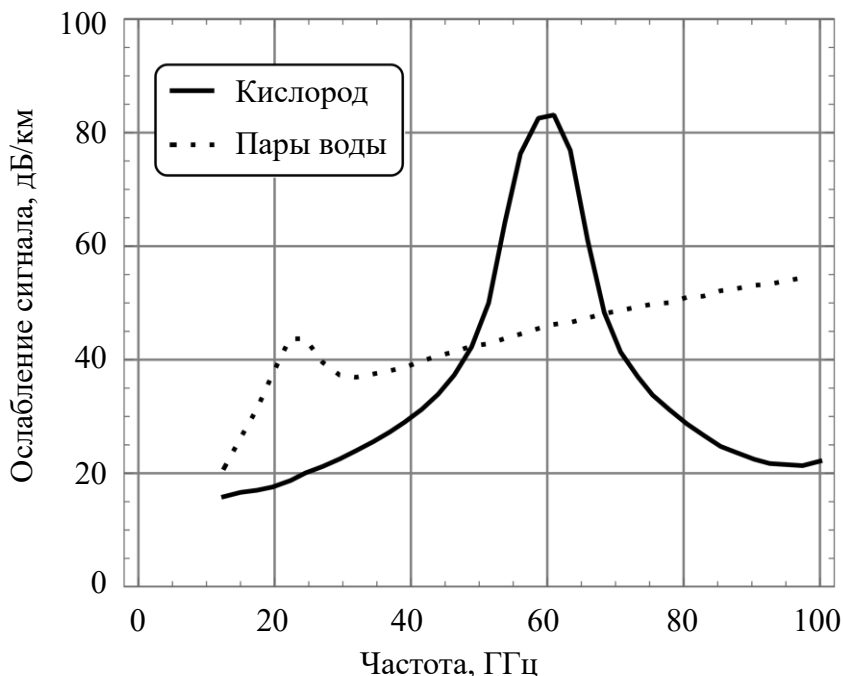


Рис. 1.5. Поглощение миллиметровых волн в атмосфере

В целом влияние поглощения несущественно как для связи внутри помещений, так и предполагаемых развертываний в городских сотовых сетях, где расстояние между БС составляет порядка 200 м. В этом случае поглощение фактически выгодно, так как позволяет уменьшать помехи от соседних БС.

1.3.3. Динамическая блокировка

Так как миллиметровые волны характеризуются меньшей дифракцией, блокировка прямой видимости между БС и АУ ведет к гораздо большим потерям мощности по сравнению с сетями доступа, работающими на частотах 3–6 ГГц.

В частности, динамическая блокировка вносит дополнительные потери порядка 15–40 дБ. Процесс динамической блокировки телом человека на частоте 60 ГГц показана на рис. 1.6. Нетрудно заметить, что ухудшение сигнала происходит достаточно резко, занимая около 20–100 мс. Длительность же блокировки зависит от плотности блокаторов, а также скорости их передвижения и может составлять от 100 мс и более.

Отметим, что при наличии блокировки прямой видимости приемник и передатчик могут использовать сложные пути распространения сигнала, например, используя для установления связи отражения от поверхностей. Однако эффективность использования таких путей сильно зависит от типа отражающих поверхностей. Так, например, отражение от шершавых поверхностей, в частности бетона или кирпича, может ослаблять сигналы миллиметрового диапазона на 40–80 дБ, что значительно выше потери при блокировке.

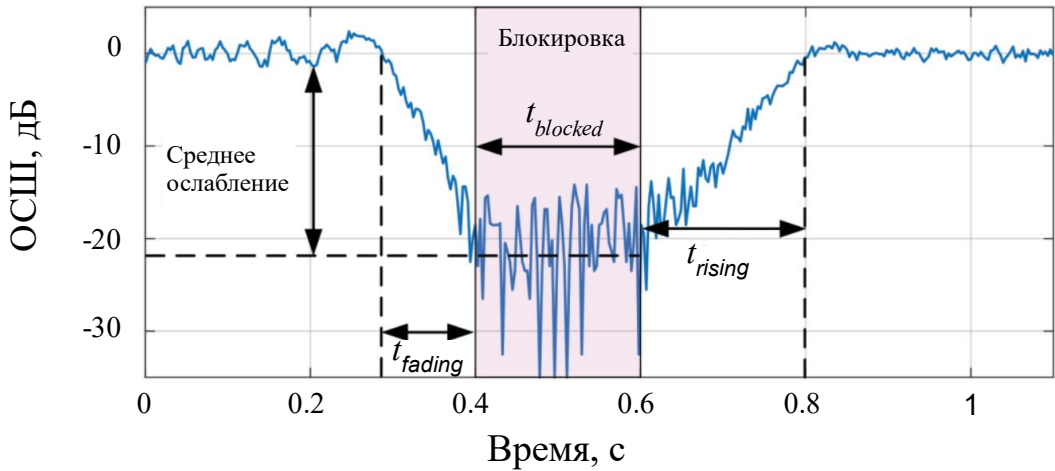


Рис. 1.6. Временная диаграмма блокировки

1.4. Программируемость и виртуализация сетевых функций

В следующих поколениях мобильных сетей ожидается высокая неоднородность сетей доступа (т.е. одновременная поддержка нескольких технологий последней мили), которая будет только повышаться из-за внедрения технологии NR, работающей в миллиметровом диапазоне. Присущая работе NR прерывистость, вызванная блокировками распространения, и новые типы высокомобильных пользователей, например, транспортные средства с выходом в интернет и беспилотные летательные аппараты, создают значительные проблемы с точки зрения управления сетью. Необходимо оперативно принимать решения, что, в свою очередь, требует новых средств для обеспечения эффективного взаимодействия и распределения сетевых ресурсов. Существующие решения ограничены тем, что все известные возможные события и соответствующие действия заранее

известны. Таким образом, внедрение сетей 5G/6G требует увеличения степени свободы и повышения интеллекта базовых сетевых элементов.

Первой важной особенностью является сетевая программируемость – важная тенденция в телекоммуникациях, которая базируется на нескольких парадигмах. Программно-конфигурируемая сеть (SDN, Software-Defined Networking) главным образом ориентирована на отделение программно-ориентированной плоскости управления от аппаратной плоскости передачи данных внутри сетевой инфраструктуры, что позволяет автоматически и динамически корректировать поведение сетевых узлов. Интеллектуальная реализация концепции SDN дает ряд преимуществ, среди которых более эффективное распределение ресурсов и облегченная интеграция множества технологий радиодоступа (RAT, Radio Access Technologies). Кроме того, SDN предлагает гибкость с точки зрения управления трафиком и контроля сети.

Еще одной важной особенностью является виртуализация сетевых функций (NFV, Network Function Virtualization), которая создает еще один уровень абстракции путем развертывания сетевых функций в качестве программных компонентов, называемых виртуальными сетевыми функциями (VNF, Virtual Network Functions). Этот подход упрощает управление сетью, обеспечивая четкое разделение между определенной функциональностью и ее фактическим развертыванием. Использование NFV позволяет снизить затраты и обеспечить обновление сети в реальном масштабе времени. NFV также повышает энергоэффективность, поскольку физический сервер, объединяющий несколько виртуальных ролей, потребляет меньше энергии.

Также важным изменением выступают повсеместные граничное кэширование и вычисления, целью которых является объединение коммуникационной инфраструктуры с возможностью промежуточного хранения данных и оперативных вычислений. Сетевая программируемость развивает синергию между этими тремя плоскостями по ряду сценариев. Это приводит к появлению принципиально новых сетевых архитектур с повышенной гибкостью.

В настоящее время ведутся многочисленные дискуссии о желательной степени интеграции программного и аппаратного обеспечения в системах 5G/6G. Существует множество проблем, требующих решения посредством исследований. По общему мнению, программируемость сети определена как ключевой компонент повышения

производительности и надежности, а также снижения затрат и эксплуатационных расходов на развертывание беспроводных сетей нового поколения. Следовательно, предполагается, что интеллектуальные сети 5G с программируемыми технологиями станут средством реализации продвинутых сетевых архитектур за счет поддержки новых приложений, таких как Тактильный интернет, Интернет навыков (IoS, Internet of Skills) и др.

1.4.1. Программируемая архитектура 3GPP 5G

Системная архитектура 3GPP 5G (рис. 1.7) подробно описана в документе TR 23.501 и представляет собой часть глобальной архитектуры NFV для виртуализации сетевых функций, разработанной ETSI. В частности, все сетевые функции рассматриваются как части программного обеспечения, которые могут работать на стандартном аппаратном обеспечении и теоретически могут перемещаться в различные места в зависимости от требований оператора связи.

В этой архитектуре роль программируемости рассматривается с точки зрения сквозного подхода, включая (i) программные решения по управлению трафиком в подсистеме доступа, (ii) роль SDN в основной подсистеме и (iii) то, как она должна взаимодействовать с архитектурой 3GPP. В целом определенные программным обеспечением правила должны применяться в физической инфраструктуре для удовлетворения определенного уровня качества обслуживания (QoS, Quality of Service), начиная с выбора узлов доступа и заканчивая управлением трафиком в транспортной сети. Использование программируемости в политике управления сетью и обеспечения QoS имеет решающее значение для изоляции трафика, а также доставки данных в плоскости пользователя. При необходимости это позволяет устанавливать приоритеты в транспортной сети.

Сетевая виртуализация и программируемость активно изучаются в контексте европейских проектов Horizon 2020/2027, представляя архитектурные решения, которые могут удовлетворять требованиям различных вариантов использования, например, 5G Car, 4G COHERENT, SONATA. Хотя существует множество подходов к построению гибкой сети 5G, представленная на рис. 1.7 архитектура объединяет компоненты, которые являются ключевыми в предоставлении QoS-совместимых сервисов в хорошо отлаженных решениях, таких как 3GPP и ETSI NFV. Она позволяет объединять

аспекты, связанные с RAN и основной сетью в контексте QoS и управления политиками, с целью обеспечения общей сквозной надежности доставки трафика, обеспечивая согласованность между применением политик на всей цепочке реализации предоставляемых сетевых услуг.

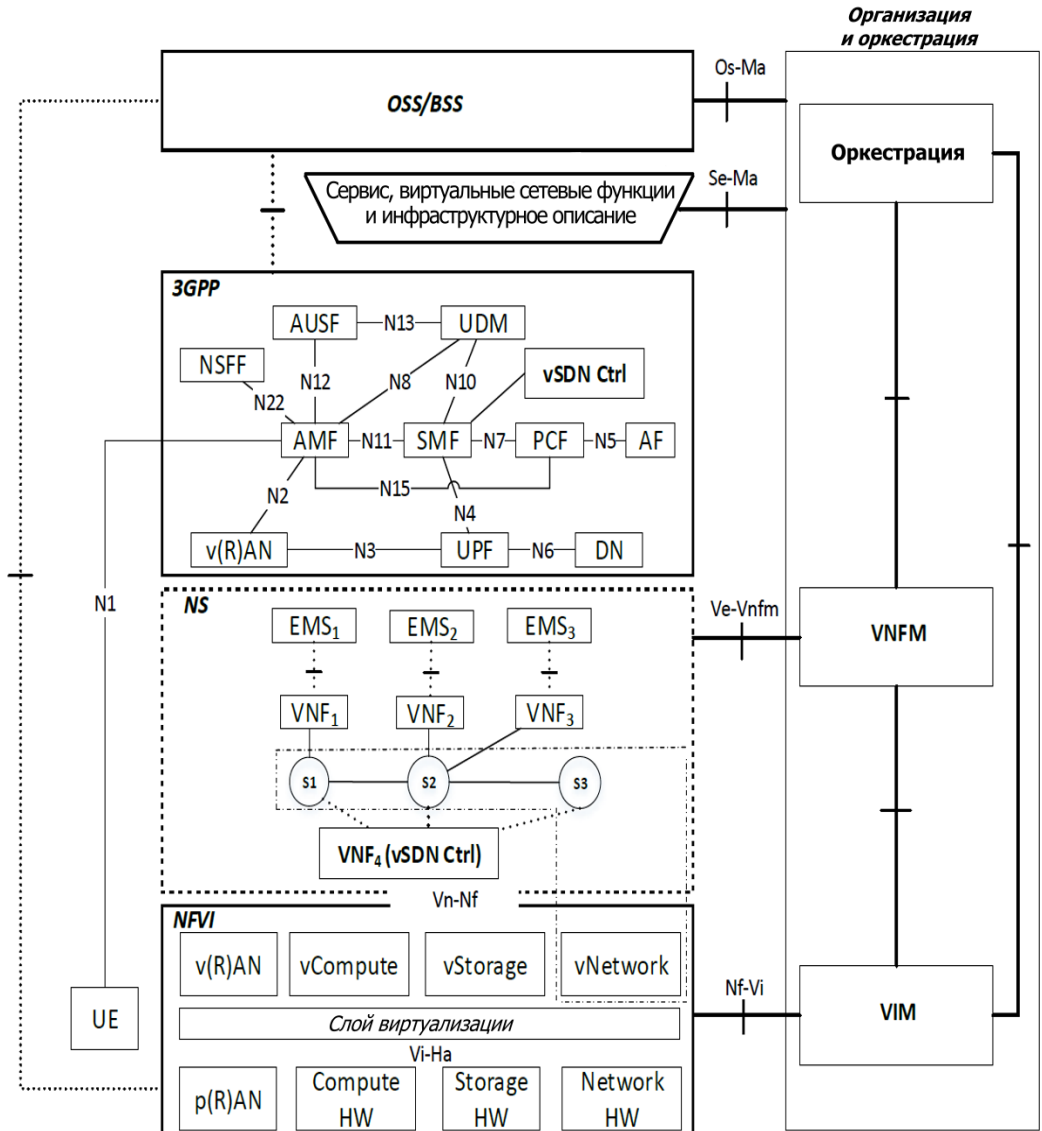


Рис. 1.7. Архитектура SDN/NVF 3GPP/ETSI

1.4.2. Модульность и принципы использования

Отличительной особенностью системной архитектуры 5G, определяемой 3GPP, является модульный принцип в проектировании сетевых функций, который обеспечивает необходимую гибкость для выполнения сетевых функций, а также позволяет реализовать концепцию сегментации сетевых ресурсов (используются также термины «нарезка» или «слайсинг», см. раздел 1.5). Модульность архитектуры 5G предназначена для поддержки развертываний с использованием как NFV, так и SDN-технологий. Предполагается, что все сетевые функции выполняются как VNF в рамках виртуализированной инфраструктуры, т.е. облачной инфраструктуры. Хотя существует множество способов инкапсуляции VNF, в том числе контейнеры, в качестве наглядного примера можно рассмотреть виртуальные машины, так как виртуализация является устоявшейся технологией, а OpenStack широко используется в облаках операторов связи.

В документе 3PPP TR 23.501 определены следующие основные сетевые функции:

- передачи данных пользователей (UPF, User Plane Function), отвечающая за обработку маршрута сессии передачи данных в плоскости пользователя и обеспечивающая интерфейс к сети передачи данных (за пределами домена 3GPP);

- управления сессиями (SMF, Session Management Function), отвечающая за установку, модификацию и запуск сеанса связи. Среди множества задач SMF наиболее важны контроль политики управления QoS и информация о маршрутизации плоскости данных в UPF;

- управления доступом и мобильностью (AMF, Access and Mobility Management Function), которая контролирует решения о доступе к сети, а также осуществляет все связанные с мобильностью процедуры. Основываясь на множестве входных данных и метрик, программное обеспечение, определяющее применение правил, позволяет AMF принять решение о наилучшей точке доступа для пользователя;

- управления политиками (PCF, Policy Control Function), отвечающая за предоставление правил политик управления соответствующим функциям плоскости управления и поддерживающая единую структуру политик по всей сети.

В рамках 3GPP было проведено много дискуссий о том, как обеспечить соответствие между правилами внедрения QoS в сетевые

функции пользовательской плоскости, например, взаимодействие между PCF, SMF и AMF, и правилами внедрения QoS в транспортной сети. С этой целью можно рассматривать механизм интерфейса от SMF к виртуальному контроллеру SDN, который позволит системе 3GPP контролировать правила передачи данных в сетевой инфраструктуре и, таким образом, удовлетворять требованиям к качеству обслуживания. В этом смысле, как только UE/eNB инициирует услугу с определенными QoS требованиями, SMF будет отвечать за применение этих правил в физической сети, а также за обеспечение согласованности между радиодоступом, ядром и транспортными сетями.

В описанной архитектуре все сетевые функции 3GPP выполняются как VNF, поэтому виртуальный SDN-контроллер является важным сетевым элементом, с которым можно связаться при помощи любой из функций 3GPP. Если говорить точнее, SMF может давать инструкции контроллеру SDN для настройки правил в виртуализированной сотовых сетей для управления QoS. Это позволяет полностью контролировать ресурсы виртуализированной сети с помощью системы 3GPP в отличие от статической конфигурации, которая предписана контроллером SDN оператора.

Существующие решения, разработанные в контексте API, такие как ONOS Intent Framework и OpenDaylight Network Intent Composition, предлагают удобную конфигурируемость сетевых интерфейсов, но в то же время отличаются ограниченной функциональностью, которой недостаточно для поддержки предполагаемого комплексного решения. В данном контексте физический SDN контроллер может управлять физической инфраструктурой, которая соединяет физические узлы Virtual Infrastructure Manager между собой, а также с внешними сетями. В свою очередь, виртуальный SDN контроллер управляет виртуализированными коммутаторами, созданными в виде VNF для поддержки виртуализированных функций 3GPP, развернутых на слейсе арендатора ресурсов сети.

На более высоком уровне архитектуры дескрипторы сетевых сервисов, согласно спецификациям, NFV используются для описания сетевых услуг и функций, в то время как OSS/BSS используется оператором NFV для целей управления и тарификации. С правой стороны схемы на рис. 1.7 слой MANO охватывает все слои архитектуры NFV и отвечает за управление инфраструктурой, мониторинг состояния сетевых функций, координацию их жизненного цикла и,

наконец, ведение каталога дескрипторов сетевых услуг и функций, которые могут быть развернуты в системе.

Для реализации сетей 5G 3GPP оператору NFV необходимо предоставить системе управления дескрипторы VNF. Эти дескрипторы содержат информацию о конфигурации и развертывании, например количество ядер процессора, оперативной памяти, хранилища и сетевых интерфейсов для каждой виртуальной машины, на которой будет размещен VNF. Дескрипторы также могут содержать дополнительную информацию, специфичную для VNF, в частности параметры конфигурации 3GPP для шлюза пакетной коммутации (PGW, Packet Gateway). Оператору нужно будет включить VNF в виртуальную машину путем добавления образов этих VNF в пул виртуального хранилища. На данном этапе дескриптор сетевой службы можно использовать для создания полной сетевой службы путем включения всех ее компонентов VNF и виртуальной сетевой инфраструктуры, которая будет соединять их вместе, чтобы предоставить функциональную совместимую с 3GPP 5G-систему.

1.5. Сетевой слайсинг

Концепция сегментации сетевых ресурсов сети так называемого слайсинга, была введена консорциумом 3GPP в релизе 14 (см., например, 3GPP TS 22.185) как одна из существенных особенностей сетей 5G, а затем уточнена в релизах 15 и 16. Ожидается, что сегментация сети радикально упростит выход на рынок операторов виртуальных мобильных сетей (MVNO, Mobile Virtual Network Operators) и предоставит дифференцированные по качеству сетевые услуги. Такая функциональность является существенным изменением в мире сотовых сетей, позволяющим создавать многоуровневые сетевые структуры, схожие с современной сетью Интернет и обеспечивающие совместное использование ресурсов с логической изоляцией между несколькими арендаторами или сетевыми операторами и/или услугами в многодоменном контексте. Опираясь на набор требований и механизмов, обеспечивающих распределение ресурсов как по радиоинтерфейсу, так и по физической сети, технология обещает виртуализировать ресурсы физической мобильной сети оператора (PMNO, Physical Mobile Network Operator), обеспечивая высокую степень логической изоляции между MVNO.

Согласно документу ассоциации GSMA NG.116, изоляция между трафиком различных слайсов должна быть как логической, так и физической. Такая изоляция на плоскости данных может быть обеспечена с помощью механизмов, работающих на различных временных шкалах, таких как планирование передачи пакетов, контроль доступа к сети (CAC, Connection Admission Control), фаза планирования сети и т.д. Однако PMNO должен обеспечить эффективное использование сетевых ресурсов посредством статистического мультиплексирования, что по своей природе противоречит требованию физической изоляции трафика.

Хотя предполагается, что сети 5G должны быть гетерогенными, позволяющими использовать несколько радиointерфейсов (см. раздел 1.8), но технология NR может быть наиболее распространенным радиointерфейсом. Несмотря на то, что NR построена для работы с широким диапазоном частот, именно использование миллиметровых волн позволит обеспечить высокие скорости при использовании сетевого слайсинга.

Согласно документу 3GPP TS 23.501, сетевой слайс является логической сетью, обеспечивающей предписанный набор сетевых возможностей, которые определяются в терминах сетевых функций. Функции – это определенные сетевые возможности, связанные либо с плоскостью данных, либо с плоскостью управления (см. раздел 1.4). В то же время, согласно GSMA NG.116, слайс – это набор сетевых функций и ресурсов, включая те, которые находятся на радиointерфейсе и в основной сети, представляющие собой существующий слайс. Слайс может рассматриваться как выделенная для определенных потребностей сеть.

Предполагается, что параметры QoS для слайса согласовываются между PMNO и внешними организациями, покупающими услугу, и определяются в рамках договора об уровне обслуживания между поставщиком услуг и оператором связи (SLA, Service Level Agreement). Хотя 3GPP неоднозначно определяет целевые параметры, шаблон соглашения о гарантированном уровне обслуживания был предложен GSMA в качестве открытого стандарта NG.116 «Общий шаблон сетевого слайса». Технически слайсы различаются по идентификатору, называемому «информация о помощи в выборе одного сетевого среза» (S-NSSAI, Network Slice Selection Assistance Information). Идентификатор определяет тип слайса/услуги (SST, Slice Service Type), описывающий глобальное применение слайса или дифференциатора слайса – опциональное поле, используемое

для различия слайсов одного и того же типа. В настоящее время определены три стандартных значения SST: eMBB, URLLC, mMTC. Примеры параметров шаблона слайса, определенные в NG.116, приведены в табл. 1.1.

Таблица 1.1

Базовые параметры слайса согласно NG.116

Параметр	Значения
Спецификация области	1: вся страна; 2: список регионов
Гарантии задержки	0: не поддерживается; 1: поддерживается
Гарантированная скорость	180 кбит/с – IP-телефония, 25 Мбит/с - игры
Максимальная скорость	Бронзовый/Серебряный/Золотой: 50/400/1000 Мбит/с
Изоляция уровня	0: без изоляции; 1: физическая; 2: логическая
Физическая изоляция	0: процессы; 1: память; 2: физическая
Логическая изоляция	0: виртуальная; 1: основанная на NF; 2: арендаторская
Максимальный размер пакета	eMBB 1500 байт, IoT 40 байт, URLLC 160 байт
Поддержка критически важных задач	0: не критически важная; 1: критически важная
Критически важные способности	1: назначение приоритета; 2: приоритетное прерывание; 3: локально
Критически важные услуги	0: MCPTT; 1: MCData; 2: MCVideo
Число соединений	10 ⁵ сессий, 10 ⁷ сессий
Число терминалов	10 ⁵ терминалов, 10 ⁷ терминалов
Точность	+/- 1 м, +/- 0,01 м
Диапазон частот	n1, n77, n38
Анализ причин неисправностей	0: N/A, 1: пассивный, 2: активный
Поддержка непрерывности обслуживания	0: N/A, 1: режим SSC 1, 2: режим SSC 2
Одновременное использование слайсов	0: любые, 2: одинаковое значение SST
Параметры качества услуг в слайсе	1: голос, 2: видео, 3: игры
Уровень приоритета	10 – IMS, 30 – онлайн-игры
Коэффициент потери пакетов	10 ⁻⁶ – данные MC, 10 ⁻² – V2X
Максимальный коэффициент потери пакетов	1%; 5%
Плотность терминалов	10 ⁵ UE/км ² , 10 ⁶ UE/ км ² , MIoT

1.6. Интернет вещей

В последние десятилетия IoT стремительно развивается: начиная с технологии радиочастотной идентификации (RFID, Radio Frequency Identification) в начале 1980-х годов и ускорения в беспроводных сенсорных сетях (WSN, Wireless Sensor Network) в конце 1990-х и в начале 2000-х годов. На сегодняшний день, несмотря на ограниченную функциональность, сегмент IoT становится все более значительным и взаимосвязанным с другими технологиями. В дополнение к недорогому оборудованию, включая датчики, механические приводы и интеллектуальные счетчики, IoT следующего поколения охватывает более продвинутое сетевые устройства с повышенной степенью сложности и неоднородности, в том числе самоуправляемые автомобили с подключением к сети Интернет, портативные гаджеты, промышленные роботы и квадрокоптеры.

Сегодня мы становимся свидетелями устойчивого и постоянного проникновения концепций IoT в автомобильную промышленность, где появляется возможность управлять отдельными транспортными средствами (далее подвижные объекты) через облачные сети, которые глубоко интегрированы с существующими системами кэширования и вычислительной техники. Например, возможны комплексные решения между множеством разнородных IoT-устройств, таких как интеллектуальные дорожные знаки, светофоры и уличные камеры. Эта новая тенденция открывает двери к колоссальным выгодам. С одной стороны, подключенные транспортные средства (объекты) становятся способными получать важные своевременные обновления и предупреждения от дорожной инфраструктуры, повышая безопасность и удобство вождения. С другой стороны, появляется возможность управлять данными от сенсоров и значительно сэкономить заряд аккумуляторов конечных устройств.

Такое развитие IoT выдвигает новую парадигму коллективного восприятия (*crowdsensing*), так называемый краудсорсинг данных датчиков с мобильных устройств, которые совместно измеряют, сопоставляют либо анализируют определенные сведения и данные. Резкий скачок вперед в гибких приложениях краудсорсинга может произойти, если люди будут соответствующим образом заинтересованы вносить свой вклад, используя возможности радиосвязи своих подключенных транспортных средств. Следовательно, автоматическая ретрансляция чувствительных к задержке данных от датчиков IoT может повысить общую надежность и энергоэффективность передачи

информации. Такое тесное взаимодействие между встроенными IoT-устройствами и транспортными средствами обещает не только поддерживать существующие приложения для краудсорсинга, но также позволяет предоставлять новые услуги в областях коллективной безопасности дорожного движения, экологичного транспорта и многих других областях. Подробно сетевые технологии поддержки IoT рассмотрены в главе 9.

1.7. Технологии прямого взаимодействия устройств

Среди попыток создать некий оптимальный баланс между покрытием и емкостью в архитектуре сотовой сети в течение нескольких десятилетий, начиная с сетей стандарта GSM, доминирует сотовый технологический ландшафт, обеспечивающий функциональность мобильной связи. Поскольку БС всегда имели связь с высокоцентрализованным ядром сети, сотовая архитектура была полностью совместима с существующими телефонными сетями общего пользования (ТфОП), гарантируя при этом покрытие почти всего земного шара и высокий уровень качества обслуживания.

На протяжении своего 40-летнего существования сотовые сети эволюционировали в направлении повышения спектральной эффективности, но при этом не смогли обеспечить масштабирование емкости сети до уровней, которые соответствовали бы постоянно растущему потребительскому спросу. Это произошло, прежде всего, потому, что емкость в сотовой системе распределяется между всеми пользователями в зоне покрытия, а при росте потребительского спроса эта область не изменяется. Было установлено, что увеличение числа сот на единицу области сотовой сети может устранить проблему. Поэтому существующие сотовые сети эволюционировали в направлении развертывания многочисленных малых сот, также называемых пико- (с покрытием около 100 метров) и фемтосотами (с Wi-Fi-подобным покрытием). Такие сети получили название гетерогенных сетей (см. параграф 1.8 и главу 11).

Однако при увеличении плотности сети с одновременным ростом пропускной способности, в расчете на одного пользователя увеличиваются и расходы на инфраструктуру, располагаемую теперь в непосредственной близости от пользователя. В то же время еще не достигнут консенсус между существующими сетевыми технологиями и технологиями сетей 5G; при этом плотность сетевой инфраструктуры не

может возрасти бесконечно. Поэтому среди множества предложений для сетей 5G появились и такие, в основе которых лежит не сетевая структура. Это, в первую очередь, прямые взаимодействия устройство–устройство (Device-to-Device, D2D) – одно из самых перспективных направлений развития сетей связи в последние годы. Фигурировавшая сперва под разными наименованиями, такими как самоорганизующиеся сети (ad-hoc networks) и ячеистые сети (mesh networks), кооперативные сети, клиентские сети с ретрансляцией (client-relay networks), а также и когнитивное радио (cognitive radio), технология D2D стала одной из фундаментальных исследовательских проблем в сотовых сетях связи. Критически важно также то, что эта технология предполагает взаимодействие клиентов сотовой сети без использования централизованной инфраструктуры, уменьшая нагрузку на нее.

Теоретически D2D предоставляет для связи лучшую энергетическую эффективность и даже более естественные потоки трафика. Вследствие этого за прошедшие несколько лет D2D-взаимодействия привлекли существенное внимание как производителей оборудования, так и исследовательских организаций. На практике, естественно, D2D приносит ряд своих проблем, но потенциальное использование ресурсов нелицензируемого спектра выглядит привлекательным. Как будет показано далее в главе 10, посвященной анализу технологии D2D, это направление – одно из самых простых способов повысить пропускную способность сотовых систем без необходимости переосмысления всей концепции.

1.8. Гетерогенные сети доступа

В настоящее время университеты, научно-исследовательские институты и предприятия промышленности придают особое значение вопросам разработки технологий, систем и сетей 5G, имея конечной целью создание полнофункциональной сотовой сети связи 5G к 2025 году. Основные требования к технологиям 5G предполагают создание системы с очень большой емкостью (в 1000 раз больше по сравнению с 4G) и обеспечение равномерного беспроводного покрытия сети с возможностью высокоскоростной передачи данных. В длительной перспективе основные игроки на рынке систем и сетей связи, включая поставщиков оборудования, производителей устройств и операторские компании, будут заняты внедрением прогрессивных технологий 5G, тренды развития которых видны уже сегодня.

Для достижения тысячекратного увеличения емкости два ключевых механизма являются необходимыми – пространственное уплотнение сети и агрегация спектра. В первом случае естественным предположением по построению сетей с большой плотностью выглядит увеличение числа сот с меньшей областью покрытия в существующей сетевой архитектуре. Во втором случае предполагается совместное использование больших участков радиоспектра в разнообразных спектральных диапазонах (лицензируемых и нелицензируемых), в перспективе включая и крайне высокие частоты.

Быстрое распространение недорогих и эффективных малых сот различного размера (микро-, пико-, фемто- и т.д.) с поддержкой разнообразных технологий радиодоступа привело к созданию так называемых гетерогенных сетей (HetNet, Heterogeneous Network). Еще большее уплотнение сети способствует появлению сетей сверхвысокой плотности (ultra-dense HetNets), базирующихся на сетевой инфраструктуре с поддержкой технологий множественных радиоинтерфейсов (например, малые соты с возможностями функционирования в различных диапазонах). В свою очередь для макросот роль «якоря» (обеспечение покрытия крупномасштабных зон) становится даже более важной, чем обеспечение необходимой емкости. Только макросота, обладающая необходимым покрытием, может обеспечить централизованное и сбалансированное распределение пользователей по «бустерам» (малым сотам с большой емкостью), что будет способствовать эффективной выгрузке трафика данных пользователей в условиях повышенной нагрузки на сеть.

Главным преимуществом рассмотренной выше архитектуры «якорь/бустер» является именно централизованное управление через макросоту. В первую очередь централизованное управление улучшает возможности координации между малыми сотами с множественным радиоинтерфейсом. Таким образом, уровень сетевой кооперации смещается от «свободных» решений по взаимодействию между разными технологиями передачи данных к централизованным или распределенным схемам с более жесткой интеграцией на уровне сети радиодоступа (RAN). Новые механизмы взаимодействия на уровне сети радиодоступа были рассмотрены в стандартах 3GPP, и эффективные решения для интеграции стандартов 3GPP и беспроводных локальных сетей (WLAN, Wireless Local Area Networks) были включены в стандарты систем LTE еще в версии 12 (Release 12). Агрегация решений WLAN/LTE с тесной интеграцией WLAN в сети операторов сотовой связи также стандартизована и для LTE версии

13 (Release 13). В целом прирост емкости и улучшение качества обслуживания QoS посредством интеграции множественных сетей связи зависят от того, как процедуры выбора и интеграции радиотехнологий скоординированы с уровнем управления доступа к каналу и сетевым уровнем.

Сегодня существует множество альтернативных радиотехнологий (GSM/EDGE, WCDMA/HSPA, LTE-FDD/TDD, NR-TDD/FDD, Wi-Fi, WiGig IEEE 802.11ad/ay и т.д.), поэтому появляющаяся кооперация на уровне сети радиодоступа должна быть более гибкой, чем возможные решения для «свободной» высокоуровневой координации, что позволит динамично управлять радиоресурсами для увеличения производительности системы. В частности, интеграция на уровне сети радиодоступа предоставляет возможности для управления потоками данных в реальном времени с целью гибкого распределения трафика пользователя между доступными радиоинтерфейсами, оптимизируя общую производительность сети.

В главе 11 рассматриваются механизмы динамического распределения ресурсов каналов доступа гетерогенных сетей, с целью обеспечения максимальной выгоды с точки зрения пропускной способности конечного пользователя/системы в целом, и/или максимизации справедливости распределения ресурсов. При этом требуется обеспечить низкую сложность решений, что необходимо для практической реализации.

1.9. Терагерцовый интерфейс сетей 6G

Разработка сетей мобильной связи 5G в миллиметровом диапазоне частот практически завершена и операторы связи тестируют первые развертывания таких сетей. Соответственно, акцент в исследованиях беспроводных сетей связи смещается в сторону определения требований и технологических возможностей будущих сетей шестого поколения (6th generation, 6G). В связи со сложностью внедрения технологии миллиметровых волн в 5G многие производители и операторы согласны с тем, что системы следующего поколения (часто называемые 5G+) в первую очередь повысят эффективность обмена данными в существующих полосах на частоте 28–73 ГГц. В то же время использование частот выше 100 ГГц планируется для решений 6G и выше в качестве ключевого фактора для новых приложений, требующих сверхвысоких пользовательских скоростей передачи данных, достигающих гигабит в секунду.

Видение 6G в настоящее время находится на начальном этапе развития и разработки. Особое внимание уделяется футуристическим сценариям, таким как повсеместное проникновение систем дополненной и виртуальной реальности (AR/VR, Augmented Reality, Virtual Reality), голографическое телеприсутствие и управление роботами. С технологической точки зрения будущие сети 6G предполагают внедрение ряда новшеств. К ним относятся полнодуплексная беспроводная связь, использование машинного обучения для оптимизации производительности, решения для множественных подключений (гетерогенные сети) и многое другое.

Что касается поддержки экстремальных скоростей передачи данных, ожидается, что системы 6G дополнят микроволновые и миллиметровые соединения с беспроводной связью на частотах выше 100 ГГц. В частности, речь идет о терагерцовом диапазоне частот (300 ГГц–3 ТГц) и использовании видимого спектра частот (visible light communications, VLC, 400–800 ТГц). Обе технологии могут предоставить полосы частот в десятки, сотни или даже тысячи гигагерц. В последние десятилетия научное сообщество, организации по стандартизации сетей связи и регулирующие органы активно изучали возможность использования таких видов связи. Для VLC эти усилия недавно привели к созданию стандарта IEEE Std. 802.15.7–2018, определяющих физический уровень и подуровень управления доступом к среде (MAC, Medium Access Control) для перспективных оптических беспроводных коммуникаций.

Стандартизация терагерцовой беспроводной связи в пределах IEEE началась в начале 2008 года, когда была создана группа Terahertz Interest Group (IGthz) под эгидой IEEE 802.15. Ключевые решения по проектированию и предварительные прогнозы производительности, сделанные членами группы, заложили основу целевой группы IEEE по *Беспроводной связи 100G* (100G Wireless, TG100G, IEEE 802.15.3d) в 2014 году. Первый стандарт IEEE для беспроводной связи в нижней части терагерцового диапазона частот (субтерагерцовом диапазоне) TG100G, IEEE Std. 802.15.3d-2017 был официально утвержден осенью 2017 года. В то время как производители оборудования в настоящее время работают над прототипами решений, непосредственно следуя спецификациям, успешное использование стандарта требует глубокого понимания его основных особенностей и существующих ограничений.

1.9.1. Приложения и сценарии использования

Стандарт 802.15.3-2017 является улучшением стандарта IEEE Std. 802.15.3-2016. Улучшенный стандарт 802.15.3-2017 определяет беспроводной физический уровень, работающий со скоростью 100 Гбит/с и поддержкой динамического снижения скорости. При этом предоставляемые скорости достаточно высоки, однако дальность связи ограничена.

Стандарт IEEE Std. 802.15.3d-2017, как один из первых шагов по стандартизации беспроводной связи на частоте 300 ГГц, в первую очередь ориентирован на приложения, которые возможно реализовать, учитывая текущий уровень развития электроники. Отметим, что рассматриваемое применение ограничивается соединениями «точка–точка» между устройствами, которые остаются статическими или «почти» статическими во время обмена данными. Поддержка мобильности узлов и многоканального доступа в настоящее время находится в стадии изучения. Далее рассмотрим потенциальные сценарии использования.

Беспроводные фронтальное и транзитное соединения

В будущих беспроводных сетях (5G и 5G+) предполагается интенсивное использование сверхплотных малых сот миллиметрового диапазона частот для удовлетворения потребностей в пропускной способности в местах большого скопления людей. Одной из основных задач является обеспечение надежного и достаточно высокоскоростного транзитного соединения (backhaul) между БС сетей 5G и проводной сетью оператора. Кроме того, когда малая сота оснащена несколькими удаленными радиоголовками (remote radio head, RRH), требуется стабильное и быстрое фронтальное (fronthaul) соединение между радиоокончанием и блоком обработки сигнала (BBU, Baseband Unit).

В настоящее время исследуется несколько вариантов решения этой задачи. Первым вариантом является использование оптоволоконных каналов связи. Однако массовое использование проводных соединений в сверхплотных сетях увеличивает расходы оператора связи, а также усложняет развертывание и обслуживание небольших сот. Чтобы избежать этих трудностей, недавно была предложена беспроводная альтернатива для дополнения оптических линий связи в сложных условиях развертывания сетей. Этот подход

называется интегрированным доступом и транзитом (IAB, Integrated Access and Backhaul) и предполагает повторное использование волн миллиметрового диапазона частот для транзитных соединений. Это решение является гибким, но его недостатком является ограниченная скорость передачи данных по транзитным каналам.

Представленный стандарт беспроводной связи IEEE Std. 802.15.3d-2017 сочетает в себе сильные стороны обоих вариантов. С одной стороны, стандарт поддерживает скорость передачи данных 100 Гбит/с, что на порядок выше, чем ультрасовременные технологии миллиметрового доступа. Благодаря этому десятки одновременных каналов от БС миллиметрового диапазона частот могут быть мультиплексированы в пределах одного транзитного канала. С другой стороны, использование стандарта IEEE 802.15.3d-2017 снимает проблему избыточного количества проводных соединений. К тому же ожидаемый диапазон в несколько сотен метров достаточен для фронтального и транзитного соединения в будущих сверхплотных установках БС.

Беспроводные соединения между центрами обработки данных

Высокоскоростные каналы связи стандарта IEEE 802.15.3d-2017 могут быть использованы для беспроводных соединений между серверами в центрах обработки данных. Постоянный рост виртуализации, облачных вычислений и онлайн-игр предъявляет строгие требования к центрам обработки данных сети (DCN, Data Center Networks). Одно из узких мест такого центра – обеспечение эффективного обмена данными между всеми вычислительными узлами и узлами хранения данных. Текущие решения в основном используют высокоскоростную технологию Ethernet с передачей по оптоволокну. Однако этот подход характеризуется ограниченной гибкостью, так как используется большое количество кабельных каналов. Потребность в кабельных каналах ставит под сомнение реконфигурируемость центра обработки данных и затрудняет охлаждение, поскольку они могут блокировать воздушные потоки.

Сегодня активно изучается альтернативное решение, предлагающее использовать проводные соединения только для передачи данных между узлами в одном и том же серверном шкафу, а для соединения одного шкафа с другим – использовать высокоскоростные беспроводные соединения *точка–точка*. Такой подход позволяет более гибко проектировать центры обработки данных, а также

сокращает количество используемых кабелей. В частности, стандарт IEEE Std. 802.15.3d-2017 призван обеспечить такую связь на расстоянии от 10 м до 100 м, тем самым дополняя оптоволоконные соединения.

Сверхширокополосный доступ

Стандарт IEEE Std. 802.15.3d-2017 также призван обеспечить высокоскоростной обмен данными для случаев, когда устройства находятся рядом. Рассмотрим две категории: (i) связь между устройствами D2D и (ii) загрузка больших объемов данных. Первый сценарий направлен на облегчение обмена большими файлами, главным образом фото- и видеоматериалами, между различными электронными устройствами, такими как смартфоны, принтеры и цифровые фотоаппараты. В данном случае стандарт IEEE Std. 802.15.3d-2017 обеспечивает гораздо более высокий уровень гибкости по сравнению с любым проводным вариантом (например, USB Type-C).

В то же время стандарт IEEE Std. 802.15.3d-2017 обеспечивает на порядок большую пропускную способность по сравнению с существующими и разрабатываемыми беспроводными D2D-решениями в микроволновом или даже миллиметровом диапазонах частот. В частности, время загрузки двухчасового фильма (900 Мб) сокращается примерно с 10 секунд с помощью IEEE 802.11ac до всего 0,1 секунды с помощью IEEE Std. 802.15.3d-2017. Решение также предлагает резервное копирование данных с мобильных устройств на домашний или офисный компьютер за считанные секунды. Дальность действия составляет несколько десятков сантиметров, что позволяет рассматривать IEEE Std. 802.15.3d-2017 в качестве «NFC на стероидах».

Связь внутри вычислительных систем

Дальнейшая миниатюризация устройств может позволить использовать связь в терагерцовом диапазоне частот для передачи данных внутри электронных систем и даже отдельных компонентов. На сегодняшний день компьютеры уже оснащены несколькими высокоскоростными проводными соединениями, например, для связи центрального процессора (CPU, Central Processing Unit) с памятью произвольного доступа (RAM, Random Access Memory), чипсета материнской платы с сетевым интерфейсом и т.д. Кроме того, крайне сложные решения используются в появляющихся многоядерных

процессорах (CPU/GPU) для соединения вычислительных ядер и общей кэш-памяти.

Быстрый рост обмена данными внутри компьютеров ставит под сомнение масштабируемость их текущих решений. Например, современные материнские платы уже имеют до 12 слоев, в то время как появляющиеся «сети-на-чипе» (NoC, Network-on-Chip) для будущих процессоров занимают более 30% процессорного пространства и потребляемой мощности. Здесь возможна альтернатива в лице стандарта IEEE Std. 802.15.3d-2017, позволяющая обеспечить связь между критически важными компонентами при скоростях передачи данных, сопоставимых со скоростями проводных шин, и одновременно упростить конструкцию и компоновку.

1.9.2. Физический уровень IEEE 802.15 3D

Для физического уровня стандарта IEEE Std. 802.15.3d-2017 определены два режима:

- 1) THz single carrier mode (THz-SC PHY);
- 2) THz on-off keying mode (THz-ООК PHY).

Первый из этих физических режимов, THz-SC PHY, предназначен для высоких скоростей передачи данных. Этот режим в первую очередь ориентирован на использование в таких приложениях, как беспроводные фронтальные и транзитные каналы и каналы связи в центре обработки данных. Второй режим, THz-ООК PHY, ориентирован на недорогие устройства, которые не могут использовать сложные модуляционные схемы и поэтому вынуждены опираться на цифровую амплитудную модуляцию. Тем не менее даже в этом режиме могут быть достигнуты скорости в десятки Гбит/с при использовании широких каналов связи.

Стандарт IEEE Std. 802.15.3d-2017 может работать на субтерагерцовых частотах в диапазоне от 252,72 ГГц до 321,84 ГГц. Имеется 69 перекрывающихся каналов (рис. 1.8). Также имеется 8 поддерживаемых каналов с полосой пропускания 2,16 ГГц, как в стандарте IEEE Std. 802.11ad-2012, и широких каналов шириной в 69 ГГц. В зависимости от варианта использования и аппаратных возможностей либо весь частотный диапазон может быть выделен для одного канала шириной 69,12 ГГц (CHNL_ID=69), либо распределен между несколькими каналами меньшего размера. Полосы пропускания меньших каналов представляют собой целые числа, кратные

2,16 ГГц. Канал номер 41 с полосой пропускания $2,16 \times 2 = 4,32$ ГГц является каналом по умолчанию.

Режим THz-SC предназначен, прежде всего, для высокоскоростной связи. Длина кадра варьируется от 2048 байт до 2 099 200 байт, не включая преамбулу физического уровня и базовый заголовок. Режим THz-SC PHY поддерживает 6 различных модуляций. Первые четыре – это амплитудно-фазовые модуляции: двоичная (BPSK), квадратурная (QPSK), 8-фазная (8-PSK) и 8-фазная асимметричная (8-APSK). Режим также поддерживает квадратурную амплитудную модуляцию с 16 и 64 точками созвездия: 16-QAM и 64-QAM. Исправление ошибок (Forward Error Correction, FEC) выполняется с помощью одного из двух кодов с проверкой четности (Low-Density Parity-Check, LDPC): высокоскоростного 14/15 LDPC (1440,1344) или низкоскоростного 11/15 LDPC (1440,1056). Модуляции BPSK и QPSK обязательны для режима THz-SC, остальные являются дополнительными.

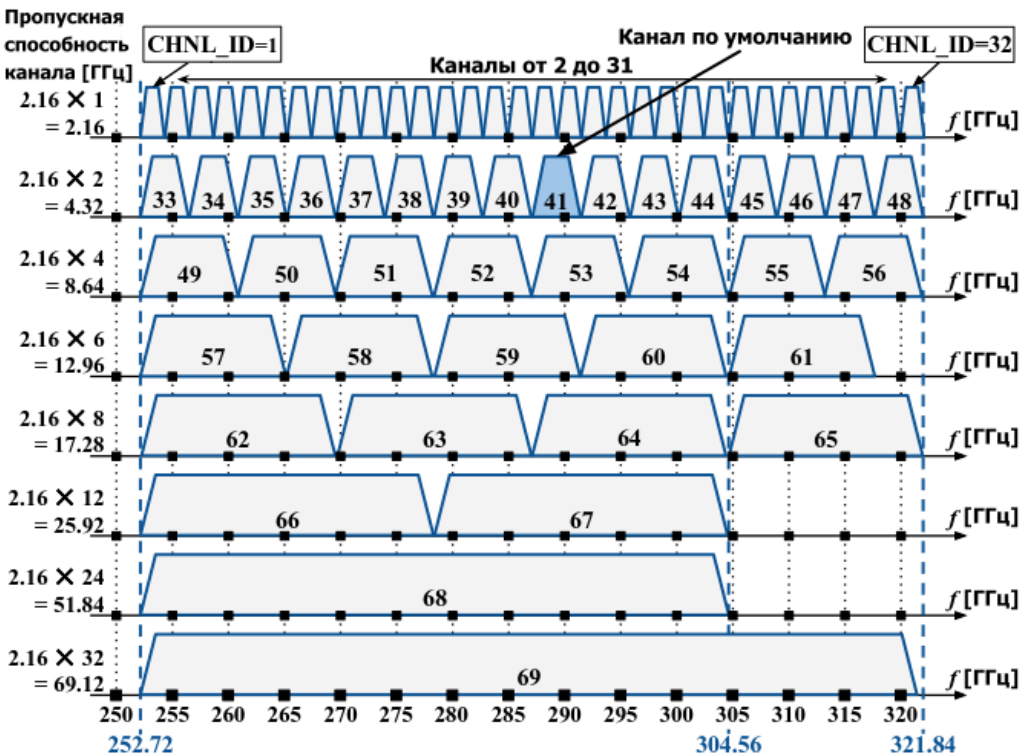


Рис. 1.8. План каналов в IEEE 802.15.3d-2017

Формат кадра, используемый в режиме THz-SC, показан на рис. 1.9. Кадр начинается с преамбулы физического уровня, используемой для обнаружения начала кадра и оценки канала на стороне приемника. Для преамбулы физического уровня определены два варианта: длинная версия передается во время работы PSP, а короткая – во время работы PAP (см. следующий раздел). За преамбулой следует заголовок физического уровня. Как заголовки протокола множественного доступа (MAC), так и заголовки физического уровня защищены последовательностью проверки заголовков (HCS, Header Check Sequence), реализованной с помощью циклического кода (CRC, Cyclic Redundancy Check), в частности CRC-16. Для повышения надежности комбинация заголовков PHY, MAC и HCS кодируется в сжатые кодовые слова расширенным кодом Хемминга. Для получения целочисленной длины порции данных в блоке могут быть добавлены дополнительные биты. Наконец, за дополнительными заголовочными битами следует полезная информация кадра.

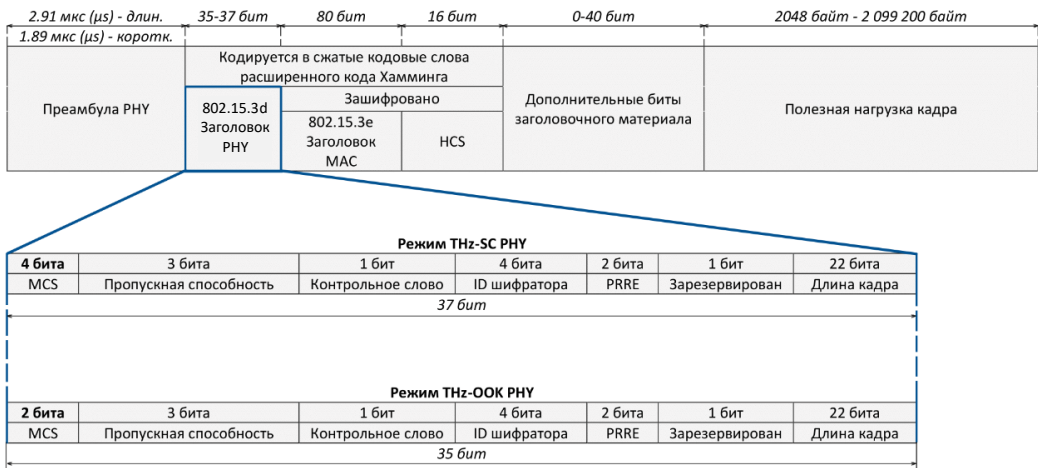


Рис. 1.9. Формат кадра IEEE 802.15.3d

В отличие от режима THz-SC, направленного на максимальную скорость передачи данных, режим THz-ООК IEEE 802.15.3d предназначен для более простых устройств. Основным примером использования этого режима – загрузка данных, максимальная дальность передачи для этого режима не превышает несколько метров. Режим THz-ООК поддерживает одну простую амплитудную манипуляцию (on-off keying, ООК) – где данные в канале представлены либо наличием несущей (логическая «1»), либо ее отсутствием (логический «0»).

Поддерживаются три метода коррекции ошибок. В частности, код Рида-Соломона (240,224) является обязательным для этого режима, так как позволяет осуществлять простое декодирование.

Структура кадра, используемая в режиме THz-ООК, аналогична структуре, используемой в режиме THz-SC, но с небольшими отличиями. В частности, поскольку существует только одна модуляция (ООК) плюс три опции FEC, для определения текущей схемы модуляции и кодирования в заголовке физического уровня используются только два бита вместо четырех битов в THz-SC. Другие поля идентичны, что делает заголовок THz-ООК PHY на два бита короче, чем в режиме THz-SC. Минимальные уровни чувствительности приемника для этих двух режимов также одинаковы и равны -67 дБм (с кодом LDPC 11/15 и наименьшей полосой пропускания 2,16 ГГц).

1.9.3. Канальный уровень IEEE 802.15 3D

В отличие от стандартов WLAN IEEE 802.11 стандарт IEEE Std. 802.15.3d-2017 поддерживает только связь вида «точка–точка». Стандарт соответствует MAC-подуровню, определенному в стандарте IEEE 802.15.3e-2017, и, таким образом, использует концепцию «парной» сети, соединяющей не более двух устройств. Характер связи «точка–точка» ограничивает диапазон возможных вариантов использования, но в то же время уменьшает проблему помех и множественного доступа к каналу.

Процедура доступа к каналу связи разделена на два периода:

- 1) установки парной связи (PSP, Pairnet Setup Period);
- 2) ассоциации парной связи (PAP, Pairnet Associated Period).

Во время периода установки парной связи одно из устройств, называемое координатором парной связи (PRC, Pairnet Coordinator), создает парную сеть и начинает периодически посылать кадры с подробной информацией о сети (рис. 1.10). Сигнальный кадр содержит информацию о количестве и продолжительности слотов доступа, которые могут быть использованы устройством, желающим подключиться к сети. Затем другое устройство, готовое к объединению в пару, называемое устройством парной связи (PRDEV, Pairnet Device), отвечает на кадр, посылая запрос на объединение в начале одного из определенных слотов доступа. После успешного приема запроса объединения PRC останавливает передачу маячков и посылает сообщение ответа на PRDEV, заканчивая PSP.

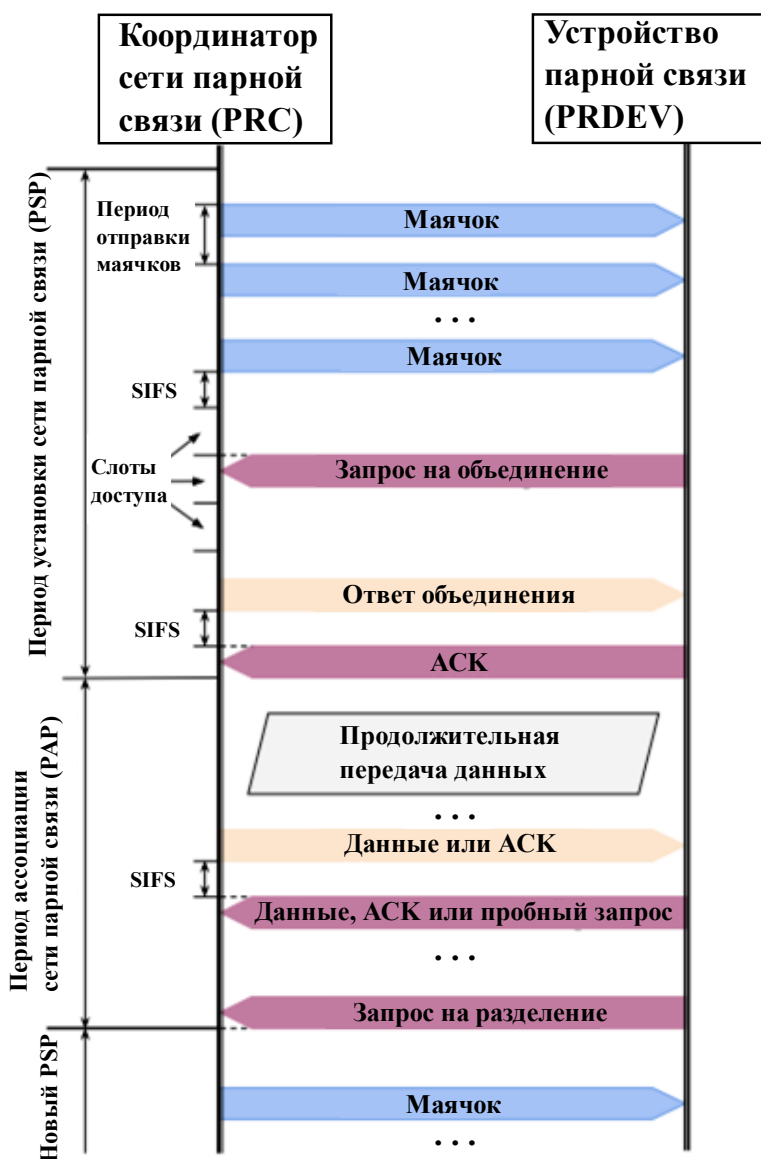


Рис. 1.10. Сигнализация для установления логической связи

Второй период – период ассоциации парной связи – начинается с успешного приема ответа ассоциации и посвящен фактическому обмену данными. В течение этого периода и PRC, и PRDEV передают кадры данных и дополнительные подтверждения, начиная с PRDEV. Кадры, передаваемые во время периода ассоциации, разделены определенным коротким межкадровым пространством (SIFS, Short Interframe Space). Когда один из узлов хочет прервать обмен данными, он посылает запрос на разрыв связи. Устройство PRC

также может прекратить период ассоциации, если в течение определенного таймаута от PRDEV не было получено никакого сообщения. В случае, если устройство PRDEV не хочет прерывать связь, но не имеет больше данных для передачи, оно может передать запрос на повторный запуск таймаута PRC. Всякий раз, когда PAP заканчивает работу (либо с помощью запроса на отключение, либо с помощью таймаута), PRC переключается обратно на PSP: продолжает передавать маячки и ожидает новых соединений.



ЧАСТЬ I

МОДЕЛИ, ХАРАКТЕРИСТИКИ И МЕТОДЫ АНАЛИЗА РАДИОКАНАЛА СЕТЕЙ 5G/6G



ГЛАВА 2

МОДЕЛИ КОМПОНЕНТОВ СЕТЕЙ 5G/6G

В этой главе рассматриваются математические модели компонентов сетей связи 5G/6G, работающих в миллиметровом или терагерцовом диапазонах частот. В параграфах 2.1–2.3 представлены модели распространения, модели фазированных антенных решеток и модели блокировки соответственно. В параграфе 2.4 предложена методика анализа параметров канала связи, которая используется в последующих главах при построении моделей обслуживания трафика в сетях 5G/6G.

2.1. Модели распространения

Известно, что полную информацию о распространении электромагнитной волны в пространстве можно получить, решая уравнения Максвелла с соответствующими граничными условиями. Однако такой метод на практике неприменим вследствие сложности среды, в которой распространяются волны. Тем не менее анализ и планирование сотовых сетей связи требуют понимания основных характеристик принимаемого сигнала на некотором расстоянии от БС. Цель моделирования распространения электромагнитных волн состоит в том, чтобы получить оценку уровня сигнала, когда заданы основные параметры системы, такие как частота, характеристики местности, высота расположения передающей и принимающей антенн и др.

В общем виде отношение *сигнал–шум плюс помеха* (ОСП, англ. – Signal-to-Interference plus Noise ratio, SINR) на приемнике определяется формулой

$$S(x) = \frac{P_T G_A G_U}{L(x)(BN_0 + M_I) C_L M_S}, \quad (2.1)$$

где B – полоса пропускания, P_T – излучаемая мощность, G_A и G_U – усиление антенн на приеме и передаче, $L(x)$ – потери при распространении сигнала на расстоянии x , N_0 – тепловой шум, C_L – потери сигнала на передатчике, M_I – помеха, M_S – замирания канала, вызываемые статичными объектами (например, зданиями) в канале связи.

Также иногда учитывают быстрые замирания канала связи, не показанные в (2.1). В указанной формуле примем, что шум N_0 является константой, равной – 174 дБм/Гц. Коэффициент C_L также является константой и зависит от технологии производства приемопередатчика. В общем случае M_S и $L(x)$ моделируются случайными величинами и зависят от местности, где развернута система. Помеха M_I определена как функция плотности абонентов и БС и также моделируется некоторой случайной величиной (см. главу 4). Усиления антенн G_A и G_U зависят от типа антенной решетки, числа антенных элементов и направленности передатчика и приемника относительно друг друга.

2.1.1. Моделирование лучей

Распространение радиоволн миллиметрового и терагерцового диапазонов частот в условиях городской застройки имеет принципиальные отличия от тех частот, на которых работают современные системы связи 4G. Причина заключается в длине волны (10 миллиметров для частоты 30 ГГц и 1 миллиметр для 300 ГГц). Например, в миллиметровом диапазоне частот значительные проблемы для радиосвязи представляют шероховатости бетонных поверхностей зданий. Поскольку размеры таких объектов сопоставимы с длиной волны, вместо отражения происходит диффузионное рассеяние, что значительно ослабляет уровень принимаемого сигнала. Такой эффект также возникает, например, при прохождении радиоволны через кроны деревьев.

Учитывая перечисленные свойства распространения, можно заметить, что распространение волн миллиметрового диапазона частот сильно зависит не только от пространственной конфигурации, будем кратко говорить «геометрии», изучаемой среды. В настоящее время для построения высокоточных моделей распространения в специфических условиях используются модели на основе трассировки лучей (RL, Ray Tracing). Метод трассировки лучей основан на методах

оптики и позволяет решить уравнения Максвелла методом имитационного моделирования. Этот метод является общим инструментом моделирования распространения, который обеспечивает оценки потерь распространения, азимута прихода луча (AoA, Angle of Arrival), азимута исхода луча (AoD, Angle of Departure), зенита прихода луча (ZoA, Zenith of Arrival), зенита исхода луча (ZoD, Zenith of Departure) и задержек сигнала. В отличие от теоретических и эмпирических моделей метод трассировки лучей не позволяет получить простые, с точки зрения вычислений, формулы для потерь распространения.

2.1.2. Эмпирические модели распространения

Модели, основанные на трассировке лучей, не позволяют получить простые выражения для потерь распространения в зависимости от расстояния между передатчиком и приемником. Это свойство не позволяет использовать их для аналитического анализа развертывания беспроводных сетей радиодоступа 5G/6G. Однако такие модели, совместно с измерениями распространения радиоволн, позволяют строить модели на основе эмпирического подхода.

Построение эмпирических моделей основано на поиске функции, описывающей измерения, проведенные в некоторой исследуемой области. Применение таких моделей обосновывается схожестью геометрии области, в которой производится развертывание системы с областью, в которой проведены измерения. Эталонным подходом считается уточнение коэффициентов теоретической модели распространения Фрииса:

$$L_{dB}(x) = 94,5 + 20\log_{10}(x) + 20\log_{10}(f_c),$$

где x – расстояние между передатчиком и приемником, км;
 f – несущая частота, ГГц.

Документ 3GPP TR 38.901 определяет несколько эмпирических моделей для волн миллиметрового диапазона частот. Основное отличие между моделями заключается в рассматриваемом сценарии моделирования. В прикладных исследованиях чаще всего используют модель городского каньона (UMi, Canyon-Street Urban Micro):

$$L_{dB}(x) = 32,4 + 21,0\log_{10}(x) + 20\log_{10}(f_c), \quad (2.2)$$

которая, как видно, существенно отличается от модели Фрииса.

Модель UMi (2.2) предоставляет усредненное значение потерь распространения на некотором расстоянии от передатчика и не учитывает дополнительных эффектов, таких как замирание,

блокировка и т.д. Также отметим, что в исследованиях беспроводных сетей доступа часто используют модель, выражающую потери распространения в линейной шкале:

$$L(x) = Ax^{-\gamma}, \quad (2.3)$$

где принято, что

$$A = 10^{-2 \log_{10} f_c + 3}, \gamma = 2,1. \quad (2.4)$$

2.1.3. Кластерная модель распространения

Сети радиодоступа 5G/6G, работающие в миллиметровом диапазоне частот, будут использовать высокую направленность антенн как на передающей, так и на принимающей стороне. Для установления связи такие системы должны использовать алгоритмы поиска луча. Для этих алгоритмов важны не только средние значения потерь распространения, но и количество путей распространения между рассматриваемыми устройствами и их дополнительные характеристики.

Стандарт 3GPP TR 38.901 определяет более сложную эмпирическую модель распространения, так называемую трехмерную кластерную модель. Согласно этой модели, принимаемая общая мощность состоит из мощностей, поступающих от некоторого количества путей распространения (кластеров), которые включают в себя путь прямой видимости (LoS, Line-of-Sight). Эта модель связывает среду распространения сигнала со следующими параметрами: а) зенит прихода/исхода луча пути распространения на приемнике и передатчике (ZoA/ZoD); б) азимут прихода/исхода луча пути распространения на приемнике и передатчике (AoA/AoD); в) задержка распространения по каждому пути; г) доли мощности, которые вносятся каждая своим путем в общую принимаемую мощность.

Построение трехмерной кластерной модели включает в себя сложные функции случайных величин, что усложняет ее использование для анализа и планирования сетей 5G/6G. Однако известны следующие взаимосвязи: а) зенит прихода ZoA $\theta_i(x)$ пути $i = 2, 3, \dots, N$, (индекс пути прямой видимости $i = 1$ рассматривается отдельно) может быть представлен распределением Лапласа со средним значением, совпадающим с AoA и ZoA, углами LoS и постоянной дисперсией; б) мощность по i -му пути является случайной величиной (с.в.) распределенной логнормально, т.е.

$$f_{\theta_i}(y; x) = \frac{1}{2b_i(x)} e^{-\frac{|y-a_i(x)|}{b_i(x)}}, i = 2, 3, \dots, \quad (2.5)$$

$$f_{P_{s,i}}(y; x) = \frac{1}{y d_i \sqrt{2\pi}} e^{-\frac{(\ln y - c_i)^2}{2a_i^2}}, i = 2, 3, \dots,$$

где $a_i(x)$, $b_i(x)$, c_i , d_i , $i = 2, 3, \dots, N$ – параметры логнормального распределения. Кроме того, известно, что $a_i(x)$ не зависит от рассматриваемого пути i , но зависит от расстояния x между передатчиком и приемником. Среднее значение ZoA для всех путей совпадает с постоянным значением ZoA пути прямой видимости. В свою очередь, $a_i(x)$, $b_i(x)$ и $c_i(x)$ зависят не от расстояния, а только от номера пути распространения.

2.2. Модели фазированных антенных решеток

Возвращаясь к выражению (2.1), определяющему ОСШ, заметим, что неизвестными в нем являются также коэффициенты усиления антенн на передающей и принимающей сторонах. Для компенсации высоких потерь распространения в миллиметровых сетях радиодоступа 5G/6G будут использованы фазированные антенные решетки (ФАР) с большим усилением в обоих направлениях. Далее в этой главе рассмотрим модели ФАР, которые позволяют определять усиления передающей и принимающей антенн как функцию количества антенных элементов и соответствующий им угол луча основного лепестка антенны.

2.2.1. Трехмерная модель

Рассмотрим модель антенны, в которой диаграммы направленности как в направлении передачи, так и в направлении приема имеют одинаковую форму, но различные вертикальные и горизонтальные направленности. Предположим, что основной лепесток антенны можно представить пирамидальной зоной, которая определяется вертикальными и горизонтальными углами, α_V и α_H (рис. 2.1). Эта модель является теоретической абстракцией диаграмм направленности. Предположим, что имеется некоторое постоянное усиления основного лепестка антенны как в направлении приема, так и в направлении передачи. Хотя модель не учитывает потери мощности на боковые и задний лепестки, этот эффект может быть учтен при помощи коэффициента распределения мощности.

Для того чтобы определить коэффициент усиления антенны G , соответствующий значениям α_V и α_H , заметим, что площадь поверхности волнового фронта равна площади сферического прямоугольника, показанного на рис. 2.1. Используя сферическую теорему косинусов, представим $\cos \chi$ следующим образом:

$$\begin{aligned} \cos \chi &= \frac{\cos\left(\frac{\pi}{2} - \frac{L_H}{2}\right) - \cos\left(\frac{\pi}{2} - \frac{L_H}{2}\right) \cos(L_V)}{\sin\left(\frac{\pi}{2} - \frac{L_H}{2}\right) \sin(L_V)} = \\ &= \frac{\sin\left(\frac{L_H}{2}\right) (1 - \cos(L_V))}{\cos\left(\frac{L_H}{2}\right) \sin(L_V)} = \tan\left(\frac{L_H}{2}\right) \tan\left(\frac{L_V}{2}\right). \end{aligned} \quad (2.6)$$

При рассмотрении сферического прямоугольника на рис. 2.1 можно заметить, что

$$\cos\left(\rho - \frac{\pi}{2}\right) = \tan\left(\frac{L_H}{2}\right) \tan\left(\frac{L_V}{2}\right), \quad (2.7)$$

где L_H и L_V – сферические геодезические линии.

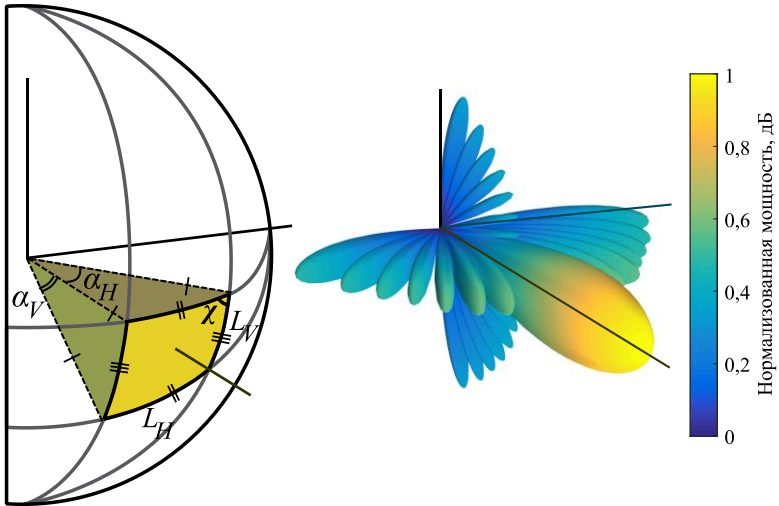


Рис. 2.1. Трехмерная модель диаграммы направленности

Сферические геодезические линии L_H и L_V соответствуют углам направленности α_H и α_V . Следовательно, площадь сферического прямоугольника имеет вид

$$S_A = 4 \arcsin\left(\tan\frac{\alpha_V}{2} \tan\frac{\alpha_H}{2}\right). \quad (2.8)$$

Коэффициент усиления антенны G , соответствующий углам α_V и α_H , имеет вид

$$G(\alpha_V, \alpha_H) = \frac{4\pi}{S_A} = \frac{\pi}{\arcsin\left(\tan\frac{\alpha_V}{2}\tan\frac{\alpha_H}{2}\right)}. \quad (2.9)$$

2.2.2. Упрощенная двумерная модель

Частным случаем трехмерной модели является двумерная модель, представляющая собой главный лепесток антенны в виде треугольника, как проиллюстрировано на рис. 2.2. Как и в случае трехмерной модели, для того чтобы полностью задать данную модель, необходимо определить усиление антенны G основного лепестка как функцию угла направленности α .

Определим спектральную плотность мощности на расстоянии r по формуле

$$\mathcal{P}_{rx} = \frac{P_{Tx}}{S_A} = \frac{P_{Tx}}{2\pi r h'}, \quad (2.10)$$

где S_A – площадь поверхности волнового фронта, определяемая площадью поверхности сферического сектора, $h' = r[1 - \cos(\alpha/2)]$ и α – углом направленности антенны.

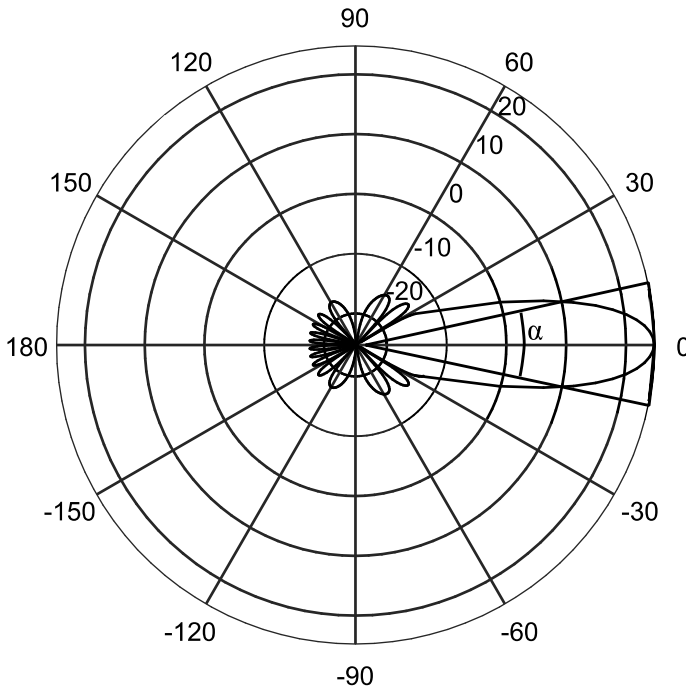


Рис. 2.2. Двухмерная модель ФАР

В соответствии с моделью Фрииса распространения сигнала в открытом пространстве (2.2) плотность мощности на расстоянии r можно представить в виде

$$\mathcal{P}_{Rx} = \frac{P_{Tx}}{S_A} = P_{Tx} \frac{G}{4\pi r^2}, \quad (2.11)$$

откуда следует, что усиление антенны G для основного лепестка в упрощенной двухмерной модели может быть определено следующим образом:

$$G = \frac{2}{1 - \cos(\alpha/2)}. \quad (2.12)$$

Заметим, что при $\alpha = 2\pi$ для идеальной всенаправленной антенны (изотропный излучатель) усиление равно единице ($G = 1$), а $S_A = 4\pi r^2$.

2.2.3. Связь с параметрами антенных решеток

Введенные выше как трехмерная, так и упрощенная двухмерная модели антенных решеток полностью определяются углом направленности главного лепестка антенны. Однако на практике этот угол является функцией количества антенных элементов, составляющих ФАР. Следовательно, нам необходимо установить связь между количеством элементов в ФАР и углом направленности основного лепестка антенны. Проиллюстрируем это с помощью антенной решетки на рис. 2.3.

Излучение антенной решетки (AF, Array Factor) может быть представлено в виде

$$AF(\theta, \beta) = \frac{\sin(N[\pi \cos(\theta) + \beta]/2)}{\sin([\pi \cos(\theta) + \beta]/2)}, \quad (2.13)$$

где N – число элементов в ФАР; β – направление основного лепестка; θ – азимут направленности основного лепестка. Примем для простоты $\beta = 0$ и определим расстояние между соседними элементами величиной $\lambda/2$, где λ – длина волны.

Определим теперь ширину луча основного лепестка антенны с помощью ширины луча половинной мощности (HPBW, Half-Power Beamwidth) антенной решетки. Напомним, что последний параметр определяет угол, при котором мощность излучения падает в два раза (т.е. на 3 дБ). Угол HPBW пропорционален количеству элементов в соответствующей плоскости и рассчитывается по формуле

$$\alpha = 2|\theta_m - \theta_{3db}|, \quad (2.14)$$

где θ_{3db} – точка, соответствующая уменьшению мощности на 3 дБ; θ_m – положение максимума антенной решетки. Последний параметр определяется следующим образом:

$$\theta_m = \arccos(-\beta/\pi). \quad (2.15)$$

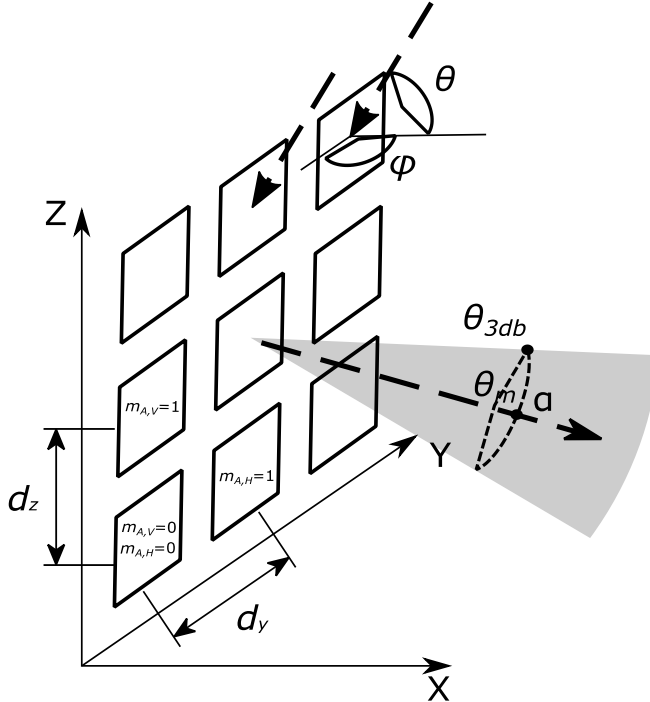


Рис. 2.3. Иллюстрация параметров антенной решетки

Предполагая $\beta = 0$, видим, что $\theta_m = \pi/2$. Следовательно, верхняя и нижняя точки, соответствующие уменьшению мощности на 3 дБ, определяются формулой

$$\theta_{3db}^{\pm} = \arccos[-\beta \pm 2,782/(N\pi)]. \quad (2.16)$$

Для $\beta = 0$ среднее усиление антенны по НРВW имеет вид

$$G = \frac{1}{\theta_{3db}^+ - \theta_{3db}^-} \int_{\theta_{3db}^-}^{\theta_{3db}^+} \frac{\sin(N\pi \cos(\theta)/2)}{\sin(\pi \cos(\theta)/2)} d\theta. \quad (2.17)$$

Значения НРВW, рассчитанные по формуле (2.15), представлены в табл. 2.1. Соответствующие им значения усиления ФАР в линейной и логарифмической шкалах приведены в табл. 2.2. Отметим, что достаточно точная аппроксимация для НРВW антенной решетки

может быть получена, используя соотношение $102^\circ/N$, где N – число антенных элементов. Грубое приближение усиления ФАР в линейной шкале может быть получено, если рассмотреть только число антенных элементов.

Таблица 2.1

Угол НРВW и его аппроксимация

Решетка	Точный расчет	Аппроксимация
64×1	1.585	1.594
32×1	3.171	3.188
16×1	6.345	6.375
8×1	12.71	12.75

Таблица.2.2

Усиление антенны для основного лепестка

Решетка	Усиление	Усиление, дБ
64×1	57.51	17.59
32×1	28.76	14.58
16×1	14.38	11.57
8×1	7.20	8.57
4×1	3.61	5.57

2.2.4. Применение моделей антенн

В зависимости от параметров системы можно использовать различные модели диаграммы направленности антенны. Например, если число элементов, формирующих диаграмму в вертикальной плоскости достаточно мало (что ожидается на ранних этапах развертывания сетей 5G/6G), а расстояние взаимодействия между устройствами велико, двухмерная модель, отличающаяся достаточно простым набором параметров, позволяет достаточно точно учитывать усиления ФАР. В других случаях, когда число антенных элементов в вертикальной плоскости достаточно велико, рекомендуется к использованию трехмерная модель. Следует подчеркнуть, что диаграмма направленности приемной антенны моделируется аналогично.

Обратим внимание, что рассмотренные модели не учитывают потери мощности на боковые и задние лепестки. Однако этот эффект достаточно просто учесть как в трехмерной, так и в двухмерной модели, предположив, что потерянная мощность равномерно рассеивается во все стороны, кроме главного лепестка. Одним из вариантов такой модели является так называемая модель «конус плюс сфера».

2.3. Модели блокировки

Блокировка путей распространения сигнала является особенностью беспроводных сетей доступа миллиметрового и терагерцового диапазонов частот. Результатом блокировки является резкое падение принимаемой мощности, что может приводить как к отсутствию связи на некоторый промежуток времени, так и резкому увеличению количества ресурсов, необходимых для поддержания заданной скорости. Натурные эксперименты показывают, что блокировка человеческим телом приводит к дополнительным потерям в диапазоне 20–40 дБ. Эксперименты, проведенные на пешеходной улице, подтвердили ослабление сигнала из-за блокировки человеческим телом до 25 дБ с длительностью периода блокировки до 200 мс. Отметим, что блокировка распространения миллиметрового сигнала автомобилями в среднем приводит к большему ухудшению качества связи.

Различают два основных типа блокировки. Первый – вызывается большими стационарными объектами, например зданиями. Этот тип блокировки не специфичен миллиметровому диапазону частот, и он также оказывает влияние на качество каналов связи в сотовых системах предыдущих поколений, работающих в диапазонах частот ниже 6 ГГц. Второй тип, который ранее не оказывал негативного влияния на качество каналов связи, вызван малыми объектами, размеры которых сравнимы или лишь в несколько раз больше, чем длина волны. Этот тип блокировки может быть вызван как динамическими объектами, людьми, автомобилями, так и статическими, в частности фонарными столбами, деревьями, зданиями и т.д. Частным случаем является самоблокировка, вызываемая поворотом пользователя вокруг своей оси, что временно преграждает путь прямой видимости.

В моделях оценки характеристик обслуживания абонентов в сетях радиодоступа 5G/6G для учета эффектов блокировки путей распространения сигнала могут быть использованы методы имитационного моделирования. Однако такой подход не является масштабируемым на реальные системы связи, в которых могут существовать десятки и сотни блокаторов. Перспективным методом является математическое моделирование блокировки как функции характеристик рассматриваемого сценария связи и плотности абонентов. В этом случае блокировка представляется вероятностью (в статическом случае) или же случайным процессом (если блокаторы или пользователь являются мобильными). В дальнейшем эта вероятность может быть

использована для построения математических моделей процесса обслуживания абонентов.

2.3.1. Статический случай

Рассмотрим статический случай, в котором как АУ, так и блокираторы являются стационарными объектами. В этом сценарии одна БС и одно АУ находятся на расстоянии r_0 друг от друга, а их высоты равны h_T и h_R соответственно.

Размещение статических блокираторов моделируется пуассоновским точечным процессом (PPP, Poisson Point Process) с плотностью λ_B $1/m^2$. Блокаторами являются человеческие тела, представляемые цилиндрами с высотой h_B и диаметром основания d_B . Высота h_B обычно выбирается равной среднему росту человека, т.е. $h_B = 1,7$ м.

Определим зону блокировки LoS в виде прямоугольника со сторонами d_B и r , как показано на рис. 2.4.

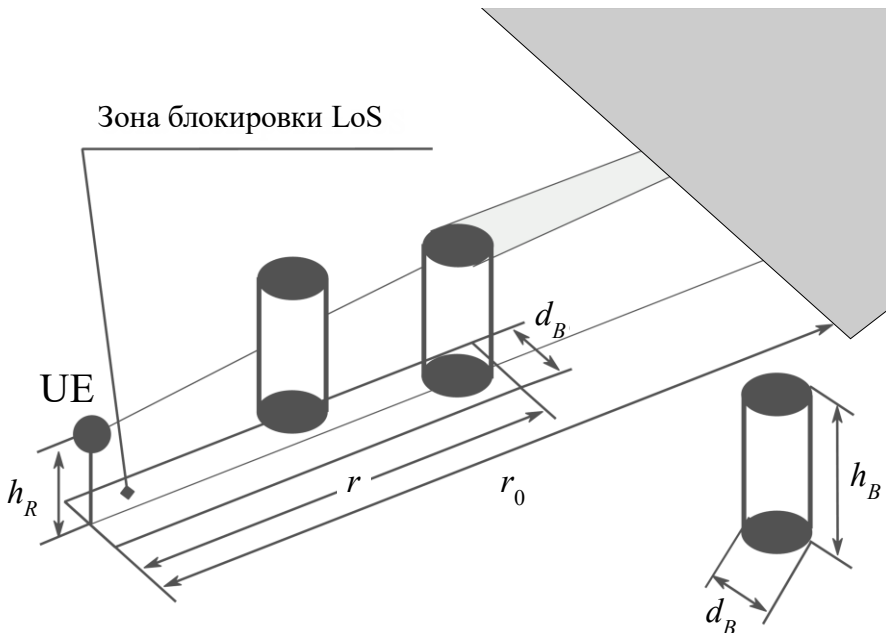


Рис. 2.4. Иллюстрация зоны блокировки LoS

Наличие хотя бы одного блокиатора в зоне вызывает блокировку сигнала, если высота блокиатора выше, чем линия LoS в данной точке. Используя свойства PPP, можно получить вероятность отсутствия блокировки p_{nB} как вероятность отсутствия блокираторов в зоне,

что определяется с учетом расположения центров цилиндров по формуле

$$p_{nB} = e^{-\lambda_B d_B \left[r_0 \frac{h_B - h_R}{h_T - h_R} + \frac{d_B}{2} \right]}, \quad (2.18)$$

где $d_B \left[r_0 \frac{h_B - h_R}{h_T - h_R} + \frac{d_B}{2} \right]$ – площадь зоны блокировки.

Из (2.18) видим, что вероятность блокировки повышается экспоненциально с ростом плотности λ_B , т.е. плотности блокаторов в зоне блокировки. Кроме того, на вероятность блокировки оказывают значительное влияние высоты АУ, БС и блокаторов. В том числе чем больше разница между высотами БС и АУ, тем меньше зона блокировки LoS, и, следовательно, тем меньше вероятность того, что путь LoS будет заблокирован. Отметим, что в частном случае прямого взаимодействия устройств (D2D, Device-to-Device, см. главу 3 и 10) длина пути LoS совпадает с расстоянием между устройствами.

2.3.2. Подвижность пользователя

Перейдем к случаю, когда блокаторы движутся, а АУ остается статичным (рис. 2.5). В данном случае, искомыми характеристиками являются распределение с.в. времени, проведенного АУ в состоянии блокировки, а также зависимость между состояниями АУ в различные моменты времени. Предположим, что БС расположена в точке P , начальной точкой движения АУ считается точка O , а конечной точкой перемещения АУ является точка M . Следует отметить, что O и M могут быть не только точками отсчета времени, когда одно АУ перемещается из точки O в точку M , но и некоторыми точками расположения двух АУ в пространстве, в частности когда два соседних приемника расположены в точках O и M . В обоих случаях задача состоит в том, чтобы найти вероятность p_{ij} того, что АУ находится в состоянии j в точке M при условии, что в точке O АУ находилось в состоянии i , где состояния 0 и 1 соответствуют незаблокированному и заблокированному состояниям соответственно.

В рассматриваемом сценарии возможны события в точке M – АУ заблокирован и АУ не заблокирован, обусловленные состоянием АУ в точке O . Заметим, что $p_{00} = 1 - p_{01}$ и $p_{10} = 1 - p_{11}$. Очевидно, что должно выполняться следующее соотношение:

$$p_{nB,M} = p_{00}p_{nB,O} + p_{10}p_{B,O}. \quad (2.19)$$

Поэтому, чтобы получить p_{ij} , достаточно найти p_{00} .

Для определения искомых параметров две области блокировки LoS для точек O и M необходимо разделить на дополнительные зоны. Как и ранее, ширина зоны блокировки равна диаметру блокатора d_m , а длина равна расстоянию между P и O (или M). Как показано на рис. 2.5, следует определить четыре основные зоны, при этом первую из них (зона 1) можно исключить из рассмотрения из-за ее относительно небольших размеров.

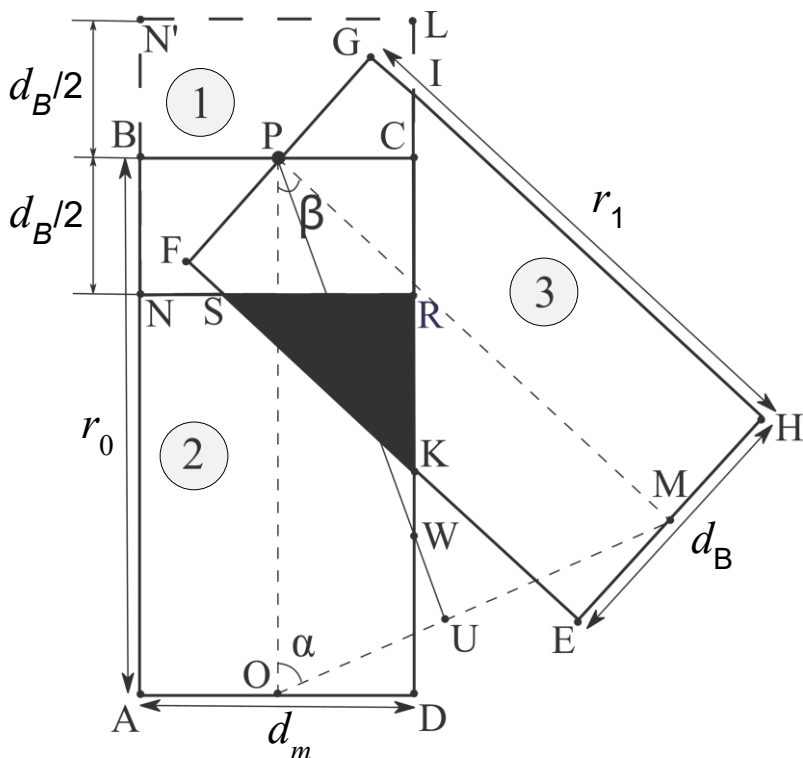


Рис. 2.5. Геометрическая интерпретация блокировки прямой видимости при движении АУ

Зоны 2 и 3 влияют на состояние АУ в точках O и M соответственно. В частности, вероятность p_{00} может быть определена как условная вероятность

$$p_{00} = P[\text{нет блок. в } M | \text{нет блок. в } O] = \frac{P[\text{нет блок. в } M \cap \text{нет блок. в } O]}{P[\text{нет блок. в } O]}, \quad (2.20)$$

где компоненты могут быть вычислены из геометрических соображений, показанных на рис. 2.5.

На рис. 2.6–2.8 показана зависимость вероятности блокировки в момент t_1 от высоты БС, высоты АУ и скорости перемещения АУ, при условии, что в момент t_0 АУ находится в состоянии LoS.

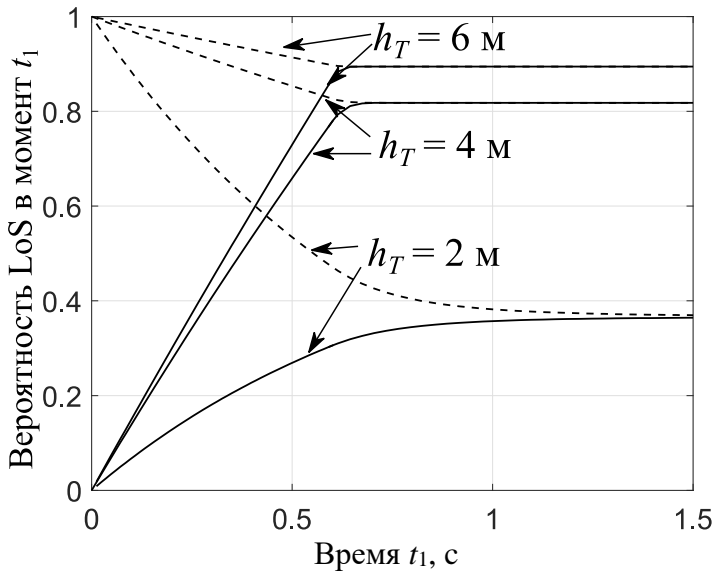


Рис. 2.6. Влияние высоты БС на вероятность блокировки

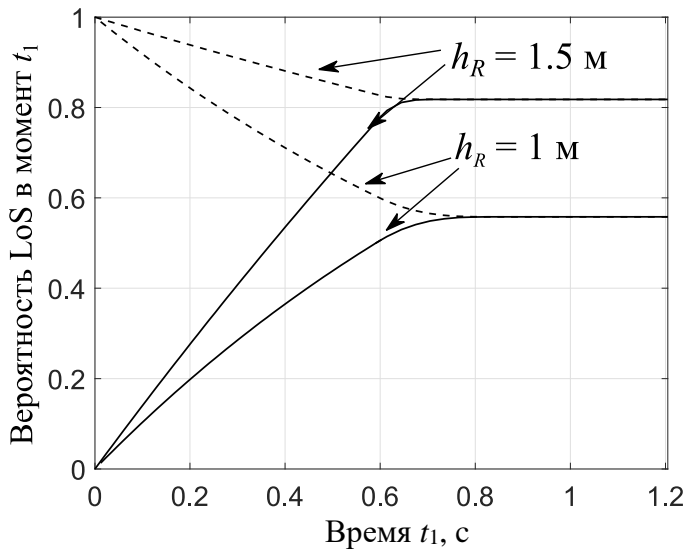


Рис. 2.7. Влияние высоты АУ на вероятность блокировки

Отметим, что в зависимости от исследуемых параметров влияние подвижности сохраняется в течение как минимум 0,5 секунды, но может продолжаться и более 1 секунды. Такое поведение значительно

влияет на планирование передач от активных АУ на стороне БС. Таким образом, по имеющимся данным о продолжительности времени блокировки планировщик может применить оптимальную схему управления.

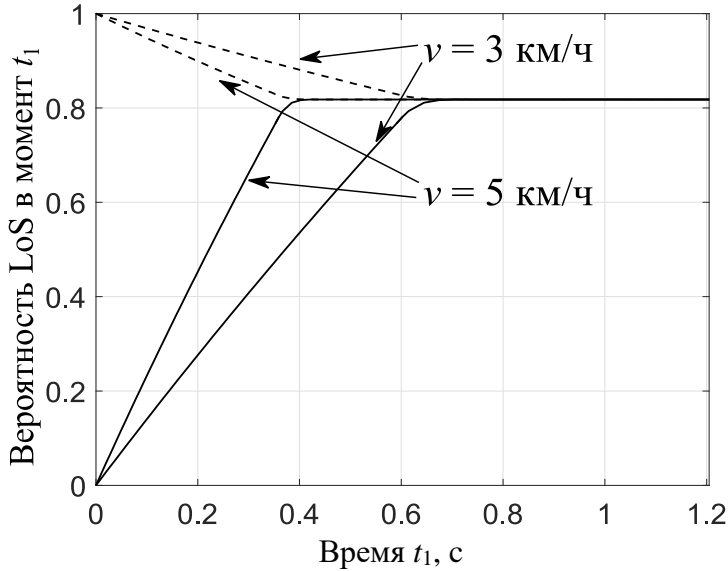


Рис. 2.8. Влияние скорости движения пользователя на вероятность блокировки

2.3.3. Движущиеся блокаторы

Проанализируем наиболее типичный случай, когда АУ, находящееся в статичном положении, окружено движущимися блокаторами. В этом сценарии основное внимание необходимо уделить изменению состояния блокировки во времени. Для описания движения случайного блуждания используем модель случайного перемещения (RDM, Random Direction Model). Согласно модели RDM, объект начинает движение в случайно выбранном месте на плоскости. Объект движется в выбранном направлении в течение экспоненциально распределенного времени с заданными средним значением и постоянной скоростью.

Аналогично предыдущему случаю, помимо времени, проводимого АУ в заблокированном и незаблокированном состояниях, представляет интерес вероятность блокировки в момент t_1 , при условии, что АУ находится в заблокированном или незаблокированном

состоянии в момент t_0 . Временная эволюция состояния АУ в виде альтернирующего процесса показана на рис. 2.9, где ξ_i – совместные длительности заблокированного и незаблокированного интервалов; ω_i – длительности незаблокированных интервалов; η_i – длительности заблокированных интервалов. Здесь предполагается, что моменты входа блокаторов в зону блокировки LoS образуют пуассоновский поток. Время, которое блокатор проводит в зоне блокировки LoS, может быть задано постоянной величиной либо получено имитационным моделированием или аналитически, исходя из свойств используемой модели движения. Проводя аналогию с ТМО, заметим, что процесс входа и выхода блокаторов в/из зоны блокировки LoS аналогичен процессу обслуживания заявок в СМО типа $M/G/\infty$, где время обслуживания соответствует продолжительности блокировки LoS одним блокатором.

Метод вычисления вероятности блокировки за время Δt основан на соотношениях

$$p_{00}(\Delta t) = \sum_{i=0}^{\infty} P \{A_i(\Delta t)\}, p_{01}(\Delta t) = \sum_{i=1}^{\infty} P \{B_i(\Delta t)\}, \quad (2.21)$$

где $A_i(\Delta t)$ – событие, когда в начальный и в конечный моменты времени АУ находится в незаблокированном состоянии с учетом того, что в течение интервала $\Delta t = t_1 - t_0$, происходит ровно $i = 0, 1, \dots$ блокировок.

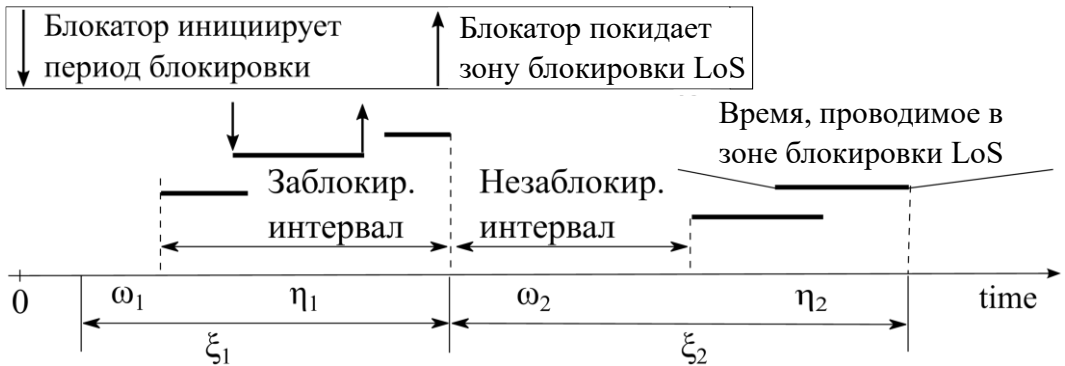


Рис. 2.9. Альтернирующий процесс восстановления

На рис. 2.10 для трех сценариев показана функция распределения (ФР) с.в. остаточного времени в заблокированном состоянии для АУ, находящегося в движущейся толпе блокаторов. Следует отметить, что отличия ФР, построенных для разных сценариев и плотностей блокаторов, незначительны.

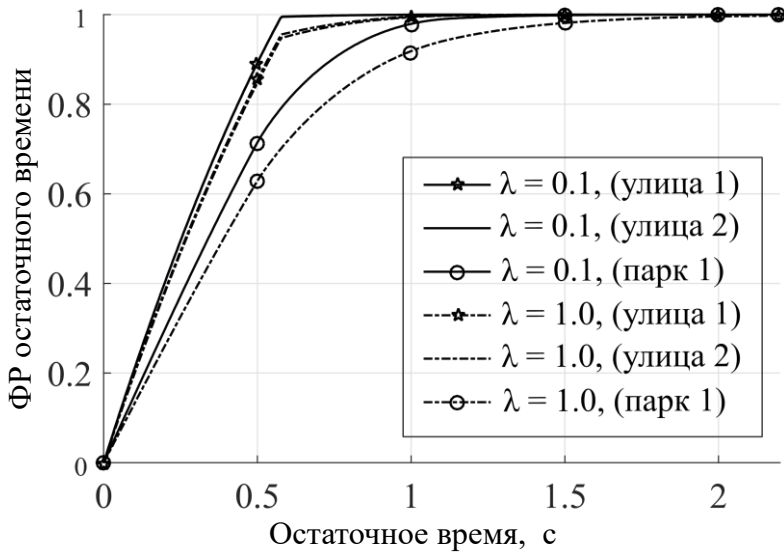


Рис. 2.10. ФР остаточного времени в заблокированном состоянии

2.4. Модели параметров канала связи

Определив модели компонентов системы связи, можно перейти к вычислению следующих параметров: а) радиус d_E области покрытия БС; б) ФР $F_R(x)$ объема ресурсов, запрашиваемых на БС сеансом связи; в) интенсивность α изменения состояния АУ с незаблокированного на заблокированное и наоборот. Далее изложены методы, позволяющие определить искомые параметры модели. Используемые в модели обозначения приведены в табл. 2.3.

Таблица 2.3

Параметры сети радиодоступа 5G/6G	
Обозначение	Определение
f_c	частота
d_E	радиус покрытия БС
h_A, h_U	высота БС и АУ
λ_B	пространственная плотность блокаторов
h_B, r_B, v	высота, радиус и скорость движения блокатора
τ	время движения блокатора
ζ	константа распространения
L_B	потери блокировки
$L(x), L_{dB}(x)$	потери распространения
A_1, A_2	константы распространения
$\theta_m, \theta_{3dB}^\pm, \beta$	параметры антенных решеток
ω	НРВW луча
P_A	излучаемая мощность БС

Обозначение	Определение
$M_{S,B}$	граница затухания в заблокированном состоянии
p_c	вероятность отсутствия связи на границе зоны
$\sigma_{S,B}, \sigma_{S,nB}$	среднеквадратическое отклонение (СКО) границ затухания
S_{min}	чувствительность приемника
S, S^{dB}	взвешенное ОСШ
S_B, S_{nB}	ОСШ в заблокированном и незаблокированном состояниях в линейной шкале
S_B^{dB}, S_{nB}^{dB}	ОСШ в заблокированном и незаблокированном состояниях в шкале децибел
C_L	потери на передатчике
M_I	граница помехи
N_F	порог шума
χ_1, χ_2	коэффициенты моделей распространения
$\pi_B(x), \pi_B$	вероятность блокировки
s_j	границы ОСШ для модуляционно-кодовой схемы (MCS, Modulation and Coding Scheme)
m_j	вероятность использования MCS j
$g(x, y)$	двумерная плотность расположения блокатора
$\eta_i(x, y)$	вероятность, что блокатор пересекает зону блокировки

2.4.1. Блокировки в моделях распространения

Проиллюстрируем блокировки в модели распространения, используя эмпирическую модель, определенную 3GPP в стандарте TR 38.901. Пусть L_B – затухание в состоянии блокировки. Тогда модель распространения можно записать в виде

$$L_{dB}(x) = \begin{cases} 32,4 + 21 \cdot \log(x) + 20 \log f_c, & \text{неблок,} \\ 47,4 + 21 \cdot \log(x) + 20 \log f_c, & \text{блок,} \end{cases} \quad (2.22)$$

где f_c – несущая частота (ГГц); x – расстояние между БС и АУ, км.

Напомним, что потери распространения (2.22) можно представить в линейном виде, используя модель $A_i y^{-\zeta_i}$, где A_i и ζ_i – коэффициенты распространения, зависящие от частоты и рассматриваемого сценария применения модели. Вводя пары коэффициентов (A_1, ζ_1) и (A_2, ζ_2) , соответствующие незаблокированному и заблокированному состояниям АУ, получаем

$$\begin{aligned} A_1 &= 10^{2 \log_{10} f_c + 3,24} M_{S,nB} M_I, \zeta_1 = 2,1, \\ A_2 &= 10^{2 \log_{10} f_c + 4,74} M_{S,B} M_I, \zeta_2 = 2,1. \end{aligned} \quad (2.23)$$

Пусть $p_B(x)$ – вероятность блокировки АУ на расстоянии x от БС. Заметим, что вероятность $p_B(x)$ не ограничена сценарием, где АУ и блокаторы являются статическими, и может определяться временем, проводимым АУ в состоянии блокировки на расстоянии x от БС в поле движущихся блокаторов. Обозначим через y проекцию пути распространения сигнала на плоскость. Значение ОСШ на АУ с учетом (2.1) может быть записано в виде

$$S(y) = \frac{P_A G_A G_U}{N_0} \left[\frac{\chi_1 y^{-\zeta} [1 - \pi_B(y)]}{A_1} + \frac{\chi_2 y^{-\zeta} \pi_B(y)}{A_2} \right], \quad (2.24)$$

где χ_1 и χ_2 – с.в., определяющие теневое замирание (SF, Shadow fading) в заблокированном и незаблокированном состояниях соответственно.

2.4.2. Область покрытия точкой доступа

Чтобы БС обеспечивала покрытие некоторой зоны, расстояние d_E должно быть таким, чтобы на АУ, которые находятся в условиях блокировки на границе соты, не происходило потери связи. Пусть S_{min} – чувствительность приемника, т.е. нижняя граница ОСШ, соответствующего самой низкой MCS. Используя усредненную модель распространения, соответствующую состоянию блокировки LoS, получаем соотношение

$$S_{min} = \frac{P_A G_A G_U}{N_0 A_2} (d_E + [h_A - h_U]^2)^{-\zeta/2}, \quad (2.25)$$

где ζ – показатель потерь; h_A и h_U – высоты БС и АУ; P_A – излучаемая мощность БС; G_A и G_U – усиление антенных решеток на БС и АУ; N_0 – тепловой шум на приемнике; A_2 – постоянная распространения.

Решая (2.25) относительно d_E , получаем

$$d_E = \sqrt{\left(\frac{P_A G_A G_U}{N_0 A_2 S_{min} M_{S,B}} \right)^{\zeta/2} + (h_A - h_U)^2}, \quad (2.26)$$

где $M_{S,B}$ – граница теневого замирания в состоянии блокировки, учитывающего динамику сигнала, вызываемого большими объектами определяется по формуле

$$M_{S,B} = \sqrt{2} \sigma_{S,B} \operatorname{erfc}^{-1}(2p_C), \quad (2.27)$$

где $\operatorname{erfc}^{-1}(\cdot)$ – обратная дополнительная функция ошибок (обратная функция Лапласа); p_C – вероятность наличия связи на границе соты;

$\sigma_{S,B}$ – среднее квадратичное отклонение распределения теневого зами- рания для заблокированного состояния LoS.

Таким образом, процедура определения максимального радиуса покрытия соты сводится к следующему. На первом этапе задается вероятность наличия связи на границе соты (например, 0,95) и значение потерь, вызываемых блокировкой. Далее, используя эти значения, проводится расчет по формуле (2.26). На практике размеры соты могут ограничивать другие факторы, определенные в 3GPP TR 38.901.

2.4.3. Распределение требований к ресурсам

Следующим этапом является нахождение ФР запроса на с.в. ресурсов сеанса связи. Пусть S_{nB} – с.в. ОСШ в незаблокированном состоянии с ФР $F_{S_{nB}}(x)$. Предполагается, что местоположения АУ, генерирующих новые сессии, распределены равномерно в пределах зоны обслуживания БС. Тогда ФР расстояния между АУ и БС имеет вид

$$F_D(y) = (y^2 - (h_A - h_U)^2) / d_E^2, \\ |h_A - h_U| < y < \sqrt{d_E^2 + (h_A - h_U)^2}. \quad (2.28)$$

ФР с.в. ОСШ может быть определена с помощью ФР с.в. расстояния до БС следующим образом:

$$F_{S_{nB}}(y) = 1 - F_D(P_A G_A G_U / N_0 A_1 y^{\zeta/2}). \quad (2.29)$$

ФР с.в. S_B ОСШ в заблокированном состоянии определяется ана- логично. Для определения взвешенной ФР ОСШ требуется усреднен- ная по расстоянию вероятность блокировки p_B , и тогда, используя (2.18), получаем

$$p_B = \int_0^{d_E} p_B(x) \frac{2x}{d_E^2} dx. \quad (2.30)$$

Дополним модель эффектом теневого замирования. Заметим, что расчеты значительно упрощаются, если их проводить в логарифми- ческой шкале. Определим ФР с.в. ОСШ, измеряемого в дБ, в виде $F_{S_{nB}}^{dB}(y) = F_{S_{nB}}(10^{y/10})$. Напомним, что замирение характеризуется логнормальным распределением в линейной шкале, что соответ- ствует нормальному распределению в логарифмической шкале. Запишем с.в. ОСШ в виде

$$S_{nB,S}^{dB} = S_{nB}^{dB} + N(0, \sigma_{S,nB}), \quad (2.31)$$

где $N(0, \sigma_{S,nB})$ – нормально распределенная с.в. с нулевым средним и среднеквадратическим отклонением $\sigma_{S,nB}$, характеризующая теневое замирание в канале связи.

ФР ОСШ, учитывающую как потери распространения, так и замирания, можно определить как свертку с.в. S_{nB}^{dB} и $N(0, \sigma_S)$ по формуле

$$F_{S_{nB,S}}^{dB}(y) = \int_{-\infty}^{\infty} F_{S_{nB}}^{dB}(y+u) \frac{e^{-u^2/2\sigma_S^2}}{\sqrt{2\pi}\sigma_S} du. \quad (2.32)$$

Для упрощения вычислений по формуле (2.32), используя функцию ошибок $\text{erf}(\cdot)$, можно записать ФР ОСШ в виде

$$\begin{aligned} F_{S_{nB}}^{dB}(y) = & \frac{1}{2d_E^2} [A^{2/\zeta} 10^{-\frac{y}{5\zeta}} e^{\frac{\sigma_{S,nB}^2 \log^2(10)}{50\zeta^2}} \times \\ & \times [\text{erf}\left(\frac{50\zeta \log(A) - 25\zeta^2 \log(B) + \sigma_{S,nB}^2 \log^2(10) - 5\zeta y \log(10)}{5\sqrt{2}\zeta \sigma_{S,nB} \log(10)}\right) - \\ & - B \cdot \text{erf}\left(\frac{-10\log(A) + 5\zeta \log(B) + y \log(10)}{\sqrt{2}\sigma_{S,nB} \log(10)}\right) - (h_A - h_U)^2 \times \\ & \times \text{erf}\left(\frac{\sqrt{2}(-10\log(A) + 10\zeta \log(h_A - h_U) + y \log(10))}{\sigma_{S,nB} \log(100)}\right) + d_E^2], \end{aligned} \quad (2.33)$$

где C_L – потери сигнала на передатчике; M_I – помехи; N_F – тепловой шум. В (2.33) использованы также следующие обозначения:

$$\begin{aligned} A = & \frac{P_A 10^{G_A G_U / 10}}{f_c^2 N_0 10^{3.24 + L_B / 10} 10^{\frac{1}{10}(C_O + C_L + M_I + N_F)}}, \\ B = & d_E^2 + (h_A - h_U)^2. \end{aligned} \quad (2.34)$$

ФР $F_{S^{dB}}(y)$ можно определить взвешиванием ФР, соответствующих заблокированным и незаблокированным состояниям с вероятностями p_B и $(1 - p_B)$. Обозначим $s_j, j = 1, 2, \dots, K$, границы ОСШ, где K – число схем MCS. Обозначим также q_i вероятность того, что сеансу связи АУ назначена схема MCS i , и получаем, что

$$q_i = \text{Pr}\{s_i < s < s_{i+1}\} = F_{S^{dB}}(s_{i+1}) - F_{S^{dB}}(s_i). \quad (2.35)$$

Определив вероятности q_i , можно найти объем ресурсов, необходимых для проведения сеанса связи. Отметим, что для фиксированной скорости сеанса связи объем требуемых ресурсов получается масштабированием закона распределения $q_i, i = 1, 2, \dots, K$. Для

случайного распределения скорости ФР необходимого для сеанса объема ресурсов может быть получена с помощью свертки соответствующих им распределений.

2.4.4. Интенсивность прерываний

Важным параметром, подлежащим определению, является интенсивность ζ изменений состояния АУ, вызванных блокировкой путей прямой видимости. Следует подчеркнуть, что метод имеет существенные отличия для различных сценариев блокировки, рассмотренных ранее. Далее ограничимся рассмотрением случая со статичным АУ и движущимися блокаторами.

Для того чтобы определить интенсивность изменений состояния АУ, необходимо охарактеризовать временную динамику блокировки. Для этого необходимо выполнить следующие шаги: а) определить процесс поступления блокаторов в зону блокировки LoS; б) определить распределение с.в. времени блокировки, вызванного блокатором; в) определить интервалы прямой видимости и блокировки. Последние два параметра позволяют найти среднее значение интенсивности изменения состояний АУ.

Предположим, что поток блокаторов, входящих в зону блокировки, является пуассоновским. Обозначим $\alpha(x)$ интенсивность входа блокаторов в зону блокировки, связанную с АУ, расположенным на расстоянии x от БС. Для этого определим область вокруг зоны блокировки, как показано на рис. 2.11, где движущиеся блокаторы могут пересекать зону блокировки, перекрывая LoS между АУ и БС. Разделим область вокруг зоны блокировки на $i = 1, 2, \dots, 7$ зон, как показано на рис. 2.11. Тогда интенсивность блокаторов, пересекающих зону блокировки АУ, расположенного на расстоянии z от БС, может быть определена по формуле

$$\zeta(z) = \sum_{i=1}^7 \iint_{M_i} g_i(x, y) Pr\{E\} Pr\{T > 1\} \lambda_B M_i dx dy, \quad (2.36)$$

где M_i – площадь зоны i ; $g_i(x, y)$ – плотность вероятности местоположения блокировщика в зоне i ; $g_i(x, y) = 1/M_i$, $Pr\{T > 1\} = \exp(-1/\tau)$ – вероятность того, что блокатор движется дольше единицы времени, не меняя своего направления.

Вероятность того, что блокатор движется в направлении зоны блокировки, имеет вид $Pr\{E\} = \eta_i(x, y)/2\pi$, где $\eta_i(x, y)$ – диапазон углов движения в зоне i , который приводит к входу блокатора в зону блокировки, см. рис. 2.11. Тогда можно упростить формулу (2.36):

$$\zeta(z) = \frac{\lambda e^{-1/\tau}}{2\pi} \sum_{i=1}^7 \iint_{M_i} \eta_i(x, y) dx dy, \quad (2.37)$$

где величина $\eta_i(x, y)$ определяется следующим образом:

$$\begin{aligned} \eta_i(x, y) &= \left(\frac{[x_D - x]}{vt} \right), i = 1, 3, 5, 7, \\ \eta_i(x, y) &= 2 \cos^{-1} \left(\frac{[x_E - x]}{vt} \right), i = 2, 6, \\ \eta_4(x, y) &= 2 \tan^{-1} \left(\frac{[x - x_E]}{[y - y_E]} \right). \end{aligned} \quad (2.38)$$

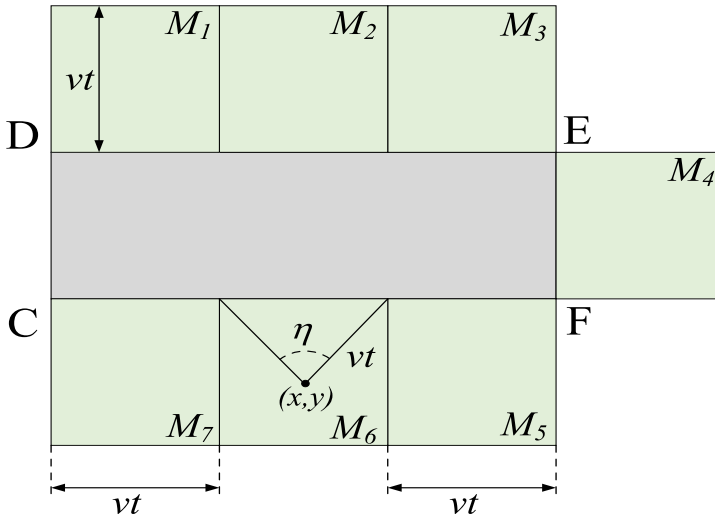


Рис. 2.11. Иллюстрация процесса входа блокировщиков в зону блокировки LoS CDEF

Можно считать, что периоды прямой видимости между BC и AU распределены экспоненциально с параметром $\zeta(z)$, а распределение интервалов блокировки совпадает с распределением периода занятости в системе массового обслуживания (СМО) типа $M/G/\infty$, который можно получить численно для произвольного распределения времени пребывания блокировщика в зоне блокировки в виде

$$\begin{aligned} F_\eta(x) &= 1 - \left[[1 - F_{T_B}(x)] \left[1 - \int_0^x (1 - F_\eta(x-z)) \times \right. \right. \\ &\quad \times \exp(-\lambda_{B,T} F_{T_B}(z)) \lambda_{B,T} dz \left. \right] + \\ &\quad \left. + \int_0^x (1 - F_\eta(x-z)) \left[de^{-\lambda_{B,T} F_{T_B}(z)} \right] \right]. \end{aligned} \quad (2.39)$$

Интенсивность вхождения блокаторов в зону блокировки LoS в зоне покрытия NR БС можно определить по формуле, учитывающей случайное расстояние до БС. Предполагая равномерное распределение АУ в зоне обслуживания круговой формы, получаем

$$\zeta = \int_0^{d_E} \zeta(z) \frac{2z}{d_E^2} dx. \quad (2.40)$$

2.4.5. Упрощенная модель прерываний

Отметим, что можно построить и более простые модели для интенсивности прерываний. Например, рассмотрим случай, когда периметр зоны блокировки определяется по формуле

$$P_L = 4r_B + \frac{4r_A(h_B - h_U)}{3(h_A - h_U)}. \quad (2.41)$$

Тогда выражение для интенсивности входа блокаторов в зону блокировки можно записать в следующем виде:

$$\epsilon = \frac{2}{5} \lambda_B v (4\pi r_A^2) \left(4r_B + \frac{4r_A(h_B - h_U)}{3(h_A - h_U)} \right), \quad (2.42)$$

где λ_B – плотность блокаторов; v – скорость движения. Далее с.в. времени между прерываниями определяется, как описано ранее.

Отметим также, что среднее время в состоянии блокировки, может быть найдено как среднее время пребывания заявок в СМО типа $M/M/\infty$, т.е.

$$E[\eta] = \frac{1}{\epsilon} \exp^{-\epsilon/v} - 1. \quad (2.43)$$

Вычислив среднюю длину интервала $[1/\epsilon + (1/\epsilon)(e^{-\epsilon/v} - 1)]/2$, определим среднее время между изменениями состояния АУ. Тогда искомую интенсивность прерываний можно получить в виде

$$\zeta = \frac{2}{1/\epsilon + (1/\epsilon)(e^{-\epsilon/\eta} - 1)} = 2\epsilon e^{-\epsilon/\eta}. \quad (2.44)$$

ГЛАВА 3

ВЕРОЯТНОСТНЫЕ МОДЕЛИ ДЛЯ АНАЛИЗА ХАРАКТЕРИСТИК СЕТЕЙ РАДИОДОСТУПА 5G/6G

В этой главе рассматриваются базовые математические модели, лежащие в основе оценки вероятностных характеристик сетей радиодоступа 5G/6G. Поскольку модели строятся с целью оценки одного или нескольких показателей эффективности объекта исследования, в параграфе 3.1, прежде всего, определяются основные характеристики производительности сетей радиодоступа. Рассматриваются два основных аппарата построения вероятностных моделей – методы стохастической геометрии (параграф 3.2) и методы ТМО, а именно модели ресурсных систем массового обслуживания (параграф 3.3). Модели в параграфах 3.2 и 3.3 называют базовыми моделями для вероятностного анализа сетей радиодоступа 5G/6G, поскольку они описывают наиболее важные особенности моделирования и оценки основных показателей качества – начальных моментов и плотности вероятности с.в. ОСШ и вероятности блокировки запроса на установление сессии из-за отсутствия необходимых для обслуживания радиоресурсов. В параграфе 3.4 ввиду важности вопроса приведен метод построения модели стохастической геометрии для анализа характеристик ОСШ в помещениях при прямом взаимодействии устройств. Анализ этой известной модели D2D для краткости не изложен, поскольку метод анализа достаточно детально описан в 3.2.

3.1. Показатели эффективности сетей радиодоступа

ОСШ является одним из важнейших показателей качества беспроводных сетей связи. Напомним, что в русскоязычной литературе также используется термин отношение сигнала к помехе (ОСП – SIR, Signal-to-Interference Ratio), а в англоязычной литературе – термин SNR (Signal-to-Noise Ratio). Разница заключается в том, что подразумевается под помехой. Это может быть только тепловой шум приемника или только помехи от передатчиков, работающих на одной и той же частоте. В ряде случаев под шумом в русскоязычной литературе подразумевается сумма помехи и теплового шума приемника, что соответствует понятию ОСШ. Часто понятие, о котором идет речь, становится очевидно из контекста.

Понятия ОСШ и ОСП также остаются актуальными в контексте сотовых сетей 5G/6G. С.в. ОСШ можно записать в упрощенном виде

$$SINR = \frac{gcl^{-\alpha}}{I + \sigma^2}, \quad (3.1)$$

где g – мощность рассматриваемого сигнала; l – расстояние между передатчиком и приемником; c и α – коэффициенты затухания радиосигнала; I – суммарная мощность мешающих сигналов; σ^2 – мощность шума.

Кроме того, чтобы учесть эффекты замирания (fading), вызываемого смещениями абонентского устройства на небольшие расстояния и теневого замирания (shadow fading), вызываемого большими объектами в канале, необходимо умножить правую часть формулы (3.1) на соответствующую случайную величину (см. главу 2). Для частотно-избирательных (frequency-selective) широкополосных каналов необходимо дополнительно интегрировать их по всей полосе частотного канала.

Применяя теорему Шеннона, ОСШ можно использовать для оценки пропускной способности с учетом помехи. Согласно формуле (3.1), ОСШ является функцией условий распространения сигнала между исследуемым и интерферирующими устройствами. При этом предполагается, что ОСШ является функцией расстояний $d_i, i = 1, \dots, N$ между исследуемым устройством и источником полезного сигнала и интерферирующими устройствами. По этой причине при оценке характеристик каналов связи сотовых систем основной задачей является определение местоположения пользователя, т.е. считается, что расстояния d_i до них являются случайными величинами.

3.2. Стохастическая геометрия и ее применение к анализу сетей радиодоступа 5G/6G

Основным методом анализа беспроводных сетей связи, получившим в последнее десятилетие наибольшее развитие, является стохастическая геометрия, с помощью которой описываются случайные события в двух- или трехмерном пространстве. В том числе для моделирования расположения БС и АУ используются пространственные точечные процессы Пуассона, Матерна и др., а также

случайные графы и тесселяции пространства, в том числе тесселяция Дирихле, приводящая к ячейкам Вороного, для моделирования зон обслуживания БС.

3.2.1. Базовая модель взаимодействия устройств

Рассмотрим сценарий с двумя беспроводными передатчиками Tx_0 и Tx_1 , находящимися на случайном расстоянии друг от друга, характеризуемом с.в. U с плотностью $f_U(u)$. Положения приемников Rx_0 и Rx_1 являются с.в., равномерно распределенными в кругах радиусов r_0 и r_1 с центрами в точках Tx_0 и Tx_1 соответственно. Для удобства изложения логическое соответствие между передатчиком и приемником будем называть кластером. Например, на рис. 3.1 показаны целевой и интерферирующий кластеры. В модели Tx_1 рассматривается как ближайший источник помех для Rx_0 , а плотность распределения расстояния L от Tx_0 до Rx_0 является с.в. с плотностью $f_L(l)$.

Без потери общности координаты Tx_0 и Tx_1 выровнены относительно горизонтальной оси. Предположим, что угол $\gamma = \angle Tx_1Tx_0Rx_0$ является с.в., имеющей равномерное распределение, хотя может быть рассмотрено и любое другое распределение. Все три с.в. U , L , и γ предполагаются взаимно независимыми.

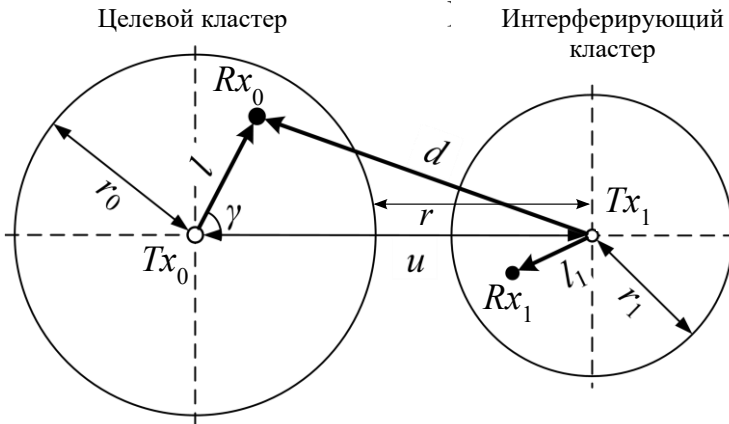


Рис. 3.1. Модель системы с целевым и интерферирующим кластерами

В качестве примера рассмотрим случай (рис. 3.1), где Rx_0 – целевой приемник, принимающий сигнал от Tx_0 , а передатчик Tx_1 , взаимодействующий с Rx_1 , создает помехи. Для круга радиуса r_0

плотность распределения расстояния между точкой, соответствующей целевому приемнику Rx_0 и равномерно распределенной внутри этого круга, и неподвижной точкой Tx_1 , расположенной на фиксированном расстоянии $u = r_0 + r, r > 0$ от центра круга, определяется по формуле

$$g(\rho) = \frac{2\rho}{\pi r_0^2} \arccos\left(\frac{r}{\rho} + \frac{\rho^2 - r^2}{2\rho(r_0 + r)}\right), \quad (3.2)$$

где $r \leq \rho \leq r + 2r_0$. Тогда плотность распределения расстояния между двумя точками, где r является с.в. с плотностью $f_r(x)$, имеет вид

$$g_2(\rho) = \int_0^\infty \frac{2\rho}{\pi r_0^2} \arccos\left(\frac{x}{\rho} + \frac{\rho^2 - x^2}{2\rho(r_0 + x)}\right) f_r(x) dx. \quad (3.3)$$

Отметим, что метод, на основе которого получена формула (3.3), можно обобщить на случай произвольного распределения с.в. расстояния от центра окружности до рассматриваемой точки.

Теперь рассмотрим общий случай, когда треугольник $\langle Tx_1Tx_0Rx_0 \rangle$ определяется распределением его двух сторон и углом между ними. Нас интересует распределение расстояния между передатчиком Tx_1 и приемником Rx_0 для заданных распределений с.в. расстояний L, U и угла γ , считая их независимыми друг от друга. Считаем также, что с.в. величина угла γ равномерно распределена на $[0, \pi]$. Метод, изложенный далее для нахождения плотности распределения расстояния между передатчиком Tx_1 и приемником Rx_0 , основан на последовательном применении функциональных преобразований случайных величин.

Для простоты изложения введем следующие с.в.:

$$\xi_1 = L, \xi_2 = U, \xi_3 = \gamma, \eta_1 = D, \quad (3.4)$$

с плотностями распределений:

$$w_{\xi_1, \xi_2, \xi_3}(x) = f_{L,U,\gamma}(x), W_{\eta_1}(y_1) = f_D(y_1), \quad (3.5)$$

где $f_{L,U,\gamma}(x_1, x_2, x_3)$ – это совместная плотность с.в. L, U и γ ; $f_D(y_1)$ – искомая плотность распределения. Из свойства независимости в совокупности с.в. L, U и γ получаем, что

$$f_{L,U,\gamma}(x) = f_L(x_1)f_U(x_2)f_\gamma(x_3). \quad (3.6)$$

Применив теорему косинусов, получим функциональную зависимость

$$\eta_1 = f(\xi_1, \xi_2, \xi_3) = \sqrt{\xi_1^2 + \xi_2^2 - 2\xi_1\xi_2\cos\xi_3}, \quad (3.7)$$

из которой можно получить плотность распределения с.в. η_1 .

Введем две вспомогательные переменные $\eta_2 = \xi_2$ и $\eta_3 = \xi_3$ и получим, что

$$\begin{cases} y_1 = f(x_1, x_2, x_3) = \sqrt{x_1^2 + x_2^2 - 2x_1x_2\cos x_3}, \\ y_2 = x_2, \\ y_3 = x_3. \end{cases} \quad (3.8)$$

Тогда искомое распределение с.в. D может быть получено из следующего выражения:

$$\begin{aligned} f_D(y_1) &= \\ &= \sum_{j=1,2} \int_{Y_j} w_{\xi_1, \xi_2, \xi_3}(\phi_j(y_1, y_2, y_3), y_2, y_3) \left| \frac{\partial \phi_j(y_1, y_2, y_3)}{\partial y_2} \right| dy_3, \end{aligned} \quad (3.9)$$

где ϕ_j – обратная функция правой части для y_1 в (3.8) с двумя ветвями обратного преобразования:

$$\begin{cases} \phi_1(y_1, y_2, y_3) = y_2 \cos y_3 + \sqrt{y_1^2 - y_2^2 + y_2^2 \cos^2 y_3} \\ \phi_2(y_1, y_2, y_3) = y_2 \cos y_3 - \sqrt{y_1^2 - y_2^2 + y_2^2 \cos^2 y_3} \end{cases}. \quad (3.10)$$

Заметим, что область интегрирования Y_j из (3.9) получается путем решения системы уравнений

$$\begin{cases} \phi_j(y_1, y_2, y_3) \geq 0, i = 1, 2, \\ y_1 \geq 0, \\ y_2 \geq 0, \\ 0 \leq y_3 \leq \pi. \end{cases} \quad (3.11)$$

Несложно показать, что в результате получаются области интегрирования вида

$$\begin{aligned} Y_1^1 &= \begin{cases} 0 \leq y_2 \leq y_1, \\ 0 \leq y_3 \leq \pi, \end{cases} \\ Y_1^2 &= \begin{cases} y_2 \geq y_1, \\ 0 \leq y_3 \leq \arccos \sqrt{\frac{y_2^2 - y_1^2}{y_2^2}}, \end{cases} \end{aligned} \quad (3.12)$$

$$Y_2 = \begin{cases} y_2 \geq y_1, \\ 0 \leq y_3 \leq \arccos \sqrt{\frac{y_2^2 - y_1^2}{y_2^2}}, \end{cases}$$

где $Y_1 = Y_1^1 \cup Y_1^2$.

Таким образом, финальное выражение для плотности распределения с.в. D имеет вид

$$f_D(y_1) = \sum_{j=1,2} \int_{Y_j} f_L(\phi_j(y_1, y_2, y_3), y_2, y_3) f_U(y_2) \frac{1}{\pi} \frac{y_1}{\sqrt{y_1^2 - y_2^2 + y_2^2 \cos^2 y_3}} dy_3 dy_2. \quad (3.13)$$

Ясно, что, даже для данной частной задачи решение (3.13) в явном виде можно получить только для некоторых конкретных типов распределений с.в., а для расчетов можно использовать численное интегрирование.

Отметим, что как следует из (3.13), с.в. D зависит от с.в. L , и поэтому для расчета характеристик ОСШ необходимо их знать совместное распределение. Следуя описанному выше методу, это распределение может быть найдено в виде

$$W_{L,D}(y_1, y_2) = \sum_{j=1,2} \int_{Y_{3,j}} \frac{f_L(y_1) f_U(\phi_j(y_1, y_2, y_3)) f_Y(y_3) y_2}{\sqrt{y_2^2 - y_1^2 + y_1^2 \cos^2 y_3}} dy_3, \quad (3.14)$$

где ϕ_j – обратная функция (3.8) по отношению к y_2 :

$$\begin{cases} \phi_1(y_1, y_2, y_3) = y_1 \cos y_3 + \sqrt{y_2^2 - y_1^2 + y_1^2 \cos^2 y_3} \\ \phi_2(y_1, y_2, y_3) = y_1 \cos y_3 - \sqrt{y_2^2 - y_1^2 + y_1^2 \cos^2 y_3} \end{cases} \quad (3.15)$$

а $Y_{3,j}$ – области интегрирования для y_3 :

$$\begin{aligned} Y_{3,1} &= Y_{3,1}^1 \cup Y_{3,1}^2, \\ Y_{3,1}^1 &= \begin{cases} 0 \leq y_2 \leq y_1, \\ 0 \leq y_3 \leq \frac{1}{2} \arccos \frac{y_1^2 - 2y_2^2}{y_1^2}, \end{cases} \\ Y_{3,1}^2 &= \begin{cases} y_2 \geq y_1, \\ 0 \leq y_3 \leq \pi, \end{cases} \\ Y_2 &= \begin{cases} 0 \leq y_2 \leq y_1, \\ 0 \leq y_3 \leq \frac{1}{2} \arccos \frac{y_1^2 - 2y_2^2}{y_1^2}. \end{cases} \end{aligned} \quad (3.16)$$

Например, предположим, что расстояние от случайно выбранной точки до i -й соседней точки в Пуассоновском поле источников помех подчиняется гамма-распределению с плотностью

$$f_i(x) = \frac{2(\pi\lambda)^i}{(i-1)!} x^{2i-1} e^{-\pi\lambda x^2}, x > 0, i = 1, 2, \dots \quad (3.17)$$

Используем модель принимаемой мощности на приемнике в упрощенном виде $P(x) = Ax^{-\zeta}$, где x – расстояние между приемником и передатчиком, ζ – константа затухания, A – константа, зависящая от свойств передатчика, приемника и несущей частоты. Обратная функция и ее производная имеют вид

$$G(y) = P^{-1}(x) = \left(\frac{y}{A}\right)^{-1/\zeta}, G'(y) = \frac{\left(\frac{y}{A}\right)^{-\frac{1}{\zeta}-1}}{A\zeta}. \quad (3.18)$$

Тогда плотность распределения с.в., определяющей мощность помехи от i -го источника, представима в виде

$$f_i(y) = \frac{2(\pi\lambda)^i}{(i-1)!} \left(\frac{y}{A}\right)^{-(2i-1)/\zeta} e^{-\pi\lambda\left(\frac{y}{A}\right)^{-2/\zeta}} \frac{\left(\frac{y}{A}\right)^{-\frac{1}{\zeta}-1}}{A\zeta}, y > 0, i = 1, 2, \dots \quad (3.19)$$

Для достаточно простых сценариев, например для случая треугольника с одной парой приемник-передатчик и одним источником помех, или для фрагмента сотовой сети с шестью источниками помех, или в условиях внутренних помещений, также с ограниченным количеством источников помех, использование описанного выше метода позволяет получить распределения базовых характеристик каналов связи, таких как ОСШ, ОСП, спектральная эффективность и скорость передачи данных.

Численный пример

Рассмотрим типичный сценарий прямого взаимодействия устройств D2D. Местоположения передатчиков определяются точечным процессом PPP с плотностью λ . Соответствующие приемники распределены равномерно внутри единичного круга с центром в точке размещения передатчика. Задача состоит в том, как при случайном выборе передатчика найти распределение расстояния до приемников, которым этот передатчик будет мешать. Известны плотности распределения определенных выше с.в. L и U :

$$f_L(l) = \frac{2l}{r_0^2}, 0 \leq l \leq r_0, \quad (3.20)$$

$$f_U(u) = 2\pi\lambda u e^{\pi\lambda r_0^2 - \pi\lambda u^2}, \quad u \geq r_0, \quad \lambda > 0.$$

Учитывая, что $r_0 = 1$, и применяя предложенный выше подход, находим плотность распределения $f_D(y_1)$ с.в. D расстояния между интерферирующими приемником и передатчиком.

Вид плотности распределения $f_D(y_1)$ при различных значениях плотности передатчиков λ показан на рис. 3.2. Здесь расстояния показаны в условных единицах, например в метрах или в километрах. На графике видно смещение плотности вправо. Можно сказать, что полученное численно распределение характеризуется «коротким хвостом».

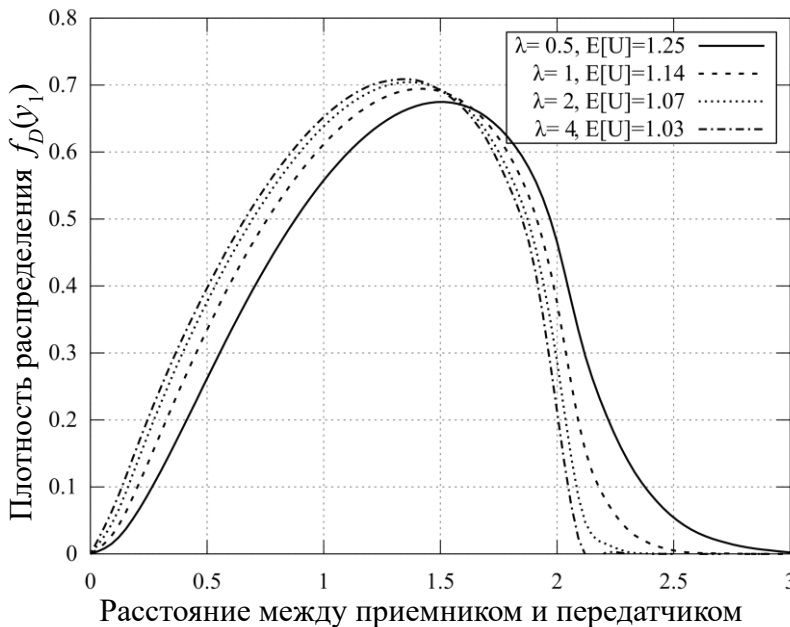


Рис. 3.2. Плотность распределения расстояния между интерферирующими приемником и передатчиком

Для рассматриваемого D2D-сценария с ограничениями по помехам можно оценить среднее значение ОСП, т.е. ОСШ в случае незначительного шума. Для анализа выберем данные, близкие к реальным, где местоположение приемника распределено равномерно внутри круга радиусом 30 метров, а интерферирующий передатчик находится вне этого круга. Выберем модель затухания радиосигнала, пропорциональную расстоянию с параметрами распространения $\alpha = 2,61116$ и $c = 0,000201944$. Мощность передатчиков считается единичной величиной.

Применив для расчета n -го начального момента с.в. ОСШ формулу (3.14), получим

$$E[SIR^n] = \int_{0 \leq y_1 \leq r_0} \int_{y_2 \geq 0} \left(\frac{y_2}{y_1}\right)^{n\alpha} W_{L,D}(y_1, y_2) dy_2 dy_1. \quad (3.21)$$

По формуле (3.21) путем численного интегрирования получены значения математического ожидания и среднеквадратичного отклонения ОСШ в зависимости от расстояния до источника помех. Графики результатов вычисления показаны на рис. 3.3.

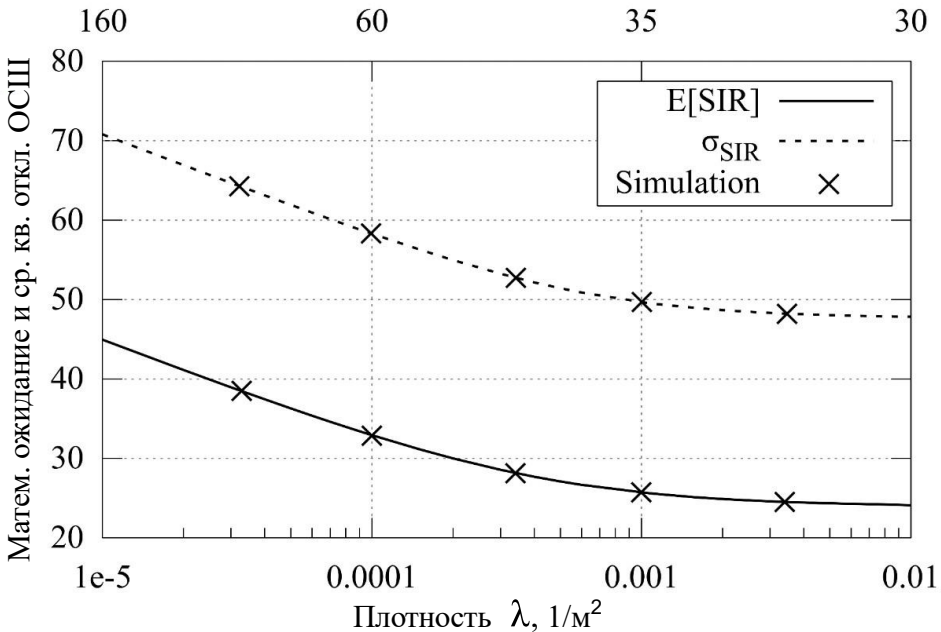


Рис. 3.3. Пример вычисления математического ожидания и дисперсии ОСШ

3.2.2. Применение теоремы Кэмпбелла

В тех случаях, когда число источников помех является случайной величиной, использование метода нелинейного преобразования случайных величин не позволяет получить замкнутые аналитические выражения искомых характеристик. Предположим, что передатчики, работающие на одной и той же частоте, распределены на плоскости согласно РРР с плотностью λ . Приемники, ассоциированные с передатчиками, расположены на фиксированном расстоянии r_0 . Рассмотрим случайно выбранный приемник и некоторую

характеристику, например ОСШ, которая в упрощенном виде может быть представлена следующим образом:

$$S(x) = \frac{C_0}{C_1 + I}, \quad (3.22)$$

где C_0 – константа, определяющая уровень принимаемого полезного сигнала от передатчика, расположенного на фиксированном расстоянии r_0 от приемника; C_1 – константа, учитывающая потери на передаче и приеме; I – суммарная помеха, определяемая по формуле

$$I = C_2 \sum_{i=1}^{\infty} R_i^{-\zeta}, \quad (3.23)$$

Здесь C_2 – константа, учитывающая усиления на передаче и приеме; ζ – константа распространения, зависящая от условий распространения; R_i – расстояние от передатчиков до рассматриваемого приемника.

Используя метод преобразования с.в. (раздел 3.2.1), можно получить распределение помехи от i -го передатчика. Однако, учитывая специфику модели распространения, а также модель размещения передатчиков на плоскости, можно видеть, что общая помеха является суммой случайного числа с.в., имеющих различные распределения. Для того чтобы получить оценку начальных моментов помехи, можно использовать теорему Кэмпбелла, которая позволит получить моменты суммы случайного числа с.в.

Ограничим зону помехи некоторым радиусом D . На практике D может быть выбрано таким образом, что вклад источника помех, находящегося на большем расстоянии, чем D , меньше, чем значение других помех на приемнике. Учитывая свойства РРР, расстояние от источников помех до рассматриваемого приемника распределено одинаково и независимо с плотностью $f_X(x) = 2x/D^2, 0 < x < D$. Рассмотрим малое dr радиальное приращение величины D . Вероятность того, что источник помех находится на расстоянии r , пропорциональна приращению области dr и тогда справедливо равенство

$$\pi(r + dr)^2 - \pi r^2 = 2\pi r dr + O(1). \quad (3.24)$$

Начальные моменты общей помехи могут быть получены, используя теорему Кэмпбелла, в следующем виде:

$$E[I^n] = \int_{r_B}^D (C_2 r^{-\zeta})^n 2\lambda \pi r dr, \quad (3.25)$$

где r_B – минимальное расстояние от приемника, на котором может находиться источник помех.

Следует подчеркнуть, что результат (3.25) может быть обобщен на случай, который учитывает блокировки прямой видимости, направленности антенн, случайной высоты передатчиков и приемников и некоторые другие случаи. Примеры будут показаны в главе 4, где рассмотрим задачи вычисления помехи и ОСШ в сетях 5G/6G.

3.2.3. Оценка распределения скорости передачи

Помимо базовых характеристик, таких как ОСШ и ОСП, рассмотренные выше модели позволяют дать оценку скорости передачи на некоторой предоставляемой пользователю фиксированной полосе частот. Отметим, что в реальных системах доступная полоса частот разделяется между активными пользователями в некоторой пропорции, которая зависит от настроек системы и целей сетевого оператора.

Стохастическая геометрия также может использоваться и для оценки скорости, предоставляемой пользователю в присутствии конкурирующих пользователей. При этом необходимо вводить дополнительные предположения о разделении ресурсов между пользователями. Предположим, например, что в зоне обслуживания БС радиуса R находится некоторое количество пользователей, распределенное в круге по закону Пуассона с параметром $\lambda\pi R^2$, где λ – плотность пользователей на квадратный метр. Предположим также, что местоположение каждого пользователя равномерно распределено в зоне обслуживания БС, а доступная полоса частот B равномерно распределяется между пользователями. Данный сценарий соответствует размещению БС в Пуассоновском поле пользователей с плотностью λ .

Предположим, что случайно выбранному пользователю полоса частот размера B/i выделяется с вероятностью

$$p_i = \frac{(\lambda\pi R^2)^i}{i!} e^{-\lambda\pi R^2}, i = 1, 2, \dots \quad (3.26)$$

При этом достижимая скорость пользователя в присутствии i конкурирующих с ним пользователей определена следующим преобразованием с.в. расстояния от пользователя до БС:

$$f_X(x) = 2x/D^2, 0 < x < D. \quad (3.27)$$

Тогда в соответствии с формулой Шеннона для верхней границы скорости передачи получаем, что

$$C_i = \frac{B}{i} \log_2 [1 + S(x)], i = 1, 2, \dots, \quad (3.28)$$

где $S(x)$ – значение ОСШ на расстоянии x от БС. Среднее значение скорости может быть получено суммированием скоростей, соответствующих i пользователям в зоне обслуживания, с взвешенными вероятностями (3.26).

Отметим, что рассмотренный сценарий может быть расширен на случай различных вариантов разделения ресурсов между конкурирующими пользователями путем введения дополнительных весовых коэффициентов. Такими вариантами могут быть минимаксное разделение ресурсов, выравнивающее скорости пользователей, или пропорциональное разделение, дающее приоритет пользователям, находящимся в более выгодных условиях с точки зрения канала связи. Возможно также расширение для случая приоритетов, где оператор имеет возможность плавно управлять разделением скоростей между пользователями (см. пример в главе 11).

3.3. Ресурсные системы массового обслуживания и их применение к анализу сетей 5G/6G

3.3.1. Построение моделей ресурсных систем массового обслуживания

Для анализа процессов обслуживания абонентов в системах со случайными потоками запросов на установление сессий традиционно применяются методы теории массового обслуживания (ТМО). С появлением сетей беспроводного доступа с коммутацией пакетов помимо характеристик потока поступающих запросов (с точки зрения ТМО заявок) возникла необходимость учитывать случайные объемы требований заявок к сетевым ресурсам, связанные, например, с расстоянием пользователя (точки возникновения запроса) от БС. Учитывая такую особенность, можно предположить, что обслуживаемый трафик является эластичным (best effort). Однако сети радиодоступа 5G/6G ориентированы на приложения, весьма требовательные к пропускной способности на беспроводном участке доступа, например, передача видео высокого разрешения, приложения AR/VR. Трафик таких приложений не обладает свойством эластичности и, как правило, требует некоторой фиксированной скорости передачи в течение всей длительности обслуживания.

Одним из возможных вариантов построения моделей, учитывающих особенности обслуживания абонентов в сетях радиодоступа 5G/6G, является совместное использование методов стохастической геометрии и ТМО. Методы стохастической геометрии позволяют учесть такие особенности, как случайный характер местоположения точек возникновения запросов на установление сессий, случайный характер блокировок LoS. Модели ТМО позволяют учесть случайный объем требований к радиоресурсам, а также характеристики поступающего на обслуживание потока заявок.

Рассмотрим систему массового обслуживания (СМО) с N обслуживающими приборами, в которой каждой поступающей заявке требуется отдельный прибор. Также в СМО имеется некоторый ресурс фиксированного объема R . Предполагается, что определена некоторая минимальная условная единица Δ , которая, например, соответствует минимальному размеру первичного ресурсного блока в сетях радиодоступа 5G/6G. Поток поступающих заявок является пуассоновским с интенсивностью λ , а время обслуживания заявки (длительность сессии) распределено экспоненциально с параметром μ . Потребности заявок в ресурсах являются независимыми одинаково распределенными с.в. с функцией распределения (ФР) $F_R(x)$, которые не зависят от процессов поступления и обслуживания. В случае если ресурс предполагается дискретным и измеряется условными единицами объема, то с.в. требования к ресурсам задается рядом распределения p_j , где $j \geq 0$ – число условных единиц ресурса. Если в момент поступления ресурсов не хватает, то поступающая заявка получает отказ в обслуживании.

Требования к ресурсам в рассматриваемой СМО не остаются постоянными в течение всего времени обслуживания. Эта особенность моделируется с помощью пуассоновского потока сигналов с интенсивностью α , ассоциированного с каждой обслуживаемой заявкой. По прибытии сигнала заявка освобождает все занятые ресурсы, создает новое требование с той же ФР $F_R(x)$ и вновь пытается занять ресурсы для своего обслуживания. Если объем нового требования к ресурсам превышает объем незанятых в системе ресурсов, то заявка блокируется и сбрасывается с обслуживания.

Описанная СМО учитывает одну из основных особенностей сетей радиодоступа 5G/6G – изменение состояния АУ, связанное с блокировкой путей распространения сигнала между БС и АУ. В качестве сценариев можно рассматривать БС с некоторым радиусом покрытия как стационарных, так и мобильных абонентов в поле стационарных

или движущихся блокаторов. Отметим, что в условиях стационарных абонентов и стационарных блокаторов необходимость в потоке сигналов, описывающем процесс изменения состояния АУ с заблокированного на незаблокированный, отпадает. Такие модели более подробно рассмотрены в контексте методов повышения качества обслуживания абонентов в сетях 5G/6G в главе 5.

Для параметризации модели необходимо определить следующие величины: а) область покрытия d_E БС с заданной излучаемой мощностью и с конфигурациями передающей и приемной антенн; б) ФР $F_R(x)$ объема ресурсов, запрашиваемых у БС новой сессией; в) интенсивность α изменений состояния АУ, т.е. интенсивность потока сигналов. Методы, позволяющие рассчитывать перечисленные параметры, рассмотрены в главе 2.

В качестве показателей качества обслуживания интерес представляют параметры производительности системы и характеристики обслуживания абонентов. Одним из основных параметров производительности системы для оператора является средний объем занятых ресурсов и коэффициент их использования, т.е. доля времени, когда ресурсы заняты. К характеристикам обслуживания абонентов относятся такие показатели, как вероятность потери заявки из-за нехватки ресурса при поступлении в систему и вероятность сброса уже принятой на обслуживание заявки из-за изменения требования к объему занимаемого ресурса. Класс рассматриваемых систем в литературе принято называть ресурсными СМО (ReСМО).

3.3.2. Базовая ресурсная система массового обслуживания

Рассмотрим процесс обслуживания трафика на отдельной (stand alone) БС сети 5G/6G. Учитываются следующие характеристики и условия обслуживания: а) радиус обслуживания БС выбран таким образом, что не происходит прерываний уже установленных соединений; б) максимально возможное количество активных сессий может быть ограничено; в) ресурс ограничен и измеряется в первичных ресурсных блоках NR; г) интенсивность поступления сессий в зоне обслуживания фиксирована и известна; д) геометрическое положение пользователей с активными сессиями распределено равномерно в зоне обслуживания БС; е) трафик является неэластичным.

Рассмотрим многолинейную СМО с $N \leq \infty$ приборами, моделирующими длительность сессии, и конечным ресурсом объема R .

На систему поступает пуассоновский поток заявок с интенсивностью λ , обслуживание которых является экспоненциальным с параметром μ . Каждая заявка при поступлении в систему требует один прибор и некоторый случайный объем ресурсов. Будем считать, что распределение требований к ресурсам является дискретным с законом распределения $\{p_j\}_{j \geq 0}$, где p_j – вероятность того, что заявке требуется j единиц ресурсов. Если при поступлении заявки в системе нет свободного прибора или недостаточно ресурсов, то она теряется. По завершении обслуживания заявка покидает систему и освобождает весь объем занятых ею ресурсов. Схема рассматриваемой РеСМО изображена на рис. 3.4.

Функционирование РеСМО описывается случайным процессом (СП) $X_1(t) = \{\xi(t), \gamma(t)\}$, где $\xi(t)$ – число заявок в системе в момент t , а $\gamma(t) = (\gamma_1(t), \dots, \gamma_{\xi(t)}(t))$ – вектор объемов занятых ресурсов каждой заявкой. Считается, что заявки нумеруются в порядке их поступления в систему.

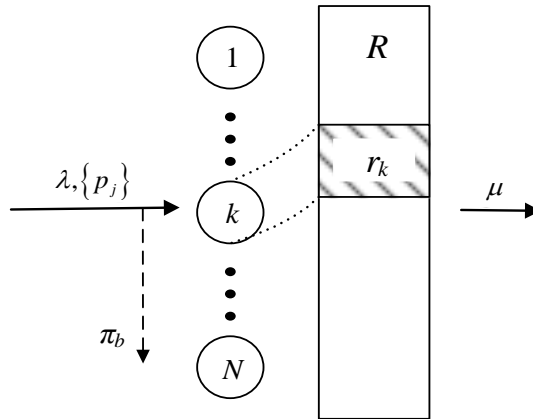


Рис. 3.4. Схема базовой ресурсной СМО

Введем стационарные вероятности СП $X_1(t)$:

$$Q_k(r_1, \dots, r_k) = \lim_{t \rightarrow \infty} P\{\xi(t) = k, \gamma_1(t) = r_1, \dots, \gamma_k(t) = r_k\},$$

$$0 \leq k \leq N, r_1 + \dots + r_k \leq R, \quad (3.29)$$

и запишем для их нахождения систему уравнения равновесия (СУР) в следующем виде:

$$\lambda Q_0 \sum_{r=0}^{\infty} p_r = \mu \sum_{r=0}^R Q_1(r);$$

$$\begin{aligned}
Q_k(r_1, \dots, r_k) & \left(\lambda \sum_{j=0}^{R-r_1-\dots-r_k} p_j + n\mu \right) = \\
& = \frac{\lambda}{k} \sum_{i=1}^k Q_{k-1}(r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_k) p_{r_i} + \\
& + \mu \sum_{i=1}^{k+1} \sum_{j=0}^{R-r_1-\dots-r_k} Q_{k+1}(r_1, \dots, r_{i-1}, j, r_i, \dots, r_k), \quad (3.30)
\end{aligned}$$

$$1 \leq k \leq N-1, r_1 + \dots + r_k \leq R;$$

$$\begin{aligned}
N\mu Q_N(r_1, \dots, r_N) & = \frac{\lambda}{N} \sum_{i=1}^N Q_{N-1}(r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_{N-1}) p_{r_i}, \\
& r_1 + \dots + r_N \leq R.
\end{aligned}$$

Отметим, что СУР (3.30) имеет единственное решение, удовлетворяющее условию нормировки. Опуская выкладки, с помощью подстановки можно убедиться, что решение СУР имеет следующий вид:

$$\begin{aligned}
Q_k(r_1, \dots, r_k) & = Q_0 \frac{\rho^k}{k!} p_{r_1} \dots p_{r_k}, \quad 1 \leq k \leq N, \sum_{i=1}^k r_i \leq R, \\
Q_0 & = \left(1 + \sum_{k=1}^N \sum_{r_1+\dots+r_k \leq R} \frac{\rho^k}{k!} p_{r_1} \dots p_{r_k} \right)^{-1}, \quad (3.31)
\end{aligned}$$

где $\rho = \frac{\lambda}{\mu}$ – предложенная нагрузка.

Следует подчеркнуть, что размерность пространства состояний СП $X_1(t)$ растет с ростом N и R , что приводит к существенным сложностям при вычислении стационарных вероятностей и расчете характеристик производительности системы, несмотря на наличие аналитических формул (3.31). Более того, СП $X_1(t)$ включает информацию, которая в большинстве прикладных сценариях избыточна. Как правило, для анализа показателей эффективности исследуемой системы достаточно знать только общее число заявок в системе и суммарный объем занятого ими ресурса. Введем стационарные вероятности укрупненных состояний

$$P_k(r) = \lim_{t \rightarrow \infty} P\{\xi(t) = k, \sum_{i=1}^k \gamma_i(t) = r\}. \quad (3.32)$$

И тогда, просуммировав соответствующие вероятности в (3.31), получим

$$P_k(r) = P_0 \frac{\rho^k}{k!} p_r^{(k)}, \quad 1 \leq k \leq N, r \leq R,$$

$$P_0 = \left(1 + \sum_{k=1}^N \sum_{r=0}^R \frac{\rho^k}{k!} p_r^{(k)} \right)^{-1}, \quad (3.33)$$

где $\{p_r^{(k)}\}_{r \geq 0}$ – k -кратная свертка распределения $\{p_j\}_{j \geq 0}$, причем $p_r^{(k)}$ можно интерпретировать как вероятность того, что k заявок суммарно запрашивают r ресурсов. На практике дискретные свертки удобно вычислять при помощи следующего рекуррентного соотношения:

$$p_r^{(k)} = \sum_{i=0}^r p_i p_{r-i}^{(k-1)}, \quad k \geq 2, \quad (3.34)$$

где $p_r^{(1)} = p_r, r \geq 0$.

Используя вероятности укрупненных состояний $P_k(r)$, можно получить формулы для основных характеристик модели – вероятности потери π_b и среднего объема занятого ресурса \bar{R} в следующем виде:

$$\pi_b = 1 - P_0 \sum_{k=0}^{N-1} \frac{\rho^k}{k!} \sum_{r=0}^R p_r^{(k+1)}, \quad (3.35)$$

$$\bar{R} = P_0 \sum_{k=1}^N \frac{\rho^k}{k!} \sum_{r=1}^R r p_r^{(k)}. \quad (3.36)$$

Более детально модели РеСМО рассматриваются в главе 5, где исследуются методы и модели улучшения показателей качества обслуживания абонентов в сетях связи 5G/6G с использованием совместного применения методов стохастической геометрии и ТМО.

3.4. Пример построения модели стохастической геометрии прямого взаимодействия устройств

В заключение данной главы рассмотрим пример применения стохастической геометрии для определения базовых характеристик сетей радиодоступа – ОСП и ОСШ.

В сетях 5G/6G ожидается активное использование новых технологий, большинство которых могут даже играть роль сетевой инфраструктуры в обеспечении беспроводной связи. К примеру, пользовательские устройства будут применять для передачи данных технологию D2D с поддержкой приложений ретрансляции различных типов медиа-услуг. Важность применения подобных технологий крайне велика, например, в случаях чрезвычайных ситуаций возникновения природных катаклизмов и техногенных катастроф, когда рабочие узлы сотовой связи частично или полностью выведены из строя. Например, в течение всего времени при проведении

восстановительных работ D2D-устройства позволяют поддерживать рабочее состояние сети. Предполагается, что технологии D2D также могут стать полезными для борьбы с перегрузками на участках беспроводной сети с большим числом и высокой плотностью пользовательских устройств, например, на массовых мероприятиях, стадионах, в крупных торгово-развлекательных центрах и т.д.

Учет влияния помехи на качество предоставления услуг является важнейшим фактором в сетях беспроводной связи. В данном разделе оценка влияния помехи определяется в предположении, что шум на беспроводном канале равен нулю, т.е. $\sigma^2 = 0$. Поэтому анализ проводится для ОСП по формуле $SIR = S/I$, где S – мощность полезного сигнала от рассматриваемого нас передатчика, I – мощность мешающего сигнала от других передатчиков. Затем, зная ОСП, по формуле Шеннона может быть вычислена спектральная эффективность и пропускная способность канала.

С помощью стохастической геометрии для вычисления помехи в интересующей точке может быть проведен анализ характеристик с.в. ОСП, когда исследуемые объекты случайно распределены на плоскости. В таком исследовании получение формул для распределения с.в. расстояния между взаимодействующими устройствами, т.е. для вычисления помехи в интересующей точке, тем более в замкнутом аналитическом виде, является весьма трудоемкой задачей. Для этого часто применяются известные модели PPP, для которых известны формулы распределения с.в. расстояния от заданной до любой соседней точки. Например, известна модель для анализа с.в. SIR -взаимодействия одного передатчика и нескольких приемников, когда устройства расположены на плоскости в пределах круга с заданным радиусом. Передатчик расположен в центре круга, а приемники равномерно распределены внутри этого круга. Примером такого сценария могут быть занятия руководителя с подчиненными с использованием соответствующих устройств и применением технологии D2D.

Ниже рассматривается пример построения модели для анализа с.в. SIR для типичного сценария применения технологии D2D. Исследуются взаимодействие устройств, расположенных в прямоугольных кластерах с заданными размерами, в которых равномерно распределены D2D-устройства. Схема взаимодействия изображена на рис. 3.5.

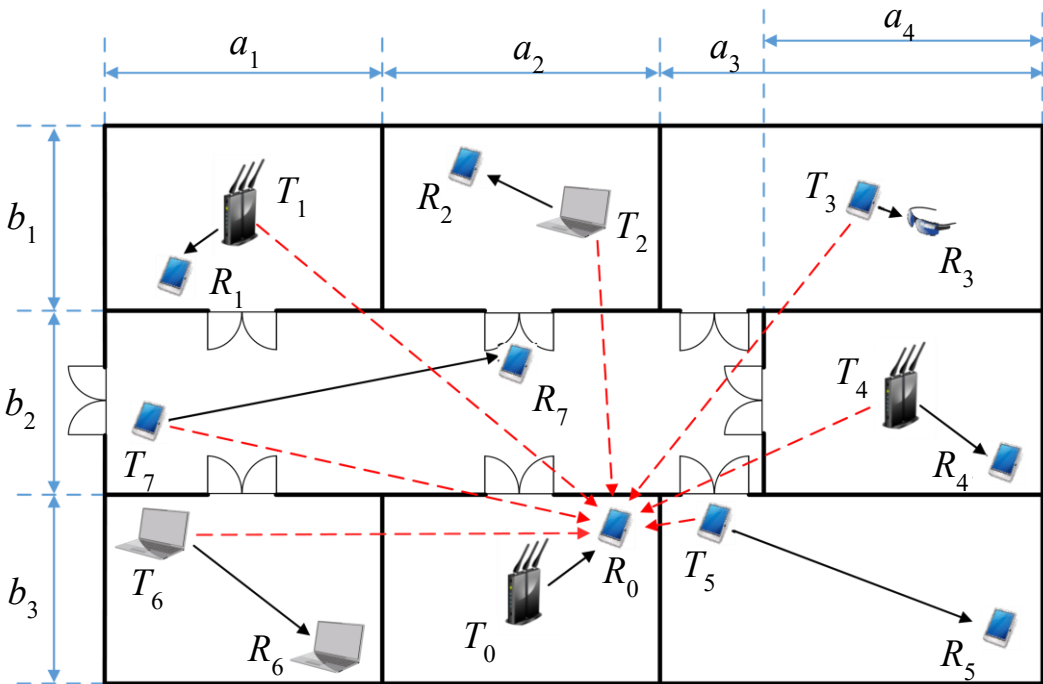


Рис. 3.5. Схема расположения беспроводных устройств в прямоугольных кластерах внутри помещений

Рассмотрим два соседних кластера, в которых могут быть расположены не более одной активно взаимодействующей пары устройств. Такая модель типична, например, для взаимодействия устройств в крупном торговом центре, где в одном из помещений находится пара T_0R_0 взаимодействующих устройств передатчик-приемник, которую будем называть целевой, а другие пары устройств T_iR_i , $i = 1, \dots, 7$, создающие помеху, будем называть интерферирующими.

Рассматриваются два сценария расположения взаимодействующих устройств. В первом сценарии приемники расположены в геометрических центрах кластеров, а передатчики равномерно распределены в пределах соответствующих кластеров. Во втором сценарии роли устройств меняются – передатчики зафиксированы в центрах кластеров, а приемники равномерно распределены в их пределах. Соответствующие сценариям модели будем называть моделями восходящей (для краткости *uplink*, см. рис. 3.6) и нисходящей (*downlink*, см. рис. 3.7) линий связи соответственно. На рис. 3.6 и 3.7 сплошными линиями показана передача данных между целевыми парами, пунктирные линии показывают мешающий (интерферирующий) сигнал.

Упростим задачу, предположив, что взаимодействие происходит в соседних квадратных кластерах (см. рис. 3.6). Построение модели и ее анализ для случая прямоугольных кластеров проводится аналогичным способом.

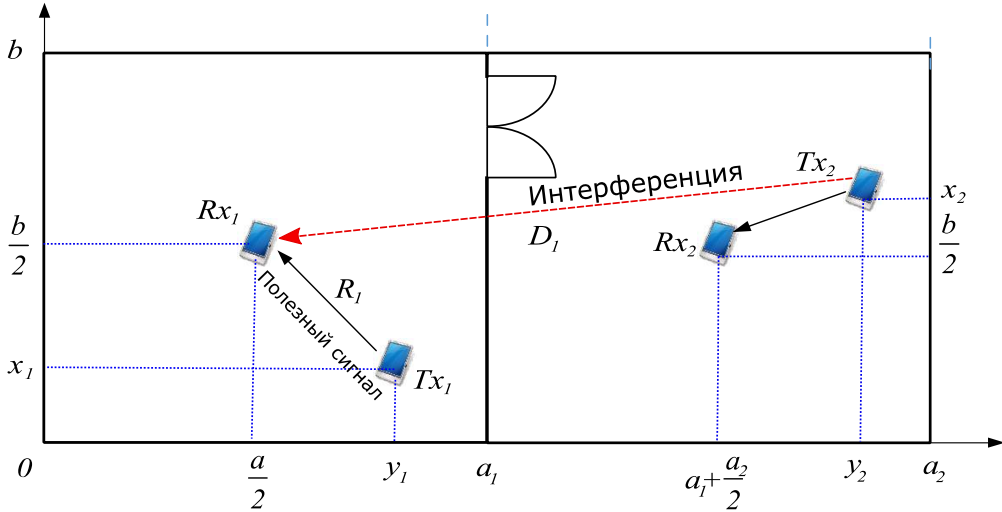


Рис. 3.6. Сценарий восходящей линии связи в квадратных смежных кластерах

В сценарии восходящей линии связи на рис. 3.6 рассматриваются «целевой» и «интерферирующий» кластер со сторонами $c = a_1 = a_2 = b$. Приемники Rx_1 и Rx_2 находятся в центрах, а передатчики Tx_1 и Tx_2 случайно распределены в соответствующих кластерах. Расстоянием между Tx_1 и Rx_1 является с.в. R_1 , а расстоянием между Tx_2 и Rx_1 является с.в. D_1 .

С.в. отношения *сигнал–помеха* (ОСП) с учетом сделанных предположений и того, что шум близок к нулю, имеет вид

$$SIR = \frac{S(R_0)}{(\sum_{i=1}^7 I_i(D_i))}, \quad (3.37)$$

где

$$S(R_0) = gR_0^{-\alpha_0}, \sum_{i=1}^7 I_i(D_i) = g \sum_{i=1}^7 D_i^{-\alpha_i}, \quad (3.38)$$

и α_0 – экспонента затухания полезного сигнала, а α_i – экспонента затухания мешающей передачи.

С учетом предположения, что мощность сигналов принимает одинаковые значения, формула (3.37) преобразуется к виду

$$SIR = \frac{gR_0^{-\alpha_0}}{g \sum_{i=1}^7 D_i^{-\alpha_i}} = \frac{R_0^{-\alpha_0}}{\sum_{i=1}^7 D_i^{-\alpha_i}}. \quad (3.39)$$

Окончательно для сценария с двумя устройствами и с обозначениями, как на рис. 3.6, с.в. ОСП определяется по формуле

$$SIR = \frac{gR_1^{-\alpha_1}}{gD_1^{-\alpha_2}} = \frac{R_1^{-\alpha_1}}{D_1^{-\alpha_2}} = \frac{D_1^{\alpha_2}}{R_1^{\alpha_1}}. \quad (3.40)$$

Заметим, что с.в. ОСП является функцией двух с.в. R_1, D_1 . Если ФР этих с.в. известны, то задача состоит в определении распределения и характеристик с.в. ОСП, определенной в (3.40). Для решения этой задачи необходимо отдельно задать плотности распределений D_1 и R_1 и применить метод, как в базовой модели из раздела 3.2. Аналогично строится и анализируется модель для сценария нисходящей линии связи (downlink) на рис. 3.7.

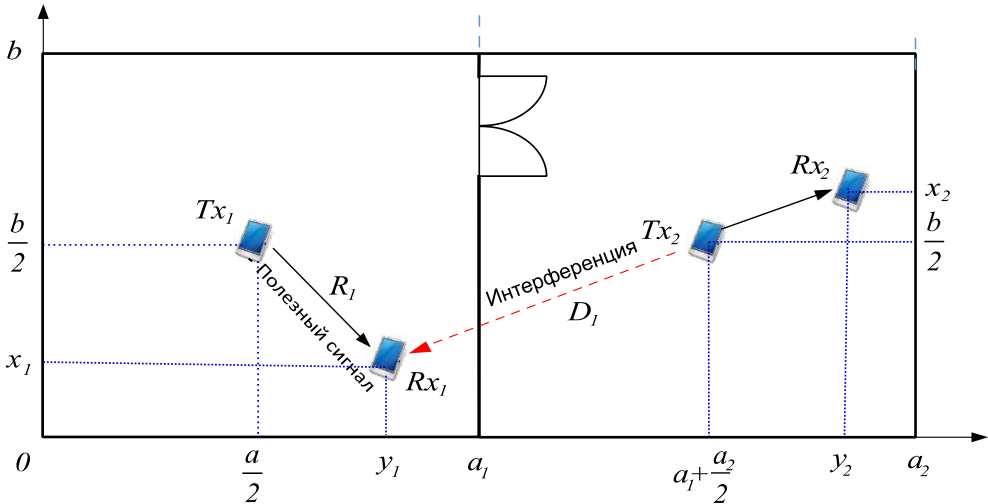


Рис. 3.7. Сценарий нисходящей линии связи в квадратных смежных кластерах

ГЛАВА 4

ПОМЕХА И ОТНОШЕНИЕ СИГНАЛ–ШУМ В СЕТЯХ 5G/6G С НАПРАВЛЕННЫМИ АНТЕННАМИ И БЛОКИРОВКОЙ

В данной главе рассматриваются методы анализа вероятностных характеристик каналов связи с учетом блокировки и направленных антенн на передатчике и приемнике. В качестве исследуемых метрик используются мощность помехи и ОСШ. Для каждой из характеристик рассмотрены двухмерный и трехмерный сценарии развертывания системы связи.

4.1. Помехи в сетях с направленными антенными решетками

Рассмотрим метод оценки первых моментов помехи в сетях доступа, работающих в миллиметровом диапазоне частот. Подробно рассмотрим двухмерный сценарий и далее кратко изложим анализ трехмерного сценария. Напомним, что модели компонентов системы введены в главе 2, а при анализе используется теорема Кэмпбелла, введенная в главе 3. Основная искомая метрика – мощность помехи может быть записана следующим образом:

$$I = A \sum_{i=1}^N d_i^{-\zeta}, \quad (4.1)$$

где A – некоторый коэффициент, зависящий от мощности передатчика, усиления антенн в направлении передачи и приема, а также некоторых других констант; d – расстояние между интерферирующими передатчиками и приемником; N – число интерферирующих передатчиков. Заметим, что мощность помехи является с.в., характеристики которой подлежат определению, в том числе аналитическими методами.

4.1.1. Двухмерный сценарий развертывания

Формула Кэмпбелла для начальных моментов помехи

Рассмотрим метод нахождения начальных моментов с.в. помехи на примере двухмерной модели с использованием диаграммы направленности антенны без потерь сигнала на боковых и задних лепестках (см. главу 2). Для этого рассмотрим модель расположения узлов сети доступа, проиллюстрированную на рис. 4.1. Модель

включает в себя приемник, расположенный в начале координат, который находится в случайном поле интерферирующих узлов. Предполагается, что местоположение интерферирующих узлов определяется РРР с заданной интенсивностью λ . Предположим, что ориентация диаграмм направленности равномерно распределена в интервале $[0; 2\pi)$. Каждый из узлов моделирует проекцию тела человека (например, пользователя сети с АУ) в виде цилиндра с радиусом r_B . На плоскости вокруг интересующего нас приемника опишем круг радиусом R , который рассчитан так, что помеха от узлов вне рассматриваемого круга меньше порога теплового шума.

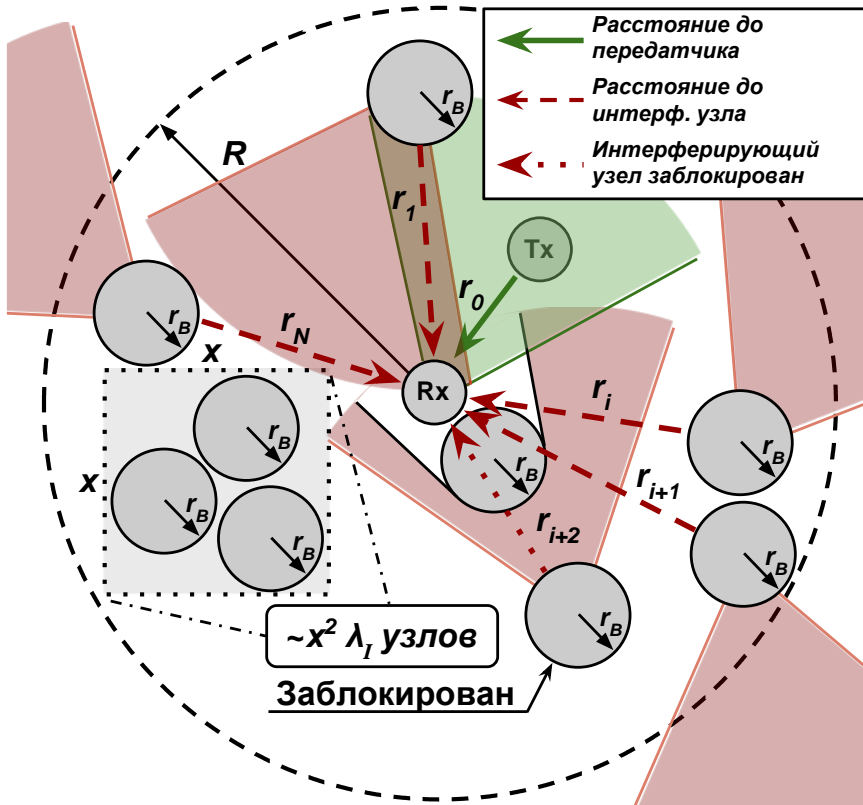


Рис. 4.1. Иллюстрация модели двумерного сценария развертывания

Рассмотрим случай, когда направленные антенны используются только на стороне передатчика или только на стороне приемника. Источник помех в пределах круга радиуса R приемника Rx не будет создавать помехи в двух случаях: а) зона покрытия передатчика Tx не включает приемник или зона покрытия приемника не включает передатчик и б) распространение помех блокируется другими

узлами, блокаторами или самим источниками помех. Вероятность последнего события определяется вероятностью блокировки, p_B . Назовем вероятность первого события вероятностью экспозиции p_C и вероятность второго события p_B вероятностью блокировки.

Вероятность экспозиции для диаграммы направленности первого типа может быть определена формулой

$$p_C = \frac{\alpha x}{2\pi x} = \frac{\alpha}{2\pi}, \quad (4.2)$$

что проиллюстрировано на рис. 4.2. Вероятность блокировки для статичного приемника и пуассоновского поля блокаторов получена в главе 2, см. формулу (2.18).

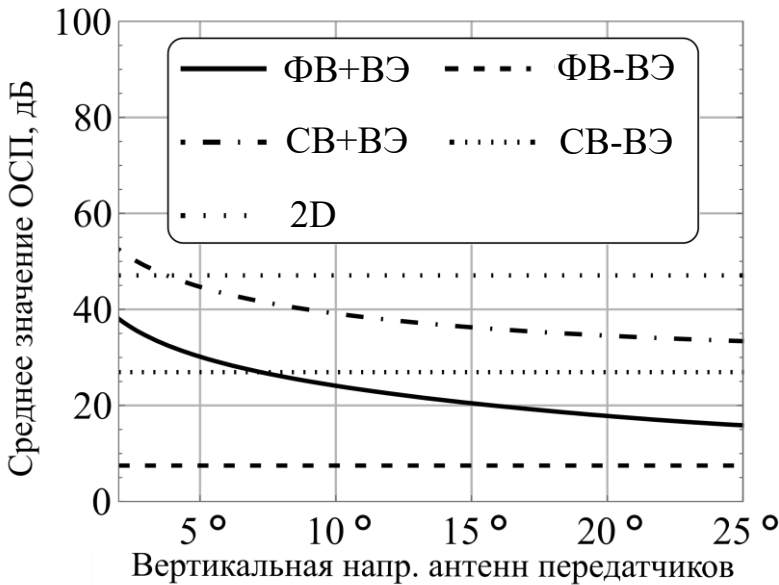


Рис. 4.2. Иллюстрация зоны охвата передатчика

Рассмотрим целевой приемник и некоторое приращение круга dr . Поскольку процесс появления узлов помех PPP является ординарным, то одновременное появление нескольких передатчиков в приращении круга dr невозможно. Следовательно, вероятность того, что источник помех расположен на расстоянии r от приемника, пропорциональна величине dr . Тогда прирост кольца может быть определен формулой

$$\pi(r + dr)^2 - \pi r^2 = 2\pi r dr + O(1), \quad (4.3)$$

из которой следует, что вероятность нахождения источника помех в кольце $(r, r + dr)$ равна $2\pi\lambda_I r dr$, где λ_I – пространственная плотность источников помех.

Если источник помех, находящийся на расстоянии r от приемника, не заблокирован и приемник находится в зоне покрытия источника помех, вклад этого узла в помехи в приемнике составляет $Ar^{-2}e^{-Kr}$ (см. главу 2), где e^{-Kr} – компонент, учитывающий эффект атмосферного поглощения (см. главу 12). Учитывая это, среднее значение $E[I]$ помехи на приемнике может быть рассчитано по формуле, являющейся частным случаем формулы Кэмпбелла:

$$E[I] = \int_{r_B}^R Ar^{-2} e^{-Kr} p_C [1 - p_B(r)] 2\lambda_I \pi r dr. \quad (4.4)$$

Подставляя в (4.4) вероятности p_B и p_C , получаем среднее значение помехи на приемнике в виде

$$\begin{aligned} E[I] &= \int_{r_B}^R Ar^{-2} e^{-Kr} e^{-\lambda_I(x-r_B)r_B} \frac{\alpha}{2\pi} 2\lambda_I \pi r dr = \\ &= A\alpha\lambda_I \theta(R, r_B, \lambda_I, K). \end{aligned} \quad (4.5)$$

Из (4.5) следует, что величина $\theta(R, r_B, \lambda_I, K)$ может быть представлена в виде

$$\begin{aligned} \theta(R, r_B, \lambda_I, K) &= e^{-\lambda_I r_B^2} Ei(-R[K + \lambda_I r_B]) - \\ &- Ei(-r_B^2[K + r_B \lambda_I]), \end{aligned} \quad (4.6)$$

где $Ei(x)$ – интегральная показательная функция.

Рассмотрим несколько частных случаев. Если не учитывать эффект блокировки, среднее значение помехи может быть записано в более простом виде:

$$E[I] = A\alpha\lambda_I \int_{r_B}^R \frac{1}{r} e^{-Kr} dr = A\alpha\lambda_I \theta_1(R, r_B, \lambda_I, K), \quad (4.7)$$

где $\theta_1(R, r_B, \lambda_I, K) = Ei(-KR) - Ei(-Kr_B)$.

Если используются всенаправленные антенны, среднее значение помехи принимает вид

$$E[I] = 2\pi\lambda_I A \theta(R, r_B, \lambda_I, K). \quad (4.8)$$

В случае всенаправленной антенны и отсутствия блокировки одновременно формула значительно упрощается:

$$E[I] = 2\pi\lambda_I A \theta_1(R, r_B, \lambda_I, K). \quad (4.9)$$

В случае, когда используются направленные антенны как на передатчике, так и на приемнике, среднее значение помехи на приемнике получается умножением соответствующих формул на величину $\alpha/2\pi$, и тогда

с блокировкой

$$E[I] = \frac{A\alpha^2\lambda_I}{2\pi} \Theta(R, r_B, \lambda_I, K),$$

без блокировки

$$E[I] = \frac{A\alpha^2\lambda_I}{2\pi} \Theta_1(R, r_B, \lambda_I, K). \quad (4.10)$$

Помеха для диаграммы направленности второго типа

Напомним, что диаграмма направленности второго типа (конус плюс сфера (см. главу 2, рис. 2.2) учитывает помехи на боковые и задние лепестки, что позволяет уточнить результаты предыдущего раздела. При этом необходимо различать три случая.

1. LoS между приемником и интерферирующим узлом заблокирована.
2. LoS между приемником и интерферирующим узлом не заблокирована, и приемник расположен в основном лепестке интерферирующего узла.
3. LoS между приемником и интерферирующим узлом не заблокирована, и приемник расположен не в основном лепестке интерферирующего узла.

Вероятность первого события обозначим $p_B(r)$. Предполагается, что вклад источника помех в совокупную помеху на приемнике принимается равным нулю. Тогда вероятность второго события равна $p_C(1 - p_B(r))$, при этом вклад источника помех узла в совокупную помеху на приемнике равен $A_1 r^{-2} e^{-Kr}$, где A_1 – коэффициент основного лепестка антенны. Наконец, вероятность последнего события равна $(1 - p_C)(1 - p_B(r))$, а вклад источника помех в совокупную помеху на приемнике равен $A_2 r^{-2} e^{-Kr}$, где A_2 – коэффициент усиления в направлении боковых и задних лепестков антенны.

Учитывая введенные события, среднее значение помехи на приемнике при использовании диаграммы направленности второго типа определяется выражением

$$E[I] = \int_{r_B}^R A_1 r^{-2} e^{-Kr} p_C (1 - p_B(r)) 2\lambda_I \pi r dr + \\ + \int_{r_B}^R A_2 r^{-2} e^{-Kr} (1 - p_C) (1 - p_B(r)) 2\lambda_I \pi r dr. \quad (4.11)$$

Вычисляя интеграл в правой части (4.11), получаем, что

$$E[I] = A_1 \alpha \lambda_I \theta(R, r_B, \lambda_I, K) + A_2 [2\pi - \alpha] \lambda_I \theta(R, r_B, \lambda_I, K). \quad (4.12)$$

Рассмотрим несколько важных частных случаев.

Если не принимать во внимание эффект блокировки путей распространения сигнала, среднее значение помехи на приемнике определяется в виде

$$E[I] = A_1 \alpha \lambda_I \theta_1(R, r_B, K) + A_2 [2\pi - \alpha] \lambda_I \theta_1(R, r_B, K). \quad (4.13)$$

Если направленные антенны используются как на приемнике, так и на передатчике, получаем, что с блокировкой

$$E[I] = \frac{A_1 \alpha^2 \lambda_I}{2\pi} \theta(R, r_B, \lambda_I, K) + \frac{A_2 [2\pi - \alpha^2] \lambda_I}{2\pi} \theta(R, r_B, \lambda_I, K),$$

без блокировки

$$E[I] = \frac{A_1 \alpha^2 \lambda_I}{2\pi} \theta_1(R, r_B, K) + \frac{A_2 [2\pi - \alpha^2] \lambda_I}{2\pi} \theta_1(R, r_B, K). \quad (4.14)$$

Отметим, что предлагаемый подход позволяет моделировать различные углы диаграммы направленности как на источниках помех, так и на приемнике.

Эффект молекулярного шума

Молекулярный шум является одним из явлений, характерных для частотных диапазонов, превышающих миллиметровые, в том числе для терагерцового диапазона частот (300 ГГц – 3 ТГц). Заметим, что существование молекулярного шума теоретически предсказано, но еще не подтверждено на практике.

Предложенный выше подход для вычисления начальных моментов помехи может быть модифицирован для учета влияния молекулярного шума. Поскольку молекулярный шум, вызванный передачей источников помех, также может вносить вклад в помеху на приемнике, то эффект можно учесть при оценке среднего значения помех на приемнике, заменяя величину $Ar^{-2}e^{-Kr}$ на Ar^{-2} . В этом

случае среднее значение помехи на приемнике для диаграммы направленности первого типа вычисляется по формуле

$$E[I] = \int_{r_B}^R Ar^{-2} e^{-\lambda_I(x-r_B)r_B} \frac{\alpha}{2\pi} 2\lambda_I \pi r dr = A\alpha\lambda_I\theta_1^*(R, r_B, \lambda_I), \quad (4.15)$$

где

$$\theta_1^*(R, r_B, \lambda_I) = e^{-\lambda_I r_B} E(-\lambda_I r_B R) - E(-\lambda_I r_B^2). \quad (4.16)$$

Для диаграммы направленности второго типа получаем, что

$$E[I] = A_1\alpha\lambda_I\theta_1^*(R, r_B, \lambda_I) + A_2[2\pi - \alpha]\lambda_I\theta_1^*(R, r_B, \lambda_I). \quad (4.17)$$

Предполагая направленные антенны как на передатчике, так и на приемнике, среднее значение помехи с учетом молекулярного шума приводит к следующим выражениям:

для диаграммы первого типа

$$E[I] = \frac{A\alpha^2\lambda_I}{2\pi}\theta_1^*(R, r_B, \lambda_I),$$

для диаграммы второго типа

$$E[I] = \frac{A_1\alpha^2\lambda_I}{2\pi}\theta_1^*(R, r_B, \lambda_I, K) + \frac{A_2[2\pi - \alpha^2]\lambda_I}{2\pi}\theta_1^*(R, r_B, \lambda_I, K). \quad (4.18)$$

4.1.2. Трехмерный сценарий развертывания

Перейдем к рассмотрению трехмерного сценария развертывания системы, модель которого проиллюстрирована на рис. 4.3. Предполагается, что расположение приемников R_i моделируется РРР с интенсивностью λ и каждый приемник соответствует своему передатчику T_i . Передатчики равномерно распределены в круге радиуса R_T с центром в точке расположения соответствующего приемника. Предполагается, что все передатчики используют один и тот же частотный канал, что является причиной помех для целевого приемника. С.в. высот расположения передатчиков и приемников H_T и H_R имеют экспоненциальное распределение с параметрами μ_T и μ_R соответственно. Заметим, что такое предположение моделирует возможность АУ находиться не только в руке пользователя, но и на его теле. Передатчик и приемник связаны с объектами, которые моделируются в виде цилиндров с радиусом r_R и высотой H_B , которая предполагается экспоненциально распределенной с.в. с параметром μ_B .

Среди пар передатчик–приемник выбирается целевая пара (R_0, T_0) с ограниченной областью помех вокруг целевого приемника в виде круга радиуса R_I . Помехи, создаваемые передатчиками, расположенными вне круга, считаются незначительными, то есть ниже уровня шума. Величина R_I рассчитывается на основе модели распространения сигнала, мощности передатчика и направленности антенны. На рис. 4.3 показаны взаимодействие целевой пары (R_0, T_0) , а также передачи, создающие помеху. Также предположим, что целевые объекты также могут перекрывать LoS между передатчиками и приемниками. Помехи от некоторого передатчика могут воздействовать на целевой приемник только в том случае, если его антенна ориентирована в направлении целевого приемника и не заблокирована. Напомним, что модели компонентов рассматриваемого сценария, включая модели антенн, распространения и блокировки введены в главе 2.

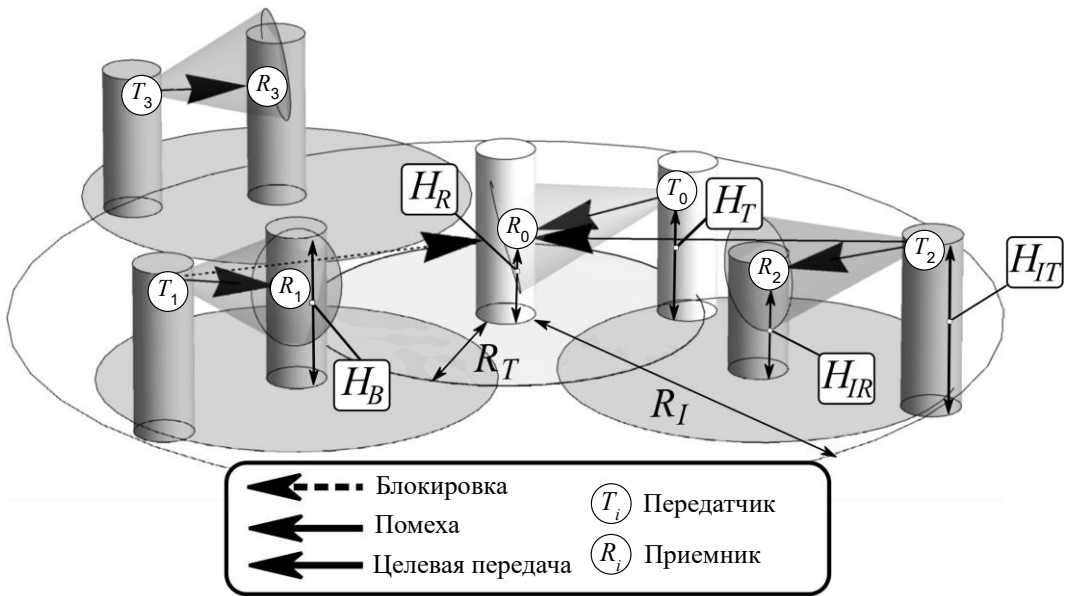


Рис. 4.3. Иллюстрация модели трехмерного сценария развертывания

В случае трехмерного развертывания начальные моменты суммарной помехи могут быть получены аналогично двумерному сценарию в следующем виде:

$$E[I^n] = \int_0^{R_I} E[I_1^n(r)] p_C(r) [1 - p_B(r)] 2\lambda\pi r dr, \quad (4.19)$$

где $2\lambda\pi r dr$ – вероятность наличия узла помех в приращении круга dr ; $p_c(r)$ – вероятность такого взаимоположения передающих и принимающей антенн, что источник помех вносит вклад в совокупную помеху на приемнике (вероятность экспозиции); $p_B(r)$ – вероятность того, что LoS блокируется; $E[I_1^n(r)]$ – начальные моменты помехи от одного источника помех на расстоянии r .

Найдем теперь неизвестные величины в формуле (4.19).

Утверждение 1. Начальные моменты мощности P_R^n сигнала на приемнике могут быть рассчитаны по формуле

$$E[P_R^n] = A^n 2^{\frac{1}{2}-\zeta n} [W(\mu_T, \mu_R) + W(\mu_R, \mu_T)] \times \\ \times \frac{\pi^{\frac{3}{2}} \csc\left(\frac{\pi\zeta n}{2}\right) \sec\left(\frac{\pi\zeta n}{2}\right) R_T^{-\left(\frac{n\zeta-5}{2}\right)}}{\mu_R^2 \mu_T^2 (\mu_R + \mu_T) \Gamma\left(\frac{n\zeta}{2}\right) \Gamma\left(\frac{n\zeta-1}{2}\right)}, \quad (4.20)$$

где

$$W(x, y) = x^3 \left(\begin{aligned} & 2\sqrt{2} y^{\zeta n} R_T^{\frac{n\zeta+1}{2}} + R_T^2 2^{\frac{\zeta n}{2}} y^{\frac{n\zeta+3}{2}} \Gamma\left(\frac{n\zeta-1}{2}\right) \times \\ & \times \left(\cos\left(\frac{\pi\zeta n}{2}\right) H_{\frac{3-n\zeta}{2}}^{yR_T} - J_{\frac{n\zeta-3}{2}}^{yR_T} - \sin\left(\frac{\pi\zeta n}{2}\right) J_{\frac{3-n\zeta}{2}}^{yR_T} \right) \end{aligned} \right). \quad (4.21)$$

Доказательство. Запишем с.в. мощности принимаемого сигнала в виде

$$P_R = A \left(\sqrt{(H_T - H_R)^2 + r^2} \right)^{-\zeta}, \quad (4.22)$$

где H_T , H_R , и r являются с.в.

Учитывая, что величина $|H_T - H_R|$ может быть найдена как модуль разности экспоненциально распределенных случайных величин, получаем следующую формулу для ее плотности вероятности в виде

$$f_{|H_T - H_R|}(y) = \frac{(e^{-y\mu_R} + e^{-y\mu_T}) \mu_R \mu_T}{\mu_R + \mu_T}, \quad y > 0. \quad (4.23)$$

Используя (4.23), начальные моменты $E[P_R^n(r)]$ мощности сигнала на приемнике могут быть представлены в интегральной форме

$$E[P_R^n] = \int_0^{R_T} \int_0^\infty \frac{A^n (e^{-y\mu_R} + e^{-y\mu_T}) \mu_R \mu_T 2r}{(r^2 + y^2)^{\frac{n\zeta}{2}} (\mu_R + \mu_T) R_T^2} dy dr. \quad (4.24)$$

Вычисляя интегралы в (4.24), получаем (4.20) и утверждение доказано.

Сформулированный результат приводит к следствию, устанавливающему первое неизвестное $E[I_1^n(r)]$ в (4.19).

Следствие 1. Начальные моменты помехи для расстояния r между источником помехи и приемником имеют вид

$$E[I_1^n(r)] = \frac{[W_1(\mu_T) + W_1(\mu_R)] \left[(\mu_R + \mu_T) \Gamma\left(\frac{n\zeta}{2}\right) \right]}{2^{\frac{n\zeta+1}{2}} A^{n\pi^2} \mu_R \mu_T}, \quad (4.25)$$

где

$$W_1(x) = \left[\frac{r}{x} \right]^{\frac{1-n\zeta}{2}} \times \\ \times \left[2 J_{\frac{n\zeta-1}{2}}^{rx} \csc(n\pi\zeta) - J_{\frac{1-n\zeta}{2}}^{rx} \sec\left(\frac{n\pi\zeta}{2}\right) + \csc\left(\frac{n\pi\zeta}{2}\right) H_{\frac{1-n\zeta}{2}}^{rx} \right]. \quad (4.26)$$

Для того чтобы вычислить вероятность блокировки LoS, определим зону блокировки прямой видимости $ABCD$, как показано на рис. 4.4.

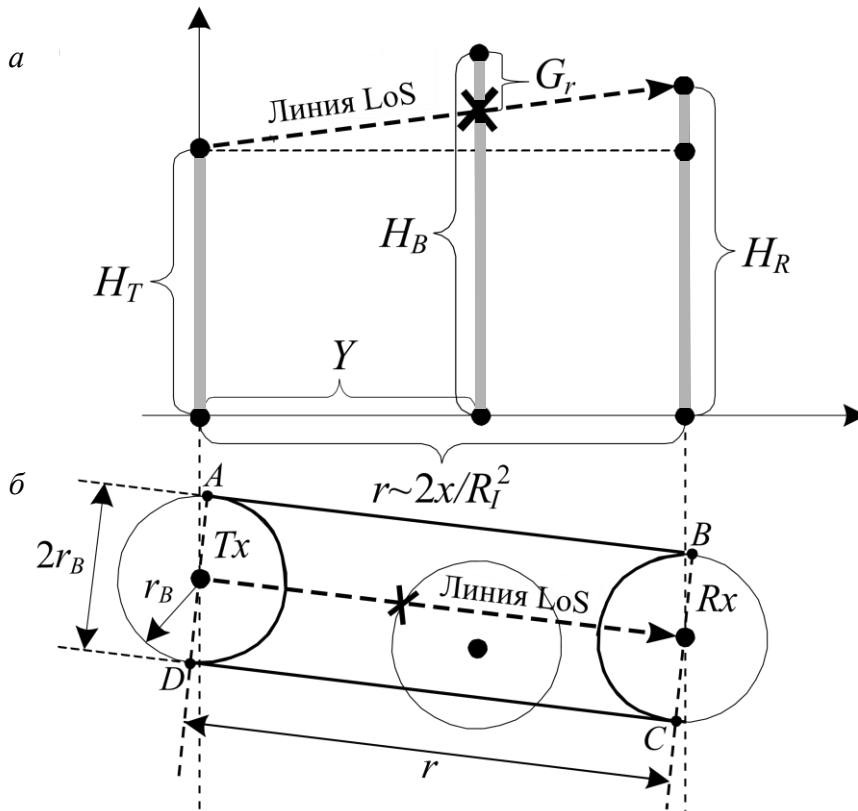


Рис. 4.4. Иллюстрация зоны блокировки LoS

Учитывая, что, если хоть один центр блокатора попадает в зону $ABCD$, LoS прерывается и тогда вероятность блокировки $p_B(r)$ можно определить следующим образом.

Утверждение 2. Вероятность блокировки линии прямой видимости в пуассоновском поле блокаторов с экспоненциально распределенными высотами определяется по формуле

$$p_B(r) = 1 - \left(\frac{\mu_R \mu_T}{(\mu_B + \mu_R)(\mu_B + \mu_T)} \right)^{\frac{2r_B r \lambda \mu_R \mu_T}{\mu_B(\mu_B + \mu_R + \mu_T)}}. \quad (4.27)$$

Доказательство. Пусть G_r , $0 < r < R_I$, является с.в., определяющей разницу между высотой LoS и высотой блокатора на расстоянии r от передающего узла. Для простоты предположим, что центры блокаторов, находящихся в зоне блокировки LoS, распределены равномерно на $[0, R_I]$. Полагая $H_R > H_T$, получаем

$$G_r = \frac{(H_R - H_T)Y}{r} + H_T - H_B, \quad (4.28)$$

где Y – расстояние между передатчиком и блокатором, как показано на рис. 4.4.

Вероятность, что блокатор, расположенный на расстоянии r от приемника, перекрывает линию прямой видимости, определяется по формуле

$$p_{B,1}(r) = 1 - Pr \left\{ \frac{(H_R - H_T)Y}{r} + H_T - H_B > 0 \right\}, \quad (4.29)$$

следовательно, можно найти вероятность общей блокировки по формуле

$$\begin{aligned} p_B(r) &= 1 - \sum_{i=0}^{\infty} \frac{(2\lambda r_B r)^i}{i! e^{2\lambda r_B r}} [1 - p_{B,1}(r)]^i = \\ &= 1 - e^{-2\lambda r_B r} - \sum_{i=1}^{\infty} \frac{(2\lambda r_B r)^i}{i! e^{2\lambda r_B r}} [1 - p_{B,1}(r)]^i, \end{aligned} \quad (4.30)$$

где $p_{B,1}(r) = Pr\{G_r - H_B > 0\}$ – неизвестная величина.

Вероятность $p_{B,1}(r)$ может быть найдена, используя преобразование

$$\begin{aligned} p_{B,1}(r) &= 1 - \int_0^r \int_0^\infty f_{\eta_1 \eta_4}(y_1, y_4) dy_1 dy_4 = \\ &= \frac{\mu_R \mu_T \log \left(\frac{\mu_R \mu_T}{(\mu_B + \mu_R)(\mu_B + \mu_T)} \right)}{\mu_B(\mu_B + \mu_R + \mu_T)}. \end{aligned} \quad (4.31)$$

Подставляя (4.31) в (4.30) и упрощая, получим (4.27).

Перейдем к определению вероятности экспозиции $p_c(r)$ (см. рис. 4.3), которая для расстояния r между приемником и источником помехи имеет вид

$$p_c(r) = p_V(r)p_H(r), \quad (4.32)$$

где $p_H(r)$ – вероятность того, что источник помехи оказывает влияние на приемник в горизонтальной плоскости, $p_V(r)$ – вероятность того, что источник помехи оказывает влияние на приемник в вертикальной плоскости.

Как можно видеть на рис. 4.5б, вероятность горизонтальной экспозиции определяется по формуле

$$p_H(r) = \frac{\alpha_{T,H}r}{2\pi r} \frac{\alpha_{R,H}r}{2\pi r} = \frac{\alpha_{T,H}\alpha_{R,H}}{4\pi^2}, \quad (4.33)$$

где $\alpha_{T,H}$ и $\alpha_{R,H}$ – углы направленности антенн передатчиков и приемников в горизонтальной плоскости. Иллюстрация вертикальной экспозиции показана на рис. 4.5а.

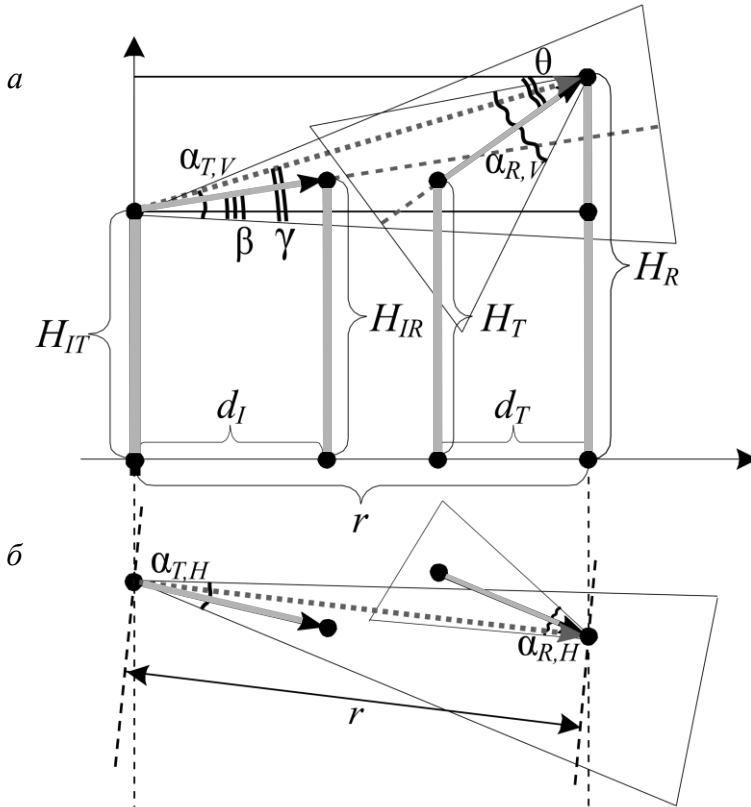


Рис. 4.5. Иллюстрация экспозиции в трех измерениях: а – вертикальная экспозиция; б – горизонтальная экспозиция

Утверждение 3. Вероятность вертикальной экспозиции в системе с направленными антеннами на передатчиках и приемниках определяется по формуле

$$p_V(r) = \int_{-\infty}^{\infty} \int_{y_4 - \frac{\alpha_{T,V}}{2}}^{y_4 + \frac{\alpha_{T,V}}{2}} \int_{y_4 - \frac{\alpha_{R,V}}{2}}^{y_4 + \frac{\alpha_{R,V}}{2}} f_{\gamma, \theta, \beta}(y_1, y_4, y_6) dy_1 dy_6 dy_4, \quad (4.34)$$

где $f_{\gamma, \theta, \beta}(y_1, y_4, y_6)$ – совместная плотность вероятности с.в. углов $\{\theta, \gamma, \beta\}$, $\alpha_{T,H}$ и $\alpha_{R,H}$ – углы вертикальной направленности антенн на передатчиках и приемниках.

Доказательство. Рассмотрим систему с.в. $\{\theta, \gamma, \beta\}$, где θ – угол между линией, соединяющей передатчик и приемник, и горизонтом, γ – угол между линией, соединяющей источник помехи и приемник, и β – угол между линией, соединяющей источник помехи и его приемник, и горизонт. Плотность распределения $f_{\gamma, \theta, \beta}(y_1, y_4, y_6)$ этих с.в. определяется с использованием метода преобразования с.в. и может быть записана в виде

$$f_{\eta}(y) = \iiint_{R^3} f_{\xi}(\phi_1(y^n), \dots, \phi_n(y^n)) |J| dy_2 dy_5 dy_3, \quad (4.35)$$

где Якобиан $|J|$ имеет вид

$$J = -ry_3 y_5 \sec^2(y_1) \sec^2(y_4) \sec^2(y_6). \quad (4.36)$$

Таким образом, подынтегральное выражение в (4.35) принимает вид

$$f_{\xi}(\phi^n(y^n)) = \frac{4y_3 y_5 \mu_R^2 \mu_T^2}{R_T^4} e^{-y_2 \mu_R - \mu_T(y_2 + y_3 \tan[y_1])} \times \\ \times e^{-\mu_T(y_2 + x \tan[y_4]) - \mu_R(y_4 - y_5 \tan[y_6])}. \quad (4.37)$$

и утверждение доказано.

4.1.3. Численные примеры

Рассмотрим теперь численные примеры, иллюстрирующие среднее значение помехи на приемнике при различных параметрах системы.

Двухмерная модель

На рис. 4.6 показана связь между коэффициентом поглощения K и средним значением помехи на приемнике, что обосновывает необходимость учета влияния молекулярного поглощения при

построении моделей помехи. Устанавливая угол диаграммы направленности в некоторое фиксированное значение, можно заметить, что среднее значение помехи должно уменьшаться с увеличением коэффициента поглощения. Аналогично, зафиксировав значение коэффициента поглощения, мы наблюдаем увеличение среднего значения помех на приемнике с увеличением направленности.

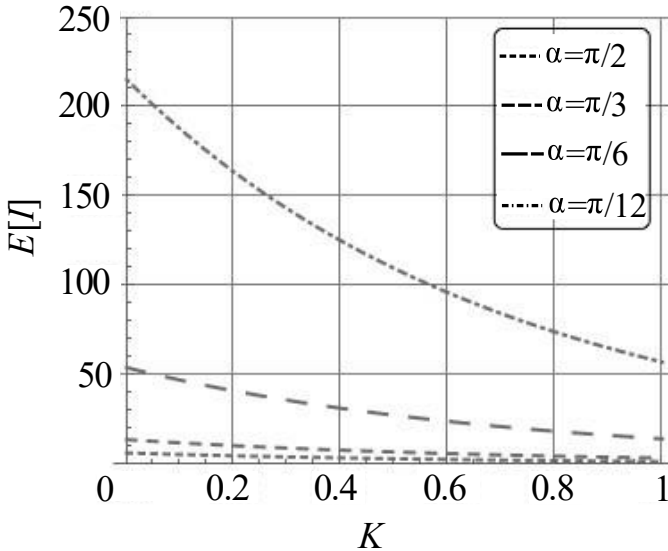


Рис. 4.6. Зависимость среднего значения помехи от коэффициента поглощения

На рис. 4.7 показана зависимость среднего значения помехи от угла направления для двух диаграмм направленности. Так как диаграмма направленности первого типа всю излучаемую мощность фокусирует в главном лепестке, график среднего значения помехи на приемнике для диаграммы направленности первого типа выше соответствующего графика для диаграммы второго типа. Однако заметим, что диаграмма направленности второго типа является более гибкой, так как зависит от коэффициента K , который определяет долю мощности, излучаемой в направлении боковых лепестков. Зависимость среднего значения помехи на приемнике от угла диаграммы направленности α для диаграммы направленности второго типа имеет принципиально иной вид, чем для диаграммы направленности первого типа. При уменьшении коэффициента K среднее значение помехи для диаграммы направленности второго типа приближается к среднему значению помехи для диаграммы направленности первого типа.

Величина коэффициента K может значительно изменяться в зависимости от реализации конкретной антенны. Во многих исследованиях коэффициент K принимает значения из диапазона $[0,1-0,2]$. Следовательно, использование диаграммы направленности второго типа является предпочтительным для точной оценки помех в миллиметровых и терагерцевых сетях связи.

В то же время использование диаграммы направленности первого типа значительно упрощает анализ и финальные выражения и, следовательно, является предпочтительным для грубых оценок.

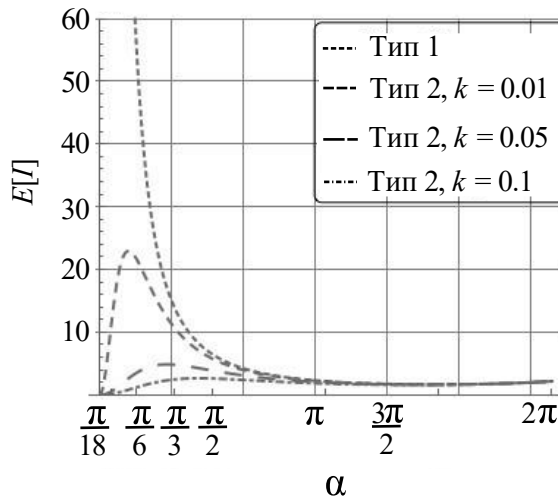


Рис. 4.7. Среднее значение помехи для двух диаграмм направленности

Анализируя полученные результаты, сделаем следующие выводы. Во-первых, среднее значение помехи для диаграммы направленности первого типа увеличивается с уменьшением угла излучения. Во-вторых, среднее значение помехи для диаграммы направленности первого типа уменьшается с увеличением коэффициента поглощения. В-третьих, поведение среднего значения помехи для диаграммы направленности второго типа принципиально отличается от поведения средней помехи диаграммы направленности второго типа. В то же время, когда параметр k стремится к нулю, разница между результатами также стремится к нулю.

Трёхмерная модель

Вероятность p_v вертикальной экспозиции является важным параметром, который отличает двухмерные и трёхмерные модели друг

от друга. Так как вычисление вероятности p_v является одним из сложных этапов расчета, ниже оценим точность ее возможных приближений с помощью имитационного моделирования.

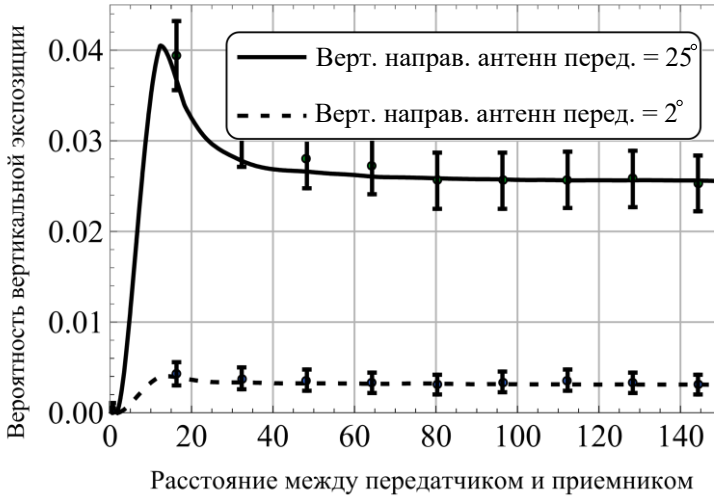


Рис. 4.8. Вероятность вертикальной экспозиции и сравнение с имитационным моделированием

Значение вероятности вертикальной экспозиции в зависимости от расстояния между передатчиком и приемником, а также от вертикальной направленности антенны передатчика показано на рис. 4.8, где также отражены результаты имитационного моделирования. Из графиков видно, что модель находится в согласии с данными имитационного моделирования, а значения располагаются в пределах доверительных интервалов. Также видно, что доверительные интервалы увеличиваются с уменьшением направленности антенны.

4.2. Отношение сигнал–шум в сетях с направленными антенными решетками

Рассмотрим метод получения характеристик для оценки качества канала связи, в том числе ОСШ, где в качестве шума выступает помеха. Сначала рассмотрим более простой для анализа двумерный сценарий развертывания, а затем перейдем к трехмерному сценарию. Рассмотренные ниже методы могут использоваться также и для оценки скоростей передачи в каналах связи в плотных развертываниях сетей 5G/6G.

Для оценки ОСШ требуется найти плотность распределения с.в. помехи. На первом этапе методом преобразования случайных

величин находим распределение помехи для одного узла $Ar_i^{-2}e^{-Kri}$, а затем распределение суммарной помехи на приемнике $\sum_{i=1}^N Ar_i^{-2}e^{-Kri}$ как функцию от случайного числа N интерферирующих узлов.

При работе с миллиметровым и терагерцовым диапазонами частот применение описанного выше подхода сопряжено с рядом сложностей. Во-первых, распределение незаблокированных интерферирующих узлов в круге не является однородным, так как вероятность блокировки $p_B(x)$ имеет зависимость от расстояния между принимающим и интерферирующим устройствами. Следовательно, вероятность блокировки интерферирующего узла снижается в зависимости от увеличения расстояния между интерферирующим и приемным устройствами. Эта особенность изменяет структуру распределения для оценки помехи и не позволяет получить аналитические выражения в замкнутой форме.

Ниже для оценки моментов ОСШ предлагается метод, заключающийся в применении разложения функции случайных величин в ряд Тейлора.

4.2.1. Двухмерный сценарий развертывания

Как и в случае оценки помехи, рассмотрим двухмерный сценарий развертывания системы, иллюстрированный на рис. 4.1. Для получения оценки ОСШ предлагается использование разложения функционального преобразования с.в. в ряд Тейлора. Напомним, что математическое ожидание с.в. $Y = g(X)$, где X с.в. со средним значением μ_0 и дисперсией $\sigma^2[X]$, может быть представлено в виде

$$E[Y] = g(\mu_0) + \frac{g''(\mu_0)}{2}, \quad (4.38)$$

а дисперсия имеет вид

$$\sigma^2[Y] = [g'(\mu_0)]^2 \sigma^2[X] - \frac{1}{4} [g''(\mu_0)]. \quad (4.39)$$

Тогда ОСШ может быть определено как

$$g(x) = \frac{C}{x}, \quad (4.40)$$

где $C = Ar_0^{-2}e^{-Kr_0}$ – мощность полезного сигнала на приемнике. Напомним, что первые две производные функции (4.40) имеют вид

$$g'(x) = -\frac{C}{x^2}, g''(x) = \frac{2C}{x^3}. \quad (4.41)$$

Дисперсия помехи

Так как математическое ожидание $E[I]$ помехи было получено выше, найдем дисперсию $\sigma^2[I] = E[I^2] - (E[I])^2$. Далее для простоты детально рассмотрен случай направленности антенны передатчика без учета эффекта блокировки. Для остальных случаев даны только финальные выражения.

Для случая без блокировки второй начальный момент помехи определяется по формуле

$$\begin{aligned} E[I^2] &= \int_{r_B}^R \left(A r^{-2} e^{-Kr} e^{-\lambda_I(x-r_B)r_B} \frac{\alpha}{2\pi} \right)^2 2\lambda_I \pi r dr = \\ &= \frac{A^2 \alpha^2 \lambda_I}{2\pi} \Phi(R, r_B, \lambda_I, K), \end{aligned} \quad (4.42)$$

где

$$\begin{aligned} \Phi(R, r_B, \lambda_I, K) &= \\ &= \frac{4R^2(K+\lambda_I r_B)^2 e^{2R(K+\lambda_I r_B)} Ei(2R(r_B \lambda_I + K)) + 2KR + 2\lambda_I R r_B - 1}{2R^2 e^{-2\lambda_I r_B(r_B - R)} - 2KR} \Big|_{r_B}^R. \end{aligned} \quad (4.43)$$

В (4.43) обозначение $\Big|_{r_B}^R$ определяет пределы первообразной функции.

Далее приведены формулы для вычисления дисперсии с применением коэффициента $\Phi_1(R, r_B, K)$, который вычисляется по формуле

$$\Phi_1(R, r_B, K) = 2K^2 Ei(-2KR) + e^{-2KR} \left(\frac{K}{R} - \frac{1}{2R^2} \right) \Big|_{r_B}^R. \quad (4.44)$$

С учетом приведенных выше формул дисперсия помехи без блокировки для всенаправленных антенн на передающей и приемной сторонах может быть рассчитана по формуле

$$\sigma^2[I] = 2\pi A^2 \lambda_I \Phi_1(R, r_B, K), \quad (4.45)$$

а дисперсия помехи с блокировкой для всенаправленных антенн имеет вид

$$\sigma^2[I] = 2\pi A^2 \lambda_I \Phi(R, r_B, \lambda_I, K). \quad (4.46)$$

Дисперсия помехи без блокировки для направленной антенны на приемной либо передающей стороне имеет вид

$$\sigma^2[I] = A^2 \alpha \lambda_I \Phi_1(R, r_B, K). \quad (4.47)$$

Дисперсия помехи с блокировкой для направленной антенны на приемной либо передающей стороне имеет вид

$$\sigma^2[I] = A^2 \alpha \lambda_I \Phi(R, r_B, \lambda_I, K). \quad (4.48)$$

Дисперсия помехи без блокировки для направленных антенн в обеих передающей и приемной сторонах имеет вид

$$\sigma^2[I] = \frac{A^2 \alpha \lambda_I}{2\pi} \Phi_1(R, r_B, K). \quad (4.49)$$

Дисперсия помехи с блокировкой для случая направленных антенн как на передающей, так и приемной стороне имеет вид

$$\sigma^2[I] = \frac{A^2 \alpha \lambda_I}{2\pi} \Phi(R, r_B, \lambda_I, K). \quad (4.50)$$

Перечисленные выражения используются далее для получения оценки ОСШ.

Математическое ожидание ОСШ

Математическое ожидание ОСШ в случае с блокировкой для направленных антенн либо на приемной, либо на передающей стороне имеет вид

$$E[S] = \frac{e^{-Kr_0 r_0^{-2}}}{\frac{\alpha^2 \lambda_I}{2\pi} \theta(R, r_B, \lambda_I, K)} + \frac{e^{-Kr_0 r_0^{-2}} \left[\frac{\alpha \lambda_I}{2\pi} \Phi(R, r_B, \lambda_I, K) - \left[\frac{\alpha^2 \lambda_I}{2\pi} \theta(R, r_B, \lambda_I, K) \right]^2 \right]}{\left[\frac{\alpha^2 \lambda_I}{2\pi} \theta(R, r_B, \lambda_I, K) \right]^3}. \quad (4.51)$$

В случае без блокировки математическое ожидание ОСШ определяется по формуле

$$E[S] = \frac{e^{-Kr_0 r_0^{-2}}}{\frac{\alpha^2 \lambda_I}{2\pi} \theta_1(R, r_B, \lambda_I, K)} + \frac{e^{-Kr_0 r_0^{-2}} \left[\frac{\alpha \lambda_I}{2\pi} \Phi_1(R, r_B, K) - \left[\frac{\alpha^2 \lambda_I}{2\pi} \theta_1(R, r_B, \lambda_I, K) \right]^2 \right]}{\left[\frac{\alpha^2 \lambda_I}{2\pi} \theta_1(R, r_B, \lambda_I, K) \right]^3}. \quad (4.52)$$

Аналогично математическое ожидание ОСШ для случая направленных антенн как на передающей, так и на приемной стороне имеет вид

$$E[S] = \frac{e^{-Kr_0 r_0^{-2}}}{\alpha \lambda_I \theta(R, r_B, \lambda_I, K)} + \frac{e^{-Kr_0 r_0^{-2}} \left[\alpha \lambda_I \Phi(R, r_B, \lambda_I, K) - \left[\alpha \lambda_I \theta(R, r_B, \lambda_I, K) \right]^2 \right]}{\left[\alpha \lambda_I \theta(R, r_B, \lambda_I, K) \right]^3}. \quad (4.53)$$

Без учета эффекта блокировки математическое ожидание ОСШ имеет вид

$$E[S] = \frac{e^{-Kr_0 r_0^{-2}}}{\alpha \lambda_I \theta_1(R, r_B, K)} + \frac{e^{-Kr_0 r_0^{-2}} \left[\alpha \lambda_I \Phi_1(R, r_B, K) - \left[\alpha \lambda_I \theta_1(R, r_B, K) \right]^2 \right]}{\left[\alpha \lambda_I \theta_1(R, r_B, K) \right]^3}. \quad (4.54)$$

4.2.2. Трехмерный сценарий развертывания

Разложение в ряд Тейлора

Используем разложение Тейлора для функции $g(x, y) = P_R/I$ для получения математического ожидания с.в. ОСШ. Получаем приближение второго порядка на основе разложения $g(x, y)$ в окрестности $\boldsymbol{\mu} = (E[P_R], E[I]) = (\mu_{P_R}, \mu_I)$ в виде

$$E[g(\boldsymbol{\mu})] \approx g(\boldsymbol{\mu}) + \frac{g''_{xx}(\boldsymbol{\mu})\sigma_{P_R}^2 + 2g''_{xy}K_{P_R,I} + g''_{yy}(\boldsymbol{\mu})\sigma_I^2}{2}, \quad (4.55)$$

где $K_{P_R,I}$ – ковариация между P_R и I ; $\sigma_{P_R}^2$ и σ_I^2 – дисперсии P_R и I соответственно.

Замечая, что

$$g''_{xx}(x, y) = 0, \quad g''_{xy}(x, y) = -y^{-2}, \quad g''_{yy}(x, y) = \frac{2x}{y^2}, \quad (4.56)$$

приходим к следующей приближенной формуле:

$$E[P_R/I] \approx \frac{\mu_{P_R}}{\mu_I} - \frac{K_{P_R,I}}{\mu_I^2} + \frac{\sigma_I^2 \mu_{P_R}}{\mu_I^3}. \quad (4.57)$$

Таким образом, необходимо определить два неизвестных – второй начальный момент помехи и ковариацию между принимаемым сигналом и помехой. Отметим, что в отличие от двухмерного сценария между этими с.в. существует зависимость.

Начальные моменты и ковариация

Аналогично двумерному сценарию развертывания второй начальный момент суммарной помехи получаем с использованием формулы Кэмпбелла. При этом используем следующее выражение при $n = 2$:

$$E[I^n] = \int_0^{R_I} E[I_1^n(r)] p_C(r) [1 - p_B(r)] 2\lambda\pi r dr. \quad (4.58)$$

В формуле (4.58) $2\lambda\pi r dr$ – это вероятность того, что присутствует интерферирующее устройство в приращении круга dr , и $p_C(r)$ – вероятность экспозиции. Величина $p_B(r)$ – это вероятность возникновения блокировки LoS некоторыми узлами, а величина $E[I_1^n(r)]$ является начальным моментом мешающего сигнала от одного источника помех, определенного расстоянием между интерферирующим источником и приемником.

Еще одним неизвестным, требующим определения, является ковариация между помехой и мощностью полезного сигнала, определяемая формулой

$$K_{P_R, I} = E[P_R I] - \mu_{P_R} \mu_I, \text{ где } E[P_R I] = E[AX_0^{-\zeta} \sum_{i=1}^N AX_i^{-\zeta}], \quad (4.59)$$

где X_0 – расстояние между целевым передатчиком и приемником; X_i , $i = 1, 2, \dots, N$, – расстояние между целевым приемником и i -м интерферирующим передатчиком; N – число передающих устройств.

Применяя тождество Вальда, получаем

$$E[P_R I] = A^2 E[N] E[(X_i X_0)^{-\zeta}], \quad (4.60)$$

где $E[N] = \lambda \pi R_I^2 p_C$, $E[(X_0 X_i)^\zeta]$ – неизвестные величины.

Величину $E[(X_0 X_i)^\zeta]$ можно представить в виде

$$E\left[\left[(H_R - H_T)^2 - r_0\right]\left[(H_R - H_i)^2 - r_i\right]^{-\frac{\zeta}{2}}\right], \quad (4.61)$$

где r_0 – константа, и тогда

$$\begin{aligned} E[P_R I] &= A^2 \lambda \pi R_I^2 p_C E[(X_i X_0)^{-\zeta}] = \\ &= \int_0^{R_I} \int_0^{R_T} \int_0^\infty \int_0^\infty \int_0^\infty \frac{[(x_1 - x_2)^2 + x_3^2]^{-\frac{\zeta}{2}}}{[(x_1 - x_3)^2 + x_4^2]^{\frac{\zeta}{2}}} f(x_1, \dots, x_5) dx_1 \dots dx_5, \end{aligned} \quad (4.62)$$

где $E[(X_0 X_i)^{-\zeta}]$ может быть получено численным интегрированием.

Очевидно, что двухмерная модель является частным случаем трехмерной модели, но при этом трехмерная модель содержит ряд вычислительных сложностей.

Упрощенные модели могут подходить для приближения ОСШ лишь в определенных диапазонах параметров системы. Для таких моделей чувствительность и направленность антенн не будут такими высокими в силу малого числа антенных элементов. Если ввести предположение, что все принимающие устройства оснащены все-направленными антеннами, то можно дополнительно упростить сложность вычислений. Кроме того, отметим, что во многих реальных развертываниях систем высоты взаимодействующих объектов фиксированы и известны заранее, что значительно упрощает модель. Если же считать, что высоты являются случайными, то оценку также можно получить с применением их средних значений.

4.2.3. Численные примеры

Двухмерный сценарий

На рис. 4.9 показана зависимость ОСШ от плотности интерферирующих узлов при наличии блокировки. Значение ОСШ снижается при росте числа интерферирующих узлов для всех исследуемых типов антенн как на передатчике, так и на приемнике. Заметим, что направленные антенны оказывают сильное влияние на значение ОСШ во всем диапазоне плотности интерферирующих узлов.

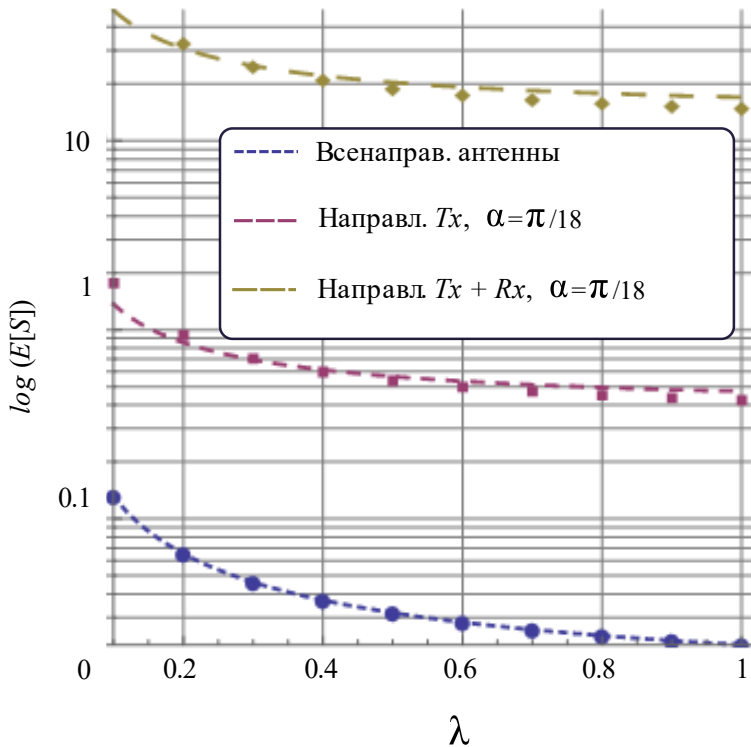


Рис. 4.9. Зависимость ОСШ от плотности интерферирующих узлов

Из графиков на рис. 4.10 видно, что ОСШ существенно улучшается при уменьшении угла диаграммы направленности. Иными словами, использование антенных систем с более высокой направленностью позволяет существенно улучшить ОСШ на приемнике. Аналогичные зависимости наблюдаются и в отсутствие блокировки.

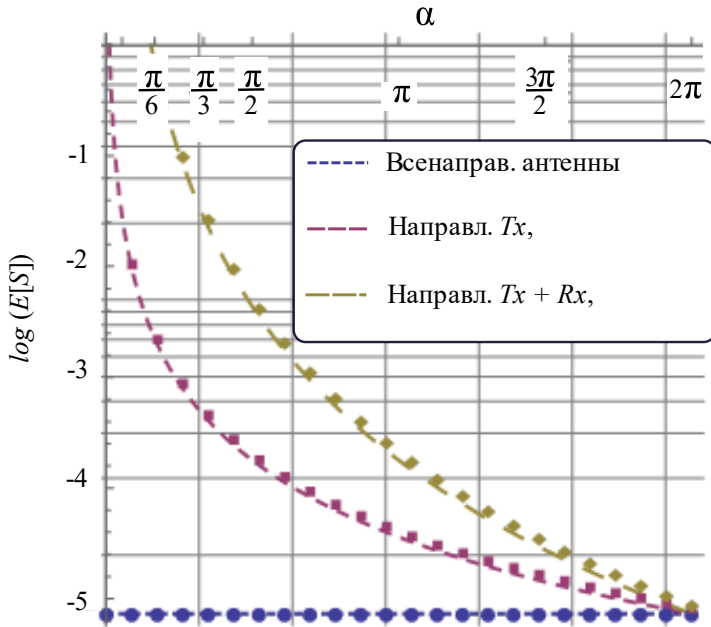


Рис. 4.10. Зависимость ОСШ от угла диаграммы направленности

Трёхмерный сценарий

Ниже рассмотрены комбинированные модели, которые будем обозначать следующим образом. Модели с фиксированными или случайными значениями высот – ФВ/СВ, модели, которые могут как учитывать, так и не учитывать вероятность вертикальной экспозиции, – ±ВЭ. Следуя таким обозначениям, трёхмерную модель можно записать как СВ+ВЭ. Напомним, что предположение о ФВ и/или исключение ВЭ позволяет рассмотреть большее количество частных случаев.

На рис. 4.11-4.13 проиллюстрированы средние значения ОСШ для разных типов моделей в зависимости от двух параметров – вертикальной направленности антенн передатчиков и интенсивности (плотности) пар взаимодействующих устройств. Из графиков на рис. 4.11 видно, что с ростом плотности взаимодействующих пар λ снижается среднее значение ОСШ для всех рассмотренных моделей, причем во всем диапазоне значений λ ни одна из упрощенных моделей не обеспечивает приближения величины ОСШ с достаточной точностью.

Наиболее близкой к трёхмерной модели в рассматриваемом диапазоне переменных является модель, которая не учитывает ВЭ, но

учитывает СВ объектов. Модель, которая является наименее точной, не учитывает ВЭ и имеет ФВ.

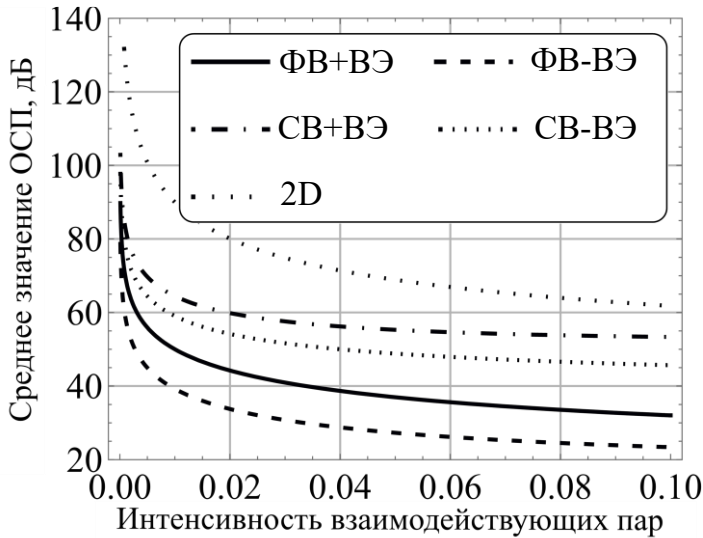


Рис. 4.11. Эффект изменения плотности пар взаимодействующих устройств

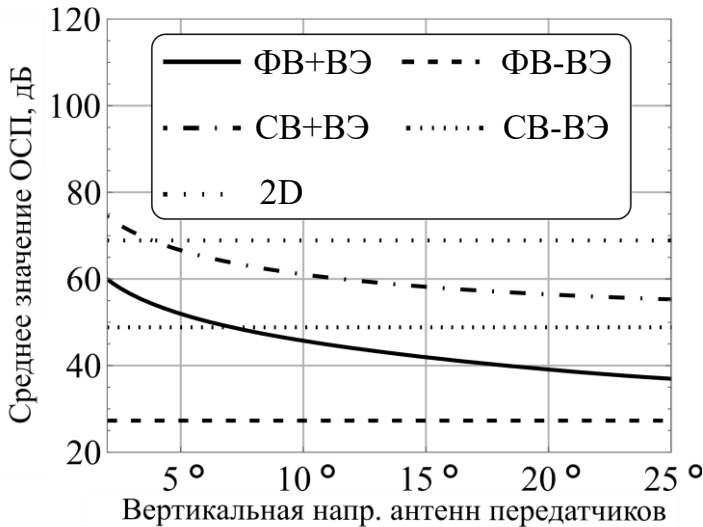


Рис. 4.12. Эффект изменения углов направленности антенн $\alpha_{T,V}$, $\alpha_{T,H} = 2^\circ$

Обратим внимание на рис. 4.12, где проиллюстрирована оценка ОСП как функция вертикальной направленности антенн для двух значений горизонтальной направленности антенн передатчиков, $\alpha_{T,H} = 2^\circ$ и $\alpha_{T,H} = 25^\circ$ соответственно. Здесь для всех значений вертикальной направленности антенн модели с ФВ дают резкие

отклонения от данных, полученных для модели СВ+ВЭ. Для двух рассматриваемых значений антенн горизонтальной направленности эти отклонения составляют не менее 10 дБ. Анализ графиков на рис. 4.12 показывает, что модель, которая не учитывает вероятность ВЭ, предоставляет достаточное приближение при больших значениях вертикальной направленности антенн в диапазоне (1–5°). Такое предположение верно до тех пор, пока направленность антенн не достигает 20° и более. В этом диапазоне другие упрощенные модели также не позволяют получить оценки ОСШ, обладающие достаточно высокой точностью.

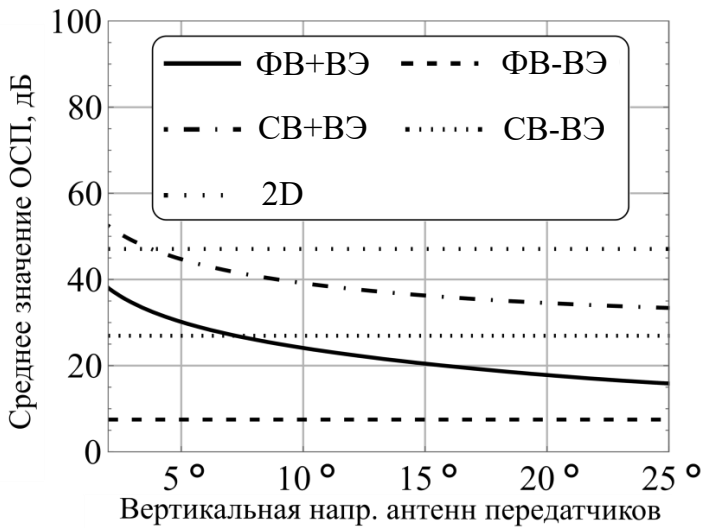


Рис. 4.13. Эффект изменения углов направленности антенн $\alpha_{т,в}, \alpha_{т,н} = 25^\circ$


Обратим внимание на двухмерную модель, которая позволяет получить хорошую оценку верхней границы значения ОСШ во всем диапазоне параметров системы. Отметим, что модель с заранее известными высотами всегда дает оценку нижней границы значений ОСШ, хотя и не учитывает вертикальную направленность. Разница для ОСШ между двумя такими моделями может достигать 60 дБ. Следовательно, ни одна из моделей не может обеспечить точную границу оценки для ОСШ. Модель с фиксированной высотой и наличием вертикальной направленности позволяет получить разумную погрешность, хотя для некоторых диапазонов параметров отклонение по ОСШ может достигать более 10 дБ.

Представленные результаты численного анализа дополняют выводы в предыдущих параграфах данной главы о зависимости

среднего значения помехи от параметров системы. В частности, если угол диаграммы направленности уменьшается, то среднее значение помехи растет. Эти результаты должны привести к монотонному уменьшению среднего значения ОСШ, когда направленность антенны увеличивается. Однако проведенный анализ также показал, что эффекты распространения миллиметровых и терагерцовых сигналов в двух- и трехмерных сценариях развертывания оказывают более сильное влияние на полезный сигнал, чем на помеху, что приводит к увеличению ОСШ.


Используя модель трехмерного сценария развертывания, мы численно исследовали величину ОСШ в зависимости от параметров системы. Простые двумерные модели дают оптимистические результаты оценки ОСШ, которые могут отличаться от реальных значений на величину до 20 дБ и более.

Подводя итог, отметим, что простейшим способом повышения ОСШ в беспроводных сетях миллиметрового и терагерцового диапазонов частот является использование антенн с высокой направленностью.



ЧАСТЬ II

МОДЕЛИ И МЕТОДЫ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ОБСЛУЖИВАНИЯ АБОНЕНТОВ В СЕТЯХ 5G/6G



ГЛАВА 5

МЕХАНИЗМЫ ПОВЫШЕНИЯ КАЧЕСТВА ОБСЛУЖИВАНИЯ АБОНЕНТОВ СЕТЕЙ 5G/6G – МУЛЬТИСВЯЗНОСТЬ И РЕЗЕРВИРОВАНИЕ РЕСУРСОВ

В данной главе изложен математический аппарат, позволяющий проводить анализ показателей эффективности беспроводных сетей 5G/6G с учетом динамики поступления запросов сессий пользователей с использованием теории ресурсных СМО (ReСМО). Заметим, что для обозначения ресурсных СМО в англоязычных источниках используется англоязычная аббревиатура ReLS (Resource Loss System). Основное отличие ReСМО от классических СМО в том, что заявкам требуется не только прибор, моделирующий, как правило, время обслуживания сессий (время обслуживания заявок), но и конечный ресурс, моделирующий, как правило, случайное требование сессии к радиоресурсам, необходимым для ее обслуживания. Именно это отличие позволяет учитывать неоднородность требований сессий к частотному ресурсу БС, возникающих из-за случайного расположения устройств в зоне обслуживания, и, как следствие, случайной спектральной эффективности радиоканала.

На основе принципов моделирования компонентов сетей с направленными антеннами и блокировками, представленными в главе 2, в первом параграфе данной главы предложен метод моделирования процесса обслуживания трафика сетей 5G/6G на основе совместного применения методов стохастической геометрии и теории ReLS. Предложенный метод закладывает основу более сложных моделей, с помощью которых проводится анализ предлагаемых алгоритмов и методов повышения качества обслуживания абонентов в последующих разделах главы. Также мы рассмотрим методы

резервирования ресурсов, методы мультисвязности и совместное использование этих двух методов.

Отметим, что настоящая глава не претендует на полное изложение применения теории РеСМО. Здесь будут представлены только основные результаты, которые наиболее часто используются при анализе типичных сценариев сетей радиодоступа 5G/6G.

5.1. Базовая ресурсная модель

Продолжим анализ ресурсной модели, введенной в главе 3 и рассмотрим процесс обслуживания трафика на отдельной БС в сетях 5G/6G. Введем следующие предположения: а) радиус обслуживания БС выбран так, что не происходит прерываний уже установленных соединений; б) максимально возможное количество активных сессий может быть ограничено; в) ресурс ограничен и измеряется в первичных ресурсных блоках NR; г) интенсивность поступления запросов на установление сессий (для краткости будем также говорить «поступление сессий») в зоне обслуживания фиксирована и известна; д) расположение пользователей с активными сессиями распределено равномерно в зоне обслуживания БС; е) трафик является неэластичным.

Рассмотрим СП $X_2(t) = \{\xi(t), \delta(t)\}$, в котором, в отличие от СП $X_1(t)$, введенного в главе 3, вторая компонента $\delta(t)$ обозначает суммарный объем ресурсов, занятый всеми обслуживаемыми заявками. Такое упрощение приводит, с одной стороны, к снижению размерности пространства состояний СП, а с другой стороны, к потере части важной информации о нем. В том числе невозможно точно сказать, какой объем ресурсов освободит заявка при уходе из системы. В этом случае предлагается использовать следующий подход. Пусть система находится в состоянии (k, r) в момент перед уходом заявки из системы. Тогда вероятность того, что уходящая заявка освободит j ресурсов, можно определить по формуле Байеса: $\frac{p_j p_{r-j}^{(k-1)}}{p_r^{(k)}}$. Запишем СУР для стационарных вероятностей $q_k(r)$ процесса $X_2(t)$ в виде

$$\lambda q_0 \sum_{r=0}^R p_r = \mu \sum_{r=0}^R q_1(r),$$

$$q_k(r) (\lambda \sum_{r=0}^{R-r} p_r + k\mu) = \lambda \sum_{j=0}^r q_{k-1}(j) p_{r-j} + \quad (5.1)$$

$$+(k+1)\mu \sum_{j=0}^{R-r} q_{k+1}(r+j) \frac{p_j p_r^{(k)}}{p_{r+j}^{(k+1)}}, 1 \leq k \leq N-1, 0 \leq r \leq R,$$

$$N\mu q_N(r) = \lambda \sum_{j=0}^r q_{N-1}(j) p_{r-j}, 0 \leq r \leq R.$$

В более сложных системах, как правило, не удается получить аналитические выражения для стационарных вероятностей, что приводит к необходимости численно решать СУР. В таком случае принципиальное значение имеет размерность пространства состояний. Поэтому предложенное упрощение процесса позволяет рассчитывать характеристики системы за разумное время численно.

Следует заметить, что равенство вероятностей укрупненных состояний исходного процесса (глава 3) и вероятностей упрощенного процесса (глава 5) не является доказанным для более сложных систем. Однако это позволяет обоснованно предполагать, что данное упрощение дает хорошее приближение с допустимой погрешностью. На практике это предположение необходимо проверять с помощью имитационного моделирования.

5.1.1. Агрегирование потоков

Рассмотрим ReLS с несколькими потоками заявок. В отличие от базовой РеСМО, на систему поступает L пуассоновских потоков заявок с интенсивностями λ_l , $l = 1, 2, \dots, L$. Заявки l -го типа обслуживаются с интенсивностями μ_l , а дискретное распределение их требований к ресурсам обозначим $\{p_{l,r}\}_{r \geq 0}$. В этом случае стационарные вероятности $Q_{k_1, \dots, k_L}(r_1, \dots, r_L)$ того, что в системе k_l заявок типа l занимают суммарно r_l единиц ресурсов (для краткости будем говорить r_l ресурсов), $l = 1, 2, \dots, L$, определяются формулой

$$Q_{k_1, \dots, k_L}(r_1, \dots, r_L) = Q_0 \frac{\rho_1^{k_1}}{k_1!} \dots \frac{\rho_L^{k_L}}{k_L!} p_{1,r_1}^{(k_1)} \dots p_{L,r_L}^{(k_L)}, \sum_{l=1}^L k_l \leq N, \sum_{l=1}^L r_l \leq R;$$

$$Q_0 = \left(1 + \sum_{1 \leq k_1 + \dots + k_L \leq N} \sum_{0 \leq r_1 + \dots + r_L \leq R} Q_{k_1, \dots, k_L}(r_1, \dots, r_L)\right)^{-1}, \quad (5.2)$$

где $\rho_l = \frac{\lambda_l}{\mu_l}$, $l = 1, 2, \dots, L$. Заметим, что выражения (5.2) малоприспособны для эффективных вычислений. Далее покажем, что стационарные

вероятности ReLS с L потоками заявок равны стационарным вероятностям ReLS с одним агрегированным потоком, имеющим средневзвешенное распределение требований к ресурсам.

Для целых положительных k введем последовательность $G_{L,k,r}$, $r \geq 0$:

$$G_{L,k,r} = \sum_{\substack{k_1, \dots, k_L \geq 0 \\ k_1 + \dots + k_L = k}} \frac{\rho_1^{k_1}}{k_1!} \cdots \frac{\rho_L^{k_L}}{k_L!} \sum_{\substack{r_1, \dots, r_L \geq 0 \\ r_1 + \dots + r_L = r}} p_{1,r_1}^{(k_1)} \cdots p_{L,r_L}^{(k_L)}, \quad (5.3)$$

и рассмотрим производящую функцию (ПФ) этой последовательности:

$$G_{L,k}(z) = \sum_{r=0}^{\infty} z^r \sum_{\substack{k_1, \dots, k_L \geq 0 \\ k_1 + \dots + k_L = k}} \frac{\rho_1^{k_1}}{k_1!} \cdots \frac{\rho_L^{k_L}}{k_L!} \sum_{\substack{r_1, \dots, r_L \geq 0 \\ r_1 + \dots + r_L = r}} p_{1,r_1}^{(k_1)} \cdots p_{L,r_L}^{(k_L)}. \quad (5.4)$$

Введем $P_l(z)$ ПФ распределения требований к ресурсам заявок типа l и тогда получаем

$$G_{L,k}(z) = \sum_{\substack{k_1, \dots, k_L \geq 0 \\ k_1 + \dots + k_L = k}} \frac{\rho_1^{k_1}}{k_1!} \cdots \frac{\rho_L^{k_L}}{k_L!} P_1^{k_1}(z) \cdots P_L^{k_L}(z). \quad (5.5)$$

Наконец, обозначив $\rho = \rho_1 + \dots + \rho_L$ и воспользовавшись полиномиальной теоремой, окончательно получим

$$\begin{aligned} G_{L,k}(z) &= \\ &= \frac{\rho^k}{k!} \sum_{\substack{k_1, \dots, k_L \geq 0 \\ k_1 + \dots + k_L = k}} \frac{k!}{k_1! \cdots k_L!} \left(\frac{\rho_1}{\rho}\right)^{k_1} \cdots \left(\frac{\rho_L}{\rho}\right)^{k_L} P_1^{k_1}(z) \cdots P_L^{k_L}(z) = \\ &= \frac{\rho^k}{k!} \left(\frac{\rho_1}{\rho} P_1(z) + \cdots + \frac{\rho_L}{\rho} P_L(z)\right)^k. \end{aligned} \quad (5.6)$$

Из (5.6) вытекает, что у системы с несколькими входящими потоками совместное распределение суммарного числа заявок в системе и объемов занятых ими ресурсов является таким же, как у системы с одним входящим потоком и средневзвешенными требованиями к ресурсу. Иными словами, стационарные вероятности $Q_k(r)$ того, что в системе обслуживается k заявок всех типов, которые вместе занимают r единиц ресурсов, определяются формулой

$$\begin{aligned} Q_k(r) &= Q_0 \frac{\rho^k}{k!} \bar{p}_r^{(k)}, \quad 1 \leq k \leq N, r \leq R, \\ Q_0 &= \left(1 + \sum_{k=1}^N \sum_{r=0}^R Q_k(r)\right)^{-1}, \end{aligned} \quad (5.7)$$

где $\rho = \rho_1 + \rho_2 + \dots + \rho_L$, а вероятности \bar{p}_r имеют вид

$$\bar{p}_r = \sum_{l=1}^L \frac{\rho_l}{\rho} p_{l,r}. \quad (5.8)$$

С учетом (5.8) вероятность блокировки заявки типа l может быть вычислена по формуле

$$\pi_{b,l} = 1 - \sum_{k=0}^{N-1} \sum_{r=0}^R Q_k(r) p_{l,R-r}. \quad (5.9)$$

5.1.2. Рекуррентный алгоритм

Вычисление характеристик РеСМО осложняется необходимостью расчета многократных сверток распределения требований заявок к ресурсам. Для снижения сложности расчетов ниже предложен рекуррентный алгоритм, позволяющий вычислять характеристики модели без непосредственного расчета стационарных вероятностей.

Для целых неотрицательных n и r введем функции $G(n, r) = \sum_{k=0}^n \frac{\rho^k}{k!} \sum_{j=0}^r p_j^{(k)}$. Согласно этому определению и формуле (5.7), нормирующая константа принимает вид $Q_0 = G(N, R)^{-1}$. Нетрудно убедиться, что функции $G(n, r)$ удовлетворяют рекуррентному соотношению

$$G(n, r) = G(n-1, r) + \frac{\rho}{n} \sum_{j=0}^r p_j (G(n-1, r-j) - G(n-2, r-j)) \quad (5.10)$$

с начальными условиями

$$G(0, r) = 1, \quad G(1, r) = 1 + \sum_{j=0}^r p_j, \quad r \geq 0. \quad (5.11)$$

При помощи функций $G(n, r)$ можно вычислять характеристики РеСМО. Например, вероятность π_b блокировки заявки и среднее число \bar{R} занятого ресурса вычисляются по формулам

$$\pi_b = 1 - \frac{1}{G(N, R)} \sum_{j=0}^R p_j G(N-1, R-j), \quad (5.12)$$

$$\bar{R} = R - \frac{1}{G(N, R)} \sum_{j=1}^R G(N, R-j). \quad (5.13)$$

5.2. Ресурсная модель с сигналами

В данном параграфе мы рассмотрим РеСМО с сигналами, которые моделируют блокировки LoS в сетях радиодоступа миллиметрового диапазона. Сигналы моделируют внешний поток событий, приводящий к изменению состояния прямой видимости (наличие или

отсутствии) и, следовательно, к изменению объема требуемого частотного ресурса. Остальные предположения остаются неизменными.

Таким образом, в отличие от модели, рассмотренной в предыдущем разделе, теперь каждая заявка, находящаяся на обслуживании, порождает пуассоновский поток сигналов с интенсивностью α , при поступлении которого она освобождает весь занимаемый ею ресурс, чтобы попытаться занять новый объем ресурса, имеющий то же вероятностное распределение.

При построении и анализе модели мы будем использовать предложенный ранее метод, состоящий в анализе только суммарного объема занятого ресурса. В сделанных предположениях поведение системы описывается СП $X(t) = \{\xi(t), \delta(t)\}$, в котором, как и прежде, первая компонента означает число заявок в системе, а вторая – суммарный объем занятого всеми заявками объема ресурса в момент t .

Пространство состояний СП определяется множеством

$$\Psi = \bigcup_{k=0}^N \Psi_k, \quad \Psi_k = \{(k, r) : 0 \leq r \leq R, p_r^{(k)} > 0\}. \quad (5.14)$$

Упорядочив состояния в множествах Ψ_k по возрастанию объема ресурсов, введем функции $I(k, r)$, значения которых равны порядковому номеру состояния (k, r) в множестве Ψ_k . Эти функции используются в формулах следующего параграфа.

5.2.1. Стационарное распределение

Введем дополнительное обозначение, которое облегчит понимание данного и последующих разделов. Пусть $\theta_i(k, r)$ – вероятность того, что заявка занимает i ресурсов при условии, что k заявок суммарно занимают r ресурсов. Тогда, согласно формуле Байеса:

$$\theta_i(k, r) = \frac{p_i p_{r-i}^{(k-1)}}{p_r^{(k)}}. \quad (5.15)$$

Для построения инфинитезимальной матрицы СП $X(t)$, который по построению является марковским процессом (МП), рассмотрим возможные переходы между состояниями системы. Пусть в некоторый момент времени система находится в состоянии (k, r) . С вероятностью p_j в систему может поступить заявка, которая займет

j единиц ресурса, если $j \leq R - r$ и система перейдет в состояние $(k + 1, r + j)$. При завершении обслуживания заявка высвобождает i ресурсов с вероятностью $\theta_i(k, r)$. В таком случае система перейдет в состояние $(k - 1, r - i)$.

В момент поступления сигнала одна из заявок системы сначала освобождает занимаемые ею i единиц ресурса с вероятностью $\theta_i(k, r)$ и затем запрашивает j единиц ресурса с вероятностью p_j . Если $j \leq i$, то заявка гарантированно принимается в систему. При этом если $j < i$, то система переходит в состояние $(k, r - i + j)$ с вероятностью $\theta_i(k, r)p_j$, а в случае $j = i$ состояние системы не меняется. Вместе с тем, если новый объем требований заявки превышает предыдущий объем, то заявка остается в системе только в случае $j \leq R - r + i$, в противном случае заявка сбрасывается и система переходит в состояние $(k - 1, r - i)$.

Можно показать, что инфинитезимальная матрица A МП $X(t)$ является блочной трехдиагональной с диагональными блоками D_0, D_1, \dots, D_N , наддиагональными блоками A_1, \dots, A_N и поддиагональными блоками M_0, \dots, M_{N-1} и имеет вид

$$A = \begin{pmatrix} D_0 & A_1 & & & & \\ M_0 & D_1 & A_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & M_{N-2} & D_{N-1} & A_N & \\ & & & M_{N-1} & D_N & \end{pmatrix}. \quad (5.16)$$

В (5.16) начальные блоки имеют вид

$$D_0 = -\lambda \sum_{j=0}^R p_j, \quad (5.17)$$

$$A_1 = (\lambda p_0, \dots, \lambda p_R), \quad (5.18)$$

$$M_0 = (\mu, \dots, \mu)^T, \quad (5.19)$$

причем из блока A_1 необходимо исключить нулевые элементы, чтобы число столбцов в ней, так же как и число строк в блоке M_0 , было равно числу состояний в подмножестве Ψ_1 . Остальные блоки определяются выражениями

$$D_n(I(n, i), (n, j)) = \begin{cases} - \left(\lambda \sum_{j=0}^{R-i} p_j + n\mu + n\alpha \left(1 - \sum_{m=0}^i \theta_m(n, i) p_m \right) \right), & i = j, \\ n\alpha \sum_{m=0}^i \theta_m(n, i) p_{j-i+m}, & i < j, \\ n\alpha \sum_{m=i-j}^i \theta_m(n, i) p_{j-i+m}, & i > j, \end{cases} \\ (n, i), (n, j) \in \Psi_n, 1 \leq n \leq N-1; \quad (5.20)$$

$$D_n(I(N, i), (N, j)) = \begin{cases} - \left(N\mu + N\alpha \left(1 - \sum_{m=0}^i \theta_m(N, i) p_m \right) \right), & i = j, \\ N\alpha \sum_{m=0}^i \theta_m(N, i) p_{j-i+m}, & i < j, \\ N\alpha \sum_{m=i-j}^i \theta_m(N, i) p_{j-i+m}, & i > j, \end{cases} \\ (N, i), (N, j) \in \Psi_n; \quad (5.21)$$

$$\Lambda_n(I(n-1, i), (n, j)) = \begin{cases} \lambda p_{j-i}, & i \leq j \leq R, \\ 0, & j < i, \end{cases} \\ (n-1, i) \in \Psi_{n-1}, (n, j) \in \Psi_n, \quad 2 \leq n \leq N; \quad (5.22)$$

$$M_n(I(n+1, i), (n, j)) = \begin{cases} (n+1)\mu \theta_{i-j}(n+1, i), & j \leq i \leq R, \\ 0, & j > i, \end{cases} \\ (n+1, i) \in \Psi_{n+1}, (n, j) \in \Psi_n, \quad 1 \leq n \leq N-1. \quad (5.23)$$

Заметим, что стационарные вероятности $Q_k(r)$ МП $X(t)$ являются единственным решением СУР с условием нормировки, записанной в матричном виде следующим образом:

$$\mathbf{Q}\mathbf{A} = \mathbf{0}, \mathbf{Q} \cdot \mathbf{1} = 1, \quad (5.24)$$

где \mathbf{Q} – вектор-строка стационарных вероятностей, а $\mathbf{1}$ – вектор-столбец из единиц соответствующей длины. Решать СУР (5.24) можно любым численным методом, в том числе использующим блочную структуру построенной выше инфинитезимальной матрицы.

5.2.2. Показатели эффективности

Перейдем к анализу показателей эффективности системы, к которым, прежде всего, относятся вероятность π_b блокировки заявки и средний объем \bar{R} занятого ресурса. Эти характеристики можно вычислять через стационарные вероятности ReLS по формулам

$$\pi_b = 1 - \sum_{k=0}^{N-1} \sum_{r=0}^R Q_k(r) \sum_{j=0}^{R-r} p_r, \quad (5.25)$$

$$\bar{R} = \sum_{k=0}^N \sum_{r=0}^R r Q_k(r). \quad (5.26)$$

Кроме того, интерес представляет вероятность π_t сброса уже принятой заявки из-за потока сигналов, изменяющего требования заявки к ресурсам. Для этого найдем интенсивность ν сброса принятых заявок, что происходит в случаях, когда из-за поступления сигнала новый объем требуемых ресурсов превышает объем незанятых ресурсов. Эта интенсивность может быть вычислена по формуле

$$\nu = \alpha \bar{N} \sum_{k=1}^N \sum_{r=0}^R Q_k(r) \sum_{j=0}^r \theta_j(k, r) (1 - \sum_{i=0}^{R-r+j} p_i), \quad (5.27)$$

где \bar{N} – среднее число заявок в системе

$$\bar{N} = \sum_{k=1}^N \sum_{r=0}^R k Q_k(r). \quad (5.28)$$

Вероятность π_t можно определить как предел отношения числа сброшенных заявок, которые были приняты в систему, к общему числу принятых заявок за время T , т.е.

$$\pi_t = \lim_{T \rightarrow \infty} \frac{\nu T}{\lambda(1-\pi_b)T} = \frac{\nu}{\lambda(1-\pi_b)}. \quad (5.29)$$

Отметим, что из введенных показателей эффективности наиболее критичным фактором является именно вероятность сброса принятой заявки. Поэтому в следующих разделах мы рассмотрим механизмы, резервирования ресурсов и мультисвязности, позволяющих снизить эту вероятность.

В заключение раздела напомним, что для ReLS с сигналами упрощение, позволяющее отслеживать только суммарный объем занимаемых ресурсов, дает возможность получать не точное, а приближенное решение. При использовании изложенных методов для прикладных расчетов необходимо проверять, что данное приближение дает приемлемую относительную погрешность в заданном диапазоне исходных данных.

5.3. Резервирование ресурсов

Резервирование части емкости сети для поддержания сессий, которые в результате блокировки LoS требуют больше ресурсов, является одним из механизмов, увеличивающих показатели непрерывности сессий в сетях радиодоступа, использующих миллиметровый диапазон частот. В этом случае новым поступающим сессиям доступна только часть общего ресурса системы, а остальная часть предназначена для поддержания уже принятых сессий. В данном параграфе рассмотрена модель такого механизма для БС 5G/6G с использованием РеСМО с сигналами, а также проведен анализ ее характеристик.

5.3.1. Построение модели

Итак, рассматривается РеСМО с N обслуживающими приборами и конечным объемом R ресурсов, часть которых $R_0 = (1 - \gamma)R$, $0 < \gamma < 1$, доступна вновь поступающим заявкам. Здесь γ является параметром механизма резервирования и означает долю зарезервированных ресурсов. Поведение системы, как и ранее, описывается СП $X(t) = \{\xi(t), \delta(t)\}$, где $\xi(t)$ – число заявок в системе, а $\delta(t)$ – суммарный объем занятого ресурса.

В отличие от РеСМО, описанной ранее, при поступлении заявки в рассматриваемую систему ей доступна только часть общего ресурса, $R_0 < R$. Иными словами, если система находится в состоянии (k, r) и поступает заявка, требующая j ресурсов, то: 1) если $r > R_0$, то любая новая поступающая заявка блокируется; 2) если $r \leq R_0$, то заявка блокируется при $j > R_0 - r$; 3) если $r \leq R_0$ и $j \leq R_0 - r$, то заявка принимается в систему. Однако при изменении объема занимаемого ресурса в результате поступления сигнала заявке становится доступен весь объем R ресурсов системы. Иначе говоря, если система находится в состоянии (k, r) , при поступлении сигнала заявка освобождает i ресурсов и требует занять j ресурсов, то она будет сброшена только в случае, если $j > R - r + i$.

В остальном система функционирует аналогично системе, рассмотренной ранее. Интенсивности переходов для типичного состояния системы проиллюстрированы на рис. 5.1.

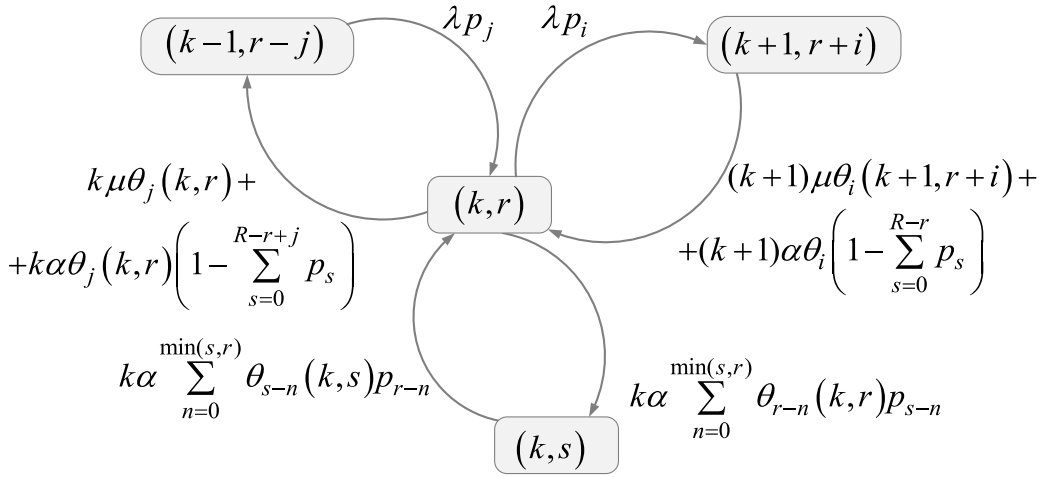


Рис. 5.1. Фрагмент диаграммы переходов состояний СМО с резервированием ресурсов

5.3.2. Система уравнений равновесия

С учетом введенных обозначений, а также функции

$$\theta(R_0 - j) = \begin{cases} 0, & j > R_0, \\ 1, & j \leq R_0. \end{cases} \quad (5.30)$$

можно выписать СУР

$$\begin{aligned} \lambda Q_0 \sum_{j=0}^{R_0} p_j &= \mu \sum_{j:(1,j) \in \Psi_1} Q_1(j) + \alpha \sum_{j:(1,j) \in \Psi_1} Q_1(j) \left(1 - \sum_{s=0}^R p_s\right), \\ &\left(\theta(R_0 - j) \lambda \sum_{j=0}^{R_0-r} p_j + k\mu + k\alpha \right) Q_k(r) = \\ &= \theta(R_0 - j) \lambda \sum_{j \geq 0: (k-1, r-j) \in \Psi_{k-1}} Q_{k-1}(r-j) p_j + \\ &+ (k+1) \mu \sum_{j \geq 0: (k+1, r+j) \in \Psi_{k+1}} Q_{k+1}(r+j) \theta_j(k+1, j+r) + \\ &+ (k+1) \alpha \left(1 - \sum_{s=0}^{R-r} p_s\right) \end{aligned}$$

$$\begin{aligned}
& \sum_{j \geq 0: (k+1, r+j) \in \Psi_{k+1}} Q_{k+1}(r+j) \theta_j(k+1, j+r) + \\
& + k\alpha \sum_{j \geq 0: (k, j) \in \Psi_k} Q_k(j) \sum_{i=0}^{\min(j, r)} \theta_{j-i}(k+1, j+r) p_{r-i}, \quad (5.31) \\
& 1 \leq n \leq N-1, 0 \leq r \leq R;
\end{aligned}$$

$$\begin{aligned}
N(\mu + \alpha) Q_N(r) = & \theta(R_0 - j) \lambda \sum_{j \geq 0: (N-1, j) \in \Psi_{N-1}} Q_{N-1}(r-j) p_j + \\
& + N\alpha \sum_{j \geq 0: (N, j) \in \Psi_N} Q_N(j) \sum_{i=0}^{\min(j, r)} \theta_{j-i}(N, j) p_{r-i}, \quad 1 \leq r \leq R.
\end{aligned}$$

Используя СУР (5.31) вместе с условием нормировки, можно найти стационарные вероятности состояний системы.

Инфинитезимальную матрицу для введенного процесса можно получить аналогично предыдущим разделам из СУР (5.31).

Показатели эффективности

Перейдем к анализу показателей эффективности системы с резервированием. Так как рассматриваемая система отличается от системы из предыдущих разделов только алгоритмом принятия новой поступающей заявки на обслуживание, то меняется лишь формула для расчета вероятности π_b блокировки новой заявки

$$\pi_b = 1 - \sum_{k=0}^{N-1} \sum_{r=0}^{R_0} Q_k(r) \sum_{j=0}^{R_0-r} p_r, \quad (5.32)$$

Остальные характеристики, такие как средний объем занятых ресурсов \bar{R} и вероятность блокировки принятой заявки π_t , определяются по формулам (5.26) и (5.29).

5.4. Мультисвязность и резервирование ресурсов

Еще одним механизмом повышения показателей непрерывности сессий является метод мультисвязности (англ. multiconnectivity), который заключается в поддержке соединения АУ с несколькими БС, когда данные передаются по одному каналу, а при переходе канала в состояние блокировки LoS АУ переключается на другое соединение с другой БС. Механизм определен в стандарте 3GPP TS 37.340.

В данном параграфе будет изложен метод решения задачи с совместным использованием механизмов мультисвязности и резервирования. Заметим, что в отличие от предположений о процессе обслуживания трафика, принятых ранее, здесь рассматривается набор БС, каждая из которых доступна каждому пользователю. В качестве примера можно рассмотреть размещение БС на периметре городской площади.

5.4.1. Построение модели мультисвязности

Рассмотрим ресурсную сеть массового обслуживания (СеМО), состоящую из K узлов. Узел k имеет N_k приборов и $R_{k,1}$ ресурсов. На k -й узел поступает пуассоновский поток заявок с интенсивностью λ_k , $k = 1, 2, \dots, K$, и $\lambda = \sum_{k=1}^K \lambda_k$. Заявка, поступающая на k -й узел, требует для обслуживания случайный объем ресурсов с дискретным распределением $\{p_{0,r}\}_{r>0}$. Для вновь поступающих сессий доступен не весь объем ресурсов узла, а только его часть, $R_{k,0} = R_{k,1}(1 - \gamma)$, $0 < \gamma < 1$. Как и ранее, если поступающая заявка обнаруживает отсутствие свободного прибора или недостаточный объем ресурсов, то заявка сбрасывается. Времена обслуживания заявок распределены экспоненциально с параметром μ .

Заявка, находящаяся на обслуживании в k -м узле, порождает пуассоновский поток сигналов, моделирующих переходы канала между состояниями LoS и nLoS, с параметром α_k , $k = 1, 2, \dots, K$. При поступлении сигнала заявка освобождает ранее занимаемый объем ресурсов и генерирует новый объем требований в соответствии с дискретным распределением $\{p_{1,r}\}_{r>0}$. Если на k -м узле достаточно ресурсов для удовлетворения новых требований, то заявка занимает требуемый объем ресурсов и продолжает обслуживание. В противном случае, если на других узлах ресурсов нет, заявка прерывает обслуживание и сбрасывается. Наконец, если $K > 1$, то заявка покидает k -й узел, генерирует новый объем ресурсов в соответствии с распределением $\{p_{0,r}\}_{r>0}$ и пытается перейти на любой другой узел из оставшихся с вероятностью $1/(K - 1)$. Такие заявки будем называть вторичными, а уровнем вторичной заявки – число переходов между узлами СеМО. Вторичные заявки образуют дополнительный пуассоновский входящий поток на каждый узел с интенсивностями φ_k , $k = 1, 2, \dots, K$. Вторичные заявки могут занимать весь доступный объем ресурсов на узле. Если же на узле, на который перешла заявка

в результате поступления сигнала, нет свободного прибора или недостаточно ресурсов, то она блокируется. Заметим, что благодаря свойству отсутствия памяти экспоненциального распределения, остаточное время обслуживания вторичной заявки имеет то же экспоненциальное распределение с параметром μ . Схема модели изображена на рис. 5.2.

Для анализа модели применяется метод декомпозиции, являющийся типичным при расчете сложных СеМО. Ключевое предположение здесь состоит в том, что процесс обслуживания заявок на каждом узле не зависит от процессов обслуживания на других узлах сети.

На основе метода декомпозиции может быть получен рекуррентный алгоритм расчета характеристик системы, схема которого показана на рис. 5.3. Начинаем вычисления, считая интенсивность поступления вторичных заявок равной нулю, $\varphi_k = 0, k = 1, 2, \dots, K$. Далее вычисляются стационарные вероятности каждого узла, из которых определяются показатели сети в целом.

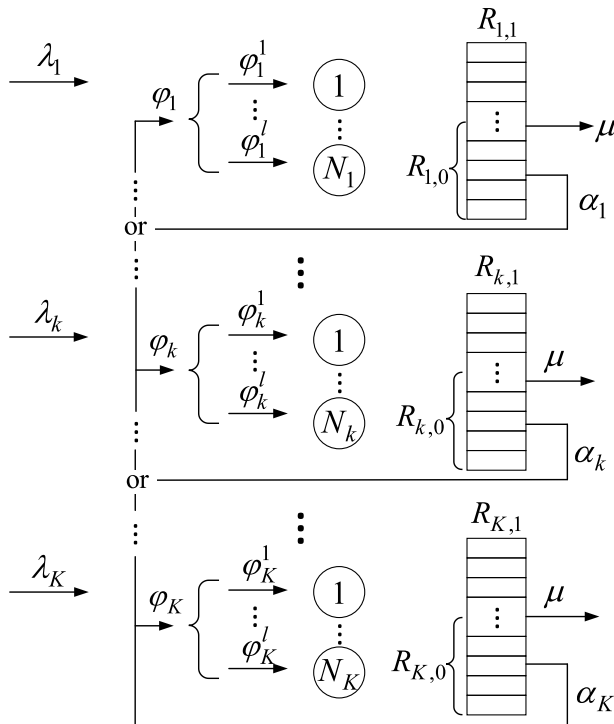


Рис. 5.2. Схема ресурсной сети массового обслуживания

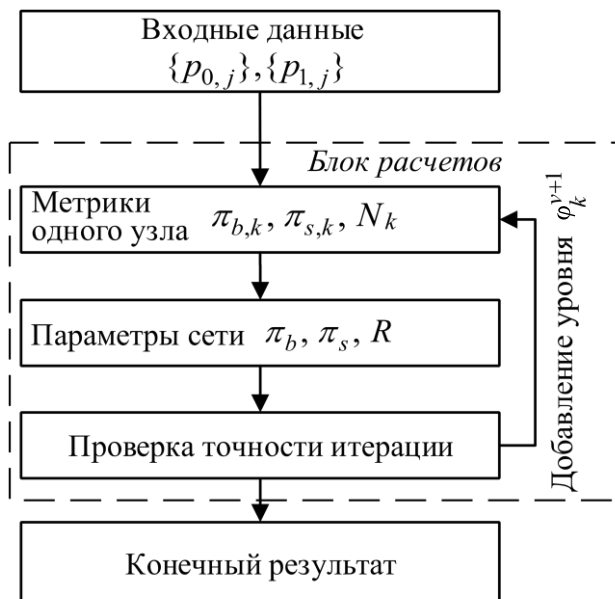


Рис. 5.3. Рекуррентная схема расчета характеристик ресурсной сети массового обслуживания

5.4.2. Модель отдельного узла сети

В данном разделе проводится анализ функционирования отдельного k -го узла сети. На узел поступают два пуассоновских потока первичных и вторичных заявок с интенсивностями λ_k и φ_k соответственно. Согласно свойству экспоненциального распределения, время пребывания заявки на узле имеет также экспоненциально распределено с параметром $\mu + \alpha_k, i = 1, 2, \dots, K$.

При построении модели мы применяем принцип упрощения системы, при котором во внимание принимается только суммарный объем занятого ресурса. Тогда поведение системы описывается трехмерным процессом СП $X(t) = \{\xi_1(t), \xi_2(t), \delta(t)\}$, где $\xi_1(t)$ – число первичных заявок в системе в момент времени t , $\xi_2(t)$ – число вторичных заявок в момент времени t , а $\delta(t)$ – суммарный объем занятого всеми заявками ресурса. Отметим, что $X(t)$ по построению является марковским процессом, его пространство состояний имеет вид

$$\Psi = \bigcup_{0 \leq n_1 + n_2 \leq N} \Psi_{n_1, n_2},$$

$$\Psi_{n_1, n_2} = \left\{ (n_1, n_2, r) : 0 \leq r \leq R_1, \sum_{i=0}^{\min(r, R_0)} p_{0,i}^{(n_1)} p_{1,r-i}^{(n_2)} > 0 \right\}, \quad (5.33)$$

где, как и прежде, верхний индекс в скобках означает кратность свертки распределения.

Стационарные вероятности $Q_{n_1, n_2}(r)$ состояний $(n_1, n_2, r) \in \Psi$ определяются следующим образом:

$$Q_{n_1, n_2}(r) = \lim_{t \rightarrow \infty} P\{\xi_1(t) = n_1, \xi_2(t) = n_2, \delta(t) = r\}. \quad (5.34)$$

Рассмотрим объем ресурсов, который освобождает заявка при уходе из узла. Обозначим $\beta_{0,j}(n_1, n_2, r)$ вероятность того, что в результате ухода первичной заявки j единиц ресурсов высвобождаются при условии, что система находится в состоянии (n_1, n_2, r) . Для случая $r \leq R_0$ вероятности $\beta_{0,j}(n_1, n_2, r)$ могут быть вычислены с помощью формулы Байеса

$$\beta_{0,j}(n_1, n_2, r) = \frac{p_{0,j} \sum_{i=0}^{r-j} p_{0,i}^{(n_1-1)} p_{1,r-j-i}^{(n_2)}}{\sum_{i=0}^r p_{0,i}^{(n_1)} p_{1,r-i}^{(n_2)}}. \quad (5.35)$$

Аналогично определим вероятность $\beta_{1,j}(n_1, n_2, r)$ того, что в результате ухода вторичной заявки j единиц ресурсов освобождается при условии, что система находится в состоянии (n_1, n_2, r) , т.е.

$$\beta_{1,j}(n_1, n_2, r) = \frac{p_{1,j} \sum_{i=0}^{r-j} p_{0,i}^{(n_1)} p_{1,r-j-i}^{(n_2-1)}}{\sum_{i=0}^r p_{0,i}^{(n_1)} p_{1,r-i}^{(n_2)}}. \quad (5.36)$$

Определение вероятностей $\beta_{0,j}(n_1, n_2, r)$ и $\beta_{1,j}(n_1, n_2, r)$ в случае $r > R_0$ несколько сложнее. Отметим, что эти вероятности зависят от порядка поступления первичных и вторичных заявок, что не может быть определено по состоянию процесса. Так как моменты поступления этих заявок независимы друг от друга и принимая во внимание известное число заявок каждого типа, любую их перестановку по времени поступления в систему можно считать равновероятной. Вероятность того, что пришедшая последней первичная заявка занимает k -е место, равна

$$\binom{k-1}{n_1-1} / \binom{n_1+n_2}{n_1}, \quad (5.37)$$

где $\binom{k-1}{n_1-1}$ равно числу способов поставить $n_1 - 1$ первичную заявку по первым $k - 1$ местам (так как последняя займет k -е место),

а $\binom{n_1 + n_2}{n_1}$ – число способов расставить n_1 первичных заявок по $n_1 + n_2$ местам. Вероятность того, что n_1 первичных заявок и $k - n_1$ вторичных заявок занимают $i \leq R_0$ ресурсов, вычисляется по свертке

$$\sum_{s=0}^i p_{0,s}^{(n_1)} p_{1,i-s}^{(k-n_1)}. \quad (5.38)$$

Далее, вероятность того, что n_1 первичных и n_2 вторичных заявок занимают r ресурсов при условии, что пришедшая последней первичная заявка занимает k -е место, также вычисляется как свертка

$$\sum_{i=0}^{\min(r, R_0)} p_{1,r-i}^{(n_2+n_1-k)} \sum_{s=0}^i p_{0,s}^{(n_1)} p_{1,i-s}^{(k-n_1)}. \quad (5.39)$$

Наконец, вероятность того, что n_1 первичных и n_2 вторичных заявок вместе занимают r ресурсов, имеет вид

$$\sum_{k=n_1}^{n_1+n_2} \frac{\binom{k-1}{n_1-1}}{\binom{n_1+n_2}{n_1}} \sum_{i=0}^{\min(r, R_0)} p_{1,r-i}^{(n_2+n_1-k)} \sum_{s=0}^i p_{0,s}^{(n_1)} p_{1,i-s}^{(k-n_1)}. \quad (5.40)$$

Поскольку конечные формулы достаточно громоздкие и их вывод теперь не представляет труда, заинтересованному читателю предлагается проделать этот процесс самостоятельно.

Вычисление вероятностей $\beta_{0,j}(n_1, n_2, r)$ и $\beta_{1,j}(n_1, n_2, r)$ в случае $r > R_0$ также является непростой задачей. Поэтому для уменьшения вычислительной сложности предлагается использовать формулу (5.41) для всего диапазона значений параметра r , пренебрегая порядком поступления заявок. Практика показывает, что для реальных значений доли зарезервированного ресурса $\gamma < 0,2$ такое допущение не приводит к существенным погрешностям.

$$\beta_{0,j}(n_1, n_2, r) = \frac{p_{0,j} \sum_{i=0}^{\min(r-j, R_0-j)} p_{0,i}^{(n_1-1)} p_{1,r-j-i}^{(n_2)}}{\sum_{i=0}^{\min(r, R_0)} p_{0,i}^{(n_1)} p_{1,r-i}^{(n_2)}},$$

$$\beta_{1,j}(n_1, n_2, r) = \frac{p_{1,j} \sum_{i=0}^{\min(r-j, R_0)} p_{0,i}^{(n_1)} p_{1,r-j-i}^{(n_2-1)}}{\sum_{i=0}^{\min(r, R_0)} p_{0,i}^{(n_1)} p_{1,r-i}^{(n_2)}}. \quad (5.41)$$

Получив вероятности $\beta_{0,j}(n_1, n_2, r)$ и $\beta_{1,j}(n_1, n_2, r)$, можно приступить к выводу СУР для МП $X(t)$, которая в итоге представлена формулами (5.42). Интенсивности переходов состояний ресурсной СеМО проиллюстрированы на диаграмме рис. 5.4.

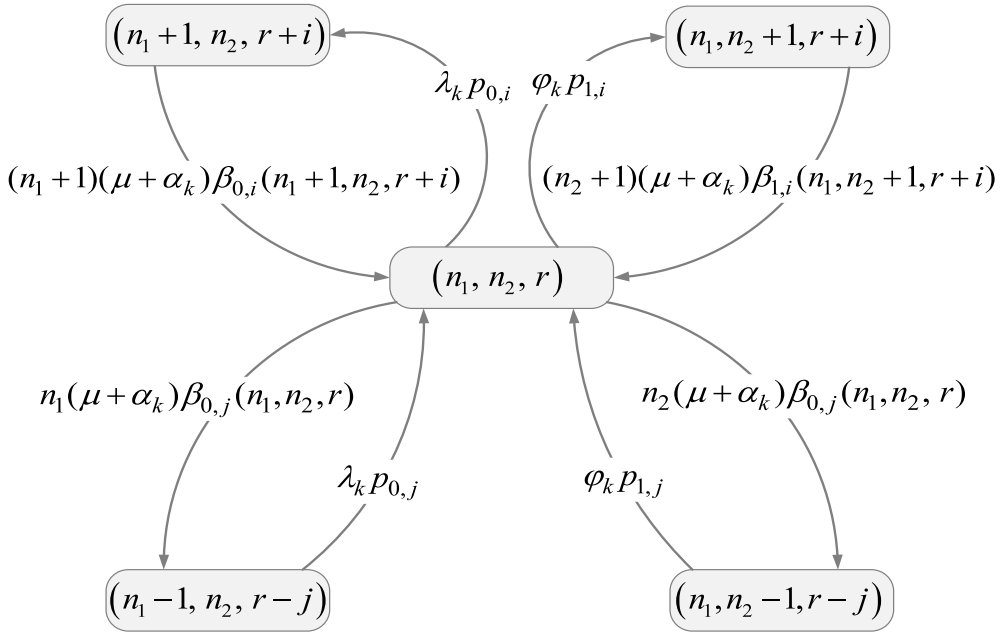


Рис. 5.4. Фрагмент диаграммы интенсивностей переходов состояний ресурсной СеМО

$$\begin{aligned}
 & Q_0 \left[\lambda_k \sum_{j=0}^{R_0} p_{0,j} + \varphi_k \sum_{j=0}^{R_1} p_{1,j} \right] = \\
 & = (\mu + \alpha_k) \left[\sum_{j:(1,0,j) \in \Psi_{1,0}} Q_{1,0}(j) + \sum_{j:(0,1,j) \in \Psi_{0,1}} Q_{0,1}(j) \right]; \\
 & Q_{n_1, n_2}(r) \left[\lambda_k \sum_{j=0}^{R_0-r} p_{0,j} + \varphi_k \sum_{j=0}^{R_1-r} p_{1,j} + (n_1 + n_2)(\mu + \alpha_k) \right] = \\
 & = \lambda_k \sum_{j:(n_1-1, n_2, r-j) \in \Psi_{n_1-1, n_2}} p_{0,j} Q_{n_1-1, n_2}(r-j) + \\
 & + \varphi_k \sum_{j:(n_1, n_2-1, r-j) \in \Psi_{n_1, n_2-1}} p_{1,j} Q_{n_1, n_2-1}(r-j) + \\
 & + (n_1 + 1)(\mu + \alpha_k) \sum_{j:(n_1+1, n_2, r+j) \in \Psi_{n_1+1, n_2}} Q_{n_1+1, n_2}(r+j) \times \\
 & \quad \times \beta_{0,j}(n_1 + 1, n_2, r + j) +
 \end{aligned}$$

$$\begin{aligned}
& +(n_1 + 1)(\mu + \alpha_k) \sum_{j:(n_1+1, n_2, r+j) \in \Psi_{n_1+1, n_2}} Q_{n_1+1, n_2}(r+j) \times \\
& \quad \times \beta_{0,j}(n_1 + 1, n_2, r+j) + (n_2 + 1)(\mu + \alpha_k) \times \\
& \times \sum_{j:(n_1, n_2+1, r+j) \in \Psi_{n_1, n_2+1}} Q_{n_1, n_2+1}(r+j) \beta_{1,j}(n_1, n_2 + 1, r+j), \\
& \quad n_1 + n_2 < N, r \leq R_0. \tag{5.42}
\end{aligned}$$

$$\begin{aligned}
& Q_{n_1, n_2}(r) \left[\varphi_k \sum_{j=0}^{R_1-r} p_{1,j} + (n_1 + n_2)(\mu + \alpha_k) \right] = \\
& = \varphi_k \sum_{j:(n_1, n_2-1, r-j) \in \Psi_{n_1, n_2-1}(r-j)} p_{1,j} Q_{n_1, n_2-1}(r-j) + \\
& +(n_1 + 1)(\mu + \alpha_k) \sum_{j:(n_1+1, n_2, r+j) \in \Psi_{n_1+1, n_2}} Q_{n_1+1, n_2}(r+j) \times \\
& \quad \times \beta_{0,j}(n_1 + 1, n_2, r+j) + \\
& +(n_2 + 1)(\mu + \alpha_k) \sum_{j:(n_1, n_2+1, r+j) \in \Psi_{n_1, n_2+1}} Q_{n_1, n_2+1}(r+j) \times \\
& \quad \times \beta_{1,j}(n_1, n_2 + 1, r+j), \\
& \quad n_1 + n_2 < N, r > R_0. \\
& (n_1 + n_2)(\mu + \alpha_k) Q_{n_1, n_2}(r) = \\
& = \lambda_k \sum_{j:(n_1-1, n_2, r-j) \in \Psi_{n_1-1, n_2}} p_{0,j} Q_{n_1-1, n_2}(r-j) + \\
& \quad + \varphi_k \sum_{j:(n_1, n_2-1, r-j)} p_{1,j} Q_{n_1, n_2-1}(r-j), \\
& \quad n_1 + n_2 = N, r \leq R_0.
\end{aligned}$$

В левой части первого уравнения СУР величина $\lambda_k \sum_{j=0}^{R_0} p_{0,j} + \varphi_k \sum_{j=0}^{R_1} p_{1,j}$ представляет собой интенсивность выхода из состояния $(0,0,0)$. В правой части $(\mu + \alpha_k)$ есть интенсивность перехода из всех состояний вида $(1,0,j)$ и $(0,1,j)$ в состояние $(0,0,0)$, что соответствует завершению обслуживания и переходу заявки на другой узел.

Второе уравнение СУР справедливо для $n_1 + n_2 < N, r \leq R_0$. Выражение в левой части $\lambda_k \sum_{j=0}^{R_0-r} p_{0,j} + \varphi_k \sum_{j=0}^{R_1-r} p_{1,j} + (n_1 + n_2)(\mu + \alpha_k)$

соответствует интенсивности выхода из состояния (n_1, n_2, r) . Далее, первое слагаемое $\lambda_k \sum_{j:(n_1-1, n_2, r-j) \in \Psi_{n_1-1, n_2}} p_{0,j} Q_{n_1-1, n_2}(r-j)$ есть интенсивность переходов из состояния (n_1, n_2, r) в состояние $(n_1 - 1, n_2, r - j)$, вызванная поступлением первичной заявки на узел k . Второе слагаемое представляет собой интенсивность перехода в (n_1, n_2, r) , связанную с направлением на узел k вторичной заявки. Следующие два слагаемых определяют переходы в (n_1, n_2, r) , связанные с поступлением сигналов.

Третье уравнение СУР справедливо для $n_1 + n_2 < N$ и $r > R_0$. Логика построения уравнения аналогична предыдущему.

Четвертое уравнение СУР описывает переходы состояний, в которых $n_1 + n_2 = N$, $r \leq R_0$. Первое слагаемое в левой части $(n_1 + n_2)(\mu + \alpha_k)$ – это интенсивность переходов из состояния (n_1, n_2, r) в состояния $(n_1 - 1, n_2, r - j)$ или $(n_1, n_2 - 1, r - j)$, вызванные завершением обслуживания или поступлением сигнала соответственно. В правой части выражение $\lambda_k \sum_{j:(n_1-1, n_2, r-j) \in \Psi_{n_1-1, n_2}} p_{0,j} Q_{n_1-1, n_2}(r-j)$ есть интенсивность переходов из состояния $(n_1 - 1, n_2, r - j)$ в состояние (n_1, n_2, r) в связи с поступлением первичной заявки. Второе слагаемое отражает переходы, вызванные поступлением вторичной заявки.

Пятое уравнение СУР выводится аналогично четвертому.

5.4.3. Итерационный алгоритм

Система уравнений равновесия, определенная в (5.42), решается численно. В связи с тем что число уравнений в системе может достигать $N(N + 1)R_1/2$, для операций с инфинитезимальной матрицей рекомендуется использовать специальные библиотеки для разреженных матриц, а для решения СУР приближенные итерационные методы, например метод Гаусса–Зейделя.

Процесс расчета стационарных вероятностей и показателей эффективности системы является итерационным, так как на каждой итерации добавляется следующий уровень вторичных заявок. Процедура расчета заканчивается в момент достижения требуемой точности характеристик. На первой итерации интенсивности поступления вторичных заявок принимаются равными нулю $\varphi_k = 0, k = 0, 1, \dots, K$. Далее выполняется следующее.

– Используя $\lambda_k, \mu, \alpha_k, \varphi_k$ и распределения $\{p_{0,j}\}, \{p_{1,j}\}$, вычисляются показатели эффективности каждого узла по отдельности: вероятность блокировки новой заявки $\pi_{b,k}$; вероятность $\pi_{s,k}$ того, что переход на другой узел вызовет блокировку вторичной заявки; среднее число заявок \overline{N}_k и среднее число единиц занятого ресурса \overline{R}_k .

– Вычисляются параметры сети в целом: вероятность блокировки новой заявки π_b ; вероятность того, что переход на другой узел вызовет блокировку вторичной заявки π_s ; среднее число единиц занятого ресурса \overline{R} .

– Если достигнута необходимая точность, алгоритм останавливается. В противном случае вычисляются интенсивности поступления вторичных заявок следующего уровня φ_k^{v+1} и запускается следующая итерация.

Теперь перейдем к формулам для расчета показателей эффективности узлов и всей сети. Вероятность $\pi_{b,k}$ блокировки новой заявки на узле k и вероятность $\pi_{s,k}$ того, что переход на другой узел вызовет блокировку вторичной заявки, принимают вид

$$\begin{aligned} \pi_{b,k} &= 1 - \sum_{0 \leq n_1 + n_2 \leq N-1} \sum_{r \leq R_{0,k} : (n_1, n_2, r) \in \Psi_{n_1, n_2}} Q_{n_1, n_2}(r) \sum_{j=0}^{R_{0,k}-r} p_{0,j}, \\ \pi_{s,k} &= 1 - \sum_{0 \leq n_1 + n_2 \leq N-1} \sum_{r : (n_1, n_2, r) \in \Psi_{n_1, n_2}} Q_{n_1, n_2}(r) \sum_{j=0}^{R_{1,k}-r} p_{1,j}. \end{aligned} \quad (5.43)$$

Интенсивности вторичных заявок рассчитываются по формулам

$$\begin{aligned} \varphi_k &= \sum_{v=1}^{\infty} \varphi_k^v, \\ \varphi_k^1 &= \sum_{i=1}^K \lambda_i (1 - \pi_{b,i}) \frac{\alpha_i}{\mu + \alpha_i} \varphi_{i,k}^0, \\ \varphi_k^v &= \sum_{i=1}^K \varphi_i^{v-1} (1 - \pi_{s,i}) \frac{\alpha_i}{\mu + \alpha_i} \varphi_{i,k}^{v-1}, \quad v > 1, \end{aligned} \quad (5.44)$$

где v обозначает уровень вторичной заявки.

Далее получим формулы для вычисления характеристик по всей СеМО: вероятность π_b блокировки новой заявки и вероятность π_s того, что переход на другой узел вызовет блокировку вторичной заявки.

$$\pi_b = \sum_{k=1}^K \frac{\lambda_k}{\lambda} \pi_{b,k}, \quad \pi_s = \sum_{k=1}^K \frac{\varphi_k}{\varphi} \pi_{s,k}. \quad (5.45)$$

Еще одна важная характеристика системы это вероятность π_0 того, что принятая в систему заявка будет сброшена. Для ее вычисления применяется формула

$$\pi_0 = \lim_{t \rightarrow \infty} \frac{\varphi \pi_s t}{\lambda(1-\pi_b)t} = \frac{\varphi \pi_s}{\lambda(1-\pi_b)}, \quad (5.46)$$

где числитель представляет собой среднее число принятых заявок, заблокированных за время t , а в знаменателе находится число всех принятых в систему заявок за время t .

5.5. Численный анализ

В этом параграфе проводится численное исследование методов резервирования ресурсов и мультисвязности в сетях радиодоступа 5G/6G.

5.5.1. Модель резервирования ресурсов

Начнем с численного исследования метода резервирования ресурсов в сетях радиодоступа 5G/6G. В качестве показателей эффективности используются характеристики, ориентированные на абонента, в том числе вероятности сброса новой сессии и принятия сессии на обслуживание, а также характеристики, ориентированные на оператора сети, которые характеризуют эффективность использования системных ресурсов.

Сначала сосредоточим внимание на характеристиках, ориентированных на пользователя, включая вероятности сброса новой сессии и принятия сессии на обслуживание. В этом разделе для сравнения отклика системы рассматриваем антенную решетку 32×4 на БС, мощность передачи 2 Вт, что соответствует эффективному покрытию 355 м согласно модели распространения UMi (см. главу 2).

Вероятности сброса новой сессии и сессии, принятой на обслуживание, в зависимости от доли зарезервированных ресурсов, показаны на рис. 5.5 для скорости сессий 10 Мбит/с и плотности блокаторов 0,04 на квадратный метр. Анализируя влияние части зарезервированных ресурсов на вероятность сброса сессий, можно заметить, что большие значения γ положительно влияют на вероятность сброса текущей сессии. Как и ожидалось, существует обменное соотношение между рассматриваемыми вероятностями, и преимущества использования метода резервирования ресурсов очевидны для всех

рассмотренных интенсивностей поступления сессий. Важно отметить, что с ростом интенсивности поступления сессий влияние на вероятность сброса сессии, принятой на обслуживание, увеличивается. Таким образом, предлагаемый метод работает лучше в условиях перегрузки, что имеет важное значение для операторов сетей радиодоступа.

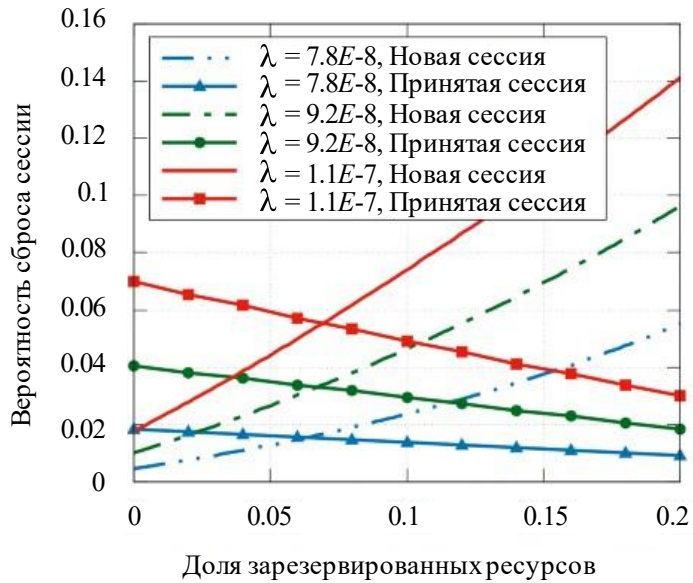


Рис. 5.5. Зависимость вероятности сброса сессии от доли зарезервированных ресурсов и интенсивности поступления сессий

Далее рассмотрим влияние скорости передачи сессий, показанной на рис. 5.6, для плотности возникновения сессий $1,1 \times 10^{-7}$ на квадратный метр и плотности блокаторов 0,04 на квадратный метр. Отметим несколько важных эффектов. Во-первых, увеличение скорости сессий отрицательно влияет на обе рассматриваемые характеристики. Основная причина заключается в том, что более высокие скорости сессий оставляют неиспользованными больше ресурсов в системе, когда новая или текущая сессия сбрасывается из-за недостатка ресурсов. Во-вторых, что более важно, преимущества резервирования ресурсов увеличиваются в ответ на растущие скорости сессий. В частности, для скорости сессии 10 Мбит/с вероятность потери сессии, принятой на обслуживание, уменьшается с приблизительно 0,42 для $\gamma = 1$ до приблизительно 0,33 для $\gamma = 0,8$. При

одинаковых значениях γ и скорости $R = 10$ Мбит/с улучшения незначительны. Следовательно, можем заключить, что метод резервирования ресурсов особенно полезен для сессий, имеющих высокие требования к скорости передачи данных. Обслуживание таких сессий является основной целью будущих сетей радиодоступа 5G/6G.

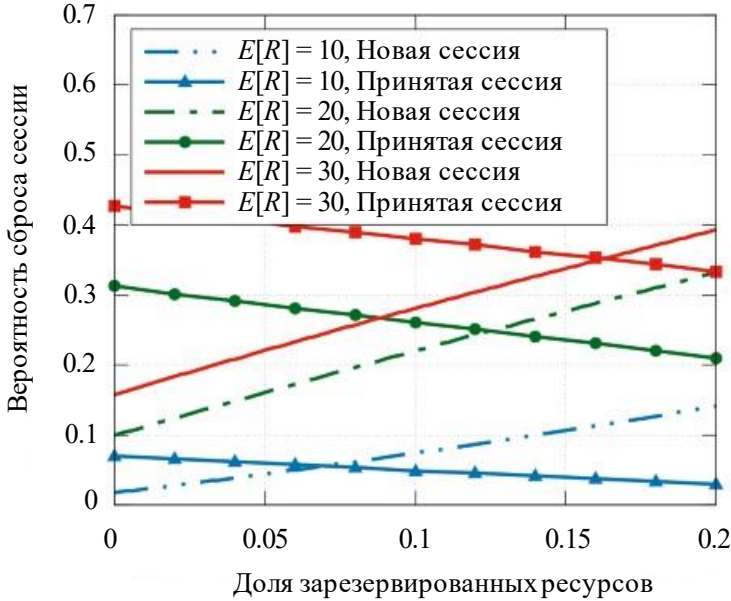


Рис. 5.6. Зависимость вероятности сброса сессии от доли зарезервированных ресурсов и скорости сессии

Известно, что явление блокировки значительно влияет на производительность сетей радиодоступа. Графики на рис. 5.7 демонстрируют влияние плотности блокаторов на вероятность сброса новой сессии и сессии, принятой на обслуживание для скорости сессий 10 Мбит/с и плотности поступления сессий $1,1 \times 10^{-7}$ на квадратный метр. Во-первых, можно заметить большую разницу между вероятностями сброса сессии, принятой на обслуживание, соответствующими разным уровням плотности блокаторов. Фактически эта вероятность возрастает с 0,03 до более чем 0,8. Соответствующая разница в вероятностях сброса новой сессии намного меньше и не превышает 0,01 во всем рассматриваемом диапазоне значений γ . Это объясняется тем, что высокая интенсивность изменения состояния на АУ вызывается более высокими значениями плотностей блокаторов. Следовательно, число изменений состояния за время обслуживания

сессии растет, что приводит к увеличению вероятности сброса сессии, принятой на обслуживание. Отметим, что качественное поведение рассматриваемых вероятностей остается одинаковым для всех рассмотренных значений плотности блокаторов.

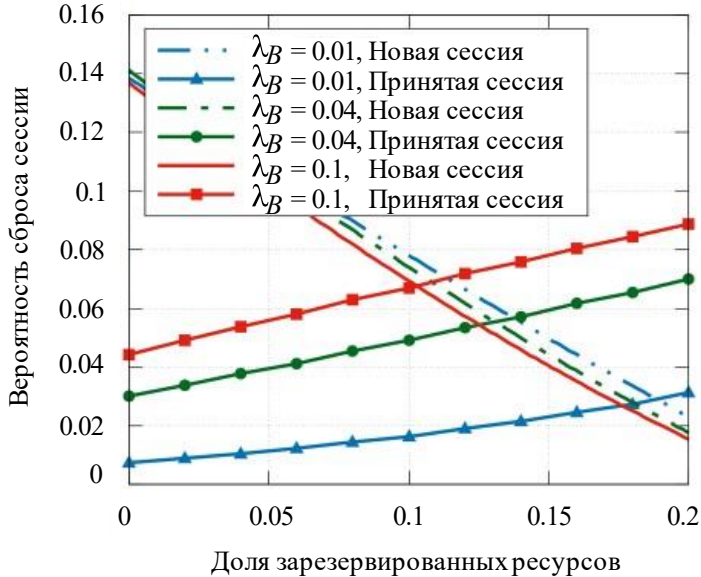


Рис. 5.7. Зависимость вероятности сброса сессии от доли зарезервированных ресурсов и интенсивности (плотности) блокаторов

Отметим, что влияние скорости перемещения блокаторов имеет характер, аналогичный влиянию плотности блокаторов. Фактически более высокая скорость приводит к более высокой частоте изменений состояния АУ, что отрицательно влияет на вероятность сброса сессии, принятой на обслуживание. Кроме того, влияния высот NR БС, АУ и блокаторов также вполне очевидны и поэтому здесь подробно не обсуждаются. В частности, большая высота NR БС или блокатора, а также меньшая высота АУ увеличивают периметр зоны блокировки LoS, тем самым приводя к более частым изменениям состояния АУ и в итоге к увеличению вероятности сброса сессии, принятой на обслуживание.

Определив, что метод резервирования ресурсов особенно полезен при высоких нагрузках, рассмотрим подробнее влияние интенсивности поступления сессий на вероятность сброса сессий на рис. 5.8. Ожидаемо, что вероятность сброса как новой сессии, так и сессии,

принятой на обслуживание, увеличивается с ростом λ . При относительно небольшой интенсивности поступления можно наблюдать, что вероятности сброса текущей сессии почти совпадают для разных значений γ . Однако с ростом λ можно заметить изменения в исследуемой вероятности. Качественно подобное поведение наблюдается и для вероятности сброса новой сессии. Однако расстояние между кривыми, соответствующими различным значениям γ , обычно больше.

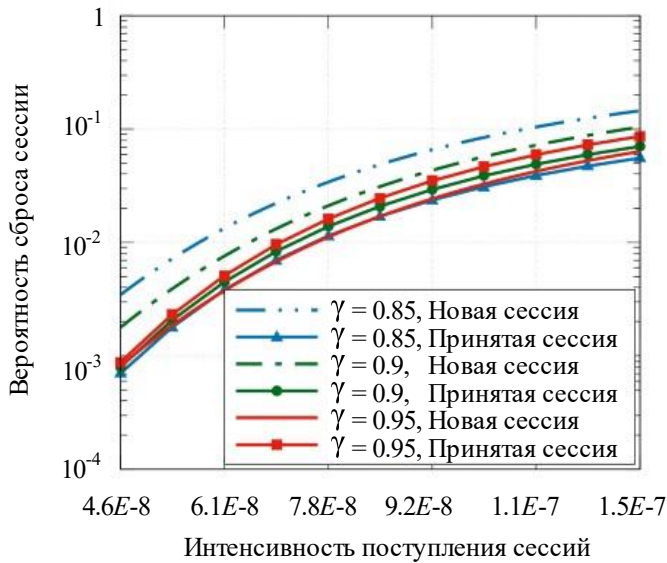


Рис. 5.8. Зависимость вероятности сброса сессии от интенсивности поступления сессий и доли зарезервированных ресурсов

5.5.2. Модель мультисвязности

Во-первых, рассмотрим характеристики обслуживания абонентов и производительность системы обслуживания, когда используется только функция мультисвязности. Графики рис. 5.9 иллюстрируют вероятности сброса новых сессий, и сессий, принятых на обслуживание, а также коэффициент использования ресурсов системы для разных значений степени мультисвязности M . Как можно заметить, увеличение степени мультисвязности приводит к лучшему использованию системных ресурсов. Для исследуемых входных параметров системы коэффициент использования ресурсов приближается к 0,8 для $M = 8$. Эта тенденция сопровождается снижением вероятности сброса текущих сессий. Это связано с тем, что по мере роста M становится доступно больше ресурсов для сессий, которые

изменяют свое состояние с незаблокированного на заблокированное. Отметим, что этот эффект вызван резким увеличением вероятности сброса новых сессий. Таким образом, заключаем, что использование только функции мультисвязности не может обеспечить желаемую гибкость в достижении компромисса между рассматриваемыми характеристиками обслуживания абонентов, однако позволяет значительно повысить коэффициент использования ресурсов системы.

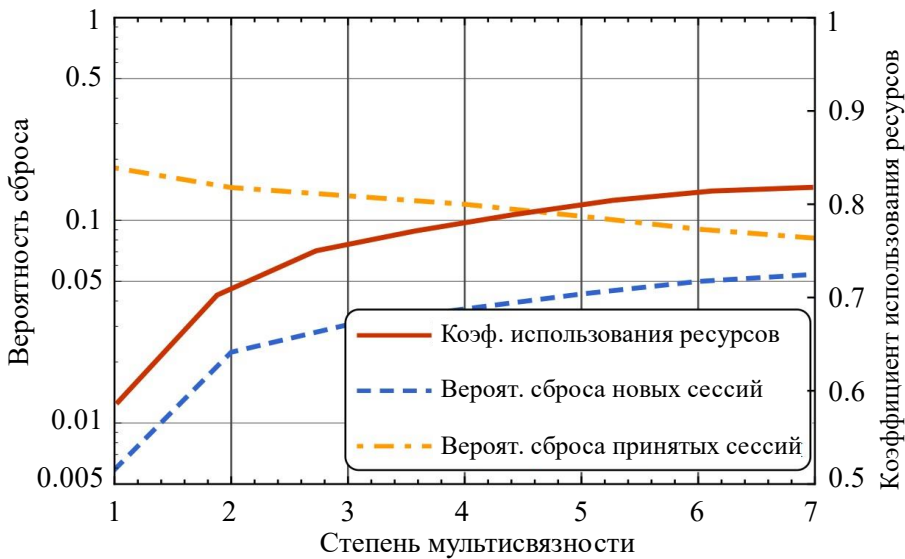


Рис. 5.9. Характеристики обслуживания в зависимости от степени мультисвязности

Важно отметить, что скорость изменения рассматриваемых характеристик обслуживания абонентов достаточно высока. Увеличивая долю зарезервированных ресурсов всего на 2%, можно уменьшить вероятность сброса сессий, принятых на обслуживание с 0,38 до 0,09 за счет увеличения вероятности сброса новых сессий с 0,06 до 0,09. Учитывая, что влияние функции мультисвязности на вероятность сброса сессий, принятых на обслуживание, значительно меньше, можно сформулировать гипотезу о том, что использование функции резервирования ресурсов может иметь чрезвычайно важное значение для улучшения непрерывности обслуживания сессий в сетях радиодоступа 5G/6G. Использование обоих механизмов позволит оператору сети получить две степени свободы для выбора рабочей точки, определяющей компромисс между вероятностями сброса новых сессий, и сессий, принятых на обслуживание,

при этом сохраняя коэффициент использования ресурсов на приемлемом уровне.

5.5.3. Совместное применение резервирования и мультисвязности

Анализируя показанный на рис. 5.10 совместный эффект применения механизмов резервирования и мультисвязности на вероятность сброса сессий, принятых на обслуживание, можно установить, что исследуемая характеристика возрастает, даже несмотря на увеличение количества потенциальных обслуживающих БС. Это вызвано отрицательным влиянием механизма резервирования ресурсов в рассматриваемой системе, т.е. даже небольшое увеличение доли зарезервированных ресурсов приводит к значительному увеличению вероятности сброса новых сессий. В частности, при степени мультисвязности $M = 1$ параметр изменяется с 0,005 при $\gamma = 0$ до 0,05 при $\gamma = 0,01$. Ухудшение характеристик обслуживания для новых сессий связано с улучшением надежности обслуживания сессий, принятых на обслуживание. В частности, как видно из графиков на рис. 5.11, более высокая степень мультисвязности приводит к снижению вероятности сброса сессий, принятых на обслуживание. Хотя этот эффект незначителен при $\gamma = 0$, он становится намного более заметным при $\gamma = 0,1$.

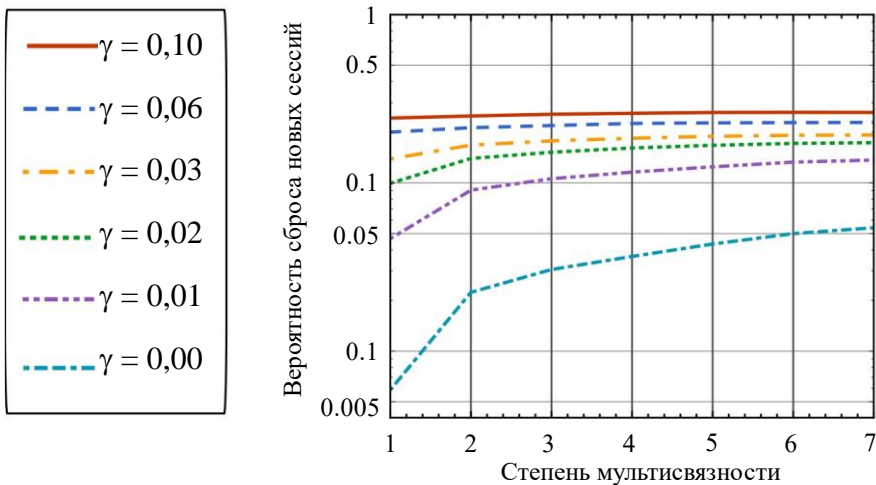


Рис. 5.10. Зависимость вероятности сброса новых сессий от уровня резервирования и степени мультисвязности

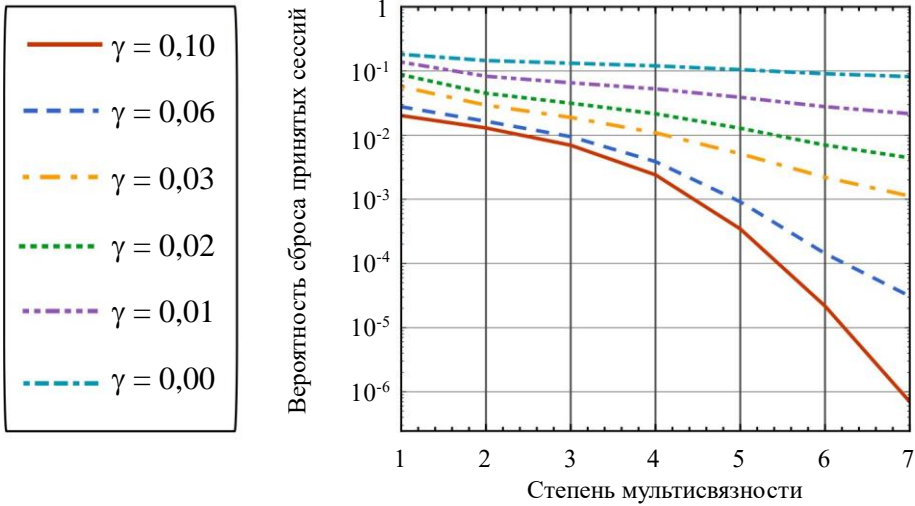


Рис. 5.11. Зависимость вероятности сброса принятых сессий от уровня резервирования и степени мультисвязности

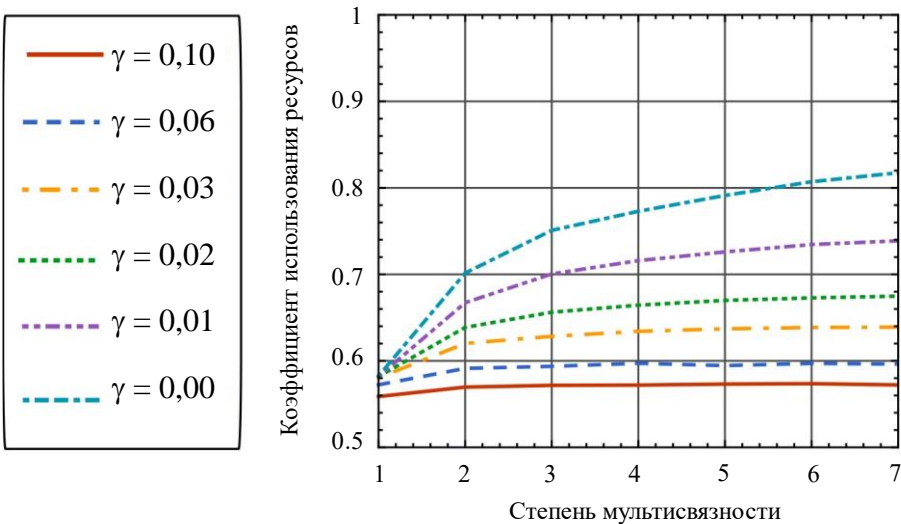


Рис. 5.12. Зависимость коэффициента использования ресурсов от уровня резервирования и степени мультисвязности

Наконец, оценим коэффициент использования ресурсов системы, представленный на рис. 5.12. Отметим, что положительный эффект мультисвязности сохраняется и в системе с резервированием ресурсов. Однако влияние этого параметра сильно зависит от доли зарезервированных ресурсов γ . В основном при относительно больших значениях γ коэффициент использования ресурсов

системы относительно мал и находится в диапазоне (0,55–0,57). Уменьшение процента резервируемых ресурсов для сессий, принятых на обслуживание, позволяет значительно улучшить эффективность использования ресурсов. В частности, коэффициент использования ресурсов системы в итоге достигает 0,75 при $\gamma = 0,01$.

Подводя итог, отметим, что совместное использование механизмов мультисвязности и резервирования ресурсов предоставляет больше степеней свободы для поиска баланса между вероятностями сброса новых сессий и сессий, принятых на обслуживание. При этом коэффициент использования ресурсов системы остается достаточно высок, что является ключевым показателем производительности системы для операторов сетей радиодоступа. При этом следует отметить, что желательно избавиться от негативного влияния доли зарезервированных ресурсов на вероятность сброса новых сессий. Этого эффекта можно добиться путем использования процедуры первоначального выбора обслуживающей БС.

ГЛАВА 6

МНОГОАДРЕСНЫЕ СОЕДИНЕНИЯ В СЕТЯХ ДОСТУПА 5G/6G

В этой главе рассматриваются вопросы применения многоадресных соединений в сетях радиодоступа с направленными антеннами и блокировками, работающими в миллиметровом или терагерцовом диапазонах частот. В параграфе 6.1 рассматривается проблема обслуживания абонентов с помощью многоадресных соединений в сетях 5G/6G и формулируется задача определения показателей качества обслуживания. Модель системы с направленными антеннами и блокировками, обслуживающая одновременно одноадресные и многоадресные соединения, предложена в 6.2. Анализ системы с помощью методов стохастической геометрии и ТМО проведен в 6.3. В качестве показателей качества выбрана вероятность потерь сессий для многоадресных и одноадресных соединений. Численный анализ показателей качества проведен в 6.4, где также обсуждаются аспекты целесообразности использования многоадресных соединений в сетях 5G/6G.

6.1. Проблемы применения многоадресных соединений в сетях 5G/6G

Использование антенных решеток на БС и на АУ рассматривается как одно из важнейших преимуществ сетей радиодоступа 5G/6G, потенциально позволяющих достичь направленности менее 1° , используя 128×4 и более антенных элементов. Эффект заключается в значительном увеличении зоны обслуживания БС, а также в возможности уменьшить негативное влияние помех. Однако, будучи полезным для сессий одноадресных соединений, использование узконаправленной антенны может привести к неэффективному использованию ресурсов при обслуживании многоадресного трафика. В частности, чем меньше угловой сектор внутри главного лепестка диаграммы направленности, в котором спектр излучаемой мощности на 3 дБ ниже пикового значения (HPBW, Half-Power Beamwidth), тем меньше число многоадресных АУ, которые могут одновременно обслуживаться одной передачей. Таким образом, для обслуживания многоадресных абонентов в системах с высокой направленностью антенн необходимо использовать несколько

передач, распространяющих одну и ту же информацию, что приводит к неэффективному использованию ресурсов БС.

Проблема многоадресной передачи в системах с направленными антеннами на БС исследована в ряде проектов. В достаточной степени изучен вопрос адаптации скорости обслуживания в системах IEEE 802.11ad при оптимизации требований к задержке передачи. Решение основано на постановке задачи минимаксной оптимизации, что приводит к решению задачи методами линейного программирования. Изучалась проблема группировки АУ на основе их близости. Соответствующие эвристические алгоритмы могут быть разработаны на последовательном тестировании различных значений НРВW, которые оптимизируют суммарную емкость системы. Кроме того, показано, что использование постоянного значения угла НРВW приводит к неоптимальному использованию ресурсов. Изучен также метод оптимизации, который не только работает с шириной полосы НРВW, но и учитывает неравное распределение мощности между лучами. Проанализирован алгоритм многоуровневого многоадресного мультиплексирования, основанный на неортогональном множественном доступе (NOMA, Non-Orthogonal Multiple Access). Однако использование подхода на основе NOMA для сетей доступа 5G/6G все еще в стадии обсуждения в организациях стандартизации. Сформулирована проблема многоадресных передач в системах с направленными антеннами, проанализированы несколько схем передачи, нацеленных на минимизацию задержки доставки пакетов.

Проблема многоадресной передачи с использованием направленных антенн признана исследовательским сообществом. Однако, насколько это известно, ни одно из исследований, выполненных до настоящего времени, не решало проблему одновременной поддержки как многоадресного, так и одноадресного режимов передачи в сетях доступа 5G/6G, работающих в миллиметровом диапазоне частот. Поэтому вопрос оптимальной конфигурации БС, обслуживающей суперпозицию многоадресного и одноадресного трафика, все еще требует подробного изучения.

Рассмотрим две БС, каждая из которых оборудована антенными решетками в каждом из трех секторов. Фокусируемся на определенном секторе, обслуживаемом одной антенной решеткой, как показано на рис. 6.1. Предполагая линейную антенную решетку на БС, линейное усиление передачи может быть реализовано числом антенных элементов N_A с направленностью $\alpha_A \sim 102^\circ/N_A$ (см. главу 2).

Можно сделать вывод, что, используя меньшее число элементов на БС, расстояние между ними, гарантирующее отсутствие пробелов в покрытии, становится меньше, но количество передач, необходимых для поддержки многоадресной услуги, уменьшается. Наоборот, при увеличении числа антенных элементов размер НРВW уменьшается, что приводит к усилению передачи на БС и, следовательно, увеличению расстояния между ними, снижая стоимость развертывания сети. Однако в то же время может потребоваться большее количество передач для обслуживания всех АУ, участвующих в многоадресной рассылке, что приводит к неэффективному использованию системных ресурсов. Решение зависит от интенсивности поступления сессий, доли многоадресных сессий, НРВW антенной решетки и усложняется из-за наличия одноадресных сессий. В конечном счете для некоторого набора параметров системы и характеристик среды должно существовать оптимальное расстояние между БС, обеспечивающее полный охват интересующей области с заданными вероятностями потерь одноадресных и многоадресных сессий.

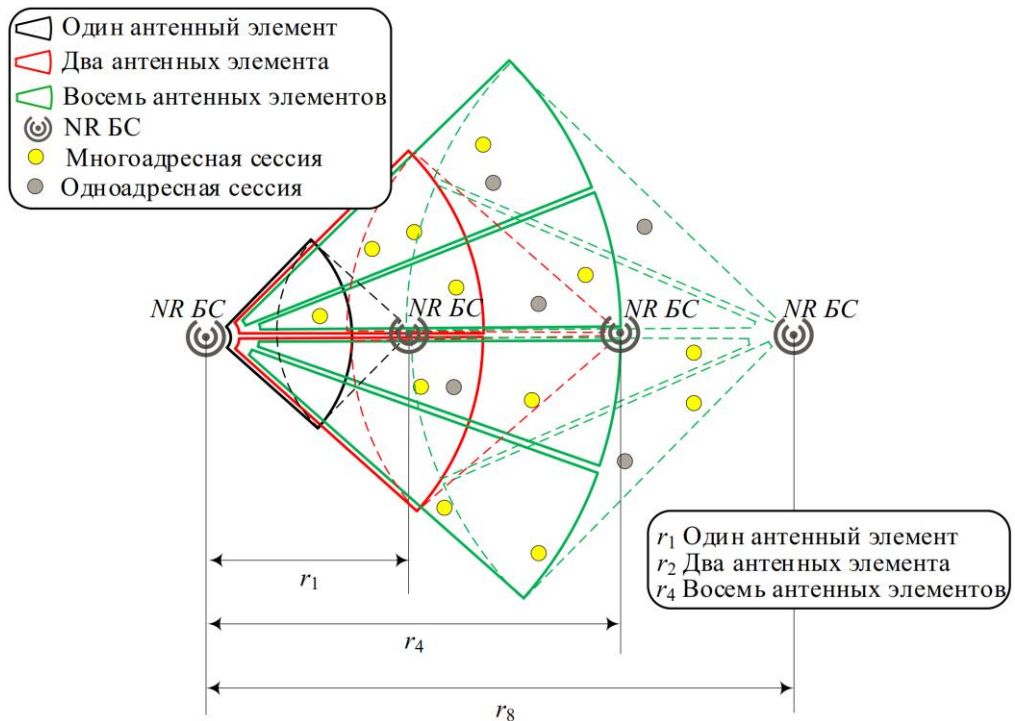


Рис. 6.1. Иллюстрация проблемы поддержки многоадресных сессий в сетях доступа 5G/6G

Далее будут рассмотрены методы поддержки многоадресных сессий, а также компромиссы, связанные с одновременным обслуживанием многоадресного и одноадресного трафика в сетях 5G/6G. Для этого, объединив методы стохастической геометрии и ТМО, строится математическая модель, учитывающая свойства распространения в миллиметровом диапазоне частот, процесс выделения ресурсов на БС и особенности обслуживания трафика многоадресных и одноадресных соединений. В качестве характеристик обслуживания абонентов рассмотрены вероятности потерь многоадресных и одноадресных сессий и коэффициент использования ресурсов. Модель позволяет исследовать взаимосвязь между плотностью развертывания БС и характеристиками обслуживания, предоставляемыми рассматриваемым типам трафика в различных условиях. Результаты численного анализа позволяют сделать вывод о нижней границе плотности развертывания БС, обеспечивающей заданные характеристики обслуживания абонентов.

6.2. Системная модель

Обозначения, используемые в этой главе, приведены в табл. 6.1. Рассматриваем БС в сотовой сети с заданной плотностью λ_B пешеходов-пользователей, как показано на рис. 6.1. Все пользователи имеют АУ, оснащенные модулем 5G/6G. Высота БС и АУ предполагается фиксированной и равной h_A и h_U соответственно. Тела пользователей моделируются цилиндрами с высотой h_B и радиусом r_B . Линия прямой видимости LoS между АУ и БС может быть перекрыта пешеходами.

Предполагается, что БС имеет круговой диапазон покрытия. Это достигается с помощью трех физических антенн, каждая из которых охватывает сектор 120° . Рассматриваем покрытие одной антенны и определяем d_E – эффективный радиус покрытия, такой, что ни одно АУ не находится в условиях потери связи, когда линия LoS заблокирована. Другими словами, существует допустимая схема модуляции и кодирования (MCS, Modulation and Coding Scheme) для АУ, находящихся на расстоянии d_E , которое вычисляется с применением модели распространения UMi, а в качестве модели антенной решетки используется двухмерная модель (см. главу 2).

Обозначения

Параметр	Определение
α_A, α_U	Направленность антенны на БС и АУ
N_A, N_U	Число антенных элементов на БС и АУ
S	Площадь области, покрываемой одной антенной
M	Число классов многоадресных сессий
K	Число классов одноадресных сессий
$R_{M,m}$	Скорость класса m многоадресных сессий
$R_{U,k}$	Скорость класса k многоадресных сессий
γ_m	Нагрузка класса m многоадресных сессий
ρ_m	Нормализованная нагрузка класса m многоадресных сессий
a_k	Нагрузка класса m одноадресных сессий
$p_{U,k}$	Вероятность, что сессия принадлежит одноадресному классу k
$p_{M,m}$	Вероятность, что сессия принадлежит многоадресному классу m
Λ	Интенсивность потока сессий от одного АУ
λ_m	Интенсивность потока многоадресных сессий класса m
μ_m	Интенсивность обслуживания многоадресных сессий класса m
b_m	Число РБ для многоадресных сессий класса m
ν_k	Интенсивность потока одноадресных сессий класса k
κ_k	Интенсивность обслуживания одноадресных сессий класса m
d_k	Число РБ для одноадресной сессии класса k
C	Число обслуживающих приборов
$p_k(n_k)$	Стационарные вероятности одноадресных сессий
s_A	Размер РБ, МГц
Δ	Защитный интервал частот, МГц
S_{th}	Граница чувствительности приемника, дБ
ϵ_j	Вероятность CQI/MCS j
s_j	Границы SNR, дБ
$q_{M,m}$	Вероятность сброса многоадресной сессии класса m
$q_{U,k}$	Вероятность сброса одноадресной сессии класса k
u	Коэффициент использования ресурсов системы

Предполагаем, что LoS между АУ и БС может быть временно перекрыта пешеходами. В зависимости от текущего состояния канала (незаблокированная или заблокированная LoS) и расстояния между БС и АУ сессия использует соответствующую MCS для поддержания надежной передачи данных. Предполагается, что пользователи движутся в соответствии с моделью движения с выбором случайного направления (RDM, Random Direction Model). При этом в течение активной сессии пользователь предполагается неподвижным. Таким образом, блокировки описываются моделью, определенной в главе 2 для статичных абонентов.

Пусть Λ – интенсивность поступления сессий от одного АУ. В произвольный момент времени каждый пользователь может инициировать сессию многоадресной или одноадресной передачи с вероятностями p_M и p_U соответственно, $p_M + p_U = 1$. Рассматриваются M классов многоадресных сессий и K классов одноадресных сессий. Сессия принадлежит классу m с вероятностью $p_{M,m}$, $m = 1, 2, \dots, M$, и классу k с вероятностью $p_{U,k}$, $k = 1, 2, \dots, K$, соответственно, причем

$$\sum_{m=1}^M p_{M,m} = p_M, \quad \sum_{k=1}^K p_{U,k} = p_U. \quad (6.1)$$

Используя свойство суперпозиции, можно считать, что процесс поступления сессий на БС, обслуживающую сектор 120° , является пуассоновским с интенсивностью $\Lambda \lambda_B \pi d_E^2 2\pi/3$ (сессий в единицу времени), где λ_B – плотность абонентов (абонентов на единицу площади). Обозначим $\lambda_m = p_{M,m} \Lambda \lambda_B \pi d_E^2 2\pi/3$ – интенсивность потока многоадресных сессий класса m и $\nu_k = p_{U,k} \Lambda \lambda_B \pi d_E^2 2\pi/3$ – интенсивность потока одноадресных сессий класса k .

Выбор АУ, который инициирует сессию, является случайным, и пользователи распределены равномерно в пределах покрытия БС. Класс многоадресных сессий m и класс одноадресных сессий k характеризуются экспоненциально распределенным временем обслуживания с параметрами μ_m и κ_k соответственно. Предполагается, что соответствующие скорости сессий постоянны и равны $R_{M,m}$ и $R_{U,k}$ Мбит/с. Объем ресурсов, измеряемых в первичных РБ, запрашиваемых сессиями многоадресной и одноадресной передачи, b_m и d_k , зависит от размера РБ, обозначенного s_A . Предполагается, что полоса частот, доступная БС, равна W Гц. Многоадресная сессия инициируется в момент первого поступления сессии этого класса в так называемый период бездействия, т.е. период времени отсутствия сессий этого класса в системе. Сессия принимается на обслуживание, если имеется достаточно ресурсов, в противном случае ей отказано в обслуживании. Все сессии класса m , которые застают при поступлении хотя бы одну сессию этого класса в системе, принимаются в систему без выделения дополнительных ресурсов.

Рассмотрим два типа дисциплин многоадресного режима передачи, которые различаются по длительности активного периода, который начинается, когда сессия принимается на обслуживание в период бездействия.

– *Дисциплина T1.* Многоадресные сессии класса m , которые поступают в систему в течение активного периода, не влияют на его

продолжительность. Ресурсы, занятые этими сессиями, высвобождаются по завершении активного периода, когда сессия, которая инициировала этот период, покидает систему. Примерами являются онлайн-игры или запланированные многоадресные рассылки.

– *Дисциплина T2.* Сессия класса m , которая поступает в систему в течение активного периода, может увеличивать его продолжительность. Активный период заканчивается, когда завершает обслуживание последняя сессия класса m . Примером является многоадресная потоковая передача видео.

Для вновь поступившей одноадресной сессии требуется новый набор ресурсов. Сессия класса k получает отказ, если для ее обслуживания нет достаточного числа РБ.

Показателями эффективности исследуемой системы являются:

- 1) вероятность сброса многоадресной сессии $q_{M,m}$;
- 2) вероятность сброса одноадресной сессии $q_{U,k}$;
- 3) коэффициент использования ресурсов системы u .

6.3. Мультисервисная модель с одноадресными и многоадресными соединениями

Модель обслуживания K классов одноадресных и M классов многоадресных сессий представлена в виде мультисервисной СМО с C обслуживающими приборами, схематично изображенной на рис. 6.2. Число обслуживающих приборов определяется по формуле $C = W/(s_A + \Delta)$, где W – доступная полоса пропускания; s_A – размер РБ; Δ – защитный интервал.

Введем СП $\{(\Psi(t), \Phi(t)), t \geq 0\}$ и его состояния (Ψ, Φ) , где $\Psi = (I_1, I_2, \dots, I_M)$ содержит индикаторы наличия многоадресных сессий в системе ($I_m = 1$, когда многоадресные сессии класса m присутствуют в системе, и $I_m = 0$ в противном случае) и $\Phi = (n_1, n_1, \dots, n_K)$, где n_k – число одноадресных сессий класса k в системе. Пространство состояний модели имеет вид

$$\begin{aligned} \tilde{Z} &= \{(\Psi, \Phi) : I_m \in \{0,1\}, m = 1, 2, \dots, M, \\ &0 \leq n_k \leq \left\lfloor \frac{C}{d_k} \right\rfloor, k = 1, 2, \dots, K, \\ &\sum_{m=1}^M I_m b_m + \sum_{k=1}^K n_k d_k \leq C\}. \end{aligned} \quad (6.2)$$

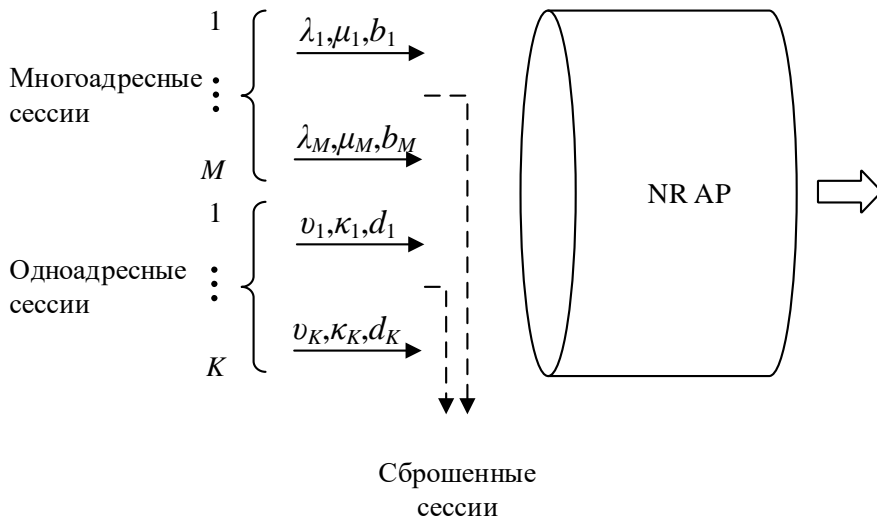


Рис. 6.2. Схема мультисервисной СМО с одноадресными и многоадресными сессиями

По построению введенный СП является марковским, и для нахождения его стационарного распределения рассмотрим сначала систему с бесконечным числом обслуживающих приборов $S = \infty$. В этом случае пространство состояний модели можно записать в виде

$$\mathcal{Z} = \{(\Psi, \Phi) : I_m \in \{0, 1\}, m = 1, 2, \dots, M, \\ n_k \in \{0, 1, \dots\}, k = 1, 2, \dots, K\}. \quad (6.3)$$

Поскольку емкость системы не ограничена, компоненты стационарного распределения СП $\vec{\Psi}(t)$ могут быть записаны в виде

$$\pi_m(I_m) = \lim_{t \rightarrow \infty} P\{\Psi_m(t) = I_m\} = \frac{\gamma_m^{I_m}}{1 + \gamma_m}, \quad (6.4)$$

где

$$\gamma_m = \begin{cases} \rho_m, & \text{если дисциплина обслуживания T1,} \\ e^{\rho_m} - 1, & \text{если дисциплина обслуживания T2,} \end{cases} \quad (6.5)$$

а величина нагрузки ρ_m рассчитывается по формуле

$$\rho_m = \left(1 + \frac{\lambda_m}{\mu_m}\right)^{p_{M,m} \lambda_{BS}} - 1. \quad (6.6)$$

Заметим, что величины γ_m соответствуют нагрузке трафика многоадресных сессий класса m , а показатель степени в формуле (6.6)

соответствует числу АУ, иницирующих многоадресные сессии. Его значение зависит от площади зоны покрытия $S = \pi d_E^2 \alpha / 3$ и конфигурации антенны, зависящей от направленности антенной решетки α_A .

Стационарное распределение $\Phi(t)$ в предположении о бесконечном числе обслуживающих приборов определяются формулой

$$p_k(n_k) = \lim_{t \rightarrow \infty} P\{\Psi_k(t) = n_k\} = \frac{a_k^{n_k} e^{-a_k}}{n_k!}, \quad (6.7)$$

где $a_k = v_k / \kappa_k$ – предложенная нагрузка одноадресных сессий и $v_k = p_{U,k} \Lambda \lambda_B S_{\text{area}}$ – интенсивность потока одноадресных сессий класса k . Теперь стационарное распределение составного случайного процесса $(\Psi(t), \Phi(t))$ может быть представлено в мультипликативном виде:

$$\tilde{\pi}(\Psi, \Phi) = G^{-1}(\mathcal{Z}) \prod_{m=1}^M \gamma_m^{I_m} \prod_{k=1}^K \frac{a_k^{n_k}}{n_k!}, \quad (\Psi, \Phi) \in \mathcal{Z}, \quad (6.8)$$

где $G(\mathcal{Z})$ – нормирующая константа.

Стационарное распределение для системы с конечным пространством состояний \mathcal{Z} можно получить из (6.8) сужением процесса $(\Psi(t), \Phi(t))$ с множества \mathcal{Z} (6.3) на множество $\tilde{\mathcal{Z}}$ (6.2), и тогда

$$\pi(\Psi, \Phi) = G^{-1}(\tilde{\mathcal{Z}}) \prod_{m=1}^M \gamma_m^{I_m} \prod_{k=1}^K \frac{a_k^{n_k}}{n_k!}, \quad (\Psi, \Phi) \in \tilde{\mathcal{Z}}, \quad (6.9)$$

с нормирующей константой

$$G(\tilde{\mathcal{Z}}) = \sum_{(\Psi, \Phi) \in \tilde{\mathcal{Z}}} \prod_{m=1}^M \gamma_m^{I_m} \prod_{k=1}^K \frac{a_k^{n_k}}{n_k!}. \quad (6.10)$$

Для организации вычислений показателей эффективности введем вспомогательные величины

$$h(n) = \begin{cases} 1, & n < 0, \\ 0, & n = 0, \\ \frac{1}{n} \sum_{k=1}^K d_k a_k h(n - d_k), & n = 1, 2, \dots, C, \end{cases} \quad (6.11)$$

с помощью которых и, с точностью до некоторой константы, можно найти вероятности того, что в системе отсутствуют многоадресные сессии класса m , все многоадресные сессии первых i классов, а одноадресные сессии занимают ровно n обслуживающих приборов в виде функций

$$f_m(i, n) = \begin{cases} 0, & i = 0, \dots, M, n < 0, \\ h(n), & i = 0, n = 0, \dots, C, \\ f_m(i-1, n) + \frac{f_m(i-1, n-b_i)}{((1-\delta_{im})\gamma_i)^{-1}}, & i = 1, \dots, M, n = 0, \dots, C. \end{cases} \quad (6.12)$$

Заметим, что величина $f_0(M, n)$ соответствует случаю, когда все одноадресные сессии и все многоадресные сессии занимают ровно n обслуживающих приборов. Теперь искомые показатели эффективности могут быть вычислены по следующим формулам:

- вероятность сброса многоадресной сессии класса m

$$q_{M,m} = \frac{\sum_{n=C-b_{m+1}}^C f_n(M, n)}{\sum_{n=0}^C f_n(M, n)}, \quad m = 1, 2, \dots, M; \quad (6.13)$$

- вероятность сброса одноадресной сессии класса k

$$q_{U,k} = \frac{\sum_{n=C-d_{k+1}}^C f_n(M, n)}{\sum_{n=0}^C f_n(M, n)}, \quad k = 1, 2, \dots, K; \quad (6.14)$$

- коэффициент использования ресурсов системы

$$u = \frac{\sum_{n=1}^C n f_n(M, n)}{\sum_{n=0}^C f_n(M, n)}. \quad (6.15)$$

6.4. Численный анализ

Изучим характеристики обслуживания абонентов многоадресных и одноадресных сессий. Сначала оценим адекватность математической модели, сравнивая аналитические результаты с результатами, полученными с помощью имитационного моделирования. Затем исследуем влияние параметров многоадресных и одноадресных сессий на характеристики обслуживания – вероятности сброса многоадресных и одноадресных сессий, а также коэффициент использования ресурсов системы. Далее численно найдем максимальное расстояние между БС для типичной сотовой структуры сети доступа 5G/6G с гарантированным обслуживанием многоадресных и одноадресных сессий. Наконец, оценим характеристики сети доступа, где многоадресный трафик передается в одноадресных сессиях.

Параметры системы, используемые при численном анализе, представлены в табл. 6.2, а табл. 6.3 содержит предварительно вычисленные соотношения между числом антенных элементов на БС, радиусом покрытия d_E и средним числом ресурсов, необходимых

для поддержания скорости сессии в 20 и 50 Мбит/с. Далее для краткости вычисления проводим для дисциплины обслуживания T1, хотя аналогичные результаты могут быть получены и для дисциплины T2. Для вычисления числа РБ используется метод, рассмотренный в главе 2 и учитывающий связь между ОСШ и набором MCS, определенных для NR.

Таблица 6.2

Параметры для численного анализа

Параметр	Значение
Рабочая частота	28 ГГц
Полоса пропускания	400 МГц
Размер РБ	1,44 МГц
Защитный интервал	0,015 МГц
Высота БС	4 м
Высота блокатора	1,7 м
Высота АУ	1,5 м
Радиус блокатора	0,4 м
Плотность блокаторов	0,5 1/м ²
Чувствительность приемника	-9,47 дБ
Излучаемая мощность	2 Вт
Константа затухания	2,1
Мощность теплового шума	-174 дБм/Гц
Затухание блокировки	15 дБ
Число антенных элементов на АУ	4 эл.
Усиление приема на АУ	5,57 дБ
Скорость сессий	{20,50} Мбит/с
Интенсивности обслуживания	30 с
Антенная решетка на БС	{4, 8, 16, 32}×{4}
Усиление на БС, G_A	{5,57, 8,57, 11,57, 14,58} дБ
Радиус покрытия БС	{107, 149, 207, 288} м
Дистанция между БС	$3d_E$ м
Число классов одноадресных сессий, K	1

Таблица 6.3

Радиус покрытия БС и требования сессий

Антенна	Усиление	НРBW	d_E	РБ для скоростей (20,50), Мбит/с
128×4	20,58	0,79	413	(10,22)
64×4	17,59	1,59	351	(8,19)
32×4	14,58	3,18	288	(7,16)
16×4	11,57	6,37	207	(6,14)
8×4	8,57	12,75	149	(5,12)
4×4	5,57	25,50	107	(5,11)

Для оценки характеристик обслуживания используем следующий подход. Параметризуем разработанную модель СМО, используя M классов многоадресных сессий и один класс одноадресных сессий, где M соответствует числу конфигураций антенны на БС (подсекторов с углом направленности α_A), необходимых для покрытия 120° сектора обслуживания. Таким образом, число классов соответствует потенциальному количеству передач, необходимых для распространения одной и той же информации среди всех пользователей многоадресной рассылки. Введем интенсивность потока сессий, λ_S , определенную как интенсивность поступления всех сессий в секторе, охватываемом одной конфигурацией антенной решетки БС, т.е. $\lambda_S = \lambda_B \pi d_E^2 \alpha_A / 3$, где α_A измеряется в радианах. Тогда доля многоадресных сессий всех классов составляет величину $\lambda_S \sum_{m=1}^M p_{M,m}$. Заметим, что равенство $p_{M,i} = p_{M,j}$ приводит к равенству $q_{M,i} = q_{M,j} = q_M$ для всех классов многоадресных сессий.

6.4.1. Оценка точности модели

Для оценки точности аналитической модели использована разработанная среда моделирования, основанная на событийном подходе, исходными данными для которой являются параметры из табл. 6.2.

Оценка точности проведена путем сравнения вероятностей сброса сессий, как показано на рис. 6.3 для случая равенства скоростей многоадресных и одноадресных сессий $R_U = R_M = 20$ Мбит/с. Рассматривается максимальный случай антенной решетки из 128×4 элементов, соответствующей усилению в 20,59 дБ и единичному времени обслуживания, и доля многоадресных сессий $p_M = 0,5$. Из графиков видно, что данные имитационного моделирования хорошо согласуются с результатами аналитической модели. Такой же уровень точности наблюдался и для других вариантов параметров системы, и поэтому далее в вычислениях применяется только аналитическая модель.

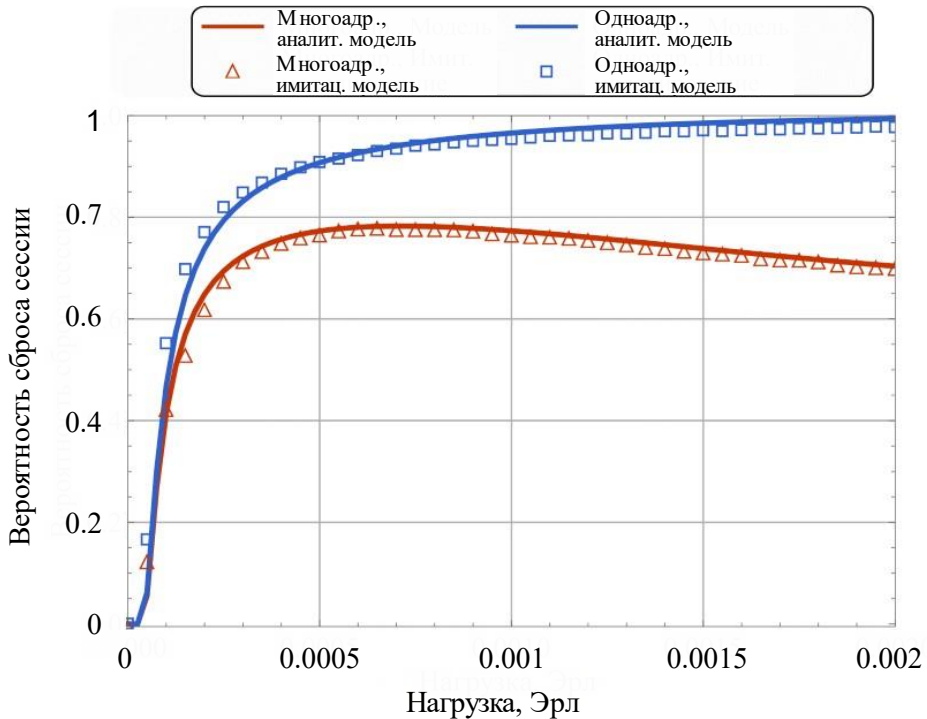


Рис. 6.3. Сравнение аналитической и имитационной моделей

6.4.2. Эффекты обслуживания многоадресных и одноадресных сессий

Сначала оценим влияние интенсивностей потоков многоадресных и одноадресных сессий, а также их времен обслуживания на характеристики системы. В расчетах используется антенная решетка 32×4 элемента, соответствующая радиусу покрытия $d_E = 288$ м и $3,18^\circ$ HPBW.

На рис. 6.4 и 6.5 показаны вероятности потерь многоадресных и одноадресных сессий, а также коэффициент использования системных ресурсов в зависимости от интенсивности поступления сессий и доли сессий многоадресной передачи для двух совпадающих значений скоростей многоадресных и одноадресных соединений 20 Мбит/с и 50 Мбит/с и среднего времени обслуживания 30 с для обоих типов сессий. На рис. 6.4 сохраняем постоянную долю многоадресных сессий, $p_M = 0,5$, и меняем интенсивность потока λ_S , тогда как на рис. 6.5 эта интенсивность является постоянной и равна 0,005 сессий на квадратный метр, а изменяется доля многоадресных сессий p_M .

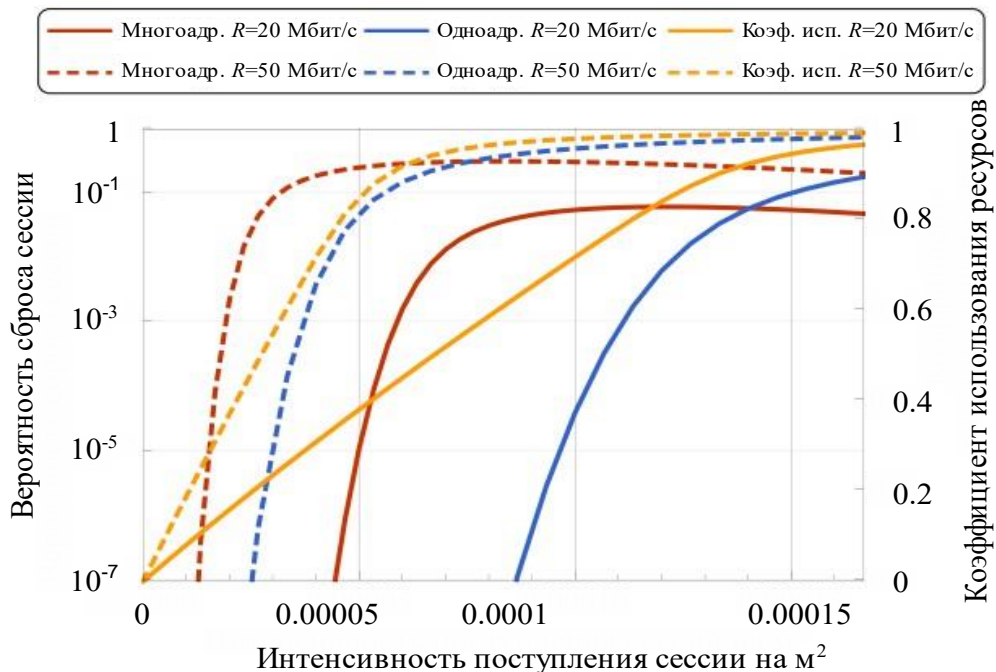


Рис. 6.4. Зависимость вероятности сброса от интенсивности поступления сессий и коэффициента использования ресурсов от интенсивности поступления сессий

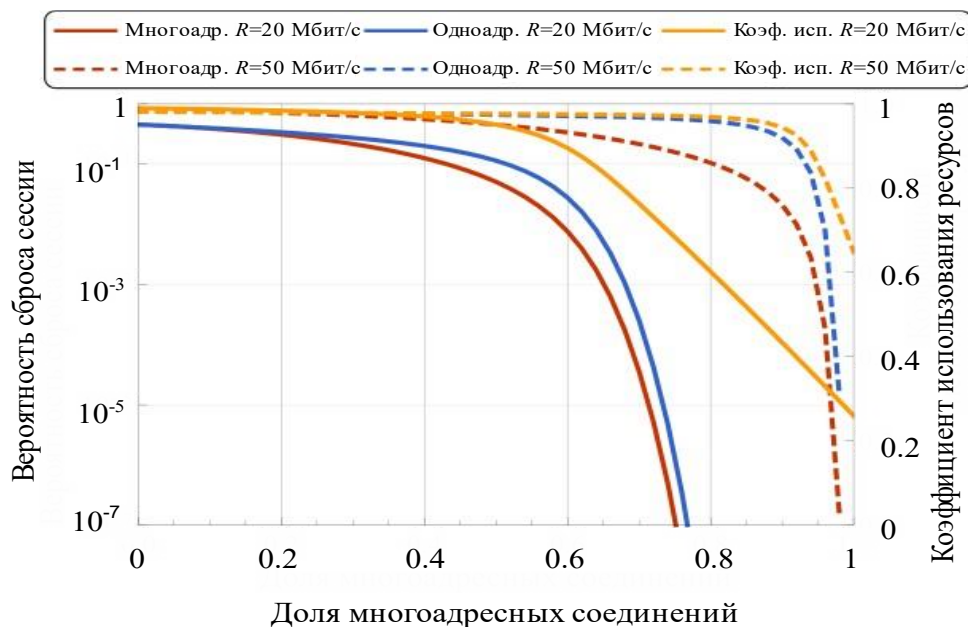


Рис. 6.5. Зависимость вероятности сброса сессий и коэффициента использования ресурсов от доли многоадресных соединений

Анализируя влияние интенсивности потока сессий на рис. 6.4, можно заметить, что для всех значений скорости увеличение интенсивности влечет увеличение вероятности сброса многоадресных сессий намного быстрее по сравнению с одноадресными сессиями. Например, для интенсивности $7,5 \times 10^{-5}$ и $R_M = R_U = 20$ Мбит/с вероятность сброса многоадресной сессии превышает величину 0,01, тогда как вероятность сброса одноадресной сессии ниже значения 10^{-7} . Однако дальнейшее увеличение интенсивности λ_s не оказывает негативного влияния на вероятность сброса многоадресных сессий, а при более высоких значениях интенсивностей даже начинает уменьшаться. Наоборот, вероятность сброса одноадресных сессий экспоненциально возрастает в рассматриваемом диапазоне интенсивностей потоков сессий. Причина заключается в том, что если в системе уже существует многоадресная сессия класса i , все новые многоадресные сессии этого класса не требуют дополнительных, что аналогично неявному приоритету обслуживания многоадресного трафика. Когда интенсивность многоадресных сессий достаточно высока, почти всегда в системе имеется хотя бы одна многоадресная сессия, и поэтому наблюдается резкое уменьшение величины q_M и значительное увеличение величины q_U .

Замеченные выше тенденции также справедливы и для скорости передачи 50 Мбит/с, разница лишь в том, что система насыщается быстрее. Это наблюдение также подтверждается поведением коэффициента использования системных ресурсов на рис. 6.4.

Теперь рассмотрим интересующие нас характеристики в зависимости от доли многоадресных сессий p_M (см. рис. 6.5). Во-первых, отметим, что вероятность сброса многоадресной сессии всегда ниже одноадресной, а разница увеличивается с ростом доли многоадресных сессий. Оба наблюдения являются следствием упомянутого эффекта «захвата» ресурсов. Фактически даже для значений $p_M > 0,4$ система всегда имеет ресурсы для обслуживания всех классов многоадресных сессий, и только оставшиеся ресурсы доступны для одноадресных сессий, что приводит к высокой вероятности сброса. Однако эффект менее заметен, когда увеличивается число элементов антенной решетки. Причина в том, что площадь области, обслуживаемой одной конфигурацией антенной решетки, уменьшается.

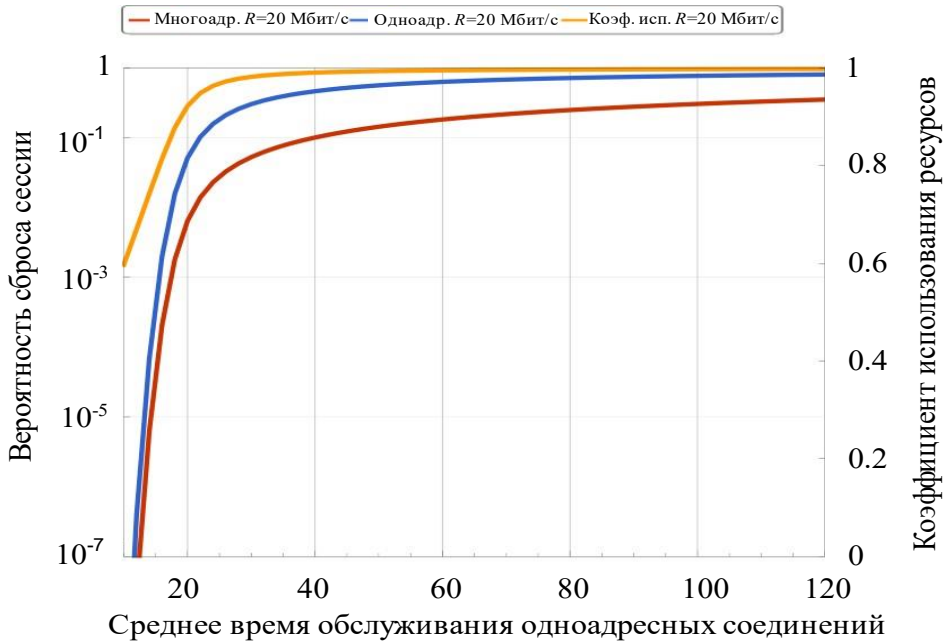


Рис. 6.6. Зависимость вероятности сброса сессии и коэффициента использования ресурсов от среднего времени обслуживания одноадресных соединений

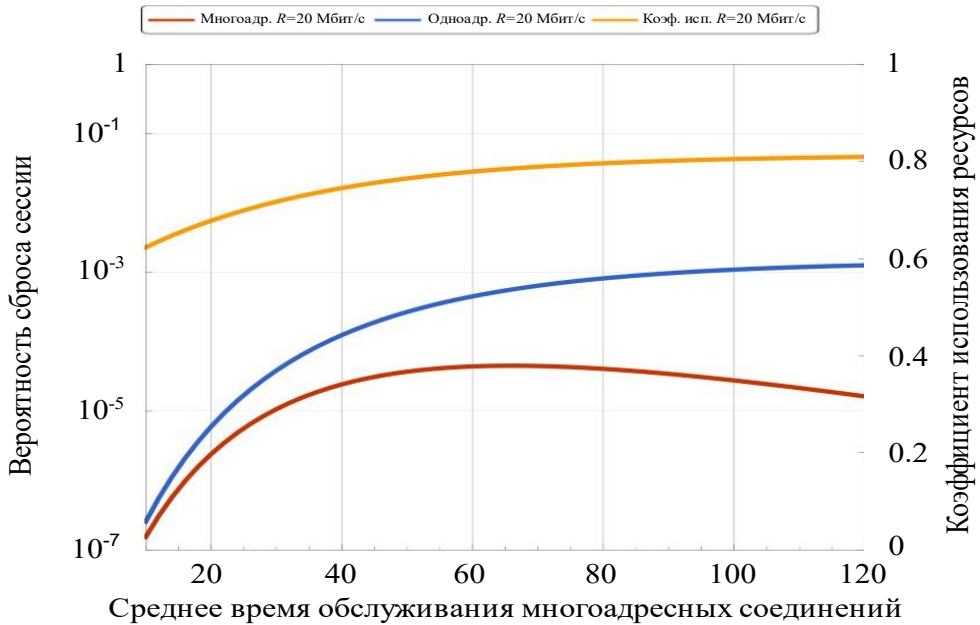


Рис. 6.7. Зависимость вероятности сброса сессии и коэффициента использования ресурсов от среднего времени обслуживания многоадресных соединений

Теперь рассмотрим влияние среднего времени обслуживания многоадресных и одноадресных сессий на вероятности сброса и коэффициент использования ресурсов системы, показанные на рис. 6.6 и 6.7. На рис. 6.6 показано влияние среднего времени обслуживания одноадресных сессий для $\lambda_S = 10^{-5}$, $p_M = 0,2$, и интенсивности поступления многоадресных сессий, определяемой средним временем обслуживания $1/\mu = 30$ с для одинаковых скоростей передачи 20 Мбит/с одноадресных и многоадресных сессий. Как и ожидалось, при увеличении среднего времени обслуживания одноадресной сессии, $1/k$ увеличивается вероятность сброса как многоадресных, так и одноадресных сессий. Заметим, что вероятность сброса одноадресных сессий приближается к 1 для $1/k = 80$ с, вероятность сброса многоадресных сессий все еще значительно ниже 1. Это можно считать эффектом неявного приоритета многоадресных сессий даже для небольшого значения их доли $p_M = 0,2$ в общем потоке сессий.

Влияние среднего времени обслуживания многоадресных сессий, $1/\mu$, проиллюстрировано на рис. 6.7. Как можно заметить, большие значения $1/\mu$ приводят к лучшим характеристикам обслуживания многоадресных сессий и увеличивают вероятность сброса одноадресных сессий. В пределе, когда $\mu \rightarrow 0$, т.е. при больших длительностях обслуживания многоадресных сессий, доля q_M приближается к 0, а значение q_U близко к 1. В этом случае система всегда занята предоставлением многоадресных услуг (32 класса для антенной решетки 32×4 на БС), почти не оставляя ресурсов для одноадресных сессий. Заметим, что это поведение не оказывает негативного влияния на коэффициент использования ресурсов, поскольку он равномерно увеличивается с ростом среднего времени обслуживания.

6.4.3. Оптимальное расстояние между базовыми станциями

Эффективный радиус покрытия БС d_E и, следовательно, расстояние между ними зависят от числа антенных элементов. Чтобы уменьшить плотность БС, требуемых для покрытия определенной области, необходимо использовать все доступные элементы антенны, увеличивая усиление в направлении передачи и приема. Однако это не всегда возможно, так как характеристики обслуживания многоадресных и/или одноадресных сессий могут становиться хуже заданных. В том числе при увеличении числа антенных элементов на БС также уменьшается ширина полосы НРВW и, таким образом, увеличивается число классов многоадресных сессий в системе.

Последнее приводит к более высокой нагрузке многоадресного трафика и может отрицательно влиять на вероятности сброса как многоадресных, так и одноадресных сессий.

Переходим к анализу соотношения между числом антенных элементов, используемых на БС, и показателями характеристик обслуживания многоадресных и одноадресных сессий. Последующее обсуждение иллюстрирует процедуру определения максимального числа элементов антенны БС и, следовательно, максимального расстояния между ними, которое может использоваться для определения гарантированных вероятностей сброса многоадресных и одноадресных сессий.

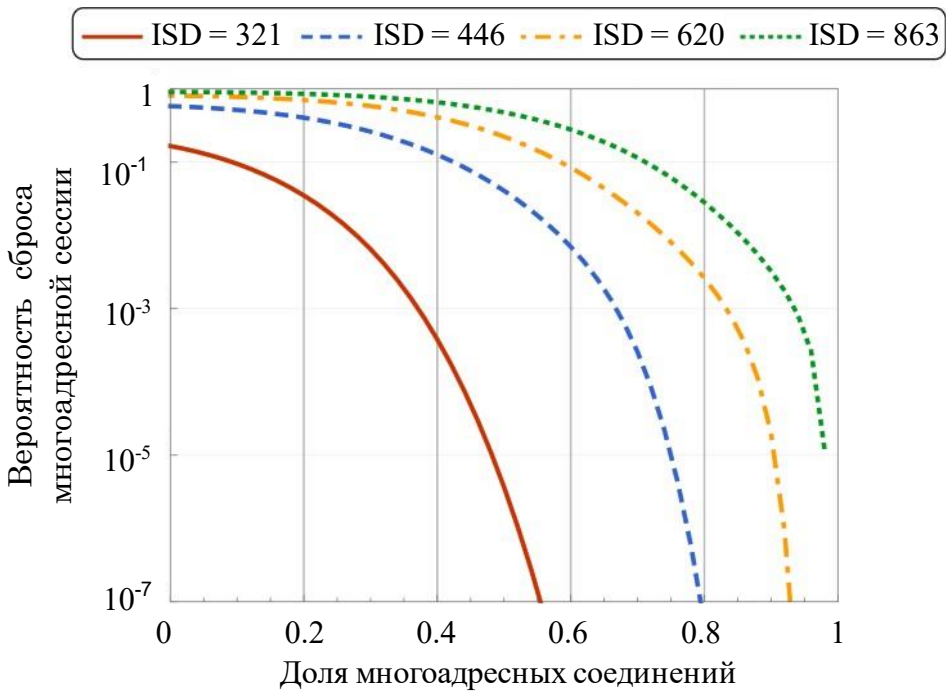


Рис. 6.8. Зависимость вероятности сброса многоадресной сессии от доли многоадресных соединений

Вероятности сброса многоадресных и одноадресных сессий в зависимости от доли многоадресных сессий p_M для различных расстояний между БС показаны на рис. 6.8 и 6.9 для средней продолжительности многоадресных и одноадресных сессий 30 с, интенсивности поступления сессий 0,0015 и совпадающих скоростей $R_U = R_M = 20$ Мбит/с, где ISD (Inter-Site Distance) – расстояние между БС в сотовом покрытии.

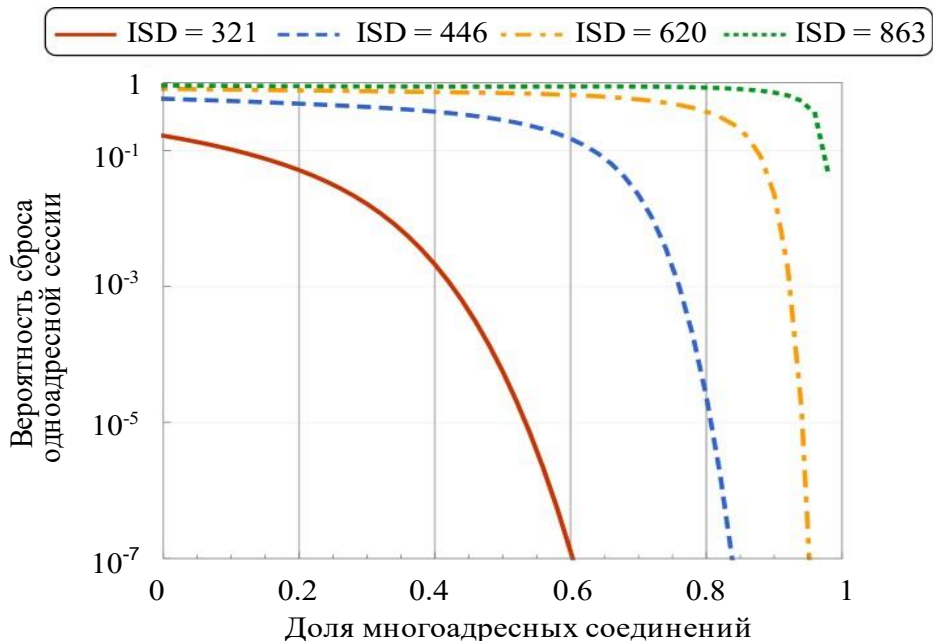


Рис. 6.9. Зависимость вероятности сброса одноадресной сессии от доли многоадресных соединений

Сравнивая результаты, показанные на рис. 6.8 и 6.9, для фиксированных значений интенсивности поступления сессий и расстояния между БС, замечаем, что увеличение доли p_M приводит к снижению вероятностей сброса обоих типов сессий. Причина в том, что большие значения p_M уменьшают фактическую нагрузку на систему. Кроме того, для небольшого числа антенных элементов (например, массива 4×4 на БС), вероятности сброса многоадресных и одноадресных сессий практически совпадают. Однако при большем числе антенных элементов и, следовательно, меньшем значении НРВW разница становится значительной, особенно при средних и больших значениях p_M . Обратим внимание, что разница может быть весьма существенной в абсолютных значениях. В частности, для $p_M = 0,8$ и антенной решетки 16×4 , соответствующей расстоянию 620 м между БС, вероятность сброса многоадресной сессии составляет всего $q_M = 0,005$, тогда как для одноадресных сессий это значение достигает значения 0,6. Основной причиной является число классов многоадресных сессий, которые необходимо поддерживать на БС. В частности, при использовании антенной решетки 32×4 , соответствующей расстоянию между БС 863 м, потребуется 32 отдельных класса многоадресных сессий.

Вероятности сброса многоадресных и одноадресных сессий в зависимости от интенсивности поступления сессий показаны на рис. 6.10 и 6.11 для различных расстояний между БС, доли многоадресных сессий $p_M = 0,5$, совпадающих скоростей $R_U = R_M = 20$ Мбит/с и средних продолжительности сессий $1/\kappa = 1/\mu = 30$ с. На рис. 6.10 можно заметить эффект захвата ресурсов многоадресными сессиями, где виден пик вероятности сброса многоадресных сессий для всех рассматриваемых значений числа элементов антенной решетки. Объяснение этого эффекта заключается в том, что при фиксированном значении доли p_M увеличение интенсивности потока одноадресных сессий приводит и к увеличению интенсивности поступления многоадресных сессий. Из-за специфики процесса обслуживания это влечет увеличение вероятности того, что по прибытии новой сессии данный класс многоадресных сессий уже присутствует на БС, тем самым уменьшая вероятность сброса многоадресных сессий. Обратим внимание, что чем меньше число антенных элементов, тем меньше максимальная вероятность сброса многоадресных сессий.

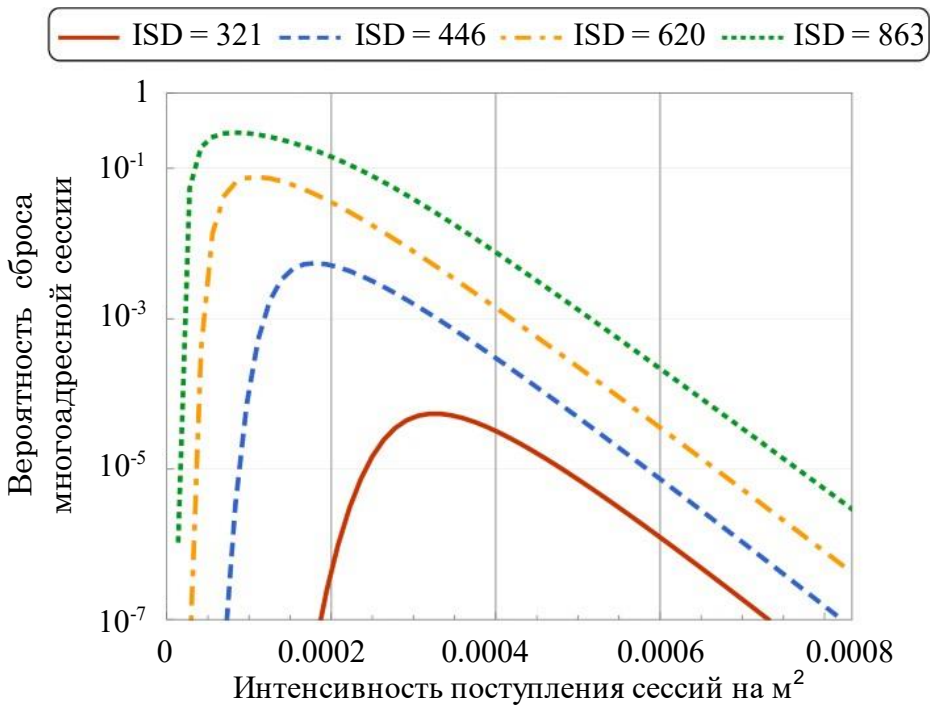


Рис. 6.10. Зависимость вероятности сброса многоадресной сессии от интенсивности поступления сессий

Как и ожидалось, вероятность сброса одноадресных сессий возрастает с увеличением интенсивности λ_s для всего диапазона расстояний между БС. Однако это происходит намного быстрее по сравнению со случаем только одноадресных сессий, поскольку на ее увеличение отрицательно влияет уменьшающаяся вероятность сброса многоадресных сессий.

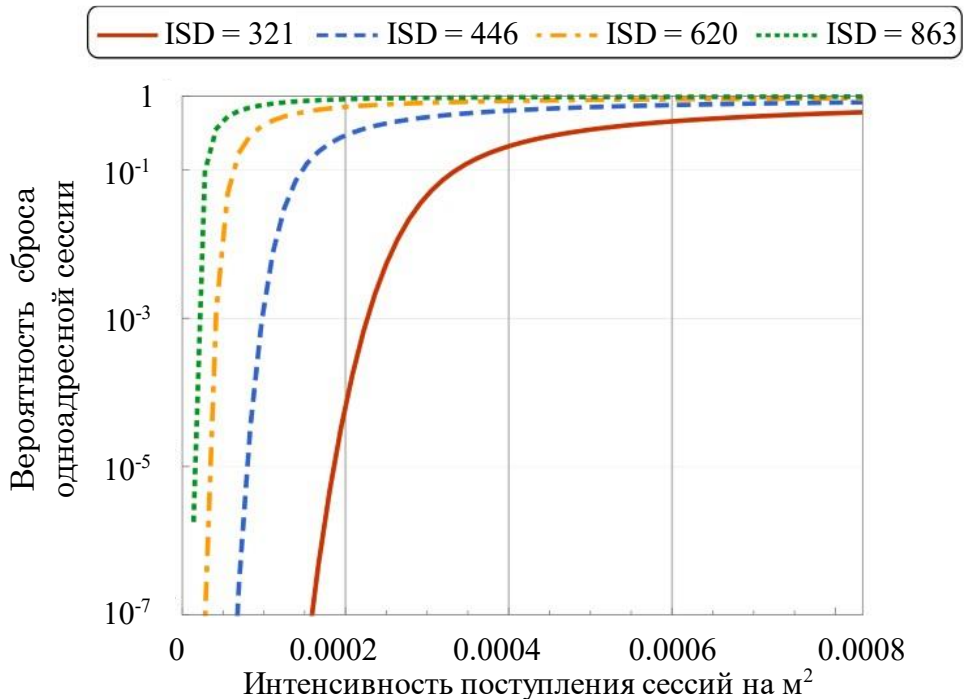


Рис. 6.11. Зависимость вероятности сброса сессии от интенсивности поступления сессий

Подводя итоги, отметим, что, хотя использование большого числа антенных элементов позволяет увеличить зону покрытия и потенциально снизить уровень помех, это также может значительно увеличить нагрузку на БС. Используя методы управления мощностью, можно уменьшить нагрузку одноадресного трафика, ограничив покрытие БС, но число классов многоадресных сессий, которые необходимо поддерживать, все еще велико из-за использования малого значения НРВВ. В то же время использование меньшего числа антенных элементов на БС для формирования диаграмм направленности антенны часто невозможно, поскольку это оставляет часть области развертывания системы непокрытой многоадресными услугами

и, следовательно, требует уплотнения сети доступа. Второе интересное наблюдение заключается в том, что увеличение доли многоадресных сессий снижает вероятность их сброса. Фактически многоадресный трафик неявно получает приоритет над одноадресным трафиком в результате специфики его обслуживания. Таким образом, в практических развертываниях систем с высокой долей многоадресных услуг операторы могут рассмотреть возможность некоторого механизма резервирования ресурсов для одноадресного трафика и/или алгоритма для явного контроля доступа к системе.

Рассмотренная методология позволяет оценивать параметры развертывания сотовых систем БС таким образом, чтобы были достигнуты заданные характеристики обслуживания. Отметим, что на практике эти результаты обеспечивают нижнюю границу плотности развертывания и должны быть уточнены с использованием специфики среды распространения.

6.4.4. Использование одноадресных сессий для многоадресной передачи

Использование антенных решеток с большим числом элементов позволяет значительно улучшить покрытие предполагаемых зон обслуживания БС, снижая затраты на развертывание сети. Однако такие антенны приводят к тому, что чрезвычайно малые значения НРВW увеличивают число многоадресных классов, которые необходимо поддерживать на БС, что приводит к неэффективному использованию ресурсов системы. Принимая во внимание неявный приоритет, получаемый многоадресными сессиями, уровень обслуживания, предоставляемый одноадресным сессиям, резко снижается. Таким образом, вопрос, на который следует ответить, заключается в том, является ли услуга многоадресной передачи все еще полезной, когда на БС используются антенные решетки с большим числом элементов, или можно полагаться исключительно на одноадресную услугу. Таким образом, сравнивается система с многоадресными и одноадресными сессиями и система, где многоадресные сессии обслуживаются с использованием одноадресного режима передачи.

Сравним вероятности сброса для многоадресной/одноадресной системы с одноадресной системой, показанные на рис. 6.12 и 6.13, в зависимости от интенсивности потока сессий для скоростей при $R_U = R_M = 20$ Мбит/с и продолжительностей сессий $1/\kappa = 1/\mu = 30$ с, двух конфигураций антенных решеток, 32×4 с НРВW $3,19^\circ$ и 16×4

с НРВW 6,37° (см. табл. 6.3). На рис. 6.12 представлены результаты вычислений для потоков с фиксированной долей многоадресных сессий, $p_M = 0,5$. Из графиков видно, что для рассматриваемого числа антенных элементов вероятность сброса в системе с одноадресной услугой всегда выше по сравнению с вероятностью сброса в системе с двумя типами соединений. Однако разница между двумя системами уменьшается, когда на БС увеличивается число антенных элементов. Причина в том, что общее число классов многоадресных сессий, которые необходимо поддерживать на БС, увеличивается. Кроме того, даже для антенны 128×4 система с одноадресной услугой все еще значительно хуже во всем рассматриваемом диапазоне интенсивностей поступления сессий.

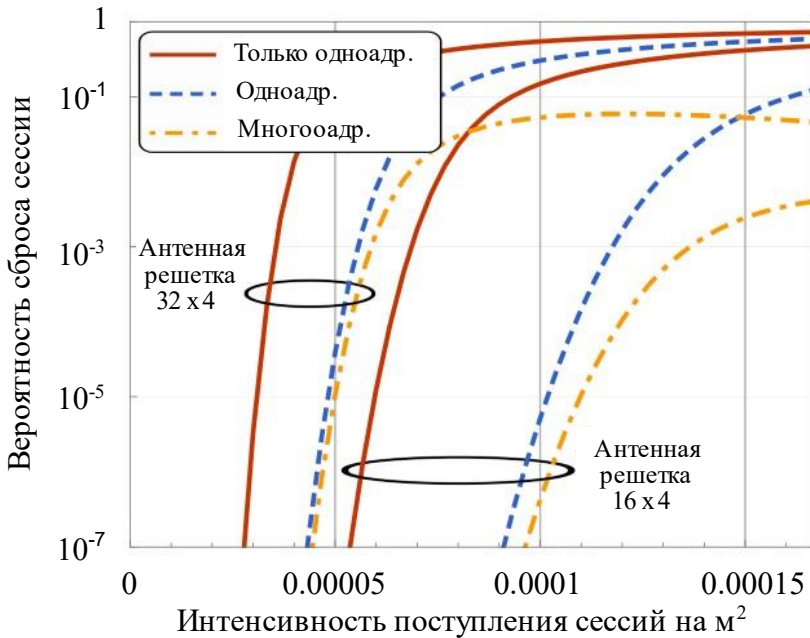


Рис. 6.12. Зависимость вероятности сброса многоадресной сессии от интенсивности поступления сессий

Данные рис. 6.13, иллюстрирующие влияние доли многоадресных сессий, подтверждают, что вышеупомянутые выводы верны практически для любой суперпозиции потоков многоадресных и одноадресных сессий. Также важно отметить, что разница между одноадресной и смешанной одноадресной и многоадресной системами возрастает с увеличением доли p_M . Можно сделать вывод, что, с точки зрения характеристик обслуживания абонентов, для современных

антенных решеток смешанная система всегда превосходит одноадресную систему. Однако этот вывод может измениться в будущем, когда появятся антенные решетки с более чем 128×4 элементами. Такие системы могут потенциально реализовывать многоадресную услугу, используя исключительно одноадресный режим передачи и, таким образом, не требуют поиска компромисса между вероятностями сброса и радиуса покрытия одной БС. Такой системой намного проще управлять, поскольку все сессии обрабатываются одинаково по сравнению со смешанной системой, где многоадресные сессии имеют неявный приоритет над одноадресными.

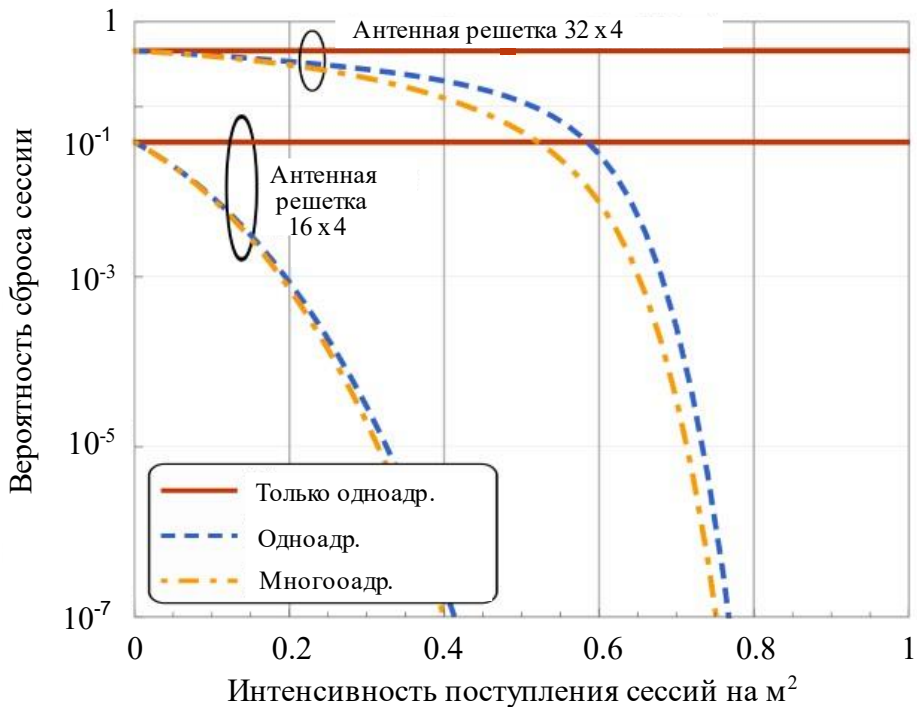


Рис. 6.13. Сравнение системы с многоадресными/одноадресными соединениями и системы с одноадресными соединениями

Проведенный анализ показывает, что наличие многоадресных сессий оказывает негативное влияние на характеристики обслуживания одноадресных сессий на БС. Это влияние усиливается, когда: а) увеличивается доля одноадресных сессий, или б) увеличивается интенсивность поступления сессий, или в) увеличивается отношение между временем обслуживания многоадресных и одноадресных

сессий. Если необходимо сбалансировать вероятности сброса многоадресных и одноадресных сессий, необходимо реализовать явную схему приоритетов на БС, например, используя резервирование для одноадресных сессий. Для заданной доли многоадресных сессий в общем потоке использование антенн с малой направленностью, характеризующихся меньшим расстоянием между ними, ведет к меньшим потерям как одноадресных, так и многоадресных сессий вследствие уменьшения необходимого числа классов многоадресных сессий.

Попытка увеличить расстояние между БС путем реализации многоадресной услуги с помощью режима одноадресной передачи приводит к гораздо худшей производительности с точки зрения вероятности сброса сессии для практических диапазонов направленности антенны на БС. Дальнейшее снижение ширины НРВW за счет увеличения числа антенных элементов, формирующих диаграмму направленности БС, позволяет уменьшить разницу между вероятностями сброса сессий в системе где многоадресная услуга предоставляется абонентам, используя одноадресные сессии, по сравнению со смешанной одноадресной/многоадресной системой.

Рассмотренная модель также дает возможность оценить плотность БС, необходимую для покрытия определенной области с заданными нагрузками трафика многоадресных и одноадресных сессий. Однако заметим, что на практике это означает обеспечение нижней границы плотности развертывания и результаты должны быть уточнены с использованием специфики среды распространения.

ГЛАВА 7

УПРАВЛЕНИЕ ПЕРЕГРУЗКАМИ В СЕТЯХ 5G/6G СРЕДСТВАМИ БПЛА

Несмотря на то, что сети 5G/6G призваны обеспечить сверхвысокую скорость на абонентском участке, динамика роста современного трафика такова, что даже в таких сетях могут возникать перегрузки и требуется решать задачи передачи потоков избыточного трафика дополнительными средствами. В этой главе мы рассмотрим управление перегрузками в сетях 5G/6G средствами подвижных сетевых элементов – БС, установленных на беспилотных летательных аппаратах (БПЛА).

В параграфе 7.1 рассмотрены основные причины пространственной и временной динамики трафика на участке доступа в сетях 5G/6G, приводящей к перегрузкам. В 7.2 кратко представлено состояние стандартизации новых элементов сети, переносимых БПЛА, включая БС и ретрансляционные узлы. Оценка плотности развертывания БС на основе БПЛА приведена в 7.3. Анализ участков доступа и транзита для таких БС представлен в 7.4, а процесс организации обслуживания на абонентском участке рассмотрен в параграфе 7.5. Наконец, в 7.6 исследован процесс выгрузки избыточного трафика при помощи ретрансляционных узлов, образующих так называемые БПЛА-мосты.

7.1. Причины возникновения перегрузок

Современные потребности в мобильном трафике характеризуются высокой степенью временных и пространственных изменений и, как ожидается, сохраняют эти свойства в ближайшем будущем. Пространственный компонент в основном обусловлен локализованными событиями, например спонтанными скоплениями людей в определенной области на местности. Кроме того, некоторые исследования показывают, что временная динамика мобильного трафика становится гораздо менее предсказуемой. Одной из причин является распространение современных приложений, в том числе AR/VR игры. Например, виртуальные объекты, появляющиеся в случайных географических местах в AR/VR играх, вызывают спонтанные перемещения абонентов и связанную с ними нагрузку.

Высокая динамика трафика во времени и пространстве делает неэффективными традиционные «статические» методы

планирования сети. Ожидается, что ситуация не изменится даже с внедрением технологии 5G/6G, работающей в миллиметровом диапазоне частот, так как эти системы доступа будут состоять из изолированных «островов», предоставляющих высокую скорость доступа на беспроводном участке. Еще одним важным свойством мобильного трафика является то, что большинство требований локализованы в пространстве, что часто делает неэффективным использование сотовой инфраструктуры.

Как поставщики сетевого оборудования, так и организации по стандартизации систем сотовой связи 5G/6G в настоящее время ищут «симметричный» ответ для обслуживания трафика с пространственной и временной кластеризацией. Использование мобильных БС, в частности устанавливаемых на автомобилях (CoW, Cell-on-Wheels) и на летающих БС (ABS, Aerial Base Stations), считается одним из возможных вариантов обслуживания такого трафика. В частности, рекомендации 3GPP по интеграции сетей доступа и механизмов выгрузки трафика (IAB, Integrated Access and Backhaul) требуют обеспечения совместного управления сетями доступа и транзитными участками, оптимизации выбора маршрута выгрузки трафика и высокой спектральной эффективности как на участке доступа, так и на транзитном участке. Такие решения характеризуются временем доставки CoW и ABS порядка нескольких минут.

Другим способом удовлетворения условиям спонтанного мобильного трафика является использование технологии D2D на основе стандартов Wi-Fi/WiGig Direct, LTE Direct и будущей технологии NR Direct (см. главу 10). Этот подход начал исследоваться, в частности, продемонстрировано, что в зависимости от услуги связи до 40 % трафика может быть локализовано непосредственно в сети доступа. Однако подчеркивается, что эффективность выгрузки трафика сильно зависит от доступности контента в непосредственной близости и готовности абонентов делиться ресурсами.

Для улучшения доступности локализованного контента и полного использования возможностей современного коммуникационного оборудования естественным расширением является использование нескольких беспроводных интерфейсов, доступных на различных устройствах для организации Mesh-сетей с несколькими технологиями доступа. Такие системы могут создаваться спонтанно по требованию в местах, где потребности в трафике временно возрастают. Такие сети потенциально могут обеспечить: а) эффективную

локализацию трафика на больших площадях за счет работы в «многошаговом» режиме передачи в процессе обслуживания требований; б) динамическое совместное использование ресурсов, включая те, что предоставляются инфраструктурой.

Далее в этой главе исследованы методы и алгоритмы выгрузки трафика из сети радиодоступа 5G/6G с применением БПЛА.

7.2. Стандартизация применения БПЛА

Далее сделан краткий обзор исследований, проводимых консорциумом 3GPP, которые играют важную роль в использовании БПЛА. Рассмотрены применение БПЛА в сотовых системах 5G/6G, архитектурные решения IAB и функции ретрансляции в сетях 5G/6G.

7.2.1. Интеграция БПЛА в беспроводную сеть

Применение БПЛА и их интеграция в беспроводные сети привлекает внимание как сетевых операторов, так и организаций по стандартизации. Стандарт 3GPP TR 36.777 суммирует проведенные исследования по поддержке БПЛА в сетях LTE. В том числе рассматриваются возможные улучшения сотовой сети для эффективного обслуживания трафика и влияния применения БПЛА на сеть, проводится оценка характеристик обслуживания с использованием БПЛА в городских и сельских условиях. Проведенное компьютерное моделирование таких систем, дополненное данными измерений, показывает, что использование БПЛА может приводить к увеличению помех как в восходящем, так и нисходящем каналах связи. Стандарт TR 36.777 также предлагает методы устранения помех. Еще одна проблема, идентифицированная в TR.36.777, связана с мобильностью БПЛА. Стандарт определяет методы предоставления дополнительной информации о развернутой наземной сети, которая может использоваться при принятии решений в процессе полета.

Начиная с 3GPP Release 16 поддержка БПЛА рассматривается как одна из ключевых особенностей инфраструктуры сотовых сетей 5G/6G. В этом контексте стандарт TR 22.829 суммирует варианты использования и анализирует функции БПЛА, которые могут требовать расширенной поддержки со стороны сетей доступа. Это включает в себя приложения для трансляции видео, услуги управления и

контроля, а также использование БПЛА в качестве подвижных БС. Последнее приложение БПЛА детально рассмотрено в стандарте TR 38.811.

Консорциум 3GPP продолжает исследования в данном направлении, им проводится работа по снижению негативных эффектов, вызванных мобильностью БПЛА, адаптации к потребностям бизнеса, обеспечению безопасности и удаленной идентификации БПЛА. С этой целью стандарт TR 22.125 определяет эксплуатационные требования для систем 3GPP. Ожидается, что 3GPP улучшит поддержку методов интеграции БПЛА в сетях связи 5G/6G в дальнейших редакциях стандартов TR 23.754 и TR 23.755. Тем не менее уже сейчас становится понятно, что БПЛА способны обеспечивать широкий спектр услуг в сетях доступа 5G/6G.

7.2.2. Архитектурное решение IAB

По сравнению с наземными развертываниями сетей 5G/6G, основным ограничением БПЛА, с точки зрения сетевой инфраструктуры, является необходимость в так называемом транзитном подключении. Обычно NR БС имеют фиксированное проводное транзитное подключение к сетевой инфраструктуре и поэтому могут предлагать очень высокие скорости передачи данных на интерфейсе доступа. Единственным доступным вариантом для БПЛА является ретрансляция трафика к ближайшей точке доступа в сеть 5G/6G, в роли которой чаще всего и выступает NR БС. Концепция использования единой технологии радиосвязи для обеспечения как доступа, так и транзитного соединения, рассмотрена в стандарте 3GPP TR 38.874. С внедрением сетей доступа 5G/6G, которые поддерживают высоконаправленные антенные решетки, БПЛА, оснащенные функциями IAB, могут упростить уплотнение сети по требованию, тем самым эффективно избегая помех, сокращая капитальные вложения в сетевую инфраструктуру, а также обеспечивая эффективное обслуживание трафика, характеризующегося высокой пространственной и временной динамикой.

Преимущества IAB для сетей 5G/6G изначально рассмотрены в документе 3GPP RP-171880. Далее концепция была принята в стандарте 3GPP TR 38.874. В настоящее время термин узел IAB определяется 3GPP следующим образом: *«Узел IAB является узлом сети*

доступа, который поддерживает беспроводной доступ на абонентском участке и осуществляет беспроводную транзитную передачу на NR БС» (см. TR 38.874). Консорциум 3GPP не предписывает какую-либо конкретную технологическую реализацию IAB, оставляя решения на усмотрение производителей оборудования и операторов сетей связи 5G/6G. С точки зрения планирования сети доступа, доступные варианты включают в себя реализацию с одним или несколькими переприемами, внутриполосную или внеполосную транзитную передачу, а также использование различных технологий на участках доступа и транзита.

Несмотря на то, что транзитная передача с несколькими переприемами может улучшить покрытие сети, она также приводит к дополнительным накладным расходам с точки зрения сигнализации. В дополнение к обычным процедурам управления сетью (случайный доступ, передача вызова, управление мощностью и т.д.) системы IAB, как с несколькими, так и с одним переприемом, должны обеспечивать специфические для ретрансляции функциональные возможности, такие как: а) обнаружение, управление и повторное установление транзитного соединения; б) выделение и координация ресурсов; в) алгоритмы устранения помех между несколькими транзитными соединениями.

Тип беспроводной технологии и используемая полоса частот на участках доступа и транзита все еще являются предметом обсуждения. Например, если оба соединения реализованы с помощью одной и той же технологии, появляется возможность использовать совместные механизмы распределения ресурсов, которые могут снизить общую пропускную способность системы, но при этом – снизить и затраты на развертывание. Кроме того, если как участок доступа, так и транзитный участок работают в одной и той же полосе частот (внутриполосная ретрансляция), существует необходимость в дополнительном управлении помехами. Согласно TR 38.874, IAB-узел также должен быть способен обеспечить функциональность множественного доступа к среде передачи.

Помимо архитектурных решений, реализация функций IAB может отличаться в зависимости от уровня интеграции транзитного участка и участка доступа. Например, можно использовать отдельные реализации физического и канального уровней для доступа и транзита или же использовать их одновременно, разделяя определенные элементы планирования в общем модуле.

7.2.3. Ретрансляция для поддержки IAB

Ретрансляция, согласно технологии NR для поддержки архитектуры IAB, рассматривается в стандарте 3GPP TR 38.874. Роль системы ретрансляции заключается в соединении АУ с донорской NR БС, которая непосредственно связана с транспортной сетью оператора. В данном контексте преимущества ретрансляции IAB для NR связаны с уплотнением сети доступа для повышения надежности обслуживания абонентов без необходимости уплотнения транспортной сети. Однако проблемы покрытия распространения миллиметровых радиоволн из-за эффекта блокировки остаются фундаментальным ограничением для будущих развертываний таких систем.

Подобно NR БС, ретрансляционный узел NR может работать в автономном или управляемом режимах, как описано в стандарте 3GPP TR 38.912. Таким образом, большинство решений для сетей 5G/6G, определенных в документах 3GPP Release 15 и 16 (см., TS 38.300), повторно используются для транзитных соединений.

Предварительные результаты компьютерного моделирования, проводимого согласно принципам стандарта 3GPP TR 38.874, демонстрируют значительные преимущества использования функций ретрансляции в архитектуре IAB. В 3GPP также исследуется поддержка многошаговой ретрансляции (см., например, 3GPP RP-1806008, 3GPP RP-1806814 и 3GPP RP-1806815). Однако ожидаемые дополнительные выгоды от использования функций переприема ставят новые вопросы, связанные с поддержкой сетевой топологии, выбором наилучшего маршрута, оптимизацией распределения ресурсов и т.д. Несмотря на это, гибкая топология многошаговой ретрансляции с несколькими узлами переприема считается одним из ключевых компонентов будущих систем связи 5G/6G.

Ожидается, что в 3GPP Release 17 исследования по поддержке функции многошаговой ретрансляции NR продолжатся. Одним из рассматриваемых сценариев, обсуждаемых в стандарте 3GPP TR 22.866, является поддержка ретрансляции для повышения энергоэффективности и улучшения покрытия сотовой сети.

7.3. Оценка производительности сетей 5G/6G на основе БПЛА

Перед анализом методов доступа и транзита в миллиметровом диапазоне частот оценим допустимую плотность развертывания

БПЛА на абонентском участке доступа, используя в качестве параметра производительности сети совокупную помеху и ОСШ. Будем использовать принципы развертывания, сформулированные в параграфах 7.2–7.4.

7.3.1. Системная модель

Рассмотрим сценарий плотного развертывания БПЛА, проиллюстрированный на рис. 7.1. Рассматривается нисходящий канал от БПЛА к АУ. На рисунке точки местонахождения АУ обозначены U_i , $i = 1, 2, \dots$. Плотность взаимодействующих пар, измеряемая далее в этом разделе числом пар БПЛА на квадратный метр, моделируется PPP с плотностью λ . Высота размещения АУ предполагается экспоненциально распределенной с.в. с параметром μ_R . Каждое АУ связано с обслуживающим его БПЛА, обозначенным T_i , $i = 0, 1, \dots$, на рис. 7.1. Позиция (точка нахождения) БПЛА, обслуживающего АУ, является с.в., равномерно распределенной в круге радиуса R_T с центром в точке расположения соответствующего АУ. С.в. высоты БПЛА считается распределенной по экспоненциальному закону с параметром μ_T .

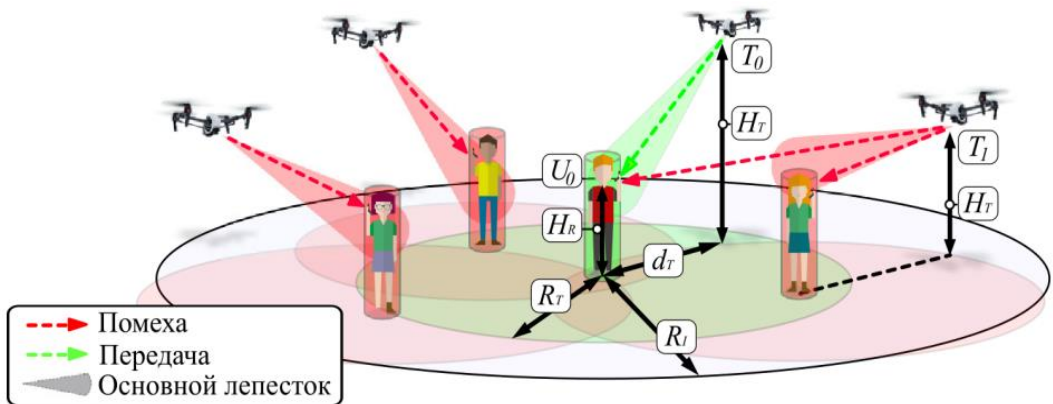


Рис. 7.1. Сценарий плотного размещения БПЛА

Предполагается, что все БПЛА работают на одной и той же частоте, тем самым являясь источниками помех для АУ. Среди пар БПЛА-АУ случайным образом выбираем произвольную пару и ограничиваем область вокруг помеченного АУ кругом с радиусом R_T . Помехи, создаваемые БПЛА, расположенными вне данного круга, предполагаются незначительными, то есть ниже минимального

уровня шума. Радиус R_l вычисляется в соответствии с моделью распространения УМi, мощности передачи и трехмерной модели антенной решетки, введенными в главе 2. Оценка совокупной помехи и ОСШ для рассматриваемой в сценарии системной модели проведена с использованием метода, предложенного в главе 4.

7.3.2. Численный анализ

Перейдем к оценке влияния плотности расположения БПЛА на уровень помех и среднее значение ОСШ. Отметим, что разработанная трехмерная модель достаточно сложна для вычислений, поскольку включает численное интегрирование. Предположение о фиксированных высотах расположения БПЛА и АУ во многом упрощает вычисления. Упрощение также может состоять в исследовании двухмерной модели, которая предполагает проекцию расположения БПЛА и АУ на плоскость. В этом случае формулы для начальных моментов с.в. ОСШ могут быть получены в замкнутом аналитическом виде (см. главу 4). Помимо оценки помехи также проведем сравнение моделей. Для краткости изложения введем обозначения для рассматриваемых далее моделей. Обозначим: ФВ – фиксированные высоты, СВ – случайные высоты, ВЭ – вертикальная экспозиция (направленность в вертикальной плоскости). Например, в таких обозначениях наиболее общая трехмерная модель, учитывающая как случайные высоты, так и вертикальную направленность, обозначается как СВ+ВЭ. Модель ФВ+ВЭ предполагает фиксированные высоты и учитывает вертикальную направленность, модель СВ–ВЭ предполагает случайные высоты, но не учитывает вертикальную направленность, и наконец, модель ФВ–ВЭ представляет простейший случай, когда высоты считаются фиксированными и мы также не учитываем вертикальную направленность антенны.

Среднее значение ОСШ, как функция угла направленности антенны α , показано на рис. 7.2 для средней высоты расположения БПЛА $E[H_T] = 40$ м, средней высоты расположения АУ $E[H_R] = 1,5$ м и плотности взаимодействующих пар $\lambda = 0,5$ на м^2 . На рис. 7.3 показано поведение ОСШ в аналогичных условиях для случая значения плотности взаимодействующих пар $\lambda = 2$ на м^2 . Во-первых, отметим, что для рассматриваемых моделей среднее значение ОСШ уменьшается с ростом α , это означает, что модели правильно отражают качественное поведение системы. При этом абсолютная разница

значений ОСШ обусловлена вероятностью экспозиции p_V , которая увеличивается по мере уменьшения угла направленности α . При этом модель с фиксированной высотой и без учета вероятности вертикальной экспозиции (ФВ–ВЭ) для значения $\lambda = 0,5$ дает ошибку оценки ОСШ около 7 дБ для $\alpha = 60^\circ$ и 27 дБ для $\alpha = 2^\circ$. Значения ОСШ в модели со случайными высотами, которая не учитывает вероятность вертикальной экспозиции (СВ–ВЭ), сходятся к значениям ОСШ для модели с вертикальной экспозицией по мере роста угла α . Лучшее приближение к модели СВ+ВЭ наблюдается для случая, который учитывает вертикальную экспозицию и предполагает фиксированные высоты БПЛА и АУ (ФВ+ВЭ). Однако это приближение зависит от угла направленности антенны. Для малых значений α модель ФВ+ВЭ дает грубую оценку сверху среднему значению ОСШ модели СВ+ВЭ.

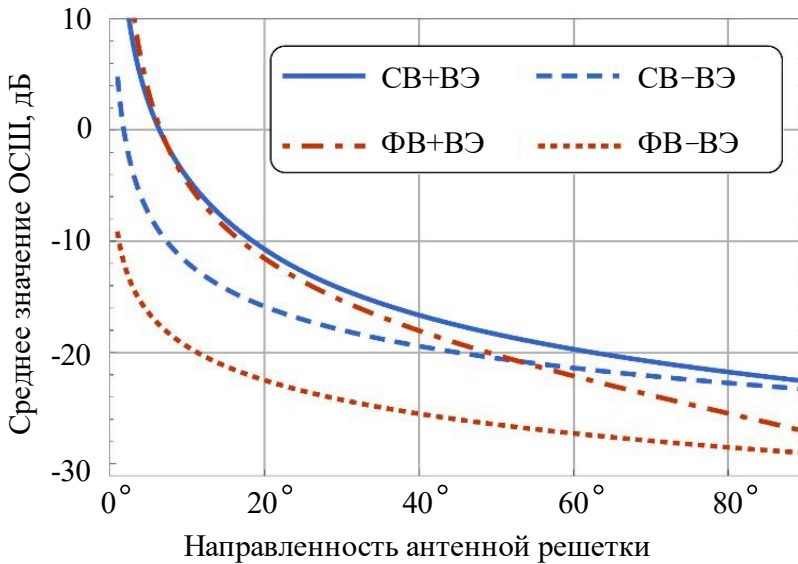


Рис. 7.2. Эффект направленности антенны на БПЛА при плотности взаимодействующих пар $\lambda = 0,5$ на m^2

На рис. 7.4 показано влияние плотности взаимодействующих пар λ и средней высоты БПЛА на ОСШ для двух значений направленности передающих антенн $\alpha = 2^\circ$ и $\alpha = 30^\circ$, для плотности $\lambda = 0,5$ и средней высоты АУ $E[H_R] = 1,7$ м. Следует подчеркнуть, что значение λ не оказывает существенного влияния на расхождение между моделями, поскольку абсолютная разница между графиками практиче-

ски одинакова во всем диапазоне значений рассматриваемых параметров. Кроме того, из графиков видно, что среднее значение ОСШ остается больше 10 дБ даже для экстремально высокой плотности взаимодействующих пар $\lambda = 0,6$ для направленности $\alpha = 2^\circ$.

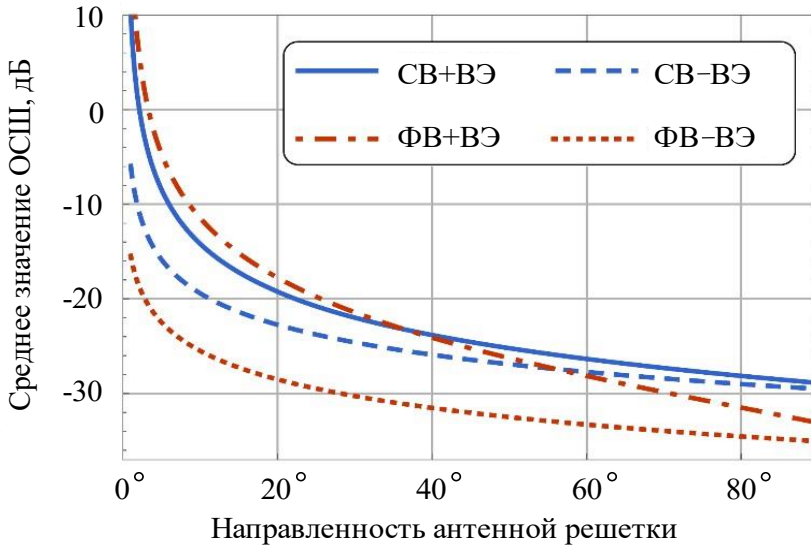


Рис. 7.3. Эффект направленности антенны на БПЛА при плотности взаимодействующих пар $\lambda = 2$ на м^2

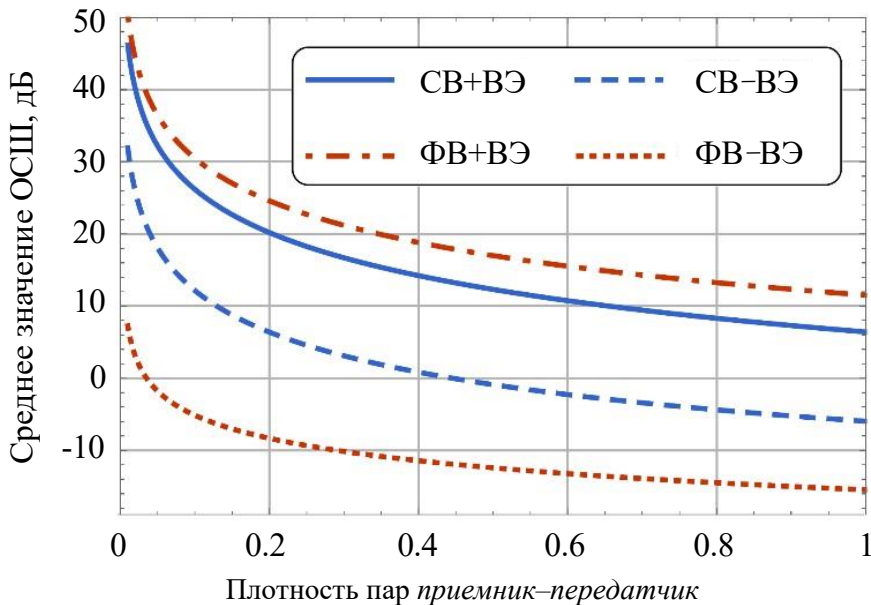


Рис. 7.4. Эффект плотности расположения взаимодействующих пар

Оценивая влияние высоты расположения на среднее значение ОСШ, на рис. 7.5 можно заметить, что для всех рассматриваемых значений высот среднее значение ОСШ остается значительно больше 10 дБ. Отметим, что модели, которые не учитывают вероятность вертикальной экспозиции, дают весьма грубую оценку снизу среднему значению ОСШ. Кроме того, такая же весьма грубая оценка снизу наблюдается и для модели с фиксированной высотой и вертикальной вероятностью экспозиции для средних значений высоты расположения БПЛА $E[H_T] > 20$.

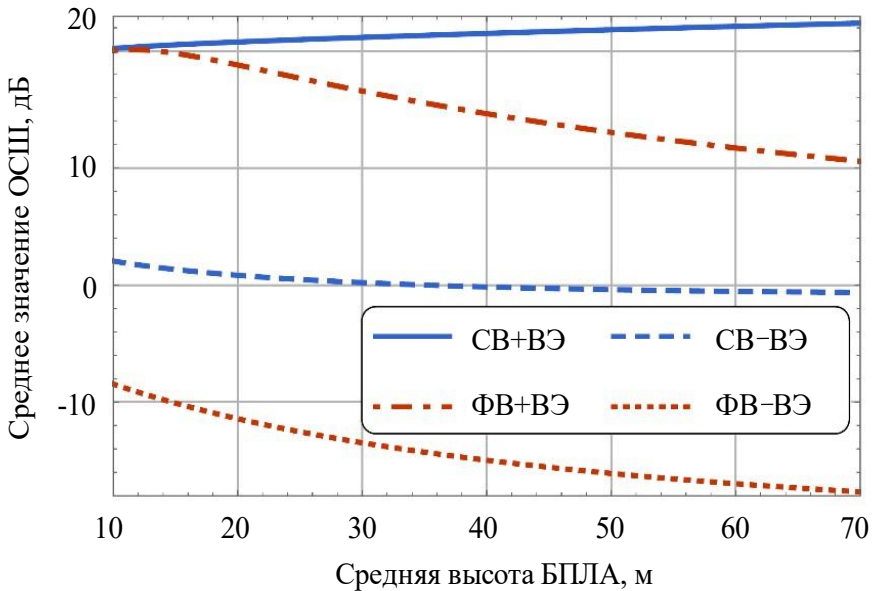


Рис. 7.5. Эффект высоты расположения БПЛА

Численные исследования позволяют сделать вывод о том, что даже при экстремально высокой плотности взаимодействующих пар БПЛА не наблюдается их существенного влияния друг на друга с точки зрения среднего значения ОСШ. Кроме того, и высота расположения БПЛА не оказывает негативного влияния на среднее значение ОСШ. На практике это означает, что при высокой направленности передающих и принимающих антенн можно размещать БПЛА практически сколь угодно плотно, наращивая емкость сети динамически.

7.4. Совместная оптимизация участков доступа и транзита

Перейдем к анализу совместной оптимизации доступа и транзита в технологии IAB с применением БПЛА в городских условиях с кластеризованными мобильными абонентами. Рассмотрим метод оптимизации участков доступа и транзита, оценим преимущества использования БПЛА при развертывании наземной сети доступа 5G/6G.

7.4.1. Системная модель

Рассмотрим квадратную область, покрытую наземными NR БС и БПЛА, предлагающими дополнительные возможности подключения АУ. Предполагаем наличие ретрансляционной топологии, где БС имеют два интерфейса для абонентского доступа и транзита с БПЛА. Схематично сценарий такой модели изображен на рис. 7.6. Аналогично БПЛА используют два интерфейса для абонентского доступа и транзита на БС. АУ принимает решение о подключении на основании значения ОСШ и могут подключаться либо к БПЛА, либо к БС. Высота расположения и скорость движения БПЛА фиксированы, и они могут изменять свое положение только в горизонтальной плоскости. Рассматриваем восходящий канал от АУ к БС или БПЛА, предполагая что на АУ генерирует эластичный трафик.

В рассматриваемом сценарии NR БС являются окончательными точками доставки трафика, поскольку трафик от АУ к БПЛА в итоге направляется к БС. Также предполагается наличие буферных накопителей конечной емкости на БПЛА и в АУ. Пакеты буферизируются на БПЛА, и, если буфер БПЛА заполнен, поступающие пакеты сбрасываются. БПЛА объединяют все принятые пакеты и отправляют их на БС, имеющую наилучшие условия канала в текущий момент времени. Предполагается использование каналов, разделенных по частоте для доступа и транзита, а также антенных решеток для каждого интерфейса. На практике это означает, что узел IAB оснащен двумя отдельными интерфейсами физического уровня, которые используют независимые планировщики канального уровня, тогда как маршрутизация, организация процесса буферизации и другие процедуры координируются общим процессом на IAB.

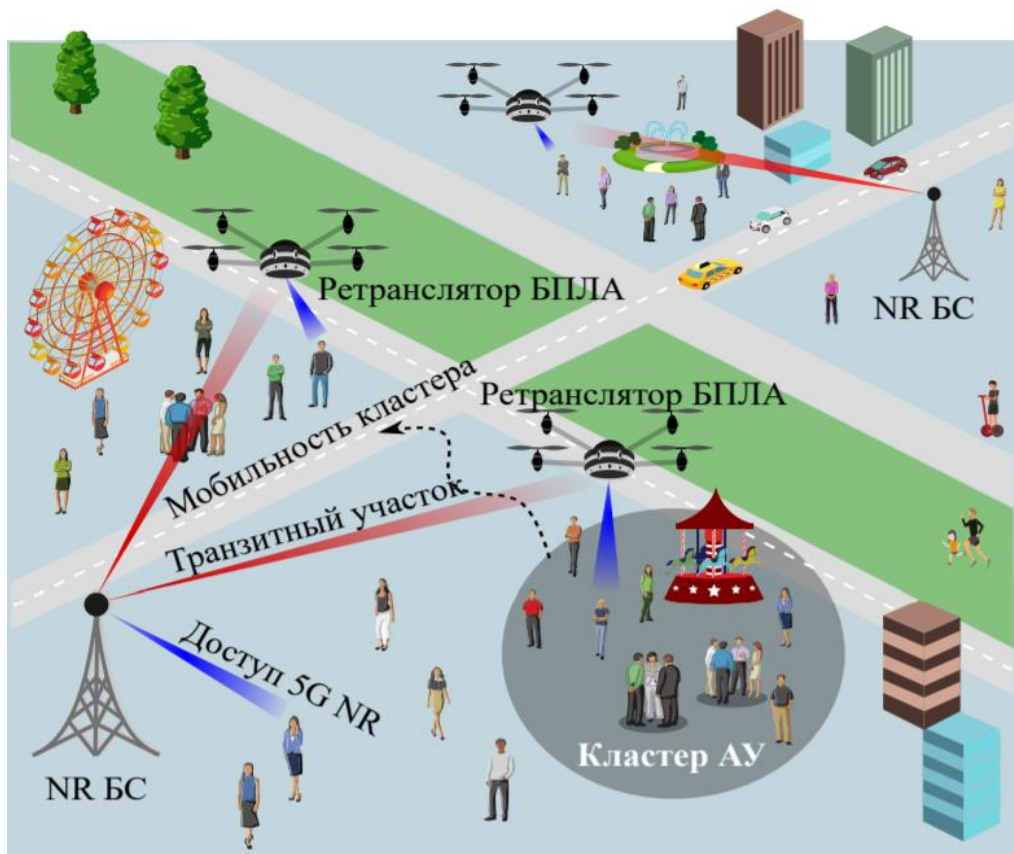


Рис. 7.6. Сценарий совместного доступа и транзита с БПЛА

В дальнейшем исследуем сценарий, схематично изображенный на рис. 7.7, где пространственная плотность пользователей изменяется во времени. Примером может служить масштабное массовое мероприятие, такое как фестиваль или концерт. В качестве модели движения пользователей используется групповая модель мобильности с лидерами (RPGM, Reference Point Group Mobility). В соответствии с этой моделью у каждой группы пользователей есть лидер, движение которого определяет направление перемещения всей группы. На интерфейсах АУ-БС и АУ-БПЛА используется α -справедливое разделение ресурсов как пример широкого класса методов разделения ресурсов, которые учитывают различные критерии справедливости (см. главу 11). Остальные параметры системы приведены в табл. 7.1.

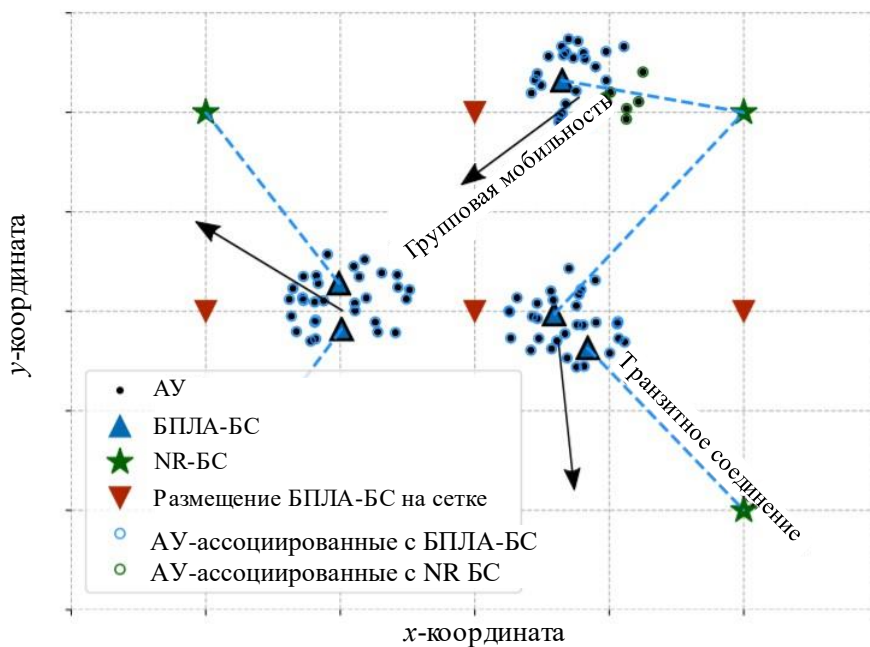


Рис. 7.7. Сценарий обслуживания массового мероприятия средствами БПЛА

Таблица 7.1

Обозначения для сценария совместного доступа и транзита

Параметр	Значение
Область	600×600 м
Число NR БС	4
Число БПЛА	2
Число АУ	100
Размещение АУ	Кластерное
Радиус расположения АУ	50 м
Антенна на NR БС	8×8 (доступ и транзит)
Антенна на БПЛА	8×8 (доступ), 4×4 (транзит)
Высота БПЛА	20 м
Высота АУ	1,5 м
Частота	28 ГГц
Полоса частот	0,56 ГГц
Мощность передатчика NR БС	24 дБм
Периодичность поиска луча	3 мс
Длительность кадра	3 мс
Размер пакета	1000 байт
Максимальная скорость	500 Мбит/с
Значение α	2

7.4.2. Метод совместной оптимизации доступа и транзита

Для оптимизации участков доступа и транзита в архитектуре IAB с применением БПЛА используем динамический алгоритм, основанный на методе оптимизации роя частиц (PSO, Particle Swarm Optimization). Метод PSO использует эвристический алгоритм с глобальным знанием, позволяющий осуществлять динамическую оптимизацию системы путем итеративного улучшения решения относительно заданной метрики. Решение ищется из множества возможных в допустимой области. Алгоритм решает оптимизационную задачу, рассматривая совокупность возможных решений и перемещая рассматриваемые элементы в пространстве поиска. Отметим, что алгоритм PSO требует глобального знания расположения БПЛА и АУ. Предполагается, что такое знание доступно в заданные моменты времени. В качестве другого решения рассмотрим статическое размещение БПЛА на регулярной сетке (см. рис. 7.7).

Численный анализ проводится методом имитационного моделирования. Среда моделирования включает возможность поддержки БПЛА в соответствии с требованиями 3GPP, определенными в стандарте TR 38.874. В качестве характеристик обслуживания рассматриваются средняя скорость АУ и справедливость разделения ресурсов между АУ. Для оценки справедливости разделения ресурсов используется индекс справедливости Джейна, который количественно оценивает «равенство» скоростей доступа, предоставляемых АУ. Если все АУ получают одинаковую скорость, индекс Джейна равен 1. Значение индекса Джейна варьируется от 1 до $1/N$, где N – число АУ.

7.4.3. Численный анализ

Проведем анализ характеристик обслуживания, предоставляемых АУ совместной реализацией доступа и транзита с учетом ограничений, накладываемых транзитными соединениями в рассматриваемом выше сценарии. Графики на рис. 7.8 иллюстрируют функцию распределения (ФР) с.в. скорости АУ для алгоритма динамической оптимизации PSO для трех случаев: а) транзит без ограничений на транзитном участке; б) транзит без знаний на ограничение скорости на транзитном участке; в) транзит с ограничением на скорость на

транзитном участке. Первые два случая наиболее часто рассматриваются в литературе, посвященной анализу систем IAB.

Анализируя результаты на рис. 7.8, отметим, что система с идеальным транзитом (без ограничений на транзитном участке) дает очень грубую оценку сверху для фактической скорости на транзитном участке, поскольку вероятность предоставления скорости менее чем 0,7 Гбит/с пренебрежимо мала. Это объясняется тем, что весь трафик, генерируемый АУ в зонах покрытия БПЛА, предполагается успешно доставленным на БС, что позволяет алгоритму оптимизации размещать БПЛА непосредственно над пользовательскими кластерами, максимизируя таким образом пропускную способность на участке АУ-БПЛА. Наоборот, вторая схема дает очень грубую оценку снизу потенциальной пропускной способности, поскольку не учитывает ограничение скорости на транзитном участке. В этом случае БПЛА располагаются, как и в первой схеме, но скорость на интерфейсе БПЛА недостаточна для передачи всего абонентского трафика. Здесь система работает в режиме с ограничением скорости на транзитном участке.

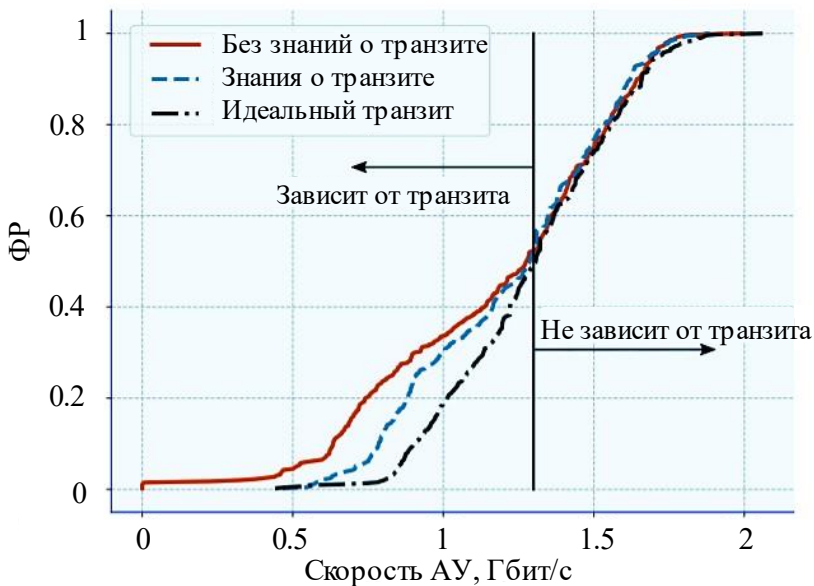


Рис. 7.8. ФР с.в. скорости АУ

Наконец, в случае ограничений как на доступ, так и транзит, значения пропускной способности лежат между нижними и верхними оценками рассмотренных выше двух случаев. Здесь позиции

расположения БПЛА больше не находятся над кластерами пользователей, поскольку требование к высокой скорости транзитного участка «сдвигает» их ближе к текущей БС. Другое интересное наблюдение, которое можно сделать, анализируя данные на рис. 7.8, заключается в том, что производительность всех рассмотренных схем отличается только для АУ, получающих низкую и среднюю скорость, так как, начиная примерно с 1,3 Гбит/с, графики совпадают. Это является следствием ограничений пропускной способности на транзитном участке. В том числе одинаковая во всех случаях пропускная способность БС приводит к тому, что для кластера пользователей, находящегося рядом с ней, ограничение на транзитном участке не оказывает существенного влияния.

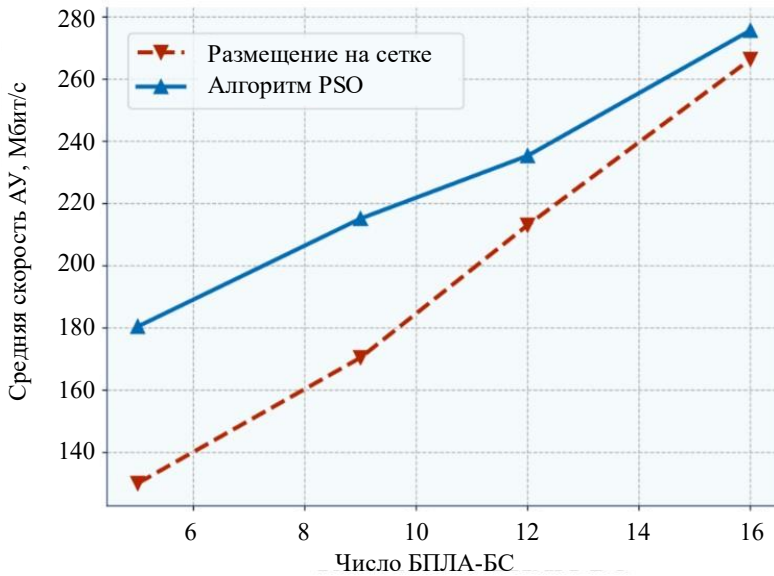


Рис. 7.9. Зависимость скорости АУ от числа БПЛА

Естественно ожидать, что позиции БПЛА и, следовательно, распределения скоростей АУ будут зависеть от числа кластеров пользователей и числа БПЛА. Изучим влияние этих параметров, используя среднюю скорость АУ в качестве исследуемой характеристики. На рис. 7.9 изменяется число БПЛА при фиксированном числе кластеров абонентов, равном 4. Этот график также показывает сравнение результатов применения алгоритма на основе динамического PSO и развертывания БПЛА на основе статической сетки (см. рис. 7.7). Как и ожидалось, для обоих типов развертывания средняя скорость АУ увеличивается с увеличением числа БПЛА. Наиболее значительное

различие между этими способами развертывания наблюдается для небольшого числа БПЛА. Объясняется это тем, что размещение на основе статической сетки не может полностью обслужить интересующую нас область, и поэтому иногда АУ либо оказываются далеко от БПЛА, либо находятся за пределами их зоны покрытия. По мере увеличения числа узлов эта разница уменьшается, поскольку развертывание на основе статической сетки теперь плотно охватывает всю зону обслуживания и все АУ всегда находятся в зоне действия по меньшей мере одного БПЛА. Другая причина заключается в том, что при большом числе БПЛА помехи начинают играть существенную роль, не позволяя алгоритму размещать БПЛА очень близко друг к другу. Однако этот недостаток может быть нивелирован за счет использования более высоконаправленных антенн на АУ, тем самым снижая негативное воздействие помех. Отметим, что даже для таких плотных развертываний использование методов динамической оптимизации позволяет увеличить скорость АУ путем оптимального позиционирования БПЛА рядом с кластерами пользователей.

Рассматриваемый алгоритм динамической оптимизации полезен для определения местоположения БПЛА при оптимизации пропускной способности АУ на основе критерия справедливости α . На рис. 7.10 показано влияние числа кластеров АУ на индекс справедливого разделения ресурсов Джейна при фиксированном количестве БПЛА для двух типов развертываний.

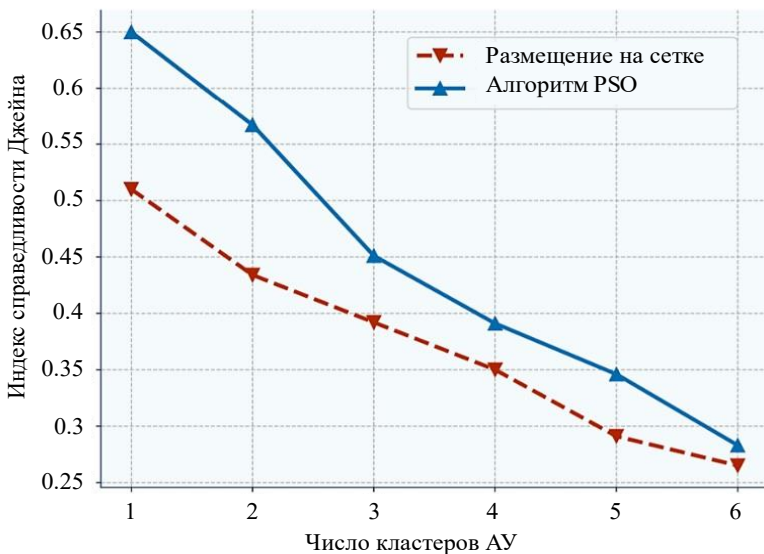


Рис. 7.10. Индекс справедливого разделения ресурсов Джейна в зависимости от количества кластеров

Следует подчеркнуть, что увеличение числа кластеров АУ приводит к снижению индекса справедливости. Основная причина в случае динамической оптимизации состоит в том, что для одного, двух или четырех кластеров существуют различные субоптимальные комбинации развертываний БПЛА, которые приводят к аналогичной производительности АУ. Например, для четырех БПЛА и двух кластеров АУ алгоритм PSO «заставляет» БПЛА отслеживать кластеры, находящиеся на противоположных сторонах (по два БПЛА на кластер). Следовательно, доступная скорость АУ для всех случаев практически одинакова, и единственное различие происходит вследствие небольших смещений АУ внутри кластеров. Когда число кластеров увеличивается, таких вариантов уже нет. Другими словами, предложенный алгоритм пытается справедливо разделить БПЛА между кластерами, но в случае, например, трех кластеров и пяти БПЛА, всегда будет существовать кластер, обслуживаемый только одним БПЛА и, следовательно, индекс справедливости ухудшается. Для размещения БПЛА на основе статической сетки можно наблюдать аналогичную тенденцию – справедливость разделения ресурсов ухудшается. Однако эта тенденция более умеренная по сравнению с алгоритмом PSO, поскольку при увеличении числа кластеров доля АУ вне покрытия уменьшается, что положительно влияет на индекс справедливости.

Наибольший выигрыш в справедливости достигается для небольшого числа кластеров АУ, а именно в случае одного или двух кластеров. Уже для трех кластеров значения справедливости для обеих схем размещения становятся близкими друг к другу. Даже несмотря на то, что решение на основе алгоритма PSO все еще демонстрирует лучшую производительность, статическая сетка размещения БПЛА составлена таким образом, чтобы максимально увеличить вероятность того, что все АУ будут в зонах покрытия по меньшей мере одним БПЛА или БС. Отметим, что два рассматриваемых алгоритма развертывания становятся все более похожими с точки зрения производительности сети, когда число кластеров больше, чем число БПЛА, т.е. шесть и пять кластеров в рассматриваемом случае. В этом случае алгоритм PSO будет стремиться разместить один из БПЛА между двумя кластерами, чтобы обеспечить покрытие, хотя это может быть и невозможно вследствие большого расстояния между кластерами. Соответственно, пользователи изолированного от обслуживания БПЛА кластера должны полагаться исключительно на покрытие БС.

7.5. Обслуживание на абонентском участке

На сегодняшний день одной из основных проблем, связанных с использованием БПЛА в качестве БС, является недостаточный запас времени полета. Следовательно, разработка эффективных стратегий управления БПЛА в процессе обслуживания абонентов является актуальной задачей. В данном параграфе исследуется метод стационарного размещения БПЛА на крышах зданий, балконах или стенах зданий при помощи специальных приспособлений, как только они прибывают в зону обслуживания абонентов. Рассматриваемая стратегия может быть предпочтительной с точки зрения увеличения времени автономной работы, поскольку основной источник энергопотребления – двигатели – можно отключать на длительные промежутки времени. Дополнительно проводится сравнение предлагаемого метода со стандартным методом, когда БПЛА находятся в режиме полета в процессе обслуживания.

7.5.1. Системная модель

Рассматривается модель системы, включающая круговую зону обслуживания (например, городскую площадь) радиуса R и станцию обслуживания и зарядки БПЛА, расположенную на расстоянии l (рис. 7.11). Абоненты предполагаются случайно распределенными в зоне обслуживания с плотностью λ пользователей на квадратный метр. АУ, ассоциированные с пользователями, расположены на высоте h_U и подключаются к ближайшему доступному БПЛА. Человеческие тела моделируются как цилиндры высотой h_B , $h_B > h_U$, и радиусом основания r_B . Абоненты одновременно являются и блокаторами LoS. Используется модель распространения для миллиметрового диапазона частот UMi, а также модель динамической блокировки со статичными пользователями и мобильными блокаторами (см. главу 2).

Рассматриваются два режима работы БПЛА – режим *полета* и *стационарный* режим. В режиме полета БПЛА находятся в воздухе в процессе обслуживания, тогда как в стационарном режиме предполагается, что БПЛА могут приземляться на близлежащие здания. В обоих режимах оператор для обслуживания абонентов имеет в общей сложности N БПЛА.

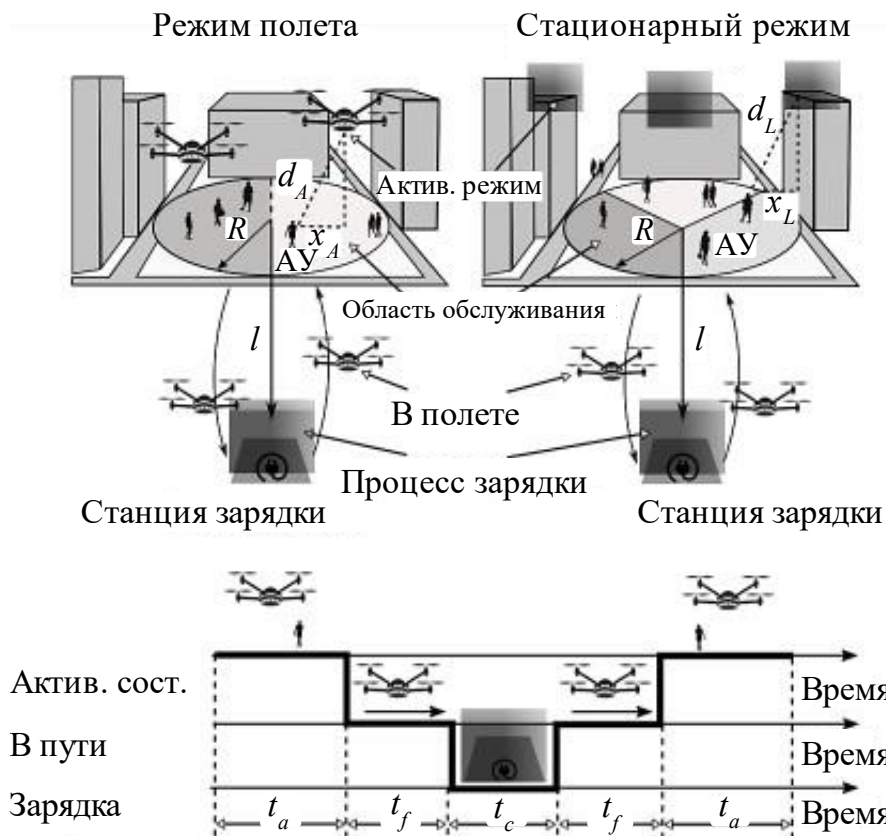


Рис. 7.11. Принцип обслуживания абонентов

Каждый БПЛА подчиняется циклу работы, состоящему из трех этапов:

1) *активная стадия*, когда БПЛА обеспечивает беспроводной доступ для абонентов в режиме полета или из стационарного положения, в зависимости от режима работы;

2) *этап зарядки*, когда БПЛА заряжает батарею на зарядной станции;

3) *этап полета*, когда БПЛА движется от зарядной станции в зону обслуживания и обратно.

В обоих режимах БПЛА покидает зону обслуживания и летит на зарядную станцию всякий раз, когда оставшегося количества энергии достаточно только для этого полета. Предполагается, что зарядная станция имеет достаточно ресурсов для одновременной зарядки всех N БПЛА. В результате такого цикла работы только определенное количество БПЛА одновременно находится в активной стадии в обоих режимах работы, $N_A < N$ и $N_L < N$, для

режима полета и стационарного режима соответственно. Предполагается, что все БПЛА, находящиеся в активной стадии, размещены на одной высоте h_A для режима полета и h_L для стационарного режима.

7.5.2. Анализ модели

Определение оптимального расположения БПЛА является сложной задачей, которая зависит от геометрии зоны обслуживания и области покрытия одного БПЛА. Напоминая, что величина ОСШ является монотонно убывающей функцией расстояния до АУ, для каждого из N БПЛА задача может быть сведена к минимизации суммы расстояний до всех потенциальных позиций АУ. Таким образом, для круговой зоны обслуживания в модели в качестве простого приближения можно использовать плотную упаковку окружностей с центрами, соответствующим положению БПЛА, в круге, соответствующем полной зоне обслуживания. Хотя математически эта проблема еще не полностью решена, оптимальное решение известно для практически значимого числа БПЛА $N_A < 20$.

Следуя аналогичным соображениям минимизации расстояния между АУ и ближайшего БПЛА, предполагается, что N_L БПЛА приземляются в круг зоны обслуживания с радиусом R на равных расстояниях друг от друга. В этом случае зоной обслуживания каждого БПЛА является сектор с углом величиной $2\pi/N_L$ (см. рис. 7.11).

Перейдем к анализу системы, сосредоточив внимание на средней скорости передачи на участке доступа от БПЛА к АУ и необходимом числе БПЛА для ее достижения. Анализ обоих рассматриваемых режимов включает три этапа. На первом этапе, предполагая постоянное число N_a БПЛА в активной стадии для обоих режимов работы, найдем функцию плотности вероятности проекции расстояния между БПЛА и АУ. Затем, на втором этапе, используя метод преобразования с.в., рассмотренный в главе 3, найдем среднее значение скорости передачи, предоставляемой одному пользователю. Наконец, на третьем этапе, оценим число активных БПЛА для обоих рассматриваемых режимов. Задачи, решаемые на первых двух этапах, большой сложности не представляют, и поэтому рассмотрим метод нахождения числа БПЛА, необходимых для поддержки заданной скорости передачи на участке доступа от БПЛА к АУ.

Задача нахождения функции плотности расстояния до БПЛА

Необходимо определить проекции расстояний от случайно выбранного АУ до БПЛА в зоне обслуживания для двух рассматриваемых режимов. Далее найдем проекции расстояний соединений сначала для режима полета, а затем для стационарного режима БПЛА.

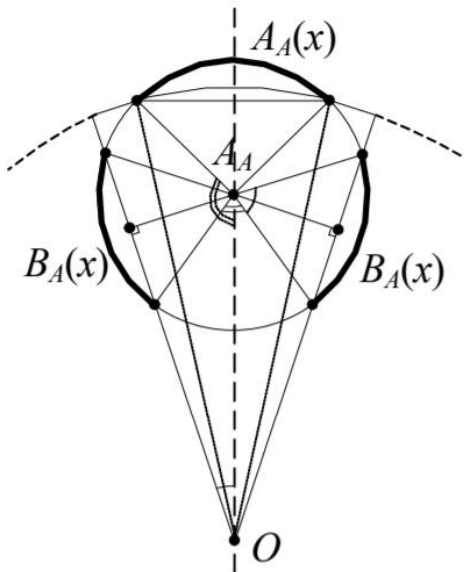


Рис. 7.12. Иллюстрация дуг $A_A(x)$, $B_A(x)$ во второй части формулы плотности вероятности $f_{x_A}(x)$

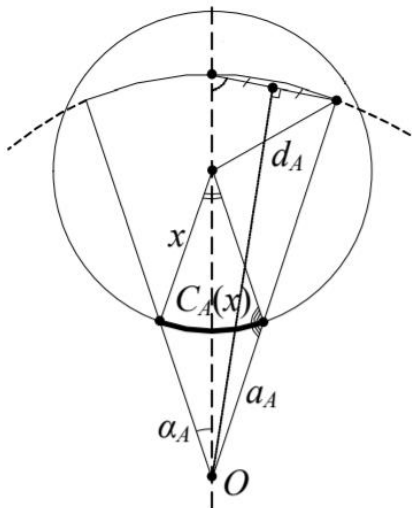


Рис. 7.13. Иллюстрация дуги $C_A(x)$ в третьей части формулы плотности вероятности $f_{x_A}(x)$

Анализ режима полета. Отметим, что плотности вероятности проекций расстояний между АУ и БПЛА для любого количества БПЛА можно получить, используя единый метод. Ниже, в качестве примера, рассмотрим один из наименее симметричных случаев $N_a = 5$.

Известно, что радиус круга с нарисованными в нем пятью окружностями (случай $N_a = 5$) может быть определен по формуле

$$r_A = R / (1 + \sqrt{2 + 2/\sqrt{5}}). \quad (7.1)$$

Плотность вероятности проекции на плоскость с.в. X_A расстояния между случайно выбранным АУ и ближайшим БПЛА можно записать в виде

$$f_{X_A}(x) = \begin{cases} 2\pi x/S, & 0 < x < r_A, \\ \frac{2\pi x - A_A(x) - 2B_A(x)}{S}, & r_A < x < d_A, \\ C_A(x)/S, & d_A < x < R - r_A, \end{cases} \quad (7.2)$$

где $A_A(x)$, $B_A(x)$, $C_A(x)$ обозначают длину соответствующих дуг, а площадь обслуживания одного БПЛА

$$S = \pi R^2 / N_a. \quad (7.3)$$

Эти величины, входящие во вторую и третью части формулы (7.2), проиллюстрированы на рис. 7.12 и 7.13. Данные рисунки соответствуют проекции на плоскость трехмерных компонент рассматриваемого сценария.

Анализируя геометрические построения для рассматриваемого сценария, можно получить следующие выражения для длины дуг:

$$\begin{aligned} A_A(x) &= 2x \left[2\pi - \cos^{-1} \left(\frac{x^2 + (R - r_A)^2 - R^2}{2x(R - r_A)} \right) \right], \\ C_A(x) &= 2x \left(\pi - \cos^{-1} \left[\frac{x^2 + a_A^2 - (R - r_A)^2}{2xa_A} \right] - \frac{\pi}{5} \right), \\ B_A(x) &= 2x \cos^{-1} \left(\frac{x^2 + a_A^2 - (R - r_A)^2}{2xa_A} \right). \end{aligned} \quad (7.4)$$

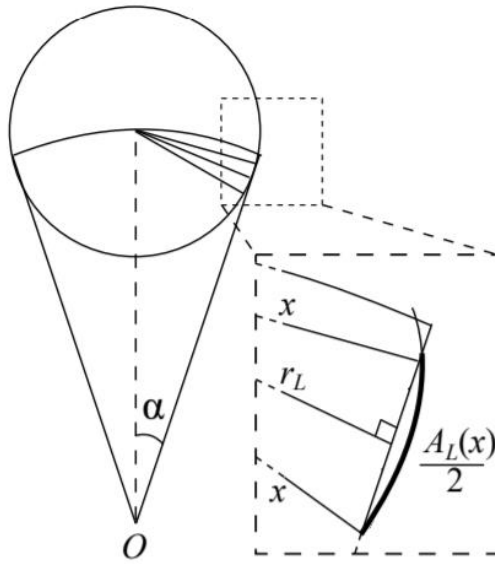


Рис. 7.14. Иллюстрация дуги $A_L(x)$ во второй части формулы плотности вероятности $f_{x_L}(x)$

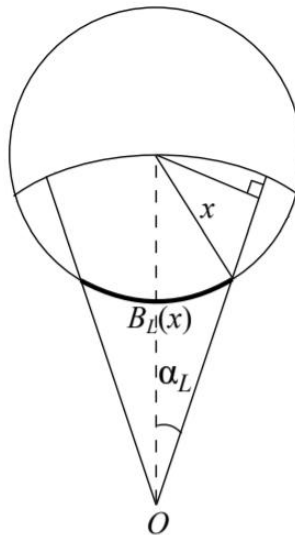


Рис. 7.15. Иллюстрация дуги $B_L(x)$ в третьей части формулы плотности вероятности $f_{x_L}(x)$

Анализ стационарного режима. Плотности вероятности проекций расстояний между АУ и БПЛА для стационарного режима получаются аналогично режиму полета. Поскольку предполагается, что БПЛА расположены на периметре обслуживаемой области и они находятся на одинаковом расстоянии друг от друга, искомая величина является с.в. X_L расстояния от точки, случайно распределенной

в секторе окружности с радиусом R и углом $2\pi/N_a$ к центру соответствующей дуги. Для $N_a = 5$ получаем следующее выражения для плотности вероятности:

$$f_{X_L}(x) = \begin{cases} [2x\cos^{-1}(x/2R)]/S, & 0 \leq x \leq r_L, \\ \frac{2x\cos^{-1}(x/2R)}{S} - \frac{A_L(x)}{S}, & r_L < x < d_L, \\ B_L(x)/S, & d_L \leq x \leq R, \end{cases} \quad (7.5)$$

где параметры, входящие во вторую и третью части функции (см. рис. 7.14 и 7.15), получаются следующим образом:

$$\begin{aligned} A_L(x) &= 4x\cos^{-1}([R\sin\alpha_L]/x), \\ B_L(x) &= x(\pi - 2\alpha_L - 2\cos^{-1}([R\sin\alpha_L]/x)), \\ r_L &= R \sin\alpha_L, \\ d_L &= \sqrt{2R^2 - 2R^2\cos\alpha_L}. \end{aligned} \quad (7.6)$$

Плотности вероятности с.в. расстояний между АУ и БПЛА в трехмерном пространстве находятся с помощью метода преобразования случайных величин. Напомним, что плотность вероятности $f(y)$ с.в. Y , являющейся функцией $y = \phi(x)$ с.в. X , имеет вид (см. главу 3)

$$f(y) = \sum_{vi} f(\psi(y))|\psi'(y)|, \quad (7.7)$$

где $\psi(y) = \phi^{-1}(y)$ – обратная функция.

Для любого количества БПЛА N_a имеем $y = \phi(x) = \sqrt{x^2 + h_T^2}$, где $h_T = |h_A - h_U|$ и x – длина проекции. Тогда обратная функция и ее производная принимают вид

$$\psi(y) = -\sqrt{y^2 - h_T^2}, \psi'(y) = -y/\sqrt{y^2 - h_T^2}. \quad (7.8)$$

Процесс перезарядки БПЛА

Стационарный режим. В этом случае соотношение времени, в течение которого БПЛА находится в активной фазе, ρ_L , можно записать как $t_{a,L}/(t_{a,L} + 2t_f + t_c)$, где $t_{a,L}$ – время в активной фазе, t_c – время зарядки, t_f – время, необходимое для полета от зарядной станции к зоне обслуживания. Емкость батареи БПЛА теперь можно записать как

$$t_{a,L}P_{AP} + 2t_fP_E = C_bV, \quad (7.9)$$

где P_{AP} – требуемая мощность БПЛА.

Обратим внимание, что $P_{AP} \gg P_A$, так как P_{AP} включает в себя потребляемую мощность приемопередатчика, обработку сигналов и обработку данных на БПЛА. P_E является суммарной потребляемой мощностью двигателей БПЛА, C_b – емкость батареи БПЛА, измеряемая в $A \cdot ч$, а V – напряжение батареи.

Запишем $t_f = l/v$, где l – расстояние между зарядной станцией и зоной обслуживания, v – скорость движения БПЛА на этапе полета. Учитывая $t_{a,L}$, из (7.9) приходим к следующей формуле:

$$\rho_L = \frac{TP_E v - 2P_E l}{TP_E v + 2l(P_{AP} - P_E) + t_c v P_{AP}}, \quad (7.10)$$

где T – время полета БПЛА на батареях.

Режим полета. Основное отличие этого режима от стационарного режима состоит в том, что БПЛА потребляет $P_{AP} + P_E$ вместо P_E в активной фазе. Таким образом, емкость батареи для БПЛА можно записать как

$$t_{a,A}(P_{AP} + P_E) + 2t_f P_E = C_b V. \quad (7.11)$$

Доля времени, в течение которого БПЛА остается в активной стадии для полета, ρ_A , может быть оценена следующим образом:

$$\rho_A = \frac{TP_E v - 2P_E l}{TP_E v + 2lP_{AP} + t_c v (P_E + P_{AP})}. \quad (7.12)$$

Найдя доли времени в активной фазе для рассматриваемых режимов, ρ_L и ρ_A , используем их, чтобы получить количество БПЛА, которые одновременно находятся в активной фазе. В общем текущее количество активных БПЛА является случайной величиной, которая зависит от расписания полетов. Применяя результаты теории расписаний, устанавливаем $N_L = N\rho_L$ и $N_A = N\rho_A$, где N_L и N_A представляют собой минимальные количества БПЛА, гарантированно находящиеся в зоне обслуживания одновременно. Эти значения будут использованы далее для определения показателей эффективности обслуживания, в том числе для определения средней скорости предоставляемой АУ при заданном значении количества БПЛА в обслуживающей группе.

Заметим, что при определенном значении расстояния от БПЛА до АУ $P(x)$ – это принимаемая мощность на АУ. $P(x)$ представляет собой смесь двух значений: $A_B x^{-\gamma}$ и $A_N x^{-\gamma}$, соответствующих заблокированному и незаблокированному состояниям, взвешенных с вероятностями $p_B(x)$ и $1 - p_B(x)$:

$$\begin{aligned}
 P(x) &= P_A G_A G_U (p_B(x) A_B x^{-\gamma} + [1 - p_B(x)] A_N x^{-\gamma}) = \\
 &= P_A G_A G_U (A_N + p_B(x) [A_B - A_N]) x^{-\gamma},
 \end{aligned}
 \tag{7.13}$$

где $p_B(x)$ – вероятность блокировки на расстоянии x .

Отметим, что $P(x)$ не имеет обратной функции в аналитической форме, что препятствует нахождению ФР ОСШ, спектральной эффективности и, далее, скорости передачи данных, используя технику преобразования случайных величин. Поэтому, чтобы определить среднюю скорость передачи данных, обеспечиваемую БПЛА в присутствии $[N_a]$ БПЛА в активной фазе, используем метод разложения в ряд Тейлора функции емкости $C_{1,N_a}(x) = B \log_2(1 + P(x)/BN_0)$, где N_0 – тепловой шум (см. главы 3 и 4). Используя этот подход, устанавливаем следующую аппроксимацию для средней скорости, предоставляемой БПЛА:

$$E[C_{1,N_a}] = C_{1,N_a}(\mu_D) + \frac{C_{1,N_a}''(\mu_D)}{2} \sigma_D^2,
 \tag{7.14}$$

где μ_D , σ_D^2 – среднее значение и дисперсия расстояния между БПЛА и АУ; $C_{1,N_a}''(x)$ – вторая производная функции скорости. Средняя спектральная эффективность получается путем подстановки $B = 1$ Гц. Наконец, средняя скорость передачи данных в рассматриваемой области может быть записана следующим образом:

$$E[C] = N_a E[C_{1,N_a}].
 \tag{7.15}$$

7.5.3. Численный анализ

Проведем численный анализ рассмотренных методов обслуживания абонентов и сравним результаты двух режимов применения БПЛА – режим полета (Р1) и стационарный режим (Р2). Значения параметров системной модели представлены в табл. 7.2.

Таблица 7.2

Параметры обслуживания на абонентском участке

Параметр	Значение
Частота	28 ГГц
Полоса пропускания	1 ГГц
Излучаемая мощность	23 дБм
Усиление антенны БПЛА	10 дБ
Усиление антенны АУ	0 дБ
Высота АУ	1,5 м

Параметр	Значение
Тепловой шум для 1 ГГц	-84 дБм
Скорость полета	40 км/ч
Емкость батареи	40 А·ч
Напряжение батареи	12 В
Потребляемая мощность двигателей	1 кВт
Потребляемая мощность БС	100 Вт
Радиус обслуживаемой области	50 м
Время зарядки БПЛА	1 ч
Число БПЛА	{1, 2, ..., 7}
Дистанция до зарядной станции	[0; 20) км

Ниже исследуются эффекты влияния основных системных параметров на характеристики системы, а также сравниваются два режима Р1 (полет) и Р2 (статика).

Рассмотрим участок доступа. Диаграммы на рис. 7.16 иллюстрируют среднюю спектральную эффективность как функцию числа N_a БПЛА, находящихся в активной фазе. Из данных на рис. 7.16 видно, что режим полета Р1 дает на 30% большую производительность, чем стационарный режим Р2 для всех возможных значений числа активных БПЛА и всех значений плотности пользователей в зоне обслуживания. Это объясняется лучшим геометрическим положением БПЛА, обеспечиваемым режимом Р1, что приводит к меньшей вероятности блокировки LoS. Противоположная тенденция наблюдается на рис. 7.17, где показана доля времени ρ_A и ρ_L как функция расстояния l между зоной обслуживания и станцией зарядки, в течение которого БПЛА может находиться в активной фазе и обслуживать пользователей. Заметим, что БПЛА в стационарном режиме характеризуются значительно меньшим энергопотреблением во время активной фазы. Поэтому, как видно из графика, доля времени ρ_L в активной фазе для стационарного режима Р2 как минимум на 30 % больше соответствующего значения доли времени ρ_A для режима полета Р2 с такими же параметрами системы.

Таким образом, можно сделать вывод о том, что наблюдаемые при численном анализе эффекты зависят от расстояния между зоной обслуживания и зарядной станцией.

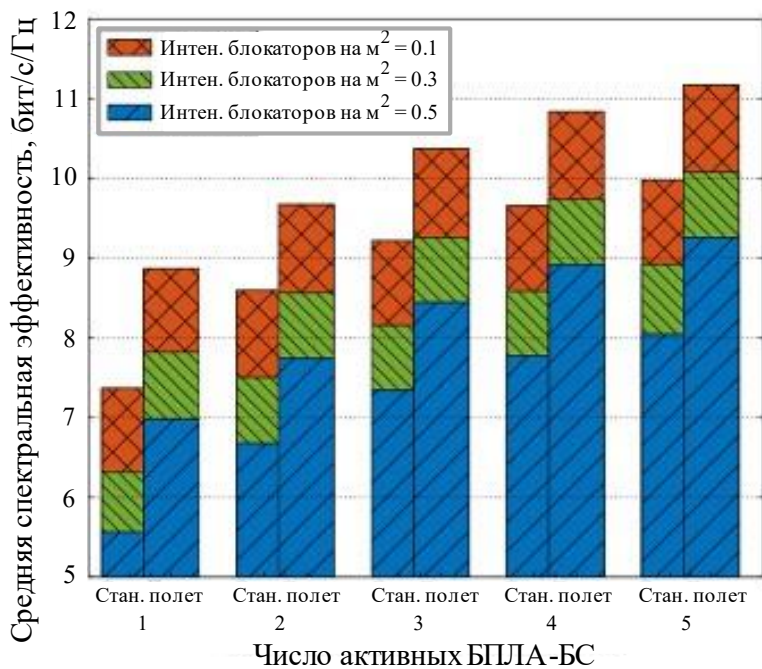


Рис. 7.16. Эффект числа обслуживающих БПЛА

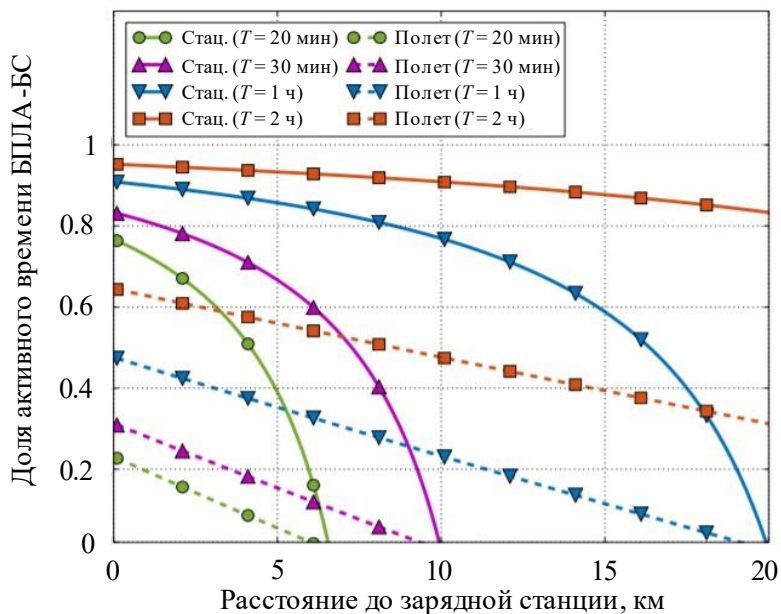


Рис. 7.17. Эффект расстояния от БПЛА до зарядной станции

Перейдем к рис. 7.18, иллюстрирующему зависимости между режимом полета и стационарным режимом. Здесь на оси ординат представлены значения, начиная с которых режим полета будет предпочтительным. В качестве исследуемой характеристики рассмотрим среднее значение пропускной способности сети. Важно отметить, что, хотя средняя пропускная способность сети для обоих режимов полета зависит от многих параметров (высота полета БПЛА, плотность пользователей в зоне обслуживания и т.д.), ответ на вопрос «Какой режим приводит к увеличению пропускной способности сети?» определяется в основном расстоянием l до зарядной станции, временем полета T на батареях и числом N БПЛА, доступных для применения оператором при предоставлении услуг связи. Эти три параметра и используются на графиках рис. 7.18, что позволяет сформулировать следующие выводы.

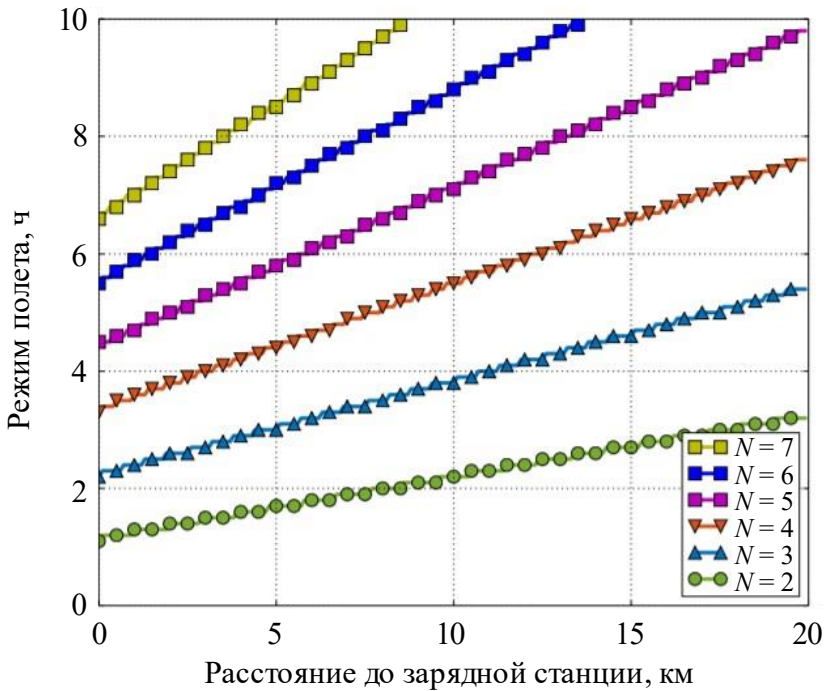


Рис. 7.18. Сравнение режима полета P1 и стационарного режима P2

Если доступны четыре БПЛА ($N = 4$), все системы, соответствующие кривым, лежащим выше красной кривой, предпочтительнее обслуживать в режиме полета P1. Все другие системы, соответствующие кривым, лежащим ниже красной кривой, обладают в среднем

большей пропускной способностью сети при применении БПЛА в стационарном режиме Р2. Заметим, что чем больше БПЛА доступно оператору и чем более жесткими являются условия их применения (емкость батареи и расстояние до станции зарядки), тем более привлекательным для оператора является режим Р2. Напротив, чем меньшее число доступных БПЛА, тем больше область значений системных параметров, в которых режим полета Р1 является предпочтительным.

7.6. Использование многошаговой ретрансляции

Ранее представленные решения характеризуются использованием одного транзитного узла на БПЛА. Однако в ряде документов 3GPP говорится о том, что использование нескольких транзитных узлов на БПЛА может значительно расширить зону покрытия и, как следствие, улучшить характеристики обслуживания пользователей.

Применение многошаговой передачи в общем случае требует использования протоколов маршрутизации, что может привести к значительным накладным расходам на передачу сигнальной информации для поддержки актуальных путей маршрутизации в динамически меняющейся топологии сети. В этом параграфе рассматривается метод использования так называемых БПЛА-мостов, которые не требуют протокола маршрутизации для выгрузки трафика с БС сетей 5G/6G, которые находятся в условиях перегрузки.

7.6.1. Концепция БПЛА-мостов

Помимо телекоммуникационного сектора предполагается, что БПЛА будут активно использоваться и в других областях, таких как транспорт, сельское хозяйство и медицина. Например, коммерческое использование БПЛА в отрасли доставки продукции может привести к повышению эффективности доставки, снижению затрат и повышению качества обслуживания клиентов. В том числе БПЛА эффективно решают дорогостоящую проблему доставки на «последней миле». Применение БПЛА может обеспечить возможность доставки по требованию в тот же день, а также позволяет избежать проблем традиционной логистики – пробок на дорогах. Решением задач доставки продукции при помощи БПЛА активно занимаются несколько индустриальных гигантов, таких как, Amazon, Walmart, Google и т.д.

Помимо основной задачи, БПЛА, осуществляющие доставку, могут выполнять и дополнительные функции. Например, Amazon запатентовал использование БПЛА, занимающихся доставкой продукции, для камер видеонаблюдения для жилых зданий. В процессе доставки, если у БПЛА есть дополнительные ресурсы, он сможет провести плановый осмотр частной собственности. Ожидается, что число таких приложений будет расти и в дальнейшем.

В данном разделе рассмотрим схему выгрузки трафика путем использования БПЛА сторонних операторов. В таком сценарии сетевые операторы заключают соглашения об уровне обслуживания (SLA, Service Level Agreement) с операторами БПЛА, определяя параметры и конфигурацию предлагаемых ими услуг, а также схему ценообразования. Вместо того чтобы полагаться на многошаговые сети с динамической маршрутизацией, предлагается использовать БПЛА-мосты, где БПЛА, регулярно следуя по одному и тому же маршруту в некоторой области, работают в многошаговом режиме ретрансляции, устанавливая топологию сети в виде цепочки, соединяющей перегруженные БС и те, которые находятся в нормальных условиях функционирования.

7.6.2. Системная модель

Обозначения рассматриваемой системной модели представлены в табл. 7.3. Рассмотрим сценарий стандартного района, как показано на рис. 7.19, состоящего из N улиц и некоторого числа зданий разной высоты. Ширина d улиц и длина d_B квартала считаются постоянными. Высота здания H_B является с.в. с плотностью распределения $f_{H_B}(x)$.

Предполагается, что в рассматриваемой области трафик пользователей обслуживается наземной инфраструктурой сети 5G/6G. Некоторые БС могут испытывать перегрузки, когда трафик пользователей превышает доступную на них скорость передачи. Также предполагается, что местоположение БС, находящейся в состоянии перегрузки, случайно и равномерно распределено на узлах сетки размера N на N , схематично изображенной на рис. 7.20. Аналогично местоположение БС, на которую может быть выгружен избыточный трафик, также случайно и равномерно распределено на той же сетке. Иными словами, все узлы перенумерованы и узел с номером i выбирается с вероятностью $1/N^2$. Трафик, который необходимо выгрузить,

передается с постоянной скоростью R Мбит/с. Для выгрузки трафика применяются БПЛА стороннего оператора, которые, пролетая над рассматриваемой областью, выполняют свою основную функцию, например доставку.

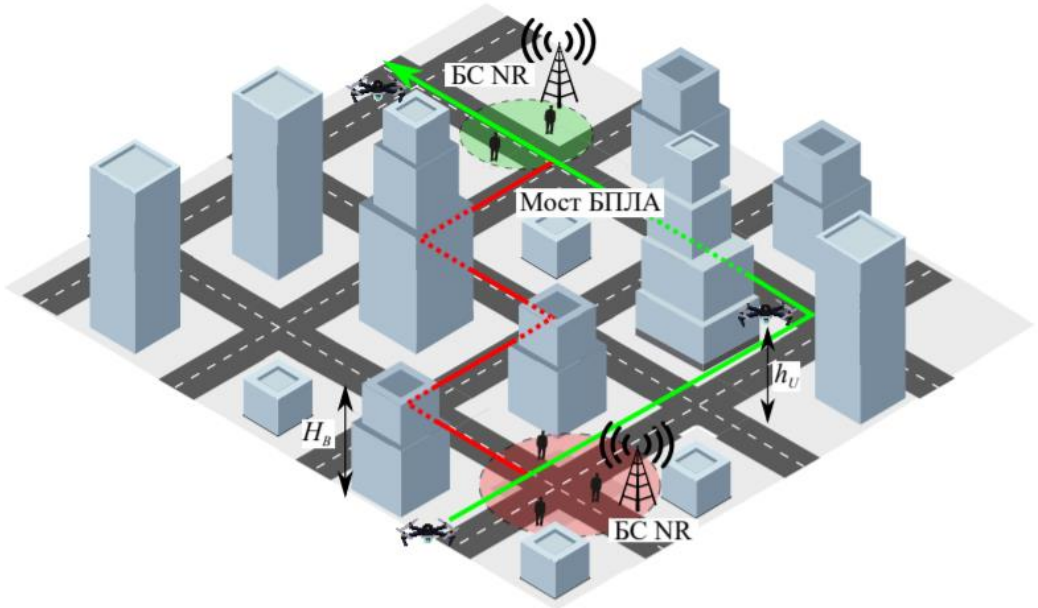


Рис. 7.19. Иллюстрация системной модели БПЛА-моста

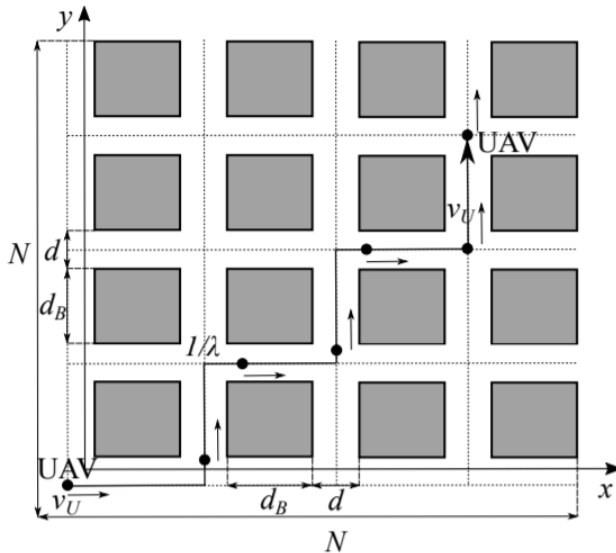


Рис. 7.20. Модель топологии городского сценария применения БПЛА-моста

Обозначения системной модели

Параметр	Определение
N	Число улиц
d	Ширина улицы
d_B	Ширина квартала
H_B	Высота здания
R	Скорость передачи
λ	Плотность БПЛА (на м)
λ_A	Плотность БПЛА в цепочке (на м)
h_U	Высота полета БПЛА
v_U	Скорость полета БПЛА
P_T	Излучаемая мощность
G_T	Усиление передающей антенны
G_R	Усиление принимающей антенны
N_T	Число антенных элементов на передаче
N_R	Число антенных элементов на приеме
L	Потери распространения сигнала
S_T	Чувствительность приемника
N_0	Тепловой шум
B	Полоса частот
f_c	Несущая частота
L_0	Максимальная дистанция связности
D_C	Дистанция между БПЛА в связанной цепочке
D_0	Дистанция между БПЛА
D_C^*	Распределение расстояния между БПЛА
C	Интервал связности
D	Расстояние между NR БС
R_{max}	Максимальная скорость передачи БПЛА-моста
ζ	Доля времени, когда БПЛА-мост функционирует
ϵ	Доля отведенной нагрузки
ν	Вероятность наличия LoS
T_0	Время непрерывной выгрузки трафика
T_1	Время непрерывного отсутствия связи
$p(i, k)$	Количество поворотов на маршруте между точками i и k
p_i	Вероятность, что путь между точкой перегрузки и точкой выгрузки состоит из i кварталов
W	Объем буферного накопителя

Далее предполагаем, что БПЛА влетают в рассматриваемую зону в углу решетки (например, в левый верхний) и перемещаются в противоположный угол по диагонали. Процесс поступления БПЛА моделируется Пуассоновским потоком с плотностью λ . Высота h_U полета и скорость v_U перемещения БПЛА фиксированы. Основным показателем качества обслуживания для оператора БПЛА является

время пролета через район развертывания услуги связи, представленный на рис. 7.20. Предполагаем, что полет возможен только вдоль улиц. Таким образом, на пути между точками входа и выхода разрешены только те повороты, которые минимизируют время полета через район. В системной модели используются модели распространения и антенной решетки из главы 2. Предполагается, что при наличии здания между соседними БПЛА соединение может обрываться вследствие блокировки LoS.

Для рассматриваемой системной модели предполагается, что трафик между БС выгружается с использованием БПЛА, управляемых оператором, в соответствии с правилами, SLA. Отметим, что предлагаемые стратегии обслуживания не должны оказывать негативного влияния на задержку пролета БПЛА через район развертывания услуги связи. С учетом этого условия рассматриваем две стратегии выгрузки трафика на БПЛА-мост:

– *C1: Мост с маршрутизацией.* В этом случае единственное требование заключается в следовании части БПЛА в соответствии с заданной схемой маршрутов на пути между пунктами отправления и назначения. Это схема, которая не требует каких-либо дополнительных условий SLA;

– *C2: Управляемый мост.* В данном случае, согласно SLA сетевому оператору разрешается не только определять маршрут следования БПЛА, но и задерживать их в точке входа в район развертывания услуги связи. Эта стратегия может повлиять на целевые характеристики оператора БПЛА, но производительность БПЛА-моста может значительно улучшиться.

Заметим, что в системной модели могут быть введены дополнительные возможности, позволяющие в целом улучшить показатели качества обслуживания. В том числе можно ввести управление, позволяющее сетевому оператору минимизировать число поворотов. Это может снизить вероятность обрыва соединения в БПЛА-мосте. Кроме того, временный обрыв соединения в БПЛА-мосте может быть компенсирован дополнительным объемом буфера для хранения передаваемых данных.

Для рассматриваемой системной модели и двух стратегий выгрузки трафика основным показателем эффективности является доля нагрузки, которая может быть выгружена на БПЛА-мост и тем самым снизить уровень перегрузки БС.

Пусть $U(t)$ – мгновенная скорость передачи, которую БПЛА-мост может поддерживать в момент t . Введем показатель, который будем называть коэффициентом выгрузки:

$$\gamma = \min\{1, \lim_{t \rightarrow \infty} (U(t)/R)\}, \quad (7.16)$$

где R – требуемая скорость передачи.

7.6.3. Анализ модели

Рассмотрим сначала мост с маршрутизацией. Для анализа этой стратегии сначала необходимо охарактеризовать свойства связности моста БПЛА, получив долю времени, когда существует многошаговая ретрансляционная цепочка между исходной и целевой БС, ζ . Для этого необходимо получить ФР с.в., определяющих длину интервала цепочки БПЛА, которые связаны друг с другом, и расстояния между исходной и целевой БС. Затем необходимо определить интервал выгрузки как разницу между этими двумя с.в. Максимальная скорость передачи моста БПЛА R_{max} – оценивается как функция максимального расстояния между БПЛА, образующими мост. Используя требуемую скорость R , скорость выгрузки определяется как $\epsilon = \min\{1, R_{max}/R\}$. Таким образом, учитывая эффект блокировки LoS, необходимо получить вероятность $1 - \nu$ того, что маршрут между БС не существует из-за блокировки на перекрестках. Наконец, коэффициент выгрузки может быть определен как $\gamma = \zeta\epsilon\nu$.

Рассмотрим теперь управляемый мост. Принципиальное отличие от моста с маршрутизацией заключается в том, что ζ является ступенчатой функцией, которая обращается в ноль для всех плотностей БПЛА λ , меньших определенного значения. Используя ζ , ν и ϵ , коэффициент выгрузки определяется аналогично предыдущему случаю, $\gamma = \zeta\epsilon\nu$.

Важным этапом является определение объема буферного пространства, необходимого для нивелирования эффекта потери связи из-за блокировки LoS на перекрестках и определения количества перекрестков на пути следования БПЛА между БС.

Доля времени выгрузки ζ

Максимальное расстояние между соседними БПЛА, при котором между ними существует связь, определяется как

$$L_O = 10^{\frac{P_T + G_T + G_R - N_0 - S_T - 32.4 - 20 \log_{10} f_c}{21}}, \quad (7.17)$$

где P_T – излучаемая мощность; G_T и G_R – коэффициенты усиления антенных решеток на передаче и приеме, соответственно; N_0 – тепловой шум; S_T – чувствительность приемника; f_c – несущая частота, измеряемая в ГГц.

Принимая во внимание максимальное расстояние связности соседних БПЛА L_O функция плотности вероятности расстояния между двумя связанными БПЛА D_C имеет вид

$$f_{D_C}(x) = \frac{\lambda e^{-\lambda x}}{\int_0^{L_O} \lambda e^{-\lambda y} dy} = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda L_O}}, \quad 0 < x < L_O. \quad (7.18)$$

Поскольку последовательные интервалы в процессе Пуассона являются независимыми с.в., число последовательных интервалов между БПЛА, образующих интервал связности, следует геометрическому распределению с параметром

$$p_C = \int_0^{L_O} \lambda e^{-\lambda y} dy = 1 - e^{-\lambda L_O}. \quad (7.19)$$

Пусть C обозначает длину интервала связности. Эта случайная величина может быть записана как сумма расстояний между БПЛА, взвешенная с геометрическими коэффициентами $p_C^{i-1}(1 - p_C)$, $i = 1, 2, \dots$. Таким образом, плотность вероятностей интервала связности может быть записана следующим образом:

$$f_C(x) = \sum_{i=1}^{\infty} p_C^{i-1}(1 - p_C) \left[f_{D_C} \underset{\text{itimes}}{\star} f_{D_C} \right](x), \quad (7.20)$$

где символ \star обозначает свертку распределений:

$$\left[f_{D_C} \star f_{D_C} \right](x) = \int_0^{2L_O} f_{D_C}(x - \tau) f_{D_C}(\tau) d\tau. \quad (7.21)$$

Чтобы получить приближение плотности вероятности интервала связности в аналитической форме, заменим ограниченное экспоненциальное распределение $f_{D_C}(x) = \lambda e^{-\lambda x} / (1 - e^{-\lambda L_O})$ экспоненциальным распределением с совпадающим средним значением:

$$\lambda_A = \frac{1}{\int_0^{L_O} \frac{\lambda}{e^{-\lambda L_O}} e^{-\lambda x} dx} = \left[\frac{1}{\lambda} - \frac{L_O}{e^{\lambda L_O} - 1} \right]^{-1}. \quad (7.22)$$

Теперь, заметив, что свертка i экспоненциально распределенных с.в. с тем же параметром является распределением Эрланга порядка i , $E(i, \lambda_A)$ с плотностью вероятности

$$f_{E(i, \lambda_A)}(x) = \frac{\lambda_A^i}{(i-1)!} x^{i-1} e^{-\lambda_A x}, x > 0, \lambda_A > 0, \quad (7.23)$$

плотность (7.20) может быть аппроксимирована следующим образом:

$$\begin{aligned} f_C(x) &= \sum_{i=1}^{\infty} (1 - e^{-\lambda L_0})^{i-1} e^{-\lambda L_0} \frac{\lambda_A^i x^{i-1} e^{-\lambda_A x}}{(i-1)!} = \\ &= e^{-\lambda_A x - \lambda L_0} \sum_{i=1}^{\infty} \frac{\lambda_A^i x^{i-1}}{(i-1)!} (1 - e^{-\lambda L_0})^{i-1} = \\ &= \frac{\lambda [e^{\lambda L_0} - 1] e^{-\lambda [\lambda L_0^2 - L_0 e^{\lambda L_0} + x [e^{-\lambda L_0} - 1] + L_0]}}{e^{\lambda L_0} - \lambda L_0 - 1}. \end{aligned} \quad (7.24)$$

Отметим, что мост между двумя БС существует, только если расстояние на решетке между ними меньше, чем длина интервала связности, C . Пусть с.в. D обозначает расстояние между БС на решетке, которое задается суммой разности координат, (X_1, Y_1) и (X_2, Y_2) , т.е.

$$|(X_1 - X_2)| + |Y_1 - Y_2|, \quad (7.25)$$

где X_i и Y_i – равномерно распределены на $(0, N)$.

Выполняя стандартные преобразования с.в., искомая дистанция имеет следующий ряд распределения:

$$\begin{aligned} p_0 &= \frac{1}{(2N+1)^2}, \\ p_i &= \frac{4(2N+1-i)}{(2N+1)^2} + \frac{(-1)^i + 1}{2} \frac{4(2N+1-i/2)^2}{(2N+1)^4} + \\ &+ \sum_{k=1}^{\lfloor i/2-1 \rfloor} \frac{4(2N+1-k)(2N+1-i-k)}{(2N+1)^4}, i = 1, 2, \dots, 4N. \end{aligned} \quad (7.26)$$

Охарактеризовав интервал связности и расстояние между исходной и целевой БС, приступим к определению средней продолжительности интервала выгрузки, когда возможна многошаговая ретрансляционная связь. Плотность вероятности искомой с.в. T_0 можно записать следующим образом:

$$f_{T_0}(x) = \frac{f_{C-D}(x)}{\int_{-\infty}^0 f_{C-D}(y) dy}, x > 0. \quad (7.27)$$

Далее можно определить плотность вероятности T_0 путем свертки C и $(-D)$. Так как распределение D является дискретным, получаем следующее выражение для ФР с.в. T_0 :

$$F_{T_0}(x) = \sum_{j=0}^i F_C(x+j)p_j. \quad (7.28)$$

Заметим, что распределение длины интервала $E[T_1]$, в котором не обеспечивается многошаговая ретрансляционная связь между исходной и целевой БС, может быть найдено аналогичным способом. Разница в том, что теперь необходимо учитывать последовательные расстояния между БПЛА, которые превышают расстояние отсутствия связи L_0 . Доля времени, в течение которого доступен мост и которое не зависит от скорости БПЛА, v_U может быть записана в следующем виде:

$$\zeta = \frac{E[T_0]}{E[C]+E[O]}. \quad (7.29)$$

Доля выгружаемого трафика

Когда найдена доля времени, в течение которого многошаговая ретрансляционная связь доступна, можно перейти к определению доли трафика, выгружаемого на мост БПЛА. Доля трафика, передаваемого ретрансляционным каналом в любой момент времени, когда он доступен, определяется как $\epsilon = \min\{1, R_{max}/R\}$, где R_{max} – максимальная скорость, предоставляемая мостом БПЛА, R – запрашиваемая скорость передачи. Поскольку последняя величина является частью спецификации системы, единственным неизвестным является скорость R_{max} .

Обратим внимание, что для моста, состоящего из i БПЛА, скорость ограничена сверху минимальной скоростью между двумя смежными БПЛА. Таким образом, задача состоит в том, чтобы определить максимум из $i-1$ экспоненциальных с.в., а затем взвесить их с вероятностью $p_C^{i-1}(1-p_C)$ наличия ровно $i-1$ БПЛА в мосте. Плотность вероятности максимума i экспоненциально распределенных с.в. имеет вид

$$\begin{aligned} f_{maxD_C}(x; i-1) &= (i-1)f_{D_C}(x)[F_{D_C}(x)]^{i-2} = \\ &= (i-1)\lambda_A e^{-\lambda_A x}(1-e^{-\lambda_A x})^{i-2}. \end{aligned} \quad (7.30)$$

Теперь плотность вероятности максимальной дистанции может быть получена в виде

$$\begin{aligned}
 f_{maxD_C}(x) &= \sum_{i=1}^{\infty} p_C^{i-1} (1 - p_C) f_{maxD_C}(x; i - 1) = \\
 &= \sum_{i=1}^{\infty} \frac{(1 - e^{-\lambda L_0})^{i-1} (i - 1) \lambda_A (1 - e^{-\lambda_A x})^{i-2}}{e^{\lambda L_0 + \lambda_A x}} = \\
 &= \frac{\lambda (e^{\lambda L_0 - 1}) e^{\lambda \left(\frac{\lambda L_0 x}{e^{\lambda L_0 - 1}} - L_0 \lambda + L_0 + x \right)}}{(e^{\lambda L_0 - 1} - L_0 \lambda) \left(e^{\lambda L_0 + e^{\lambda \frac{x}{L_0}} \frac{1}{e^{\lambda L_0 - 1} - 1}} \right)^2}.
 \end{aligned} \tag{7.31}$$

После нахождения максимального расстояния между соседними БПЛА используем метод преобразования с.в., чтобы определить плотность вероятности $f_{R_{max}}(x)$ максимальной скорости, поддерживаемой мостом БПЛА. После этого плотность вероятности R_{max}/R находится масштабированием R_{max} , то есть имеет вид $f_{R_{max}}(xR)$. Наконец, плотность вероятности с.в. $\min\{1, R_{max}/R\}$ определяется формулой

$$f_{\min\{1, \frac{R_{max}}{R}\}}(x) = \begin{cases} f_{R_{max}}(xR), & 0 < x < 1, \\ \int_1^{\infty} f_{R_{max}}(yR) dy, & x = 1. \end{cases} \tag{7.32}$$

Доля скорости выгрузки в некоторый случайный момент времени, когда существует мост БПЛА между БС, определяется по формуле

$$\epsilon = \int_0^1 f_{\min\{1, R_{max}/R\}}(x) x dx. \tag{7.33}$$

Вероятность блокировки

В дополнение к плотности БПЛА на пути между исходной и целевой БС здания также могут влиять на долю времени существования моста. Например, как показано на рис. 7.21, на перекрестках здания могут блокировать LoS между соседними БПЛА. В этом случае мост не может быть установлен, даже если это позволяет расстояние между смежными БПЛА. Тогда вероятность блокировки ν зависит от числа поворотов, высот зданий и ширины улиц.

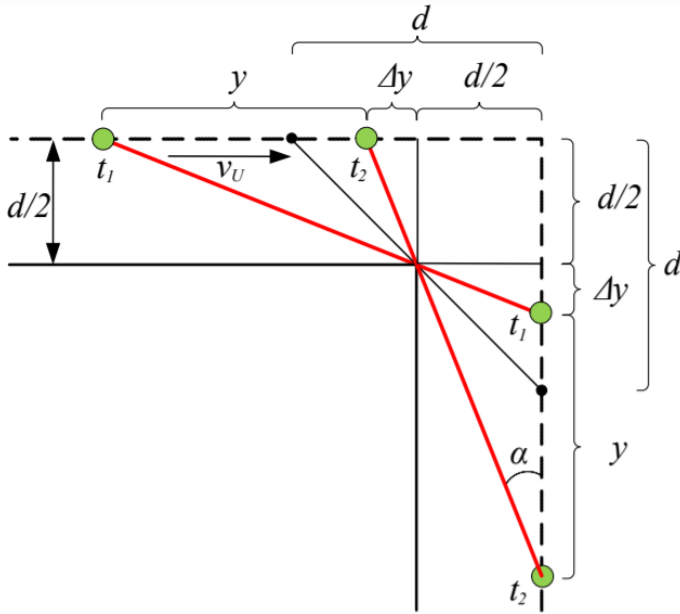


Рис. 7.21. Вероятность блокировки зданием

Пусть q_i – вероятность того, что на пути между исходной и целевой БС мост пересекает i перекрестков. Поскольку оператор БПЛА разрешает только те повороты, которые минимизируют расстояние до точки выхода из области, количество четных и нечетных поворотов, которые совершают БПЛА на пути от $(0,0)$ к (N,N) , задается следующим распределением:

$$\begin{aligned} p(N, 2k) &= 2C_{N-1}^{k-1}C_{N-1}^k, \\ p(N, 2k - 1) &= 2(C_{N-1}^{k-1})^2. \end{aligned} \quad (7.34)$$

Учитывая случайное расстояние между исходной и целевой БС, вероятность ровно i перекрестков определяется формулой $q_i = p(i, k)p_i$. Учитывая геометрическую схему перекрестка (см. рис. 7.21), заметим, что блокировка LoS отсутствует, когда высота здания H_B меньше высоты h_U полета БПЛА и когда расстояние между БПЛА меньше, чем ширина улицы d . Таким образом, процесс многошаговой ретрансляции прерывается, когда расстояние между связанными БПЛА находится в пределах $D_C^* \in (d, L_O)$, и плотность вероятности этой с.в. определяется формулой

$$f_{D_C^*}(x) = \frac{f_D(x)}{1 - [F_D(L_O) - F_D(d)]} = \frac{1}{e^{-d\lambda} - e^{-\lambda L_O}} \lambda e^{-\lambda x}, d < x < L_O. \quad (7.35)$$

Наконец, вероятность блокировки LoS определяется в виде

$$v = \sum_{i=1}^{4N} p(i, k) q_i \left(F_{H_B}(h_U) \left[1 - \int_a^{L_0} f_{D_C^*}(x) dx \right] \right)^i, \quad (7.36)$$

где h_U – высота полета БПЛА.

Подставляя $F_{H_B}(x)$, $f_{D_C^*}(x)$, а также p_i и $p(i, k)$ в (7.36), получаем искомую вероятность v .

Анализ предложенных улучшений

Один из простых способов снизить эффект блокировки LoS и повысить долю времени выгрузки трафика на БПЛА-мост – использовать функцию выбора маршрута через рассматриваемую область. Заметим, что между двумя случайно и равномерно распределенными точками на решетке всегда существует путь, имеющий только один поворот. В этом случае вероятность блокировки уменьшается до величины

$$v = F_{H_B}(h_U) \left[1 - \int_a^{L_0} f_{D_C^*}(x) dx \right], \quad (7.37)$$

которая может быть найдена в замкнутой аналитической форме.

Поскольку ожидается, что БПЛА в процессе полета будут развивать скорости, достигающие 60–80 км/ч (16–22 м/с), разрывы моста, вызванные блокировкой LoS, можно уменьшить, используя буферизацию на БПЛА. Оценим необходимый объем буферного пространства.

Пусть W – буферное пространство, необходимое для устранения эффекта блокировки зданиями, для которого справедливо соотношение

$$W = R_{max} Y / v_U, \quad (7.38)$$

где Y – с.в. времени, которое БПЛА проводит в условиях блокировки; R_{max} – максимальная скорость, предоставляемая мостом БПЛА; v_U – скорость движения БПЛА.

Используя геометрические соотношения между величинами на рис. 7.21, последовательно устанавливаем

$$\tan \alpha = \frac{\Delta y + d/2}{y + \Delta y + d/2} = \frac{\Delta y}{d/2},$$

$$(\Delta y)^2 + (d - D_C) \Delta y + d^2/4 = 0,$$

$$\Delta y = \sqrt{D_C^2 - 2D_C d} + D_C - d)/2. \quad (7.39)$$

Теперь *дистанция блокировки* может быть найдена в виде

$$Y = \frac{\sqrt{D_C^{*2} - 2D_C^* d + D_C^{*2} - d}}{2}, \quad (7.40)$$

где D_C^* – условная с.в. расстояния между двумя связанными БПЛА.

Подставляя (7.40) в (7.38), получаем

$$W = \frac{R_{max}(\sqrt{D_C^{*2} - 2D_C^* d + D_C^{*2} - d})}{2v_U}. \quad (7.41)$$

Плотность распределения с.в. W можно найти с помощью метода нелинейного преобразования случайных величин. Чтобы оценить средние требования к размеру буферного накопителя на БПЛА, достаточно применить приближение рядом Тейлора, рассмотренное ранее в главе 3 для оценки ОСШ и ОСП.

7.6.4. Численный анализ

Начнем анализ со сравнения производительности стратегий выгрузки в зависимости от параметров системы и типов городской застройки. В соответствии с рекомендацией Р.1410 МСЭ-Р различаем пригородные, городские, густонаселенные и высотные городские районы. Каждый тип застройки характеризуется средней шириной улиц d , средней длиной кварталов d_B , а также средней высотой зданий $E[H_B]$ в зависимости от плотности населения в различных районах. Значения этих параметров, используемых в численном анализе, показаны в табл. 7.5. Также в соответствии с рекомендацией Р.1410 предполагается, что с.в. высоты здания распределена по закону Рэля с параметром $\sigma = E[H_B]/\sqrt{\pi/2}$. Другие значения параметров системы представлены в табл. 7.4.

Таблица 7.4

Параметры системной модели БПЛА-мостов

Параметр	Значение
Несущая частота, f_c	73 ГГц
Полоса пропускания, B	100 МГц
Число улиц, N	10
Тип застройки	Густонаселенный район
Ширина улицы, d	10/13/20/20 м
Длина квартала, d_B	37/45/60/60 м

Окончание табл. 7.4

Параметр	Значение
Высота здания, H_B	10/19/25/63 м
Излучаемая мощность, P_T	23 дБм
Чувствительность приемника, S_T	10 дБ
Антенные эл. на передаче, N_T	4
Антенные эл. на приеме, N_R	4
Высота полета БПЛА, h_U	30 м
Скорость БПЛА, v_U	10 м/с
Требуемая скорость выгрузки, R	3 Гбит/с
Плотность потока БПЛА, λ	0,04 1/м

Таблица 7.5

Параметры городской застройки

Тип	Высота	Ширина	Ширина улицы
Пригород	10 м	37 м	10 м
Город	19 м	45 м	13 м
Густонаселенный город	25 м	60 м	20 м
Высотные районы города	63 м	60 м	20 м

Графики на рис. 7.22 иллюстрируют влияние плотности БПЛА на коэффициент выгрузки для стратегий выгрузки трафика *C1: Мост с маршрутизацией* и *C2: Управляемый мост*, для густонаселенной городской застройки, антенных решеток на передаче и приеме 4×1 , с размером решетки улиц $N = 10$, скоростью передачи $R = 3$ Гбит/с, средней высотой зданий $E[H_B] = 25$ м и высотой полета БПЛА $h_U = 30$ м. Анализируя эти данные, можно заметить, что стратегия выгрузки с маршрутизацией *C1* позволяет установить мост БПЛА, начиная с плотности потока БПЛА около 0,01 1/м, что с учетом скорости полета соответствует расстоянию между БПЛА примерно 100 м. Как видно из графиков на рис. 7.22, начиная с этого значения плотности потока БПЛА, коэффициент выгрузки γ начинает быстро расти. Обратим внимание, что, управляя выбором маршрута, требование только одного поворота между точкой выгрузки трафика и точкой перегрузки приводит к значительному улучшению поведения системы с точки зрения коэффициента выгрузки. Заметим, что стратегия *C1* может даже превзойти стратегию с управлением *C2*, начиная примерно с плотности потока БПЛА, равной 0,02 1/м.

Анализируя графики на рис. 7.22, можно также заметить, что стратегия с управлением *C2* применима в широком диапазоне значений плотности потока БПЛА. Например, выгрузка начинается

уже со значения плотности около 0,008 1/м, что соответствует дистанции между БПЛА примерно 125 м. Кроме того, выгода от выбора маршрута намного выше по сравнению со схемой 1 и достигает 0,6 Гбит/с. Это объясняется случайным выбором маршрута в стратегии С1, а также тем, что разрыв связи в мосте БПЛА в результате блокировки LoS определяется максимальным расстоянием между ними. Таким образом, можно сделать вывод о том, что формирование детерминированного потока мостов БПЛА путем принудительной задержки БПЛА в точке отправления (управляемый мост на рис. 7.22) приводит к значительному выигрышу и увеличению коэффициента выгрузки трафика.

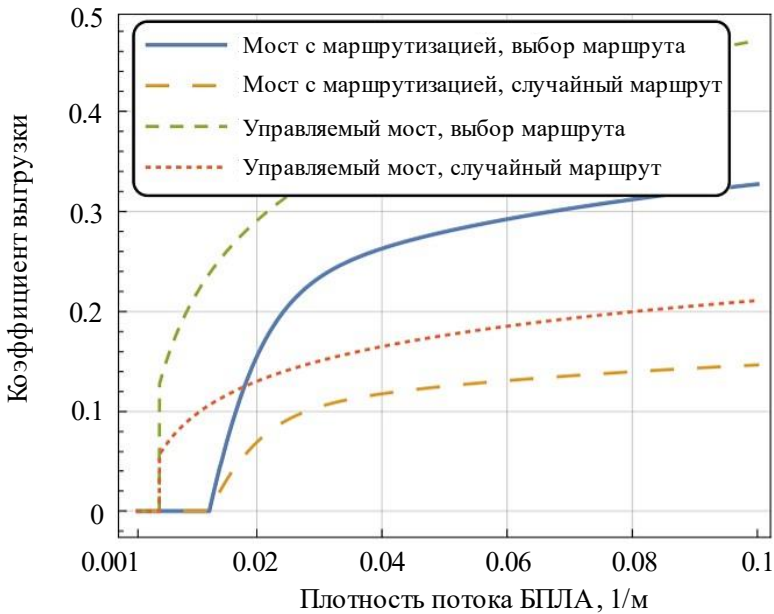


Рис. 7.22. Эффект плотности потока БПЛА

Теперь рассмотрим влияние типа городской застройки на коэффициент выгрузки, которая фактически определяется шириной зданий и улиц, а также средней высотой зданий. На рис. 7.23 показан коэффициент выгрузки для маршрутизируемой схемы с выбором маршрута в зависимости от плотности потока БПЛА λ , размерности решетки $N = 10$, высоты полета БПЛА $h_U = 30$ м и 4×1 антенных решеток на приеме и передаче. Можно заметить, что более высокое значение коэффициента выгрузки наблюдается для пригородных

застроек, в то время как высотные городские районы характеризуются худшими значениями коэффициента выгрузки. Это объясняется средней высотой зданий, которая является наибольшей для высотных районов. Эффект, кроме высоты зданий, усиливается за счет ширины улиц и зданий, которые являются наименьшими для пригородных районов и наибольшим для высотных городских районов (см. табл. 7.5).

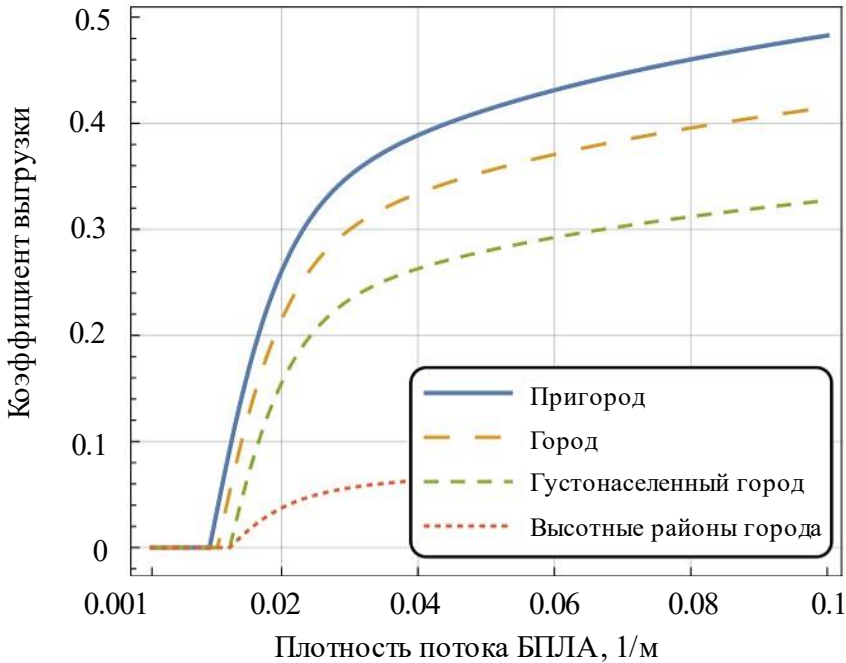


Рис. 7.23. Эффект типа городской застройки

Влияние высоты полета БПЛА для густонаселенной городской застройки на коэффициент выгрузки для стратегий С1 и С2 видно из графиков на рис. 7.24 для значений $\lambda = 0,04$ 1/м, $N = 10$, $E[H_B] = 25$ м и антенных решеток 4×1 на передающей и принимающей сторонах. Можно заметить, что увеличение высоты полета улучшает характеристики выгрузки для обеих стратегий. Важно отметить, что при увеличении h_U как маршрутизируемая, так и управляемая схемы без выбора маршрута приближаются к тем, которые используют маршруты с одним поворотом (стратегии с выбором маршрута на рис. 7.24). Причина в том, что даже при большем числе поворотов между точкой перегрузки и точкой выгрузки трафика вероятность по меньшей

мере одной блокировки LoS резко уменьшается при увеличении h_U . Таким образом, чтобы смягчить негативное влияние блокировок, можно либо принудительно выбрать маршрут либо увеличить высоту полета БПЛА. Однако чтобы полностью избавиться от блокировок, требуются дополнительные меры.

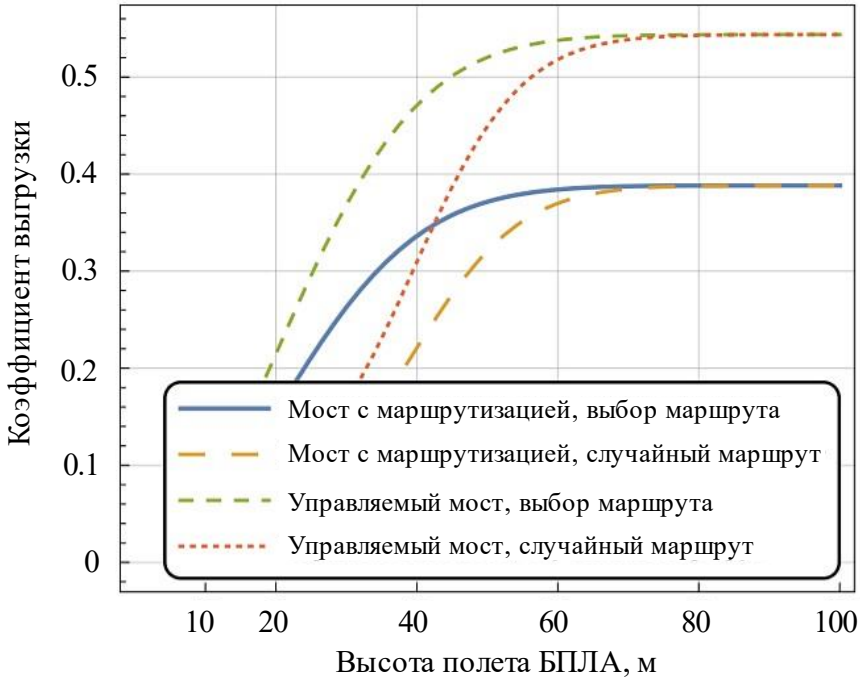


Рис. 7.24. Эффект высоты полета БПЛА

Напомним, что использование буферизации увеличивает доступность БПЛА-мостов, устраняя вызванные блокировкой разрывы в потоках данных. Перейдем к оценке получаемой выгоды от использования буферизации на БПЛА.

Оценим объем буферного накопителя, необходимого для промежуточного хранения данных на время кратковременных блокировок LoS, вызванных поворотами полета БПЛА. Следует ожидать, что на требуемый объем буфера будет влиять в основном скорость полета БПЛА. Влияние скорости БПЛА на коэффициент выгрузки стратегии S1 показано на рис. 7.25 для $N = 10$, $R = 3$ Гбит/с, антенных решеток 4×1 на принимающей и передающей сторонах и густонаселенной городской застройки. Заметим, что требования к буферному накопителю не зависят от выбора маршрута. Анализируя графики

на рис. 7.26, можно заметить, что повышение скорости полета БПЛА значительно уменьшает объем буферного накопителя, необходимого для бесперебойной работы БПЛА-моста. Кроме того, требования к объему накопителя также резко снижаются при увеличении плотности потока БПЛА. Однако даже для наихудшего случая $\lambda = 0,001$ 1/м размер буферного накопителя составляет всего 8 ГБ, что является приемлемым техническим значением. Аналогичные эффекты наблюдаются и для схемы с управлением С2. Принципиальное отличие состоит в том, что в схеме С2 не требуется буферизация, когда плотность потока БПЛА достигает значения $\lambda = 1/2d$ м⁻¹.

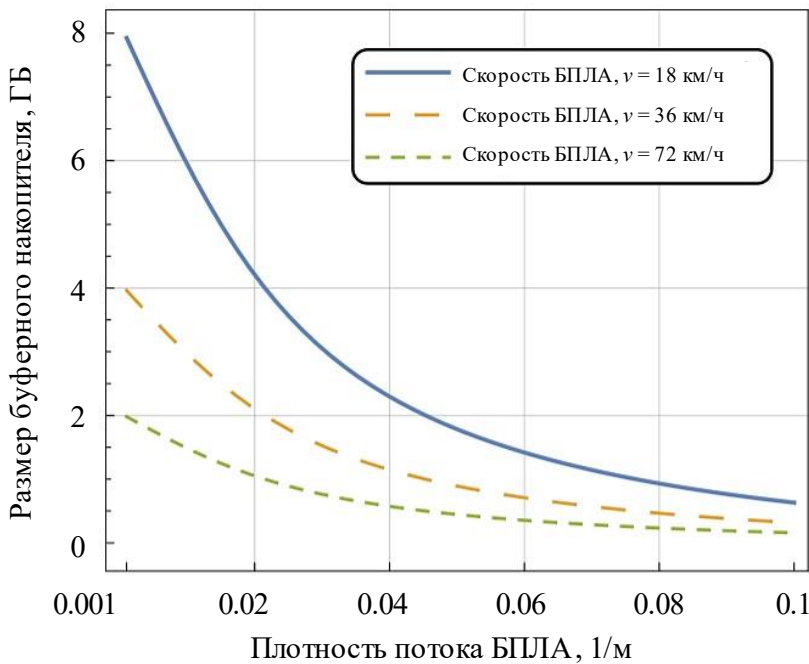


Рис. 7.25. Размер буферного накопителя

Графики на рис. 7.26 иллюстрируют коэффициент выгрузки для схем С1 и С2 с выбором маршрута, с буферизацией и без нее, для $N = 10$, $R = 3$ Гбит/с, $h_U = 30$ м, густонаселенной городской застройки и антенных решеток 4×1 на передающей и принимающей сторонах. Можно заметить, что для обеих стратегий буферизация на БПЛА значительно увеличивает коэффициент выгрузки. В том числе для рассматриваемого диапазона параметров системы стратегия С1 с буферизацией превосходит стратегию С2 без буферизации,

начиная с плотности потока БПЛА около 0,03 1/м. Кроме того, разница между схемами с буферизацией и без нее увеличивается с ростом плотности потока БПЛА. Таким образом, можно сделать вывод, что, задействуя весь потенциал ретрансляционных соединений на основе БПЛА путем использования функций буферизации, выбора маршрутов и задержки БПЛА в точке отправления, можно значительно улучшить производительность системы.

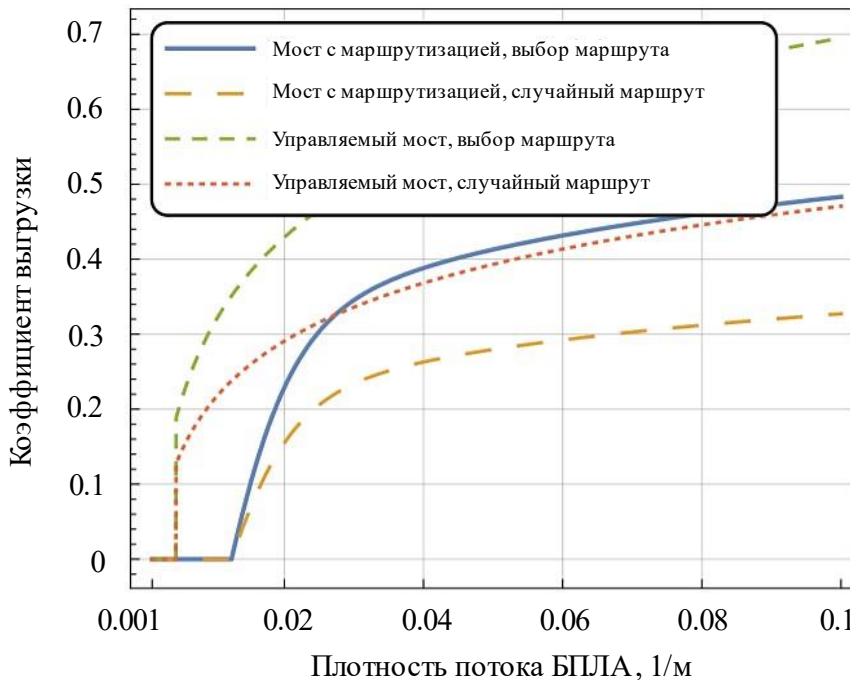


Рис. 7.26. Коэффициент выгрузки для стратегии с буферизацией данных

Одним из преимуществ схем выгрузки, рассмотренных в этом разделе, является то, что модули NR, используемые на БПЛА, представляют собой обычные модули потребительского уровня. Тем не менее было бы полезно оценить эффект от использования более совершенных модулей NR.

Перейдем к изучению влияния антенных решеток на передающей и принимающей сторонах. Влияние размера антенной решетки передатчика (числа антенных элементов, формирующих диаграмму направленности антенны в горизонтальной плоскости) показано на рис. 7.27 для стратегий C1 и C2, $\lambda = 0,04$ 1/м, $N = 10$, $R = 3$ Гбит/с, $h_U = 30$ м и густонаселенной городской застройки. Из графиков

видно, что увеличение числа антенных элементов улучшает характеристики выгрузки для рассматриваемых стратегий. Кроме того, наибольший выигрыш наблюдается для стратегии с управлением и с буферизацией. При использовании антенных решеток 32×1 вместо 4×1 коэффициент выгрузки увеличивается с 0,55 до 0,70 (примерно на 450 Мбит/с) для стратегии С2 с буферизацией и с 0,27 до 0,34 (примерно на 250 Мбит/с) для стратегии С1 с выбором маршрута. Подобные улучшения наблюдаются и при использовании антенных решеток с большим количеством элементов.

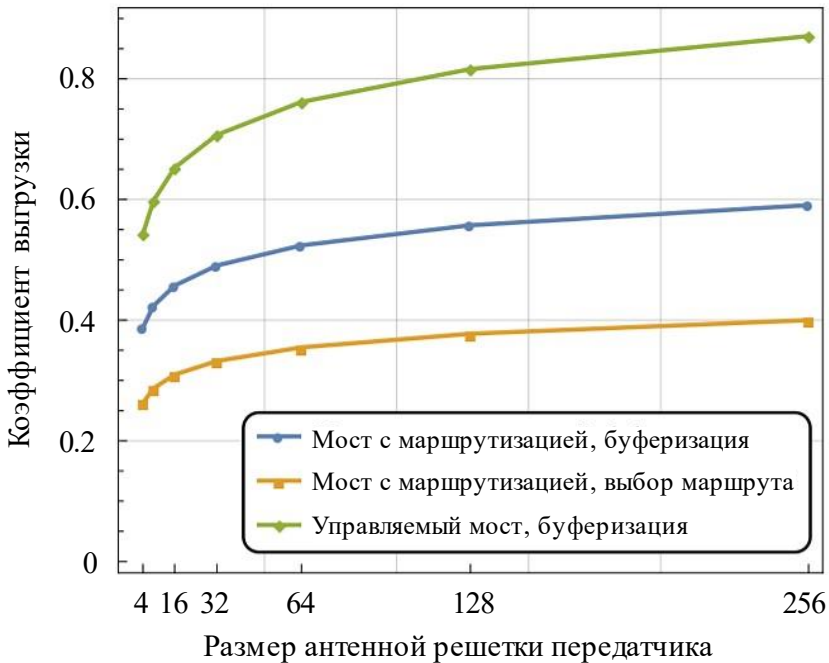


Рис. 7.27. Влияние размера антенной решетки передатчика на коэффициент выгрузки



ЧАСТЬ III

УСЛУГИ И ФУНКЦИОНАЛЬНЫЕ ВОЗМОЖНОСТИ СЕТЕЙ 5G/6G



ГЛАВА 8

УСЛУГИ СВЕРХНАДЕЖНОЙ ДОСТАВКИ С УЛЬТРАМАЛОЙ ЗАДЕРЖКОЙ

В данной главе рассматриваются аспекты поддержки услуг сверхнадежной доставки с ультрамалой задержкой (URLLC, Ultra-Reliable Low-Latency Communication) для сетей радиодоступа 5G/6G. Изложение начинается с предпосылок внедрения этих услуг в промышленной среде, описания тенденций изменения в промышленности, представлены примеры использования беспроводных сетей и роли технологии NR в Индустрии 4.0. Особое внимание уделяется совместной поддержке услуг URLLC и расширенной услуги широкополосного доступа (eMBB, enhanced Mobile Broadband) в промышленном сценарии развертывания сетей радиодоступа с направленными антеннами. Для предотвращения блокировок в таких сетях раскрываются возможности кэширования и предсказания событий блокировок.

8.1. Концепция Индустрии 4.0

Следуя требованиям рынка, за последние столетия промышленность пережила несколько революций. Во времена, когда появились ремесла, акцент делался на пожелания заказчика и уникальность товара, но за счет этого росла цена на товары. Более поздние промышленные революции были направлены на повышение производительности путем замены ручного труда машинами. Как побочный эффект, продукция перестала быть уникальной, «на заказ», и стала стандартизированной. Эта тенденция продолжилась и в XX веке, когда развитие обрабатывающей промышленности стало следовать парадигме «фордизма». В социологии «фордизм» определяется как экономическая и социальная система, основанная на индустриализированном, стандартизированном производстве и массовом потреблении.

В настоящее время рынок требует персонализации продукции при сохранении низкой цены и стабильного качества массового производства. Парадигма фордизма не может удовлетворить этим требованиям потому, что реконфигурация производственных мощностей не является достаточно гибкой. Реконфигурация требует развития новой инфраструктуры, включающей телекоммуникационные системы. Более того, реконфигурация может быть нетривиальной из-за неоднородности коммуникационных технологий и протоколов. Последнее требует значительных усилий по настройке на этапе запуска.

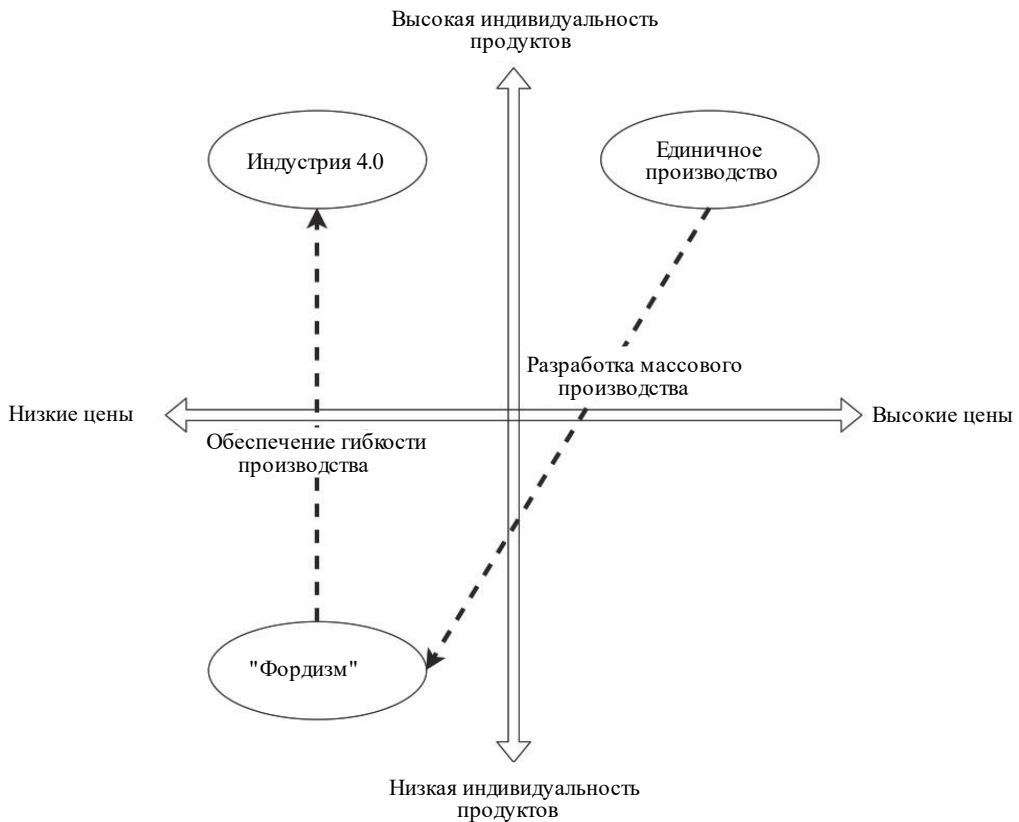


Рис. 8.1. Эволюция технологий производства

Происходящая в начале XXI века промышленная революция, получившая название «Индустрия 4.0», приводит к широкой автоматизации и гибкости производственных мощностей. Такая трансформация происходит благодаря цифровизации всех физических систем и процессов, создающих беспрецедентный уровень осведомленности и знаний о производственных процессах. При этом возможность

реконфигурации производственных мощностей «на лету» позволяет динамично адаптироваться к требованиям рынка, выпуская ограниченные серии продукции с приемлемым качеством и стоимостью.

Технологически производственные объекты в Индустрии 4.0 опираются на концепцию индустриального Интернета вещей (IIoT, Industrial IoT). Объекты такой системы состоят из датчиков и исполнительных механизмов (актуаторов), способных собирать данные и выполнять заранее определенный набор действий. Усовершенствованная беспроводная связь в данном контексте призвана обеспечить мобильность устройств в процессе управления ими. Логика управления обычно размещается в облаке или на граничных серверах (Edge Server). Такая архитектура позволяет изменять конфигурацию производственных линий по требованию.

8.1.1. Сценарии использования беспроводной связи на производстве

В соответствии со спецификацией документа 3GPP TS 22.104 ниша для беспроводной связи на производстве включает следующие основные области:

- а) замена проводных соединений, используемых устаревшим оборудованием;
- б) управление движением устройств и связь между системами управления;
- в) мобильные роботизированные платформы;
- г) мониторинг объектов и процессов;
- д) интерфейсы *человек–машина*.

Беспроводная связь должна заменить устаревшие проводные соединения, которые в настоящее время доминируют на уровне операционных технологий, например PROFINET, EtherCAT, Sercos, Modbus. Это становится особенно важно для снижения затрат на ранней стадии трансформации промышленности в Индустрию 4.0. Производственные машины старого образца, которые все еще актуальны для производства, должны быть интегрированы в новые системы управления производством там, где это возможно. Это можно осуществить путем встраивания беспроводных АУ в эти машины.

Системы управления движением отвечают за перемещение устройств или их частей. Этот процесс требует жестких гарантий доставки информации и крайне чувствителен к задержкам передачи

между датчиками, измеряющими текущее положение и/или окружающую среду, и исполнительными механизмами, управляемыми контроллером движения с использованием полученных данных. Связь между системами управления обеспечивает координацию между разнотипными промышленными контроллерами. Требования к услугам связи для таких систем более мягкие по сравнению с управлением движением. Тем не менее системы связи *управление–управление* не имеют фиксированной структуры сообщений и их периодичности, что может привести к динамически изменяющимся требованиям к передаче информации.

В современном промышленном производстве мобильная роботизированная платформа является необходимым инструментом автоматизации. Подобные устройства управляются автоматически и могут выполнять такие задачи, как транспортировка товаров, материалов или других объектов. Мобильные роботы могут требовать сетевые ресурсы в разных местах их расположения и нуждаться в беспшовном хэндовере при перемещении для обеспечения требуемого качества обслуживания. Требуемая задержка и скорость передачи могут варьироваться в зависимости от приложения. В сложных случаях, таких как совместная переноска твердых или хрупких объектов в динамически меняющейся среде, роботизированной системе может потребоваться как ультрамалая задержка, так и высокая скорость, например, для передачи видеoinформации для совместной координации действий.

Системы мониторинга объектов и процессов полагаются на большое количество сенсорных устройств, установленных на предприятии, и предоставляют информацию о текущих процессах. Разнообразие датчиков не ограничивается простыми измерительными приборами, например, термометрами или манометрами. Могут также использоваться системы контроля среды, такие как обычное или тепловое видеонаблюдение, лазерные дальнометры и т.д. Такие датчики могут внести заметный вклад в требования к пропускной способности беспроводной сети.

Предполагается, что люди будут продолжать играть существенную роль в умных производствах будущего. Взаимодействие между людьми и машинами может быть обеспечено с помощью различных интерфейсов. Традиционные панели управления и контроля, широко используемые в промышленности, будут развиваться в направлении поддержки мобильности пользователя. Кроме того,

ожидается, что производственные процессы выиграют от использования технологий дополненной реальности (AR, Augmented Reality). Устройства на базе AR с закрепленными на голове человека дисплеями могут облегчить руководство рабочими, обеспечивая бесперебойную работу и эффективное взаимодействие человека и машины.

В зависимости от конкретного развертывания суммарная скорость передачи данных может достигать нескольких Гбит/с на соту сети, что является сложной задачей для большинства существующих беспроводных технологий и требует новых решений на базе технологий 5G/6G.

8.1.2. Сети 5G/6G для индустрии будущего

Технологии 5G/6G объединяют различные полосы частот и позволяют предоставлять принципиально новые услуги, такие как массовое межмашинное взаимодействие (mMTC, massive Machine-Type Communication), услуги URLLC и eMBB. АУ в сети радиодоступа NR может динамически выбирать необходимую полосу пропускания в зависимости от текущих потребностей путем переключения между частями полосы пропускания. Одновременно для одного АУ может быть настроено до четырех различных частей полосы частот (BWP, Bandwidth Part), однако в каждый момент времени активной может быть только одна из них. Это обеспечивает гибкость сети радиодоступа при обслуживании приложений с различными требованиями, обеспечивая при этом эффективное применение радиоресурсов.

Работа над промышленными вариантами использования технологии NR в контексте сетей 5G/6G началась еще в 3GPP Release 15. Случаи использования с жесткими требованиями к задержкам, включая автоматизацию производства, определены в 3GPP Release 16. Дальнейшая работа по стандартизации в основном направлена на определение конкретных требований промышленности, которые не удовлетворяются текущим состоянием технологий сетей 5G/6G.

С точки зрения физического канала, модели, определенные в документе 3GPP TR 38.901, в настоящее время не включают промышленные сценарии использования. В то же время промышленные сценарии могут иметь особенности, влияющие на распространение сигнала. Что касается производственного сценария использования, то он может включать динамическую блокировку каналов связи миллиметрового или терагерцового диапазонов частот на предприятии, а также электромагнитные помехи, вызванные оборудованием.

Управление радиоресурсами для промышленных сценариев находится в центре внимания стандартизирующих организаций. Например, для сетей 5G/6G необходим пересмотр протоколов канального уровня для реализации услуг, чувствительных к временным задержкам (TSC, Time-Sensitive Communication), и услуг управления ресурсами во время случайных всплесков трафика eMBB. Усовершенствования протоколов, связанные с проблемой TSC, должны на участке доступа в сеть гарантировать, по крайней мере, качество обслуживания, сравнимое с технологией Ethernet. Максимальная задержка в таких каналах не должна превышать 1 мс, а время наработки на отказ – составлять как минимум десять лет. Сети 5G/6G поддерживают услугу URLLC, но скорость передачи данных для такой услуги в конкретных реализациях может оказаться слишком низкой. В промышленных сценариях скорость передачи данных услуг TSC может динамически изменяться и достигать 500 Мбит/с. Поскольку пики трафика TSC могут перекрываться с пиками трафика eMBB и вызывать отказ в обслуживании, управление радиоресурсами, а также оптимизация покрытия сетей 5G/6G в промышленной среде приобретает решающее значение. Ниже рассматриваются проблемы и стратегии совместного обслуживания трафика URLLC и трафика eMBB, основанные на приоритизации и механизмах выгрузки на соединения D2D.

8.2. Поддержка услуг URLLC

Известно несколько подходов, позволяющих обеспечить совместное обслуживание трафика URLLC и трафика eMBB при сохранении весьма строгих требований к качеству обслуживания услуг URLLC. В том числе в случае перегрузки предлагается использовать технологию неортогонального множественного доступа (NOMA, Non-Orthogonal Multiple Access) на беспроводном интерфейсе для обеспечения передачи данных URLLC совместно с передачей данных eMBB. Одним из преимуществ этого подхода является то, что данные URLLC могут быть переданы немедленно, не дожидаясь окончания текущего кадра NR, как это происходит при обычном планировании передачи. Однако такой подход влияет на вероятность корректного приема трафика eMBB и требует тщательного выбора временных слотов для передачи трафика URLLC. Известны методы совместного планирования передачи этих услуг на отдельной БС NR для доступа по технологии NOMA путем решения оптимизационных задач.

Еще один вариант одновременного обслуживания рассматриваемых типов трафика основан на резервировании ресурсов для трафика URLLC. Поскольку нагрузка трафика URLLC заранее неизвестна, для управления резервированием ресурсов необходим механизм прогнозирования. При этом необходимо учесть, что, поскольку перераспределение ресурсов может быть реализовано только в момент начала кадра NR, требование к задержке данных URLLC может быть не выполнено. Задача может быть решена и другим способом – путем приоритизации трафика URLLC относительно трафика eMBB. В этом случае, когда нагрузка трафика URLLC возрастает настолько, что ресурсов для его обслуживания недостаточно, то несколько сессий eMBB могут быть прерваны.

Далее рассмотрен метод совместной поддержки услуг eMBB и URLLC с приоритизацией трафика URLLC и построена модель для оценки плотности обслуживающих БС в миллиметровом диапазоне частот. Для построения модели, учитывающей специфику распространения миллиметровых волн и динамику трафика услуг, использованы методы стохастической геометрии и ТМО. В качестве основных метрик, влияющих на плотность БС, рассмотрены вероятность потери сессий URLLC и eMBB, а также вероятность прерывания обслуживания сессии eMBB.

8.2.1. Системная модель

Рассматривается развертывание сети 5G/6G в промышленной среде, например, на большом заводе с несколькими производственными линиями, как показано на рис. 8.2. Предполагается, что местоположение БС описывается PPP в круге R с плотностью χ $1/m^2$. Предполагается, что БС установлены в помещениях завода на высоте h_A . Ширина полосы пропускания БС равна W .

Используются два типа АУ – устройства управления движением (например, датчики, исполнительные механизмы) и устройства мониторинга (в частности камеры). Устройства управления движением связаны с роботизированным оборудованием или производственными линиями и генерируют трафик URLLC, который требует передачи как по восходящей, так и по нисходящей линии связи. Устройства мониторинга генерируют трафик eMBB и используют только восходящую линию связи. Восходящая и нисходящая линии предполагаются разделенными во времени (TDD, Time Division Duplex). В предположении, что нисходящий канал имеет достаточно ресурсов

для немедленной отправки трафика обоих типов, интерес представляет восходящий трафик. Также предполагается, что геометрическое местоположение датчиков и устройств мониторинга описывается независимыми процессами PPP.

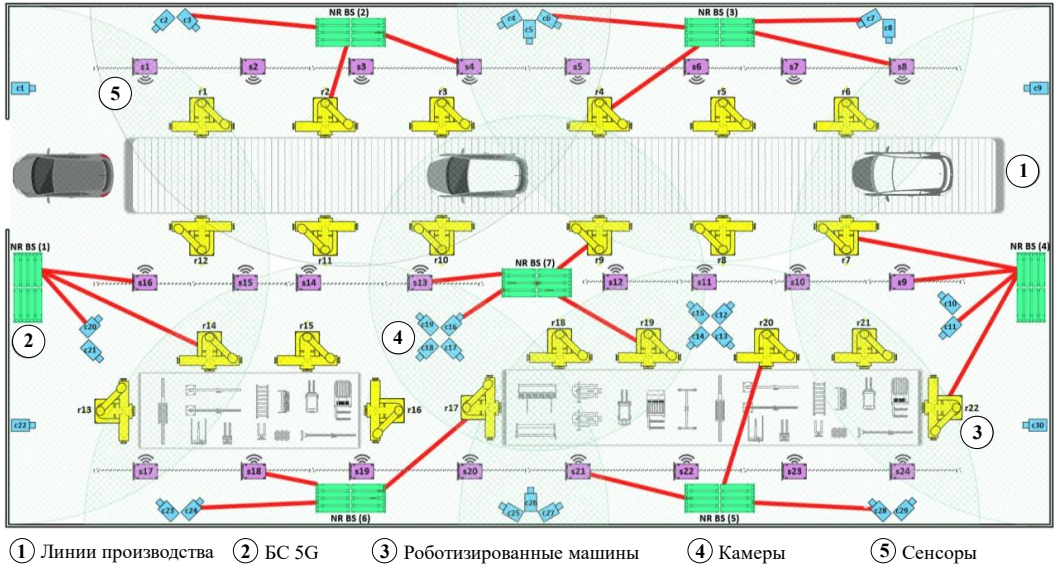


Рис. 8.2. Иллюстрация сценария развертывания услуг 5G на заводе

Пусть запросы на установление сессий мониторинга (eMBB) образуют пуассоновский поток с интенсивностью λ_M и генерируют эластичный трафик с минимальными требованиями к скорости, $c_2^{\min} \geq 1$. Эластичность трафика означает, что приложение может изменять скорость кодека для адаптации к текущим условиям сети с ограничением на ее минимальное значение. Потеря этих сессий может произойти как в момент поступления, так и во время обслуживания из-за прерывания сессией URLLC. Продолжительность eMBB сессии экспоненциально распределена со средним значением μ_M^{-1} .

Предполагается, что поток сессий от датчиков является пуассоновским с интенсивностью λ_S . Требования к скорости являются статическими, $c_1^{\min} \geq 1$, но требования к ресурсам могут меняться в зависимости от местоположения АУ. Значение c_1 определяется на основе скорости передачи данных и требуемого уровня надежности и составляет не более чем 10^{-6} ошибок на блок (BLER, Block Error Rate). Для выполнения этого требования для услуги URLLC предполагается кодирование с повторами, т.е. в заданном кадре передается несколько копий одного и того же сообщения, тем самым повышается

вероятность успешной доставки. Блокировка этих сессий может произойти только в момент поступления, когда все ресурсы системы заняты трафиком услуги URLLC. Продолжительность сессии URLLC распределена экспоненциально со средним значением μ_S^{-1} .

Поскольку сессии URLLC критичны к задержкам, мы предполагаем, что им предоставляется доступ к ресурсам передачи данных в соответствии с приоритетной дисциплиной обслуживания типа Preemptive Priority. В зависимости от сценария приоритет может реализовываться путем снижения объема занимаемого ресурса или даже путем его полного освобождения, т.е. прерывания. Например, если не хватает ресурсов для обслуживания сессий URLLC, одна или несколько сессий eMBB могут быть прерваны. Напомним, что сессии URLLC могут быть потеряны только в том случае, если их суммарная мгновенная нагрузка превышает ресурсы системы.

8.2.2. Анализ системы

Учитывая предположение, что потоки запросов на установление сессий являются пуассоновскими, а их длительности распределены экспоненциально, модель функционирования БС, обслуживающей совместно трафик услуг URLLC и eMBB, может быть описана марковским процессом (МП) $X(t) = \{N_1(t), N_2(t), t > 0\}$, где $N_1(t)$ – число сессий URLLC и $N_2(t)$ – число сессий eMBB в момент t . Пространство состояний МП $X(t)$ имеет вид

$$X = \{(n_1, n_2): n_1 \geq 0, n_2 \geq 0, n_1 b_1 + n_2 b_2^{\min} \leq C\}. \quad (8.1)$$

Обозначим $N_1 = \lfloor C/b_1 \rfloor$ – максимальное число сессий URLLC, $N_2 = \lfloor C/b_2^{\min} \rfloor$ – максимальное число сессий eMBB, $k(n_1) = \lfloor C - n_1 b_1 / b_2^{\min} \rfloor$ – максимальное число сессий eMBB, когда число сессий URLLC в системе равно n_1 . Из предположения об «эластичности» трафика сессий eMBB, т.е. равномерного распределения скорости передачи между одновременно обслуживаемыми сессиями, следует, что достижимая скорость передачи $b_2(n_1, n_2)$ сессий eMBB в состоянии $(n_1, n_2) \in X$ может быть вычислена по формуле

$$b_2(n_1, n_2) = \left\lfloor \frac{C - n_1 b_1}{n_2} \right\rfloor \geq b_2^{\min}. \quad (8.2)$$

Для понимания работы системы рассмотрим процедуру управления доступом в момент поступления запросов на установление сессий URLLC и eMBB. В момент поступления сессии URLLC возможны следующие изменения состояний системы:

– если новая сессия застает в системе не менее b_1 свободных ресурсных блоков, то эта сессия принимается, а сессии eMBB не прерываются;

– если в момент поступления сессии одновременно выполняются условия: а) свободно менее b_1 ресурсных блоков, б) число сессий URLLC на обслуживании менее N_1 , в) если на обслуживании имеются сессии eMBB, то поступившая сессия принимается, а $\left[\left(b_1 - C + (n_1 b_1 + n_2 b_2^{\min}) \right) / b_2^{\min} \right]$ сессий eMBB прерываются;

– в противном случае поступившая сессия URLLC сбрасывается (теряется).

Если в момент поступления сессии eMBB в системе имеется не менее $b_2(n_1, n_2)$ свободных ресурсных блоков, то сессия принимается в систему. В противном случае сессия eMBB сбрасывается.

Следуя сформулированным условиям, можно построить диаграмму переходов состояний марковского процесса $X(t)$, как показано на рис. 8.3.

Используя диаграмму на рис. 8.3, получаем систему уравнений равновесия (СУР) в виде

$$\begin{aligned}
 & \left[\lambda_1 I \{ n_1 < N_1, b_1(n_1 + 1) + b_2^{\min} n_2 \leq C \} + \lambda_2 I \{ n_2 < N_2, b_1 n_1 + \right. \\
 & \left. + b_2^{\min}(n_2 + 1) \leq C \} + n_1 \mu_1 + n_2 \mu_2 + \lambda_1 I \{ n_1 < N_1, n_2 > 0, b_1(n_1 + 1) + \right. \\
 & \left. + b_2^{\min} n_2 \leq C \} \right] p(n_1, n_2) = \lambda_1 I \{ n_1 > 0, b_1 n_1 + b_2^{\min} n_2 \leq C \} p(n_1 - 1, n_2) + \\
 & \quad + \lambda_2 I \{ n_2 > 0, b_1 n_1 + b_2^{\min} n_2 \leq C \} p(n_1, n_2 - 1) + (n_1 + 1) \mu_1 I \{ n_1 < \\
 & N_1, b_1(n_1 + 1) + b_2^{\min} n_2 \leq C \} p(n_1 + 1, n_2) + (n_2 + 1) \mu_2 I \{ n_2 < N_2, b_1 n_1 + \\
 & \quad + b_2^{\min}(n_2 + 1) \leq C \} p(n_1, n_2 + 1) + \lambda_1 I \{ n_1 > 0, n_2 + 1 \leq N_2, b_1(n_1 - 1) + \\
 & \quad + b_2^{\min}(n_2 + 1) \leq C, b_1 n_1 + b_2^{\min}(n_2 + 1) > C \} p(n_1 - 1, n_2 + 1) + \dots + \\
 & \quad + \lambda_1 I \{ n_1 > 0, b_1(n_1 - 1) + b_2^{\min} k(n_1 - 1) \leq C, b_1 n_1 + b_2^{\min} k(n_1 - 1) > \\
 & \quad C \} p(n_1 - 1, k(n_1 - 1)), \\
 & n_1 = 0, 1, \dots, N_1 \text{ и } n_2 = 0, 1, \dots, N_2.
 \end{aligned} \tag{8.3}$$

Обозначим $\mathbf{p} = (p(n_1, n_2))_{(n_1, n_2) \in X}$ стационарное распределение СП $X(t)$, которое можно найти, решая численно СУР:

$$\mathbf{pA} = 0, \mathbf{p1}^T = 1, \tag{8.4}$$

где \mathbf{A} – инфинитезимальный генератор, элементы которого $a((n_1, n_2), (n'_1, n'_2))$ могут быть вычислены по формулам

$$a((n_1, n_2), (n'_1, n'_2)) = \quad (8.5)$$

- λ_1 , если $n'_1 = n_1 + 1, n'_2 = n_2, n_1 < N_1, b_1(n_1 + 1) + b_2^{\min} n_2 \leq C$,
или $n'_1 = n_1 + 1, n'_2 = n_2 - l(n_1, n_2), n_1 < N_1, n_2 > 0, b_1(n_1 + 1) + b_2^{\min} n_2 > C$;
- λ_2 , если $n'_1 = n_1, n'_2 = n_2 + 1, n_2 < N_2, b_1 n_1 + b_2^{\min}(n_2 + 1) \leq C$;
- $n_1 \mu_1$, если $n'_1 = n_1 - 1, n'_2 = n_2, n_1 > 0$;
- $n_2 \mu_2$, если $n'_1 = n_1, n'_2 = n_2 - 1, n_2 > 0$;
- $*$, если $n'_1 = n_1, n'_2 = n_2$;
- 0 , иначе.

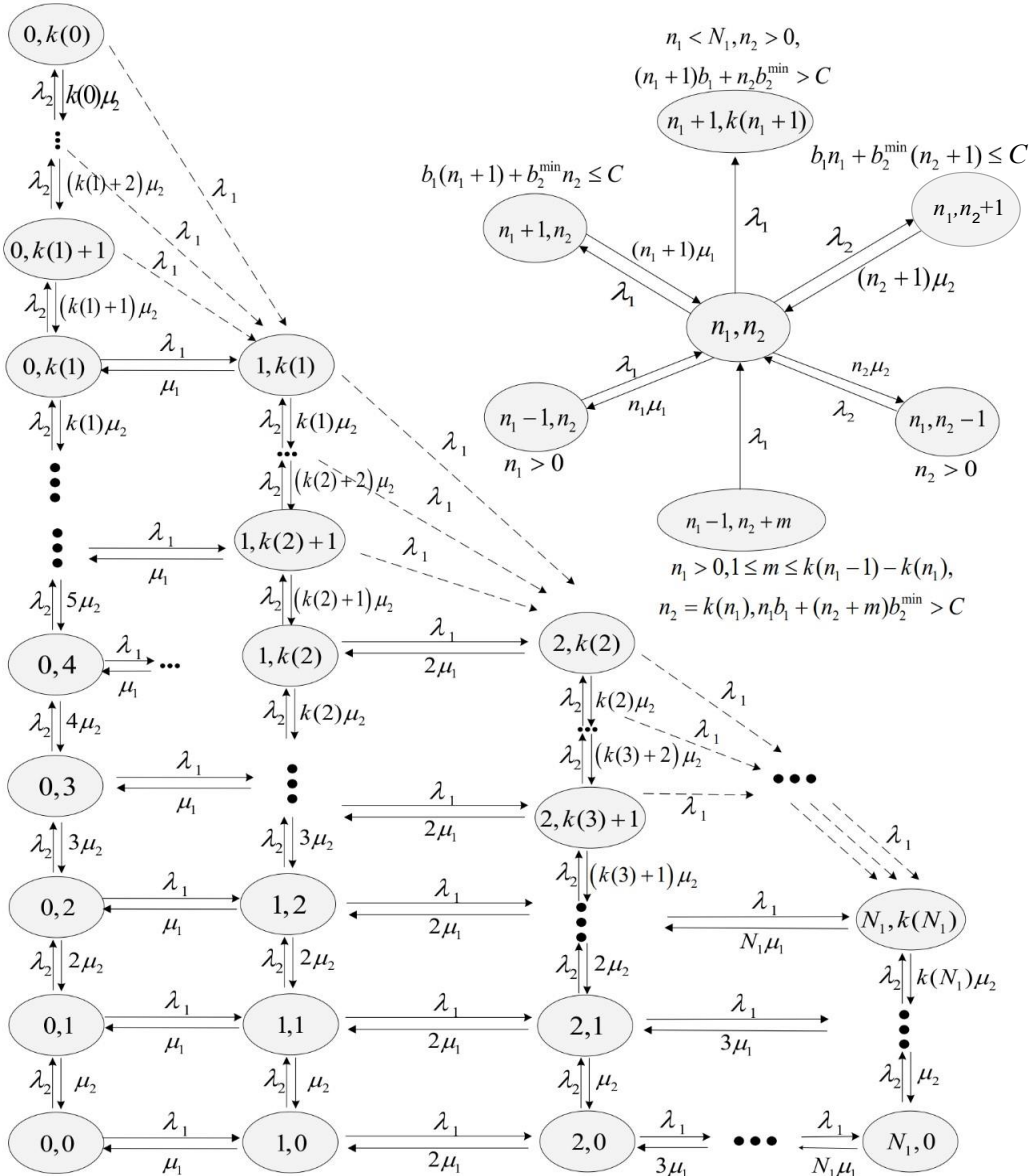


Рис. 8.3. Диаграмма переходов состояний марковского процесса $X(t)$

В формуле (8.5) символом * обозначена следующая формула:

$$\begin{aligned} * = & -[\lambda_1 I\{n_1 < N_1, b_1(n_1 + 1) + b_2^{\min} n_2 \leq C\} + \\ & + \lambda_1 I\{n_1 < N_1, n_2 > 0, b_1(n_1 + 1) + b_2^{\min} n_2 > C\} + \\ & + \lambda_2 I\{n_2 < N_2, b_1 n_1 + b_2^{\min}(n_2 + 1) \leq C\} + n_1 \mu_1 + n_2 \mu_2]. \end{aligned} \quad (8.6)$$

Найдя стационарное распределение $p(n_1, n_2)$, $(n_1, n_2) \in X$, можно вычислить следующие показатели эффективности системы.

Вероятность потери сессии URLLC

$$B_1 = \sum_{i=0}^{k(N_1)} p(N_1, i). \quad (8.7)$$

Вероятность потери сессии eMBB

$$B_2 = \sum_{i=0}^{N_1} p(i, k(i)). \quad (8.8)$$

Вероятность вытеснения сессии eMBB (вероятность того, что сессия eMBB будет прервана во время обслуживания):

$$\Pi = \sum_{i=0}^{N_1-1} \sum_{j=k(i+1)+1}^{k(i)} \frac{\lambda_1 p(i, j)}{\lambda_1 + \lambda_2 I\{j < k(i)\} + i\mu_1 + j\mu_2}. \quad (8.9)$$

Коэффициент использования ресурсов

$$U = C \sum_{i=0}^{N_1} \sum_{j=1}^{k(i)} (i + j) p(i, j) + b_1 \sum_{i=1}^{N_1} i p(i, 0). \quad (8.10)$$

8.2.3. Численный анализ

Результаты численного анализа получены для параметров системы, перечисленных в табл. 8.1. Чтобы оценить задержки передачи трафика URLLC, длительность кадра NR выбрана равной 1 мс. Для обеспечения надежной передачи данных в течение этого времени предполагается повторное кодирование, считая, что копия сообщения URLLC повторяется трижды в одном и том же кадре NR. Скорость передачи для сессии URLLC рассчитана с учетом этих повторов и при условии передачи 80 байт полезной нагрузки.

На рис. 8.4 показаны вероятность сброса сессии URLLC, вероятность прерывания сессии eMBB и коэффициент использования ресурсов в зависимости от плотности покрытия БС при минимальной скорости передачи для сессий eMBB 12,5 Мбит/с, их длительности 120 с и различных конфигураций антенных решеток на БС. Графики на рисунке показывают, что увеличение плотности покрытия БС приводит к резкому снижению вероятности прерывания сессии

eMBB. Однако для рассматриваемых значений параметров пороговая вероятность 10^{-6} потери сессии URLLC достигается только при высокой плотности БС $0,0008 \text{ 1/м}^2$ и использовании антенной решетки 32×4 . Вероятность прерывания сессии eMBB уменьшается при увеличении плотности БС, тем самым гарантируя качество обслуживания сессий URLLC. При этом значительно улучшается качество обслуживания сессий eMBB. Из графиков также видно, что использование антенной решетки 64×4 обеспечивает наилучшее использование ресурсов, что говорит в пользу применения передовых антенных решений на NR БС.

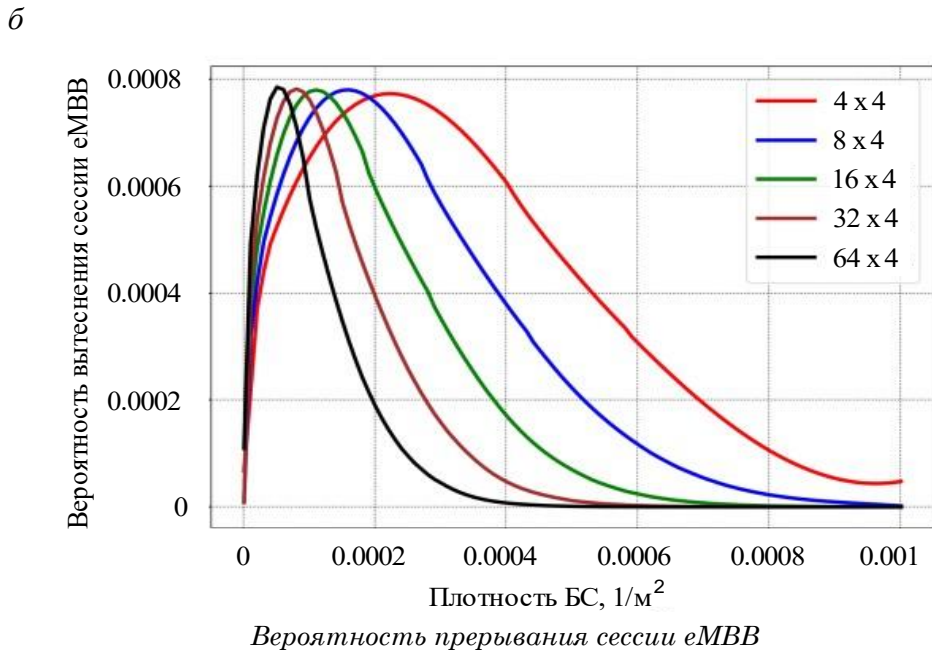
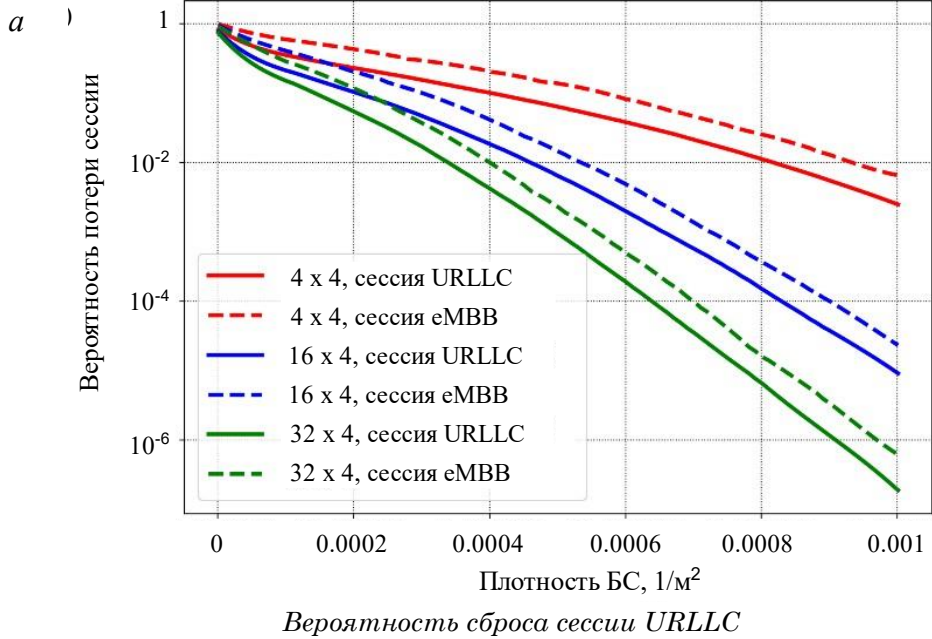
Таблица 8.1

Исходные данные для численного анализа

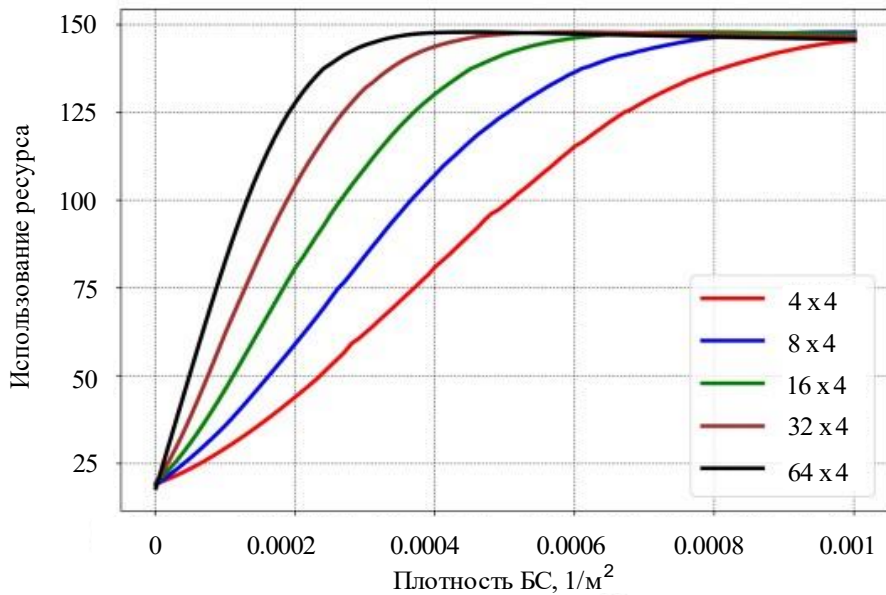
Параметр	Значение
Частотный диапазон	28 ГГц
Пропускная способность NR БС	50 МГц
Мощность передачи	0.2 Вт
Боковая антенная решетка NR БС	{4, 8, 16, 32, 64}x4
Боковая антенная решетка АУ	4x4
Высота NR БС, АУ	6 м, 1 м
Интенсивность потока сессий URLLC	0.114 с^{-1}
Интенсивность потока сессий eMBB	0.0005 с^{-1}
Средняя длительность сессий URLLC	1 мс
Средняя длительность сессий Embb	120 с, 6000 с
Скорость передачи для сессий URLLC	2 Мбит/с
Минимальная скорость передачи для сессий Embb	12.5, 25 Мбит/с

Перейдем к анализу поведения системы в случае большой длительности и минимальной скорости передачи для сессий eMBB, выбрав их значения равными 6000 с и 25 Мбит/с соответственно. Отметим, что примером такой длительности сессий могут служить сессии для постоянно включенных камер наблюдения. Результаты расчетов показаны на рис. 8.5 в зависимости от плотности покрытия БС для различных конфигураций антенных решеток. Сравнивая результаты на рис. 8.4 и 8.5, можно заметить, что обслуживание трафика URLLC практически не зависит от обслуживания трафика eMBB, что является следствием применения дисциплины приоритетного обслуживания с прерыванием сессий. Анализ поведения вероятности прерывания сессий eMBB показывает ее увеличение почти в два раза во всем рассматриваемом диапазоне плотности покрытия БС. Этот эффект является следствием достаточно высокой

нагрузки, создаваемой сессиями eMBB. Отметим также, что результаты использования ресурсов на рис. 8.5 качественно соответствуют тем, что показаны на рис. 8.4.



6

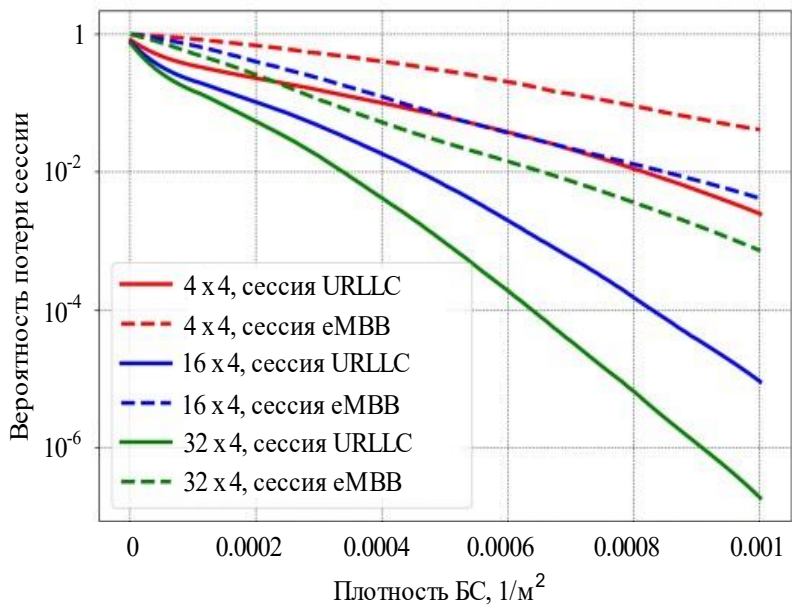


Коэффициент использования ресурсов системы

**Рис. 8.4. Вероятность сброса сессии URLLC (а);
вероятность прерывания сессии eMBB (б);
коэффициент использования ресурсов системы (в).**
Длительность сессий eMBB 120 с при минимальной скорости
передачи 12,5 Мбит/с

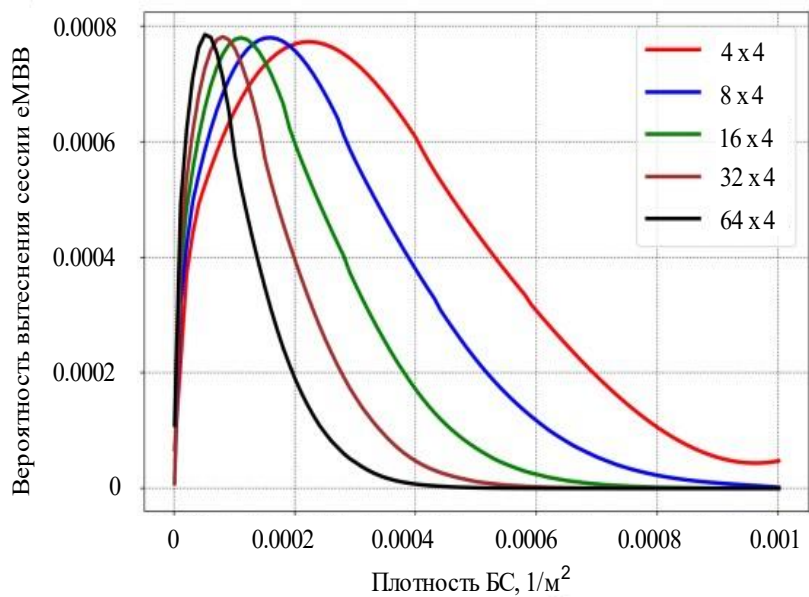
На рис. 8.6 исследуемые показатели рассматриваются в зависимости от интенсивности потока сессий URLLC для трех значений минимальной скорости передачи для сессий eMBB, плотности покрытия БС $5 \times 10^{-4} 1/m^2$, антенной решетки 16×4 и длительности сессий eMBB 120 с. Из графиков видно, что увеличение интенсивности поступления сессий URLLC незначительно влияет на вероятность их потерь, в то время как вероятность прерывания сессий eMBB резко возрастает.

а

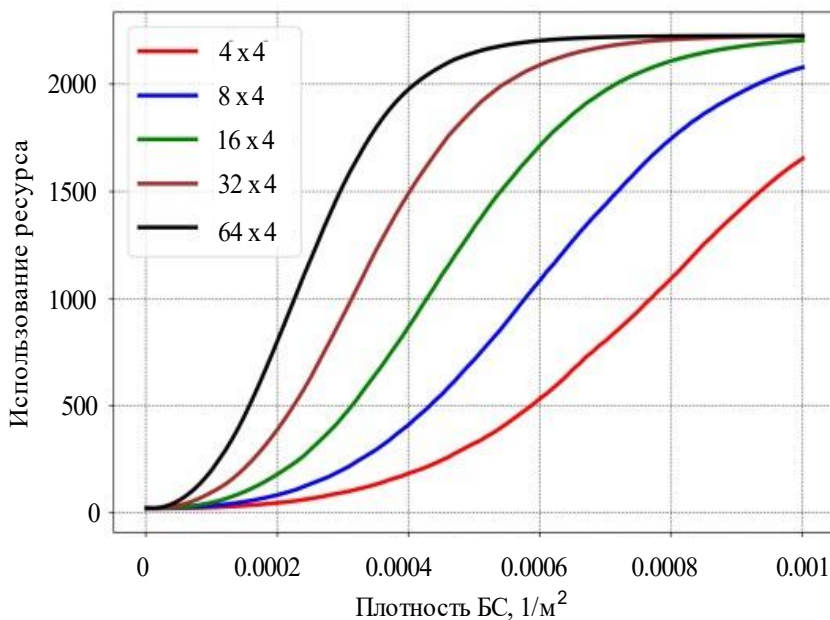


Вероятность сброса сессии URLLC

б



Вероятность прерывания сессии eMBB



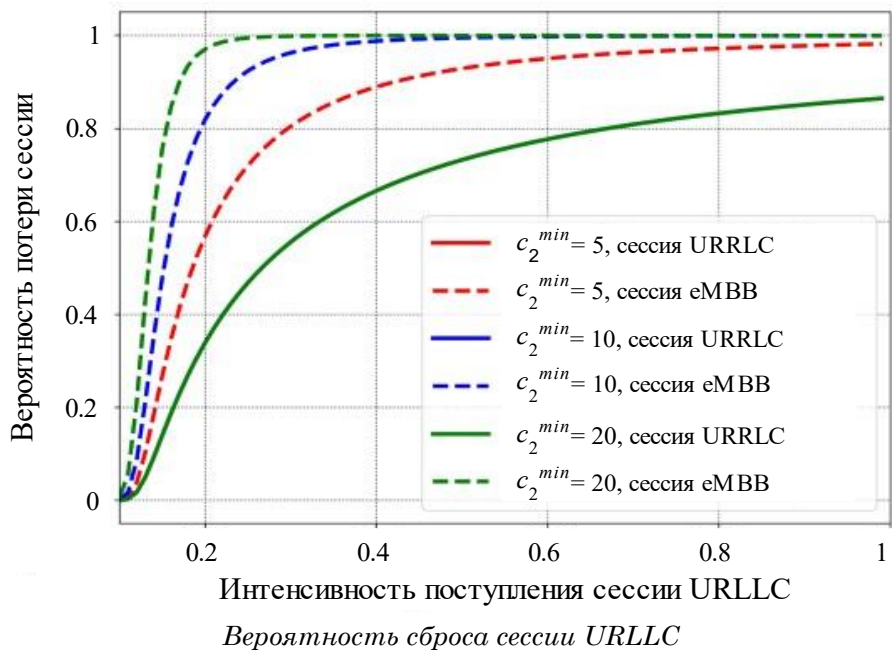
Коэффициент использования ресурсов системы

Рис. 8.5. Вероятность сброса сессии URLLC (а); вероятность прерывания сессии eMBB (б); коэффициент использования ресурсов системы (в). Длительность 600 с при минимальной скорости передачи сессий eMBB 25 Мбит/с

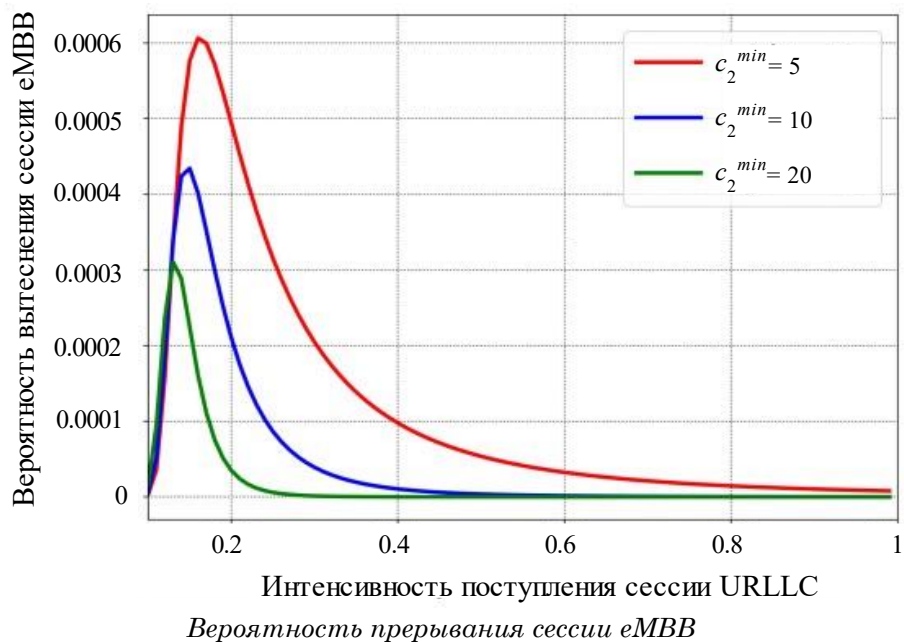
8.3. Системы с кэшированием

Промышленные варианты применения IoT обычно имеют большие масштабы и оригинальные конструкции, в которых достаточно сложно обеспечить бесперебойную работу беспроводных систем связи. Поколение 5G мобильных технологий является достаточно радикальной инновацией, так как использует не только микроволновый, но и миллиметровый диапазоны частот с помощью технологии NR. Хотя миллиметровые каналы и предлагают более высокую пропускную способность, они подвержены многим побочным эффектам, которые необходимо принимать во внимание. Одним из самых важных является динамическая блокировка путей распространения радиоволн подвижными препятствиями (рис. 8.7). Поэтому критически важные данные будущих промышленных приложений IoT будут передаваться на высокой скорости, но без дополнительных решений доступность и надежность доставки данных могут оказаться невозможными.

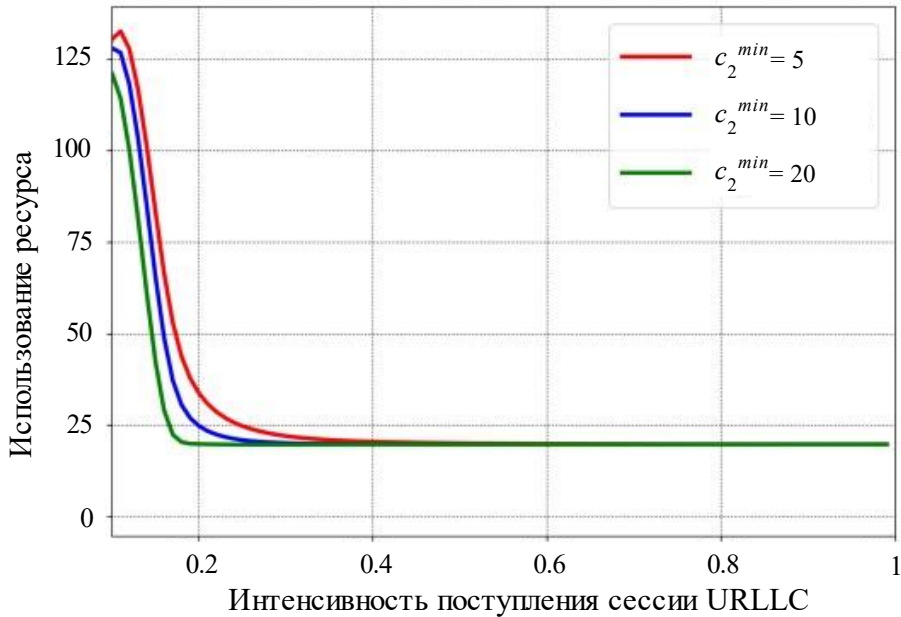
а



б



в



Коэффициент использования ресурсов системы

Рис. 8.6. Вероятность сброса сессии URLLC (а); вероятность прерывания сессии eMBB (б); коэффициент использования ресурсов системы (в). Плотности покрытия БС 5×10^{-4} 1/м², антенная решетка 16×4 и длительность сессий eMBB 120 с

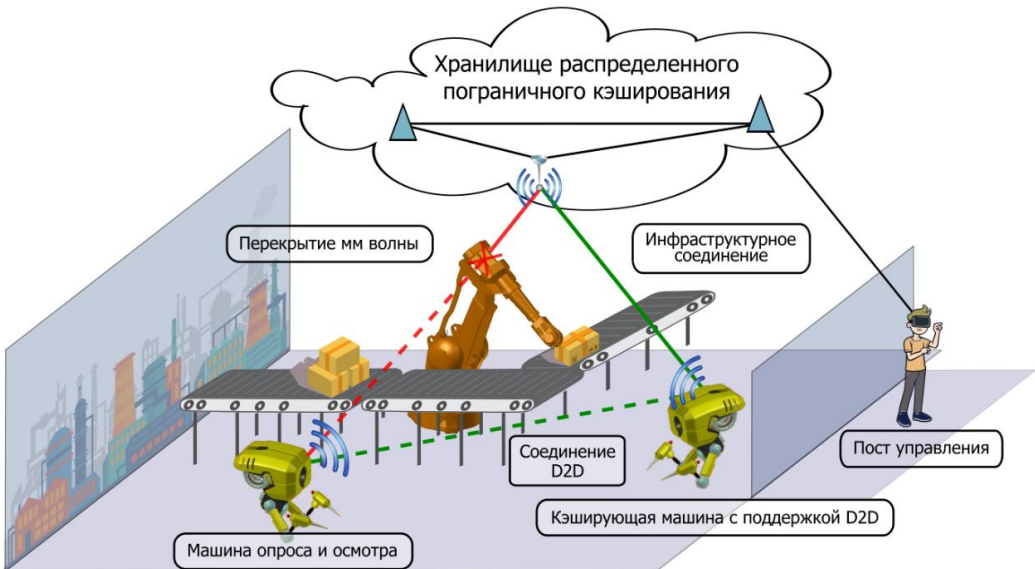


Рис. 8.7. Примеры соединений в промышленных применениях сетей 5G

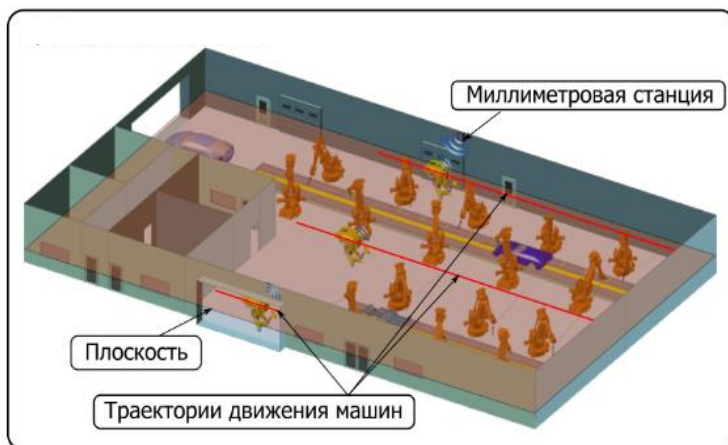
Автоматизация промышленности становится все более важной задачей в секторах производства, где требуются поэтапный сбор, тестирование или упаковка различных продуктов, в том числе автомобилестроение, бытовая электроника и др. Автоматизированные заводы и фабрики нуждаются в проведении регулярных осмотров, которые требуют быстрой и надежной передачи большого количества информации во избежание перебоев в работе. Автоматизированное производство обычно предполагает взаимодействие с множеством управляемых централизованно (например, с помощью программируемого логического контроллера) датчиков и актуаторов, расположенных в сети. Переход от надежных каналов связи проводных технологий к новым беспроводным технологиям накладывает дополнительные требования на качество обслуживания трафика на участке радиодоступа.

Одним из вариантов проведения регулярных осмотров является использование роботов. Они наблюдают за производственными объектами при помощи камер, могут анализировать информацию и сообщать о состоянии объектов операторам, находящимся в центре управления. Эти машины действуют автономно, либо ими может дистанционно управлять оператор с помощью оборудования AR/VR, создавая весьма чувствительные к пропускной способности и задержке передачи видеопотоки. Для доставки такого трафика от машин к оператору рекомендуется использовать услугу URLLC. Помимо этого, роботы могут быть объединены в команды, которые совместно собирают необходимые данные (см. рис. 8.7). Однако, чтобы эти операции выполнялись бесперебойно, необходимо наличие надежной высокоскоростной связи.

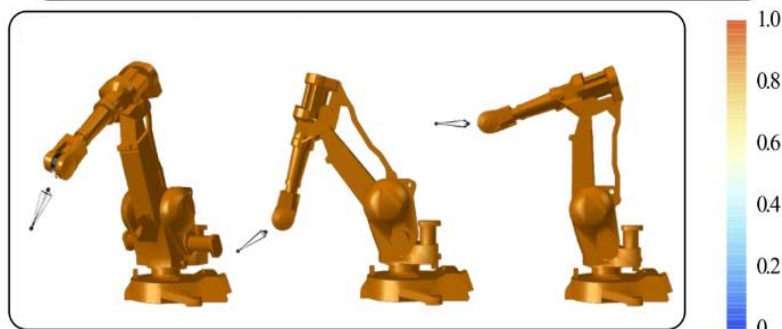
8.3.1. Прогнозируемый выбор режима с учетом мобильности

В описанных выше условиях возможны два способа передачи данных: а) инфраструктурное соединение, при котором движущееся устройство передает информацию сразу в пограничную сеть через станцию миллиметрового диапазона, которая может быть размещена, например, на стене или потолке в помещении завода, б) прямые соединения D2D, когда устройство пересылает информацию ближайшему IoT-устройству в ситуациях, когда первый способ передачи недоступен. Одна из основных причин его недоступности – блокировка LoS движущимися объектами. Поэтому в таких условиях

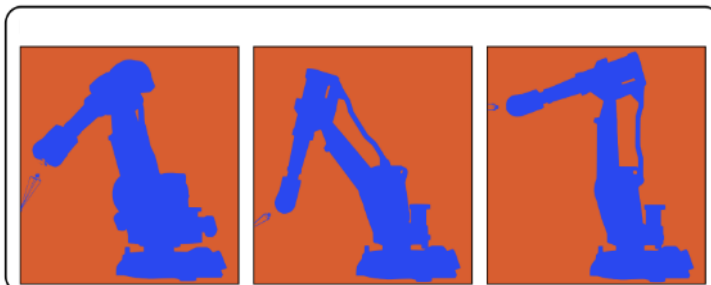
a



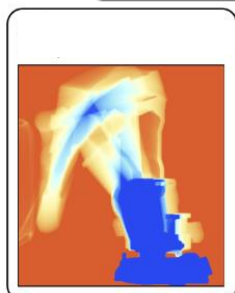
б



в



г



д

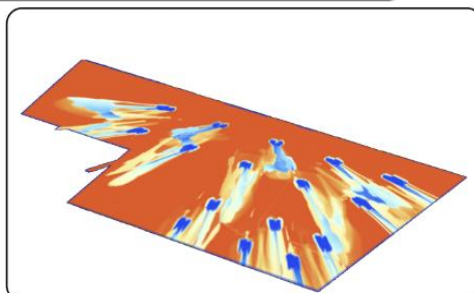


Рис. 8.8. Методология оценки динамической блокировки в промышленных развертываниях сетей миллиметрового диапазона частот: *a* – выделение плоскости; *б* – статическое отображение объектов; *в* – статическое отображение объектов с точки зрения путей видимости; *г* – динамическое отображение объектов; *д* – геометрическое преобразование

вероятность наличия LoS между передатчиком и приемником является ключевым показателем качества. Например, эта вероятность может быть использована для разработки политики выбора режима работы. Далее мы сначала воспользуемся программным обеспечением для 3D-моделирования условий распространения и рассмотрим метод, позволяющий оценить вероятность наличия прямой видимости как функцию от времени.

Рассмотрим инфраструктурное соединение, когда IoT-устройство обслуживается БС в миллиметровом диапазоне частот. Предлагаемый ниже подход к оценке вероятности блокировки дает возможность создания вероятностной карты путей LoS, которая позволяет количественно оценить вероятность наличия незаблокированного соединения от произвольной точки пространства к БС (рис. 8.8).

Ниже изложены основные этапы такой процедуры.

1. *Определение плоскости, в которой взаимодействуют устройства.* К рассматриваемой области предоставления услуги на заданной высоте добавляется горизонтальная плоскость контрастного цвета (см. рис. 8.8а).

2. *Статическое отображение объектов.* С помощью 3D-моделирования строится набор изображений, охватывающих все движения изучаемого IoT-объекта с желаемой степенью детализации во времени. На этом этапе важно правильно выбрать угол обзора, чтобы охватить всю интересующую область (см. рис. 8.8б).

3. *Статическое отображение объектов с точки зрения LoS.* Все полученные изображения переводятся в бинарную форму, используя цвет плоскости в качестве фона. Результатом являются необработанные карты путей видимости объекта (см. рис. 8.8в).

4. *Динамическое отображение объектов.* Полученные изображения объединяются, и полученные значения цвета пикселей нормализуются суммой максимальных значений для получения «карты прозрачности» объекта (см. рис. 8.8г).

5. *Геометрическое преобразование.* Проекция карты прозрачности объектов на плоскость, в которой взаимодействуют устройства для получения карты среды в пределах исходной системы координат (см. рис. 8.8д).

Полученное таким образом изображение затем используется для определения вероятности прямой видимости между движущимся

IoT-устройством, находящимся в произвольной точке рассматриваемой области, и обслуживающей его в миллиметровом диапазоне частот БС. Очевидно, что аналогичная оценка блокировки для каналов D2D требует более общего подхода для двух произвольно расположенных и движущихся устройств. Соответствующий процесс состоит из двух этапов. Во-первых, набор подготовительных шагов необходимо выполнить для обеспечения второго этапа вероятностными данными видимости для каждого из рассматриваемых объектов. Как только эффекты блокировки LoS в интересующей области определены количественно, на втором этапе для каждого типа устройств применяется следующая процедура.

1. *Составление карт с нескольких точек обзора.* С помощью 3D-моделирования создается набор изображений, которые фиксируют все возможные передвижения объекта. Камера устанавливается на выбранной высоте и направляется на интересующий объект. Также важно снимать объект под разными углами в горизонтальной плоскости с нужным шагом. Это облегчает вычисление вероятности блокировки пути прямой видимости между двумя произвольно расположенными мобильными IoT-устройствами. Примеры промежуточных изображений на этом этапе показаны на рис. 8.8в.

2. *Статическое отображение пути прямой видимости.* Каждый набор изображений обрабатывается на основе угла обзора в соответствии с шагами (2) и (3) предыдущей процедуры. Пример результирующей карты вероятности блокировки LoS для движущегося IoT-объекта показан на рис. 8.8б.

3. *Уточнение точки обзора.* Для каждого исследуемого объекта рассчитывается минимальный ограничивающий прямоугольник с учетом угла обзора.

Пример зависимости вероятности блокировки LoS от времени, рассчитанной по предложенной методике, показана на рис. 8.9 для прямолинейно и равномерно движущегося устройства в условиях развертывания, показанных на рис. 8.8а. Заметим, что даже в таком простом случае модели движения заметны значительные изменения в динамике блокировки. Рассматривая различные режимы доставки данных и используя знания о вероятности блокировки путей прямой видимости, которая изменяется в зависимости от фактических траекторий перемещения различных IoT-устройств, можно построить и

оценить стратегии распространения данных в промышленном сценарии развертывания услуг связи.

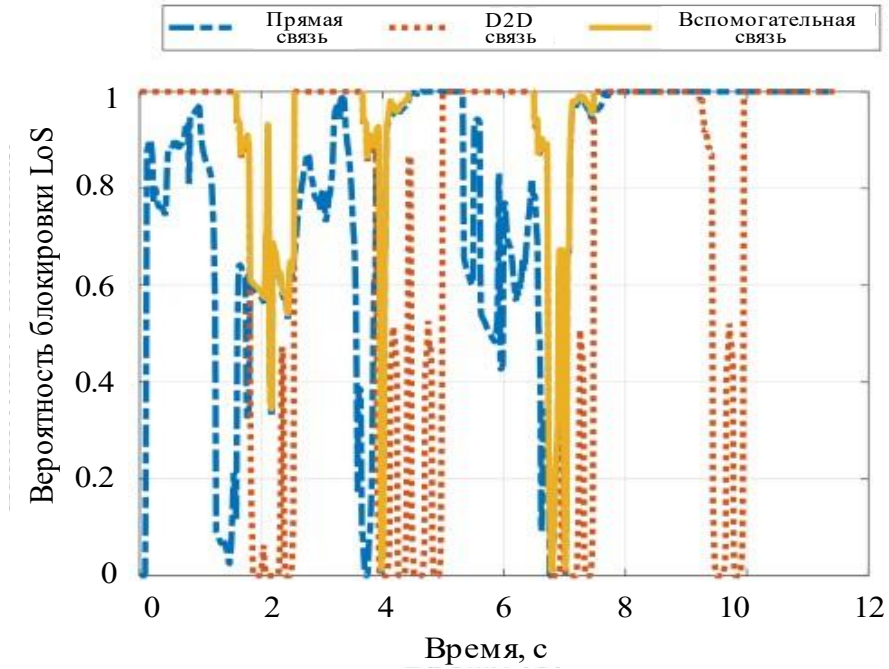


Рис. 8.9. Вероятность блокировки LoS

ГЛАВА 9

МЕТОДЫ УПРАВЛЕНИЯ РЕСУРСАМИ В СЕТЯХ NB-IoT

В данной главе рассматриваются механизмы и модели разделения ресурсов сетей, построенных по технологии NB-IoT (Narrow Band IoT), которая разработана для межмашинного взаимодействия, где данные не являются чувствительными к задержке передачи. Хотя NB-IoT была стандартизирована еще в процессе разработки сетей связи 4G, определено, что NB-IoT удовлетворяет также и требованиям МСЭ-Р к технологиям межмашинного взаимодействия для сетей 5G IMT-2020, определенным в рекомендации М.2410 МСЭ-Р.

Глава начинается с обзора в параграфе 9.1 технологий сотовой связи, которые могут быть использованы в качестве сетей доступа для IoT, а методы их анализа изложены в 9.2. Модель для анализа показателей производительности NB-IoT представлена в 9.3. Параграф 9.4 посвящен использованию и анализу методов ретрансляции в NB-IoT. Анализ моделей из параграфов 9.3 и 9.4 проведен в параграфе 9.5. Наконец, численные примеры производительности технологии NB-IoT представлены в заключительном параграфе 9.6.

9.1. Технологии сотовой связи для IoT

Низкое энергопотребление, доступная цена чипсета, большая дальность связи и возможность массовых развертываний устройств являются основными требованиями к новым решениям IoT, которые в совокупности можно назвать энергоэффективной сетью дальнего радиуса действия (LPWAN, Low-Power Wide-Area Networks). Разработка этих технологий была начата и развивалась параллельно в нескольких организациях по стандартизации, в основном в IEEE, ETSI и 3GPP, а также в рамках отдельных компаний и небольших промышленных консорциумов. Две из самых популярных технологий дальней связи LPWAN, основанных на запатентованных стандартах, это SIGFOX и LoRaWAN. В то время как оба варианта работают в нелицензируемых частотных диапазонах (ISM, Industrial Scientific and Medical) и топологически привязаны к сотовой инфраструктуре, принципы их работы принципиально различаются.

Технология SIGFOX работает с узкополосными каналами 100 Гц и 600 Гц для восходящей и нисходящей линии связи соответственно. Достижимая скорость передачи по восходящей линии связи

составляет около 100 бит/с, и, хотя полезная емкость кадра не может превышать 12 байт, эфирное время каждого кадра составляет около 1 секунды. Другими ограничениями SIGFOX являются использование всего одной схемы модуляции и кодирования (MCS, Modulation and Coding Scheme) и ограничение на количество пакетов, которые могут быть переданы по восходящей и нисходящей линии связи за один день – 140 и 4 на одно устройство соответственно. Несмотря на то, что устройства SIGFOX отличаются низким энергопотреблением и длительным сроком службы батареи, присущие им ограничения сужают спектр использования данной технологии для будущих приложений межмашинного взаимодействия.

В отличие от SIGFOX LoRaWAN поддерживает несколько схем модуляции и кодирования, результатом чего является увеличение скорости передачи до 50 кбит/с, что значительно расширяет диапазон приложений. Несмотря на то, что сети LoRaWAN, так же как и сети SIGFOX, уже развернуты и реализованы операторами связи, промышленная поддержка этих технологий со стороны телекоммуникационного сообщества довольно низка, так как ни одна из технологий не была глобально стандартизирована на международном уровне.

Наоборот, семейство технологий IoT 3GPP в основном ориентировалось на масштабные сценарии с самых ранних этапов их развития. Во-первых, уже в LTE Release 12 была определена упрощенная категория AU – Cat.0 с пропускной способностью 20 МГц и пиковыми скоростями передачи до 1 Мбит/с. У этой категории устройств примерно на 50% меньше логическая сложность по сравнению с Cat.1 AU. Новая категория пользовательских устройств, Cat. M1, была разработана для работы с полосой частот 1,4 МГц и позволила уменьшить максимальную мощность передачи до 20 дБм, при этом все еще обеспечивая существенное покрытие. Эта категория является частью LTE Release 13 и в несколько раз менее сложна, чем Cat.1 UE. Кроме того, для развивающихся стран с низким уровнем проникновения LTE было разработано основанное на GSM решение IoT под названием Extended Coverage GSM (EC-GSM), которое может обеспечить емкость соты до 50 тысяч устройств.

Согласно последним разработкам 3GPP для сотового IoT (CIoT, Cellular IoT), была проведена работа над стандартом сотовой связи для устройств телеметрии с низкими объемами обмена данными NB-IoT, как решение, которое может быть легко интегрировано в существующие сети. Полоса пропускания дополнительно сужается

до 180 кГц, обеспечивая дополнительный бюджет в 20 дБ. Процесс стандартизации NB-IoT был завершен в июне 2016 года, и технология официально включена в LTE Release 13.

В NB-IoT полоса пропускания по-прежнему достаточна для обеспечения возможности соединения с тысячами подключенных устройств. Полоса 180 кГц разделена на 12 поддиапазонов по 15 кГц, каждый с общей скоростью передачи до 250 кбит/с в восходящем и до 227 кбит/с в нисходящем канале. Это существенное масштабирование может обеспечить покрытие до 10 км в городе и до 30 км в сельских районах, что на порядок больше, чем в «ориентированном на человека» LTE. Меньшая полоса частот канала связи также способствует упрощению радиоэлектроники в устройствах конечного пользователя, что ведет к снижению затрат на оборудование пользователя. По той же причине размеры транспортных блоков как в восходящем, так и в нисходящем каналах уменьшаются и упрощаются процедуры сигнализации. Предполагается, что абонентское оборудование NB-IoT будет значительно дешевле, чем оборудование LTE, и по цене будет конкурировать с технологиями LoRaWAN и SIGFOX. Таким образом, NB-IoT является самой многообещающей технологией массового развертывания услуг IoT.

Далее рассмотрены стратегии разделения ресурсов при развертывании технологии NB-IoT на сетях LTE, а также технологические аспекты взаимодействия транспортных средств и датчиков, установленных оператором сети. Также оценены преимущества использования механизмов ретрансляции для сетей NB-IoT.

9.2. Методы анализа NB-IoT

Комбинация двух или более сетей радиодоступа увеличивает надежность системы. Например, сигнал об обнаружении чрезвычайных ситуаций, сообщаемый группой датчиков, может быть сразу проверен по визуальному каналу. Кроме того, в таких комбинациях существует ряд задач, в которых надежность доставки потоков данных играет ключевую роль. Учитывая ограниченный объем радиоресурсов, выбор стратегии их разделения между высокоскоростными видеокамерами и большим количеством низкоскоростных датчиков является важной задачей для операторов беспроводных сетей.

В этой задаче необходимо сбалансировать вероятностно-временные характеристики (VBX), например вероятность потерь для отдельных потоков данных. Для этого далее построена аналитическая

модель для исследования ВВХ обработки трафика NB-IoT при наличии конкурирующего трафика LTE. Модель позволяет сравнить три стратегии разделения радиоресурсов – статическую, динамическую и динамическую с резервированием. В качестве сценария развертывания сети NB-IoT рассматривается внутриполосный режим LTE Release 13, где БС LTE разделяет спектр частот с технологией NB-IoT. Этот вариант был выбран потому, что 1) LTE и NB-IoT – единственные на момент решения, использующие один и тот же спектр частот, 2) метод получает наибольшую поддержку со стороны операторов связи и, как ожидается, станет основным стандартом для сбора больших данных в мире IoT. В модели задействована функция оператора, реализующего службу наблюдения с возможностями технологий LTE и NB-IoT.

Рассмотрим оператора, который управляет системой зонального наблюдения. Система представляет собой гибридное решение, состоящее из 1) некоторого числа устройств для видеомониторинга, 2) большого числа датчиков для обнаружения вторжений, пожаров, контроля температуры и давления воздуха и 3) системы обработки больших объемов данных, получаемых как с видеокамер, так и с датчиков.

Поскольку проводное и Wi-Fi-соединения не всегда доступны в районах наблюдения, передача трафика от камер наблюдения производится при помощи сотовой сети LTE. Предполагается, что датчики подключены по технологии NB-IoT. Поскольку оба решения работают в лицензированных полосах частот, поставщик услуг, ориентированный на надежность связи, будет приобретать определенный частотный ресурс у оператора сети, используя механизм сетевой нарезки (Network Slicing). Такой подход может обеспечить бесперебойное обслуживание независимо от того, в каком состоянии находятся оставшиеся частоты оператора.

Описанный сценарий развертывания предполагает совместное использование предоставленных ресурсов LTE между двумя типами конечных узлов – камерами наблюдения и датчиками NB-IoT, что ставит задачу эффективного разделения ресурсов. В связи с этим характеристики надежности обслуживания потоков двух типов трафика приобретают первостепенное значение, поскольку сервер обработки не сможет принимать обоснованные решения, если часть потока станет временно недоступна. Возможным примером может служить система обнаружения пожара, развернутая в определенных зонах, где любое ложное срабатывание является чрезвычайно

дорогостоящим. Таким образом, если сервер обработки данных получает предупреждение о пожаре от датчиков дыма/температуры, он сначала перекрестно проверяет его с помощью визуального канала (видеопоток от ближайшей камеры) и только затем передает сигнал аварийной эвакуации.

Для решения задач далее рассмотрим принципы построения модели, а затем перейдем к формализации стратегий разделения ресурсов.

9.2.1. Модель соты LTE

Рассмотрим сеть LTE с БС, расположенной в центре соты. Напомним, что радиус R_L покрытия соты LTE намного меньше радиуса R_N покрытия технологии NB-IoT. Набор доступных ресурсов будем измерять в единицах каналов NB-IoT и будем называть эту единицу базовым каналом. В восходящем направлении доступны C каналов, которые можно рассчитать как $C = cS$, где S – число ресурсных блоков (РБ) и c – число базовых каналов NB-IoT в одном РБ.

Входящий поток сессий LTE является пуассоновским с интенсивностью ν , длительность сессии распределена по экспоненциальному закону со средним $1/\mu$, а минимальный объем ресурсов, необходимых для установления соединения на каждом временном интервале восходящей линии, составляет d базовых каналов. Пусть $a = \nu/\mu$ – предложенная нагрузка LTE. Запросы на соединение NB-IoT образуют пуассоновский поток с интенсивностью λ и каждый запрос характеризуется экспоненциально распределенным временем обслуживания с параметром θ , а для сеанса NB-IoT требуется b базовых каналов. Интенсивность предложенной нагрузки на передачу блоков данных от NB-IoT обозначим $\rho = \lambda\theta$.

9.2.2. Стратегии разделения ресурсов

Стратегии разделения ресурсов между видеочамерами LTE и измерительными датчиками NB-IoT схематично изображены на рис. 9.1. Максимальное число базовых каналов, которые могут быть выделены для NB-IoT и LTE, обозначим $C_N = C - R_L$ и $C_L = C - R_N$ соответственно, где R_L и R_N – минимальное число каналов, которые всегда доступны и зарезервированы для трафика LTE и NB-IoT. Далее проведено сравнение следующих трех стратегий разделения ресурсов.

1. Статическая стратегия (STAT) соответствует случаю, когда минимальное и максимальное число базовых каналов, выделенных для NB-IoT и LTE, совпадают и $R_L = C_L$, $R_N = C_N$, т.е. все ресурсы полностью разделены между NB-IoT и LTE.

2. Динамическая стратегия (DYN) применима в случае, когда минимальный объем ресурсов не назначается NB-IoT и LTE, т.е. $R_L = R_N = 0$, но максимальный объем ресурсов полностью разделен между NB-IoT и LTE, т.е. $C_L = C_R = C$.

3. Динамическая стратегия с резервированием (DYNRES), при которой максимальный объем ресурсов, доступный NB-IoT и LTE, определен так, что $C_N = C - R_L > 0$ и $C_L = C - R_N$. При этом $R_N > 0$, $R_L > 0$ определяют минимальный гарантированный объем ресурсов для NB-IoT и LTE соответственно, а остальные ресурсы разделяются динамически.

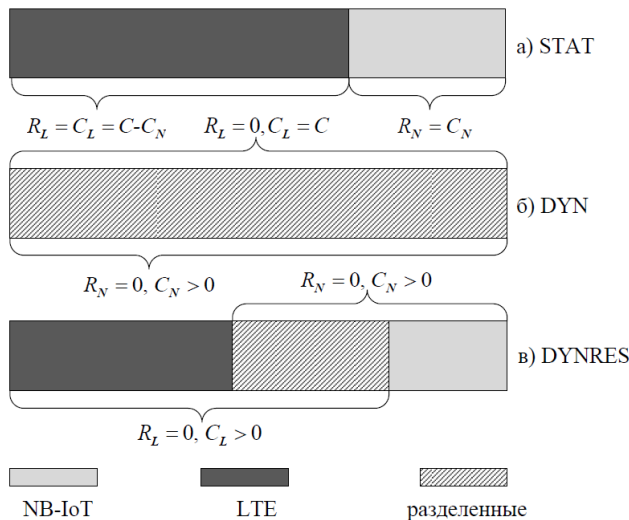


Рис. 9.1. Стратегии разделения ресурсов

Особенностью системы с двумя разнотипными входящими потоками является то, что базовые каналы последовательно выделяются с учетом особенностей внутрисполосной реализации технологии NB-IoT. Процесс управления ресурсами показан на рис. 9.2. Рассмотрим систему без активных сессий и предположим, что поступает новая сессия NB-IoT. В этом случае РБ становится доступным NB-IoT и каждый запрос на принятие данных требует ровно b базовых каналов. Таким образом, новая сессия NB-IoT занимает базовые каналы в этом РБ. Если при следующем установлении сессии NB-IoT все

каналы в данном РБ будут уже заняты, то будет выделен следующий РБ, доступный для услуг NB-IoT.

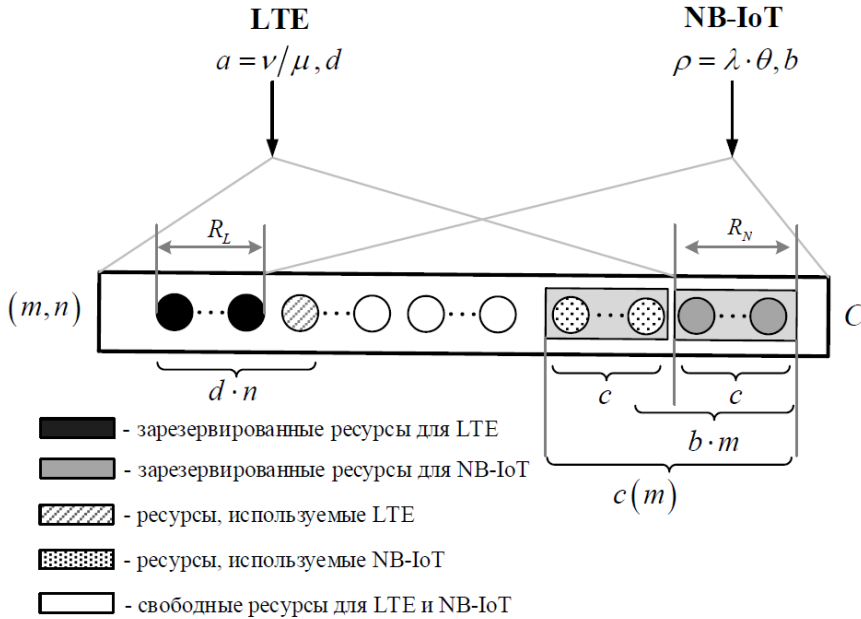


Рис. 9.2. Схема разделения ресурсов соты сети LTE

9.3. Модель разделения ресурсов

Модель разделения ресурсов может быть формализована в виде СМО с двумя входящими потоками, имеющими разные дисциплины обслуживания с выделением ресурсных блоков. Отметим, что система позволяет управлять потерями сессий как NB-IoT, так и LTE. Основное отличие модели заключается в последовательной стратегии последовательного разделения ресурсов.

Опишем пространство состояний системы. Запрос на обслуживание от LTE принимается, если в момент его поступления доступно не менее d из $C_L = C - R_N$ каналов, т.е. уменьшается число доступных базовых каналов на величину d и число РБ на величину d/c , причем число сессий LTE d/c всегда должно являться целым $d/c \geq 1$. Если при принятии запроса на соединение от NB-IoT число базовых каналов в текущем РБ больше, чем b , то сессия NB-IoT принимается в этом РБ. Если число доступных базовых каналов ниже, чем b , и новый РБ доступен для трафика NB-IoT, то этот РБ принимает на обслуживание поступивший запрос от NB-IoT, тем самым уменьшая

число базовых каналов в этом РБ на число b . В других случаях сессия NB-IoT теряется.

Пусть $m(t)$ и $n(t)$ определяют число активных сессий NB-IoT и LTE соответственно в момент t . Тогда состояние системы может быть описано СП $\{m(t), n(t), t > 0\}$ с пространством состояний

$$\mathcal{X} = \{m \geq 0, n \geq 0: nd \leq C - R_N, c(m) \leq C - R_L, nd + c(m) \leq C\}, \quad (9.1)$$

где $c(m) = cbm/M$ – число базовых каналов, занятых сессиями NB-IoT, M – максимальное число сессий NB-IoT, которые могут быть обслужены в одном РБ, $M = c/b$ – число базовых каналов, необходимых для обслуживания этих сессий.

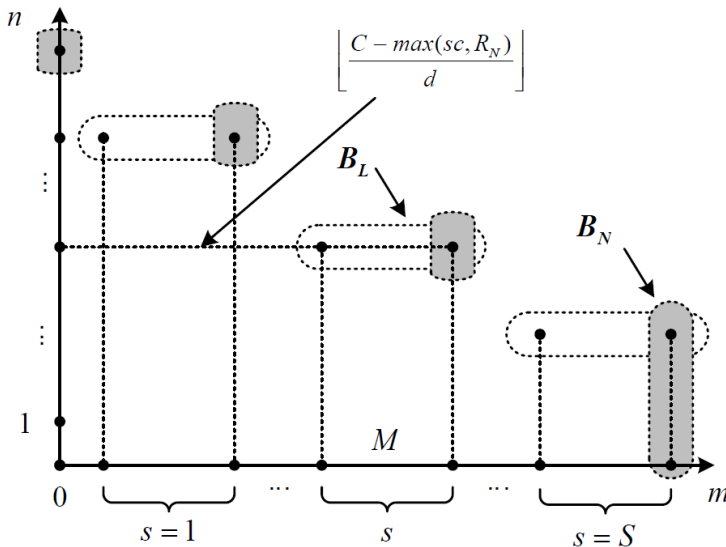


Рис. 9.3. Разбиение пространства \mathcal{X} на макросостояния

Заметим, что $\{m(t), n(t), t > 0\}$ по построению является МП и введем его стационарное распределение

$$p(m, n) = \lim_{t \rightarrow \infty} p(m, n)(t), \{m, n\} \in \mathcal{X}, \quad (9.2)$$

которое, как нетрудно убедиться, удовлетворяет уравнениям локального баланса

$$\begin{aligned} p(m, n) \frac{c(m)}{\theta} &= p(m-1, n)\lambda, \quad m > 0, (m, n) \in \mathcal{X}, \\ p(m, n)\mu n &= p(m, n-1)\nu, \quad n > 0, (m, n) \in \mathcal{X}. \end{aligned} \quad (9.3)$$

Решение этих имеет вид

$$p(m, n) = G^{-1}(\mathcal{X}) \left(\frac{\rho}{Mb} \right)^m \left(\prod_{i=1}^m \left[\frac{i}{M} \right] \right)^{-1} \frac{a^n}{n!}, \quad (9.4)$$

с нормирующей константой

$$G(\mathcal{X}) = \sum_{(m,n) \in \mathcal{X}} \frac{\rho}{Mb} \left(\prod_{i=1}^m \left[\frac{i}{M} \right] \right)^{-1} \frac{a^n}{n!}. \quad (9.5)$$

9.3.1. Характеристики обслуживания

Для организации вычислений и снижения их трудоемкости предлагается рекуррентный алгоритм, основанный на разбиении пространства состояний \mathcal{X} на макросостояния:

$$\mathcal{X} = \cup_{s=0}^S \mathcal{X}_s, \quad (9.6)$$

где $\mathcal{X}_s = \{(m, n) \in \mathcal{X} : c(m) = sc\}$, как показано на рис. 9.3. На рисунке также показаны множества состояний блокировки сессий LTE и NB-IoT. Из формул для стационарного распределения (9.4) и (9.5) и разбиения (9.6) могут быть получены формулы для расчета вероятностей блокировок сессий LTE:

$$p_L = p \left(0, \left\lfloor \frac{C - R_N}{d} \right\rfloor \right) + \sum_{s=1}^{\left\lfloor \frac{C - R_L}{c} \right\rfloor} \sum_{m=[s-1]M+1}^{sM} p \left(m, \left\lfloor \frac{C - \max[sc, R_N]}{d} \right\rfloor \right), \quad (9.7)$$

и для вероятностей блокировок сессий NB-IoT в виде

$$p_N = \sum_{n=0}^{\left\lfloor \frac{R_L}{d} \right\rfloor} p \left(\left\lfloor \frac{C - R_L}{c} \right\rfloor M, n \right) + \sum_{s=\left\lfloor \frac{R_N}{c} \right\rfloor}^{\left\lfloor \frac{C - R_L}{c} \right\rfloor - 1} \sum_{n=\left\lfloor \frac{C - (s+1)c}{d} \right\rfloor + 1}^{\left\lfloor \frac{C - \max(sc, R_N)}{d} \right\rfloor} p(sM, n). \quad (9.8)$$

Средние значения продолжительности сессии, числа базовых каналов для сессий NB-IoT и числа каналов, занятых сессиями LTE, могут быть вычислены по формулам

$$E[T_N] = \frac{\sum_{m=0}^{\left\lfloor \frac{C - R_L}{b} \right\rfloor} \sum_{n=0}^{\left\lfloor \frac{C - \max\left(\frac{m}{M}c, R_N\right)}{d} \right\rfloor} \left[\frac{m}{M} \right] p(m, n)}{\lambda(1 - p_N)}; \quad (9.9)$$

$$E[b_N] = M \sum_{m=0}^{\lfloor \frac{c-R_L}{b} \rfloor} \sum_{n=0}^{\lfloor \frac{c-\max(\frac{m}{M}c, R_N)}{d} \rfloor} \left\lfloor \frac{m}{M} \right\rfloor p(m, n); \quad (9.10)$$

$$E[b_L] = d \sum_{m=0}^{\lfloor \frac{c-R_L}{b} \rfloor} \sum_{n=0}^{\lfloor \frac{c-\max(\frac{m}{M}c, R_N)}{d} \rfloor} np(m, n). \quad (9.11)$$

9.3.2. Численный анализ

Исходные данные для численного анализа приведены в табл. 9.1.

Таблица 9.1

Данные для численного анализа

Описание	Значение
Число базовых каналов в соте LTE	100
Число базовых каналов в РБ	4
Число каналов NB-IoT	[0,...,100]
Число каналов LTE	[0,...,100]
Число каналов для сессий NB-IoT	1
Число каналов для LTE-сессий	4
Средняя скорость сессий NB-IoT	100 кбит/с
Средняя продолжительность сессий LTE	10 с
Интенсивность потока сессий NB-IoT	10 сессий в 1 мин
Интенсивность потока сессий LTE	1 сессия в 1 мин

На рис. 9.4 показаны графики вероятности блокировки сессии LTE в зависимости от общего числа подключенных датчиков NB-IoT для стратегии STAT, где общий объем радиоресурсов полностью разделен между трафиком сессий обоих типов. Для этой стратегии определен коэффициент разделения ресурсов γ , характеризующий объем ресурсов, зарезервированных для NB-IoT. Заметим, что вследствие полного разделения ресурсов между LTE и NB-IoT характеристики стратегии STAT зависят только от значения γ . Причиной является то, что фактически стратегия создает две независимые виртуальные беспроводные системы (слайсы) – одну для LTE, а другую для NB-IoT. Хотя с точки зрения реализации стратегия является самой простой, ее можно рекомендовать только для систем, где средняя общая нагрузка: 1) редко изменяется во времени, 2) известна заранее или 3) может быть точно предсказана. Такие свойства редко наблюдаются на практике.

Значения вероятности блокировки для стратегии DYN, где общий объем ресурсов доступен для обоих трафиков, сильно зависит от нагрузки трафика NB-IoT. Для малых значений числа подключенных устройств NB-IoT стратегия DYN превосходит стратегию STAT, так как позволяет переключить всю полосу частот в режим доступности для LTE в отсутствие трафика NB-IoT. Рост нагрузки трафика NB-IoT сопровождается увеличением значений вероятности блокировки для стратегии DYN, поскольку трафик NB-IoT начинает конкурировать за ресурсы с трафиком LTE. Наконец, при подключении большого числа устройств NB-IoT этот трафик начинает доминировать и в итоге вероятность блокировки будет выше, чем для стратегии STAT независимо от значения коэффициента разделения ресурсов. Стратегия DYN может быть рекомендована для систем с сопоставимой нагрузкой, создаваемой трафиком LTE и NB-IoT. В то же время стратегия остается приемлемым решением для динамичных систем, где нагрузка может сильно изменяться во времени.

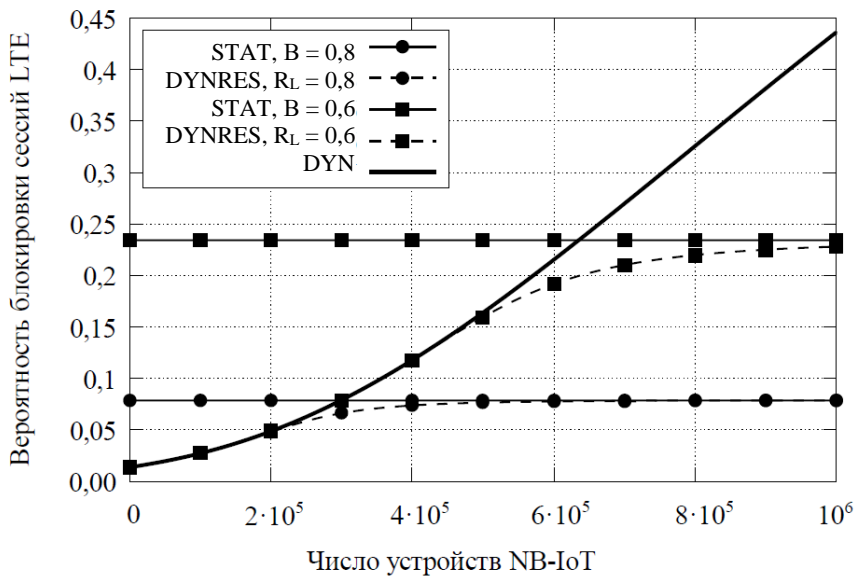


Рис. 9.4. Зависимость вероятности блокировки сессий LTE от числа устройств NB-IoT

Перейдем к третьей стратегии, DYNRES, где определенный объем ресурсов резервируется для трафика каждого типа, а остальные ресурсы динамически разделяются между LTE и NB-IoT. Значения R_L выбираются пропорционально соответствующим значениям γ ,

как в ранее рассмотренной стратегии STAT, а $R_N = 0$. Заметим, что для меньшего числа устройств NB-IoT показатели стратегии DYNRES практически совпадают с показателями стратегии DYN. При этом вероятности блокировок для DYNRES растут медленнее, чем для стратегии DYN. Причиной является то, что часть ресурсов, которые мог бы занять трафик LTE, занята динамическим разделением ресурсов. Кроме того, при более высоких нагрузках трафика NB-IoT вероятность блокировки сессий LTE для стратегии DYNRES никогда не превышает соответствующие значения для стратегии STAT. С увеличением числа устройств NB-IoT кривая для DYNRES асимптотически приближается к кривой для STAT, но не пересекает ее.

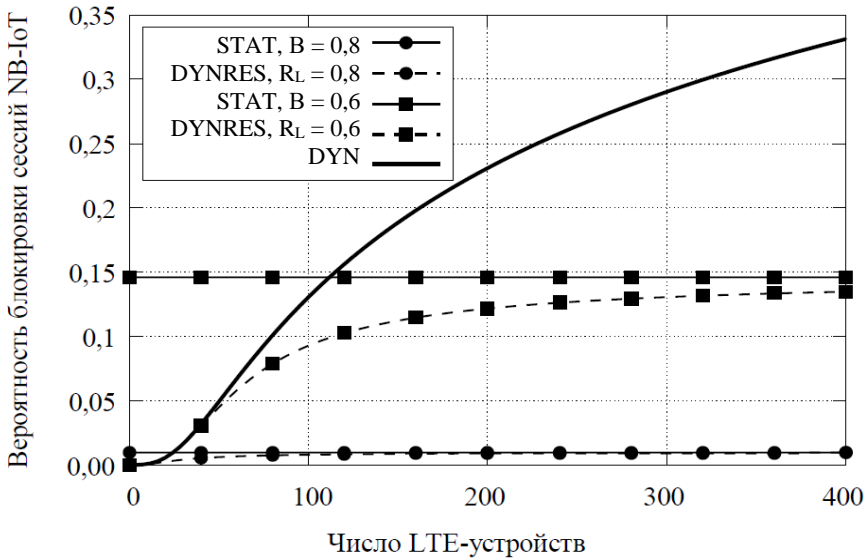


Рис. 9.5. Зависимость вероятности блокировки сессий NB-IoT от числа устройств LTE

Стратегия DYNRES показывает результаты не хуже, чем стратегии STAT и DYN. Это объясняется тем, что DYNRES сохраняет основные свойства этих стратегий. Стратегия DYNRES ведет себя так же, как и DYN при низких нагрузках и как стратегия STAT при высоких нагрузках. Следовательно, стратегия DYNRES может быть рекомендована для повышения надежности системы.

Графики на рис. 9.5 показывают вероятность блокировки сессии NB-IoT в зависимости от числа подключенных LTE-устройств. Как и в предыдущем случае, подберем параметры стратегии DYNRES так,

чтобы аналогичные ресурсы, зарезервированные для NB-IoT в STAT и DYNRES, оказались по объему равными друг другу. В этом случае значение R_N пропорционально $(1 - \gamma)$, а не γ , и $R_L=0$. Для такого симметричного случая на рис. 9.5 наблюдается то же качественное поведение кривых, что и на рис. 9.4. Тем не менее видны количественные различия – значения вероятности блокировки NB-IoT примерно на порядок ниже, чем для LTE. Это связано с более низкими потребностями в ресурсах у сессий NB-IoT, что ведет к увеличению вероятности принятия на обслуживание даже в почти полностью загруженной системе. Последнее обуславливается используемым способом выделения каналов NB-IoT. В частности, трафик NB-IoT получает полный РБ, хотя одна сессия и не занимает его полностью. Поэтому следующие несколько сессий NB-IoT гарантированно будут приняты на обслуживание, поскольку оставшаяся часть уже выделенного РБ не может быть занята каким-либо потоком LTE. Независимо от количественных значений трафика LTE система DYNRES всегда находится в выигрышном положении.

9.3.3. Оценка производительности NB-IoT

В этом разделе дана предварительная оценка производительности технологии NB-IoT по двум основным показателям производительности – вероятности потери сообщений и средней задержки передачи. Рассмотрен сценарий без ретрансляции, а сценарий с ретрансляторами, установленными на транспортных средствах, будет детально рассмотрен в параграфе 9.4, где также построена модель для анализа показателей эффективности NB-IoT в виде дискретной цепи Маркова (ЦМ). Эта же модель использована для вычислений, показанных на графиках рис. 9.6 и 9.7 в предположении, что ретрансляция отсутствует. Здесь и в параграфе 9.4 зона обслуживания БС представлена кругом радиуса R с равномерно распределенными в ней N устройствами NB-IoT, т.е. плотность распределения датчиков определяется формулой $\sigma = N/\pi R^2$ (на 1 км²). Интенсивность поступления запроса от одного датчика обозначена λ_t , а общая интенсивность поступления сообщений в слоте $\lambda_l = \lambda_t \Delta$, где Δ – длительность кадра LTE в секундах. Датчики передают сообщения размером 100 бит, и поэтому для передачи сообщения требуется один РБ NB-IoT.

Рассматривается сценарий, иллюстрированный на рис. 9.8. Предполагается, что датчик NB-IoT передает данные на БС, которые буферизируются, пока не освободится канал для их передачи. Как

только появляется возможность для передачи данных, БС выделяет один канал NB-IoT датчику и тот начинает передачу на мощности $P(d) = Ad^\gamma$, где d – расстояние от датчика до БС, A – константа, зависящая от частоты, и γ – экспонента затухания.

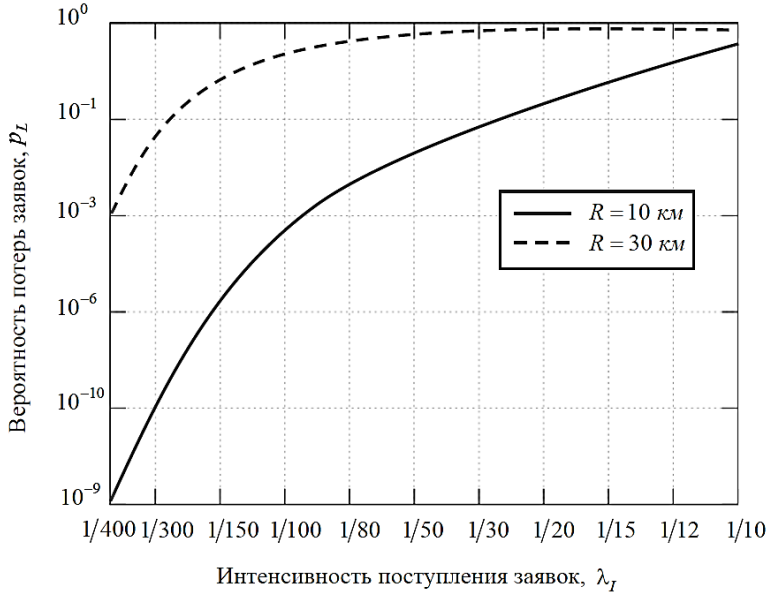


Рис. 9.6. Вероятность потери сообщения

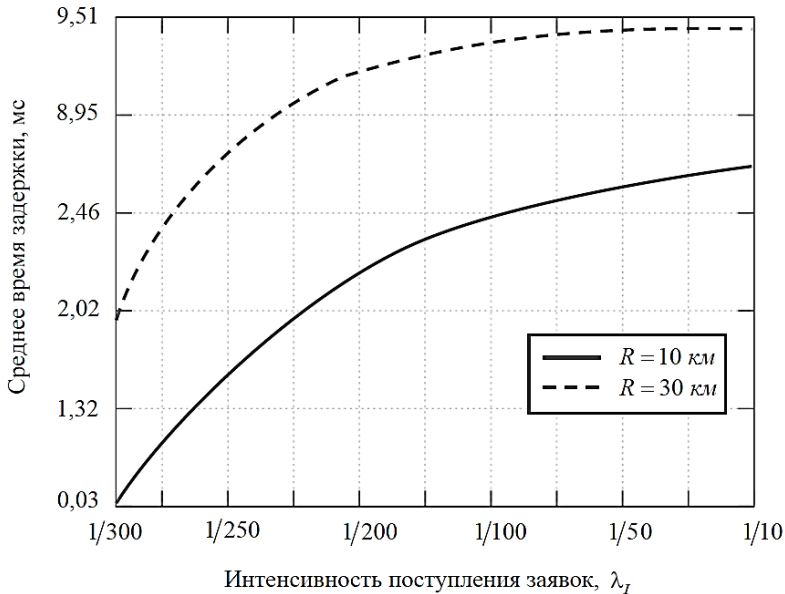


Рис. 9.7. Средняя задержка сообщения

Результаты численного анализа приведены на рис. 9.6 и 9.7 для сот радиусом 10 и 30 км. Из графиков видно, что технология NB-IoT может обеспечить надежную передачу данных (вероятность потери сообщений меньше 10^{-3}) для 100 подключенных датчиков на квадратный километр, каждый из которых генерирует до $\lambda_l = 1/300$ сообщений в час.

9.4. Использование ретрансляции в NB-IoT

9.4.1. Системная модель

Рассматривается сценарий возможности применения системы NB-IoT для приложений краудсорсинга с помощью ретрансляторов, установленных на транспортных средствах. Предполагается автономный режим работы системы NB-IoT, т.е. ее производительность не зависит от сети LTE. Модель учитывает особенности краудсорсинга на основе ретрансляторов и мобильность транспортных средств. Модель построена с целью анализа ориентированных на устройство показателей производительности – вероятность потери сообщения, задержка передачи, энергоэффективность и срок службы датчика (сенсора).

Отметим, что расширенная архитектура системы NB-IoT включает функции вспомогательных подвижных объектов, поэтому транспортные средства помогают им передавать свои данные на принимающую БС. Поскольку подключенные транспортные средства могут быть оснащены более совершенными радиомодулями и не столь ограничены с точки зрения потребления энергии, модель с ретрансляцией может потенциально привести к значительному улучшению показателей производительности.

Модель позволяет провести сравнение базового и расширенного решений NB-IoT и учитывает использование транспортных средств в качестве ретрансляторов для доставки данных IoT. Как показывает проведенный ниже численный анализ, массовое применение принадлежащих пользователю транспортных средств для приложений краудсорсинга позволяет существенно улучшить работу устройств IoT.

Итак, рассматривается зона обслуживания БС в виде круга радиуса R , в котором равномерно распределены N датчиков для двух сценариев – базового на рис. 9.8 и расширенного с ретрансляцией на рис. 9.9.

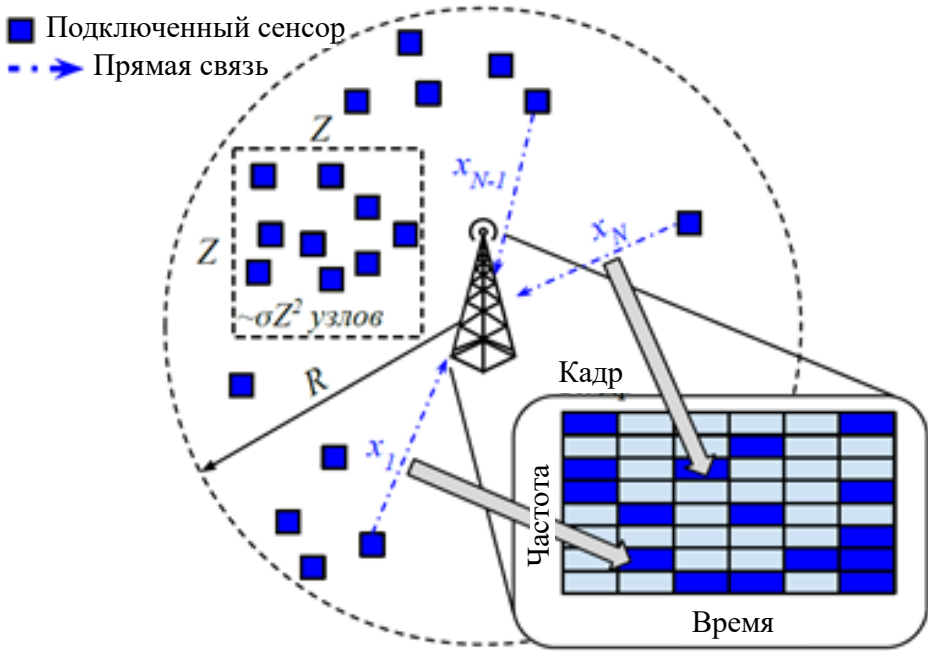


Рис. 9.8. Базовый сценарий применения NB-IoT

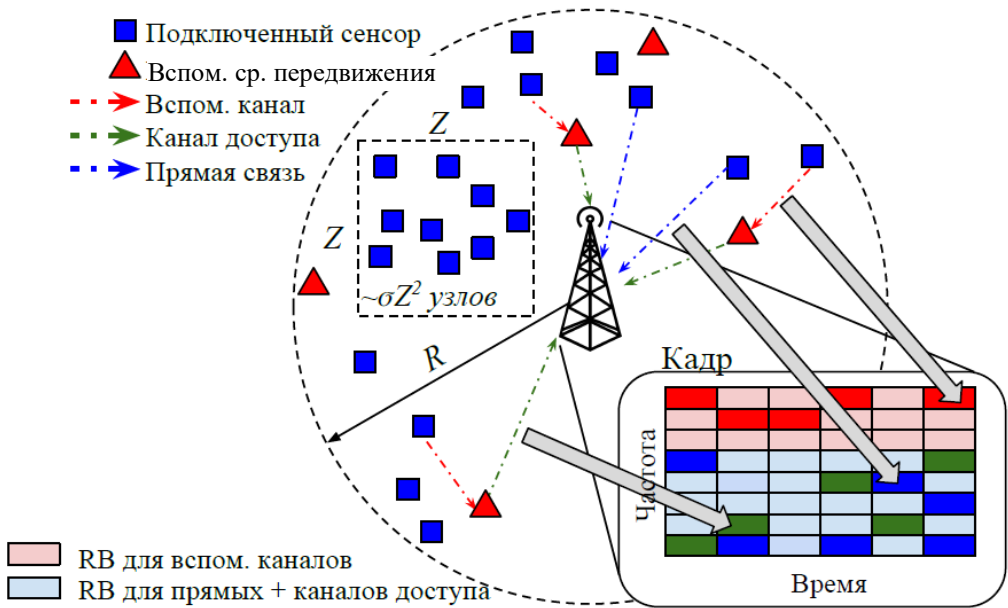


Рис. 9.9. Расширенный сценарий NB-IoT с ретрансляцией

Датчик, расположенный в зоне обслуживания, генерирует в случайный момент времени сообщение и, следуя стандартной процедуре случайного доступа, определенной спецификациями NB-IoT, уведомляет БС о необходимости передачи данных. Планировщик сети принимает запрос и направляет его в очередь для ожидания момента, когда освободится канал для передачи данных, а датчик переходит в режим ожидания передачи. Наличие буфера для очереди ожидания дает возможность контролировать максимальные задержки передачи, так как датчики крайне чувствительны к задержкам, особенно когда необходимо передать информацию критической важности. Это также позволяет контролировать расход энергии датчиком, который все время ожидания до начала передачи данных находится в режиме прослушивания канала. В случае отсутствия мест в буфере запрос на передачу сообщения считается потерянным. Как только появляется возможность для передачи данных, БС выделяет ресурсные блоки датчику и тот начинает передачу данных с заданной мощностью.

Преимущество использования подвижных объектов в расширенном сценарии заключается в том, что они могут быть расположены значительно ближе к датчикам, чем БС, что позволяет снизить мощность передачи и повысить энергоэффективность. Предполагается, что в зоне обслуживания доступны M подвижных объектов, которые, в отличие от основной стационарной БС, будем называть ретрансляционными БС.

Мобильность подвижных объектов моделируется процессом RDM с постоянной скоростью v и средним временем движения $E[\tau]$. В этом случае ретрансляционные БС функционируют как БС NB-IoT и также реализуют очередь. Чтобы уменьшить помехи, во-первых, между датчиком и ретрансляционной БС и, во-вторых, между передачами от ретрансляционной БС к основной БС, K_1 каналов выделены всем ретрансляционным БС и K_2 каналов – основной БС. При генерации сообщения датчик сканирует беспроводные каналы на БС, а выбор обслуживающей БС производится исходя из самого высокого уровня принимаемой мощности. Поскольку подвижные объекты являются мобильными, датчик не хранит информацию о предыдущих передачах, а выбор предпочтительного адресата выполняется заново каждый раз, когда необходимо передать новое сообщение.

Сравнивается производительность базовой и расширенной моделей для четырех показателей эффективности: а) вероятность того,

что сообщение будет потеряно из-за переполнения буферного накопителя (вероятность потери сообщений), б) количество энергии, необходимой для доставки сообщений датчиком, в) время жизни датчика, г) общая задержка передачи сообщения. Используемые обозначения приведены в табл. 9.2.

Таблица 9.2

Обозначения расширенной модели NB-IoT

Параметр	Описание
λ_l	Интенсивность генерации сообщений от датчика, мин^{-1}
σ	Плотность распределения датчиков в зоне обслуживания, км^{-2}
T_l	Время генерирования сообщения
$P(d)$	Мощность передачи
K	Общее число каналов системы NB-IoT
K_1, K_2	Число каналов, выделенных для ретрансляционных БС и для стационарной БС
γ_K	Доля каналов, выделенных подвижным объектам
L_1, L_2	Число мест в буферном накопителе на ретрансляционных БС и на стационарной БС
v	Скорость подвижных объектов
$E[\tau]$	Среднее время движения для модели RDM
$Q(n)$	Число сообщений в системе в момент окончания обслуживания
$A(n)$	Число сообщений, поступивших в слоте n
$B(n)$	Число потерянных сообщений в слоте n
$S(n)$	Число сообщений, переданных в слоте n
$C = [c_{ij}]$	Матрица переходных вероятностей ЦМ $Q(n)$
$\pi = [\pi_k]$	Стационарное распределение ЦМ
$p_{Loss,i}$	Вероятность потери сообщения из-за переполнения очереди i
q_B, q_R	Время обслуживания датчика на основной БС и на ретрансляционной БС
p_i	Вероятность приема сообщения в очереди i
Δ	Длительность кадра LTE
λ_i	Интенсивность потока сообщений в i -ю ретрансляционную БС
W	Случайное число тактов ожидания в очереди
T_l	Длительность состояния сна датчика
T_S, T_R, T_W, T_M	Длительности периода поиска датчиком БС, периода доступа к БС, периода нахождения сообщения в буфере датчика и периода передачи данных
P_{Rx}, P_{Tx}, P_l	Потребляемая мощность в состоянии готовности к передаче, состоянии передачи и состоянии сна, Дж/с
E_T, E_S	Энергия за транзакцию, затраченная в состоянии передачи и состоянии сна, Дж
U	Емкость батареи, мА·ч

9.5. Аналитическая модель

В этом параграфе в виде дискретной ЦМ построена модель для расширенного сценария использования системы NB-IoT. Начнем с исследования базового сценария, а затем мы модифицируем модель на случай применения подвижных БС.

9.5.1. Базовая модель

Напомним, что интенсивность пуассоновского потока сообщений от одного датчика равна λ_l и в зоне обслуживания имеется N датчиков, распределенных с плотностью σ . Заметим, что суммарный поток сообщений также является пуассоновским с интенсивностью $\lambda = N\lambda_l$.

В стандартах для NB-IoT определена выделенная радиочастотная полоса 180 кГц, которая может быть разделена на 12 диапазонов по 15 кГц, в каждом из которых имеется 6 каналов для передачи данных. Следовательно, в одном слоте, определяемом продолжительностью одного слота LTE, для NB-IoT выделяется $K = 6 \times 12 = 72$ виртуальных каналов, и в каждом слоте может быть передано до K сообщений. Модель построим в виде СМО в дискретном времени с групповым поступлением и групповым обслуживанием заявок (сообщений) с детерминированным временем обслуживания. Поскольку поток поступающих на систему сообщений является пуассоновским, то число сообщений, поступающих в слот, является с.в., распределенной по закону Пуассона с параметром $\lambda\Delta$, где Δ – длительность слота.

Время модели считаем дискретным с длительностью такта, равной длительности слота Δ . Будем считать, что состояния системы изменяются сразу после окончания слота в результате обслуживания пачки, а поступления пачек учитываются в момент, непосредственно предшествующий его окончанию. Вновь пришедшие заявки не могут поступить на обслуживание в слоте, в котором они пришли, а будут обслужены только в следующем слоте.

Введем ЦМ, $\{Q(n), n = 0, 1, \dots\}$, описывающую число заявок в системе после ухода пачки. Время ожидания передачи и время пребывания заявки в системе измеряется числом слотов, проведенных ею в системе. Заметим, что время ожидания соответствует верхней границе времени ожидания датчиком начала передачи.

Напомним, что размер буферного накопителя рассчитан на L сообщений. Поэтому, если в СМО поступает пачка из i заявок, застаюшая в системе k заявок, так, что $k + i > L$, то в систему принимаются только $L - k$ сообщений, а остальные заявки пришедшей пачки теряются. Так как поток сообщений является пуассоновским, то вероятность поступления k заявок за один слот вычисляется по формуле

$$A_k = \frac{(\lambda\Delta)^k}{k!} e^{-\lambda\Delta}, k = 0, 1, \dots \quad (9.12)$$

Нетрудно убедиться, что для введенной ЦМ справедливо следующее соотношение:

$$Q(n) = \min(Q(n-1) + A(n-1) - S(n-1), L), \quad (9.13)$$

где $S(n)$ – число обслуженных заявок; $A(n)$ – число поступивших заявок в слоте n .

Пусть $C = [c_{ij}]$, $i, j \in \{0, 1, \dots, L - K\}$ матрица переходных вероятностей ЦМ $Q(n)$, элементы которой вычисляются по следующим формулам:

$$c_{ij} = \begin{cases} \sum_{k=0}^{K-i} A_k, & 0 \leq i \leq K, & j = 0, \\ A_{K+j-i}, & 0 \leq i \leq K, & 0 < j < L - K, \\ 0, & K < i \leq L - K, & 0 \leq j < i - K, \\ A_{K+j-i}, & K < i \leq L - K, & i - K \leq j < L - K, \\ \sum_{k=L-i}^{\infty} A_k, & 0 \leq i \leq L - K, & j = L - K. \end{cases} \quad (9.14)$$

Введем стационарное распределение ЦМ

$$\pi_i = \lim_{n \rightarrow \infty} \Pr \{Q(n) = i\}$$

и обозначим $\boldsymbol{\pi} = (\pi_0, \pi_1, \dots, \pi_{L-K})$ вектор стационарного распределения числа заявок в СМО в момент времени после окончания слота, который может быть найден численно решением СУР $\boldsymbol{\pi}C = \mathbf{0}, \boldsymbol{\pi}\mathbf{e} = 1$, где \mathbf{e} – вектор из $L - K + 1$ единиц.

Вероятность потерь

Пусть $B(n) \in \{0, 1, \dots\}$ – число потерянных заявок в слоте n с распределением $f_B(i) = \Pr\{B(n) = i | A(n) \geq 1\}$, $i = 0, 1, \dots$, при условии, что в слот поступает по крайней мере одно сообщение.

Рассмотрим случай потери ровно i заявок в слоте n , когда в слоте $n - 1$ есть ровно $k = 0, 1, \dots, L$ заявок и число поступающих заявок в

слоте n равно $L - k + i$. Нетрудно убедиться, что распределение $f_B(i)$ может быть найдено по формуле

$$f_B(i) = \frac{\sum_{k=0}^{L-K} \pi_k A_{L-k+i}}{\Pr\{A \geq 1\}} = \frac{\sum_{k=0}^{L-K} \pi_k A_{L-k+i}}{1 - e^{-\lambda \Delta}}, \quad i = 1, 2, \dots, \quad (9.15)$$

где π_k , $k = 0, 1, \dots, L - K$, – вероятность того, что перед поступлением пачки в системе имеется ровно k заявок.

Заметим, что важный показатель – вероятность отсутствия потерь в слоте – можно вычислить по формуле $f_B(0) = 1 - \sum_{i=1}^{\infty} f_B(i)$, а вероятность потери сообщения имеет вид

$$p_{Loss} = \frac{1}{1 - e^{-\lambda \Delta}} \sum_{i=1}^{\infty} \sum_{k=0}^{L-K} \frac{i}{L-k+i} \pi_k A_{L-k+i} \quad (9.16)$$

где $i/(L - k + i)$ – вероятность того, что случайно выбранное сообщение будет потеряно, а в знаменателе (9.16) записана вероятность отсутствия поступления заявок в этом слоте.

Распределение времени ожидания в буферном накопителе

Рассмотрим сообщение от датчика, сгенерированное в слот n . Задержка сообщения в буфере $T_W = T_A + W + T_X$, измеряемая в слотах и состоящая из трех слагаемых, проиллюстрирована на рис. 9.10.

Пусть $W \in \{0, 1, \dots, L/K\}$ – с.в. числа полных слотов ожидания сообщением обслуживания в буфере в системе. Обозначим $f_W(i) = \Pr\{W(n) = i | A(n) > B(n)\}$ закон распределения дискретной с.в. W , который требуется найти:

$$f_W(0) = \sum_{k=0}^{K-1} \pi_k \sum_{l=1}^{K-k} A_l + \sum_{k=0}^{K-1} \pi_k \sum_{j=1}^{\infty} \frac{K-k}{K-k+j} A_{K-k+j}; \quad (9.17)$$

$$\begin{aligned} f_W(i) = & \sum_{k=iK}^{\min((i+1)K, L-K)-1} \pi_k \sum_{l=1}^{(i+1)K-k} A_l + \\ & + \sum_{k=iK}^{\min((i+1)K, L-K)-1} \pi_k \sum_{j=1}^{\infty} \frac{(i+1)K-k}{(i+1)K-k+j} A_{(i+1)K-k+j} + \\ & + \sum_{k=0}^{iK-1} \pi_k \sum_{l=1}^K \frac{l}{iK-k+l} A_{iK-k+l} + \\ & + \sum_{k=0}^{iK-1} \pi_k \sum_{j=1}^{\infty} \frac{K}{(i+1)K-k+j} A_{(i+1)K-k+j}. \end{aligned} \quad (9.18)$$

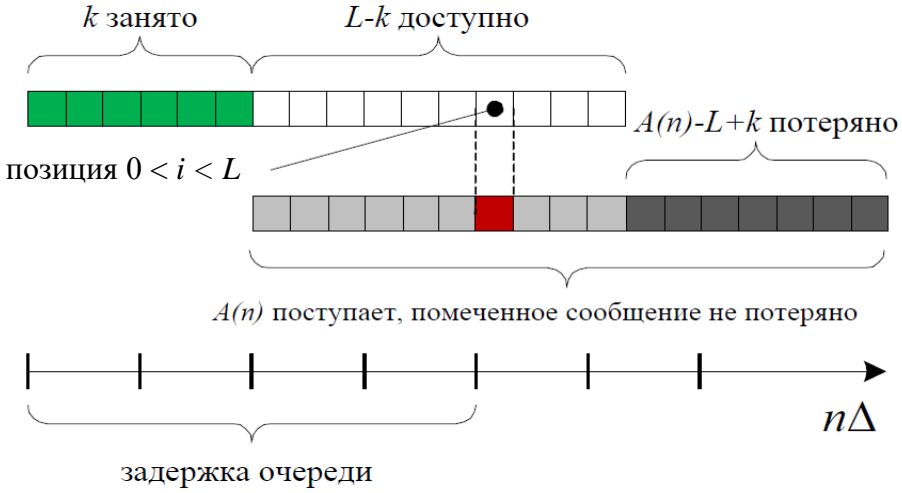


Рис. 9.10. Иллюстрация времени ожидания начала обслуживания

Заметим, что в модели задержки, измеряемые в слотах, округляются до ближайшего целого числа сверху. Чтобы учесть ошибки округления, рассмотрим компоненты T_X и T_A , проиллюстрированные на рис. 9.11. На рисунке показана величина T_X времени, прошедшего с начала слота, где происходит обслуживание сообщения во время передачи пачки. Поскольку диапазон NB-IoT составляет малую величину 180 кГц, можно считать эту величину постоянной.

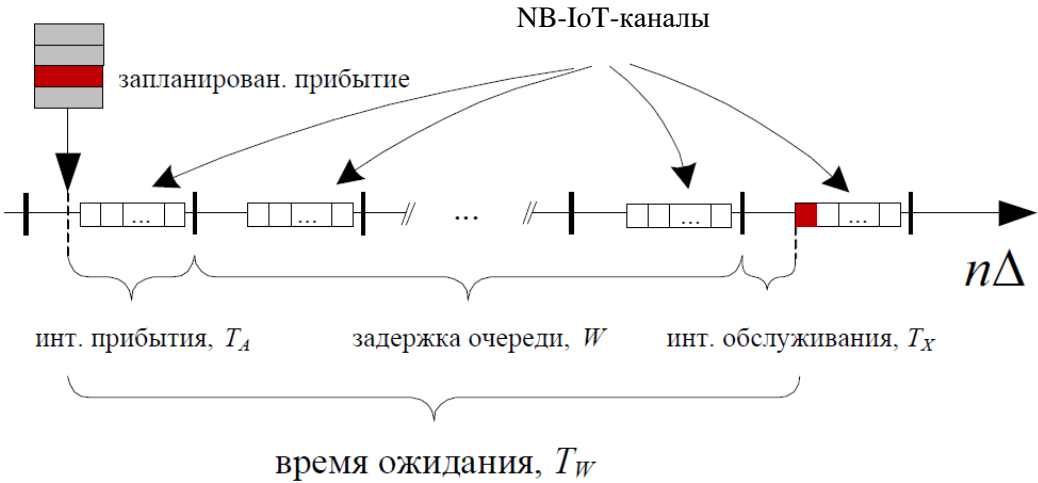


Рис. 9.11. Иллюстрация дополнительных компонент задержки

Вторым компонентом задержки является сдвиг момента поступления заявки на величину T_A , которая является непрерывной с.в., равномерно распределенной на интервале Δ . Тогда плотность распределения с.в. T_A имеет вид

$$f_{T_A}(x) = 1/\Delta, 0 < x < \Delta. \quad (9.19)$$

С учетом введенных с.в. окончательно получаем закон распределения времени ожидания заявки в очереди в виде дискретной свертки

$$f_{T_W}(x) = f_{T_A}(x) \sum_{i=0}^{\lfloor \frac{x}{K} \rfloor - 1} f_W(i). \quad (9.20)$$

Потребляемая датчиком энергия

Полученные формулы позволяют выразить распределение с.в. энергии, потребляемой датчиком для передачи сообщения (на одну транзакцию). На рис. 9.12 показаны состояния датчика и соответствующие уровни потребляемой мощности в течение так называемого «цикла обновления», представляющего собой длительность интервала между соседними моментами начала поиска БС и состоящего из периода сна между соседними транзакциями (передачами сообщения) и периода транзакции. Для простоты будем рассматривать три основных состояния датчика: а) состояние сна (I), б) состояние готовности к приему (Rx) и в) состояние передачи (Tx), при этом в периоде транзакции датчик будет находиться в одном из двух последних состояний.

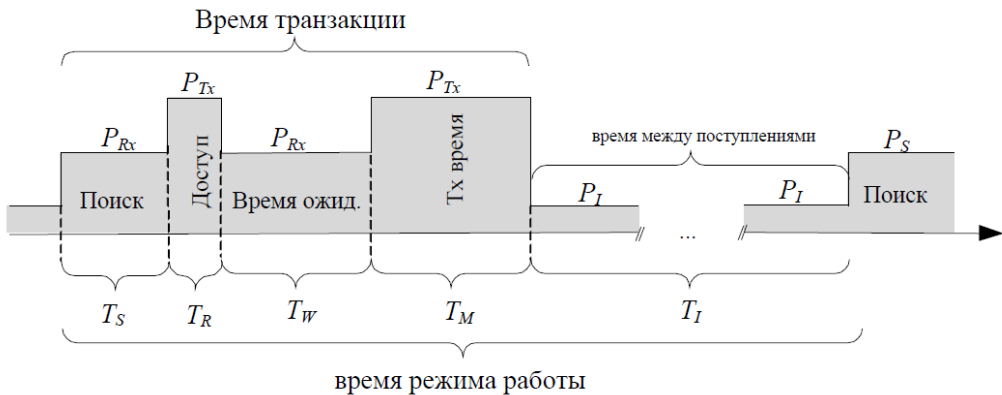


Рис. 9.12. Иллюстрация уровней потребляемой энергии для датчика NB-IoT

Для простоты предполагаем, что процедура случайного доступа в канал такова, что отсутствуют потери и дополнительные задержки. Это предположение обосновано тем, что процедура доступа в NB-IoT хорошо оптимизирована в процессе ее разработки. Оценим энергию E_T , затрачиваемую на транзакцию (передачу одного сообщения). Считаем, что уровень энергии P_{Tx} , необходимый для передачи сообщений, зависит от расстояния до БС, а требуемая для приема мощность P_{Rx} является постоянной величиной. Обозначая индексы величин T_* и P_* соответственными буквами, вводим следующие обозначения:

- поиск БС: постоянные величины T_S и P_{Rx} ;
- произвольный доступ: постоянная T_R и случайная P_{Tx} ;
- время ожидания: случайная T_W и постоянная P_{Rx} ;
- передача сообщения: постоянная T_M и случайная P_{Rx} ;
- состояние сна: случайная T_I и постоянная P_I .

Заметим, что энергия, необходимая для передачи сообщения, является с.в., состоящей из двух случайных компонентов и двух постоянных величин. Для нахождения времени жизни датчика необходимо также учитывать с.в. времени сна, которое определяется ФР $F_{T_I}(x)$ времени между моментами поступления сообщений. Таким образом, энергия, затраченная на одну транзакцию (на передачу одного сообщения, в Джоулях), может быть представлена в виде

$$E_T = \underbrace{P_{Rx}}_{\text{Конст.}} \overbrace{(T_S + T_W)}^{\text{Случ.}} + \underbrace{P_{Tx}}_{\text{Конст.}} \overbrace{(T_R + T_M)}^{\text{Случ.}}. \quad (9.21)$$

Плотность распределения $f_{T_S+T_W}(y)$ можно найти, используя формулы для линейного преобразования с.в. Если $Y = a + bX$ и $f_X(x)$ известны, то плотность $f_Y(y)$ имеет вид

$$f_Y(y) = f_X[g^{-1}(y)] \left| \frac{dx}{dy} \right| = f_X\left(\frac{y-a}{b}\right) \frac{1}{|b|}. \quad (9.22)$$

Напомним, что уровень энергии, потребляемый в состоянии T_x , зависит от расстояния D до БС. Будем использовать, как и ранее, модель распространения в линейной шкале $P_{Tx}(D) = AD^{-\gamma}$, где A – параметр, зависящий только от частоты, и γ – константа распространения. Тогда плотность распределения с.в. P_{Tx} можно получить как нелинейное преобразование $y = \phi(x)$ с.в. Y по формуле

$$f_Y(y) = f(\psi(y))|\psi'(y)|, \quad (9.23)$$

где $\psi(y) = \phi^{-1}(x)$ – обратная функция.

Известно, что плотность распределения расстояния от точки, равномерно распределенной в круге, до центра круга имеет вид $f_D(x) = 2x/R^2$, $0 < x < R$. Заметим, что функция $P_{Tx}(D) = AD^{-\gamma}$ монотонна, а обратная ей функция равна $(A/y)^{-1/\gamma}$, и тогда

$$|\psi'(y)| = \frac{(A/y)^{1/\gamma}}{\gamma y}. \quad (9.24)$$

Из (9.23) следует, что

$$f_{P_{Tx}}(y) = \frac{2(y/A)^{\frac{2}{\gamma}}}{\gamma R^2 y}, y > 0. \quad (9.25)$$

Теперь из (9.21) с учетом преобразований $f_{T_S+T_W}(y)$ и $f_{P_{Tx}}(y)$ и свертки соответствующих плотностей распределения можно найти искомую плотность распределения $f_{E_T}(x)$ потребления энергии на одну транзакцию.

Время жизни датчика

Зная плотность $f_{E_T}(x)$, можно найти функцию распределения времени жизни случайно выбранного датчика $F_{T_L}(t)$. Рассмотрим работу IoT устройства как процесс обновления, где цикл обновления представляет собой сумму с.в. времени транзакции и времени ожидания передачи. Количество энергии, затрачиваемой в течение цикла, зависит от его продолжительности. Плотность распределения энергии, потребляемой в спящем режиме, может быть получена путем линейного преобразования $E_S = P_I T_I$, времени сна T_I , где P_I – постоянная величина, определенная в (9.22).

Пусть емкость батареи равна U Дж, а $f_E(t, x)$ – плотность с.в. количества энергии, затрачиваемой за время t ; $E[E_C]$ – средняя энергия, затрачиваемая за один цикл обновления; $E[T_C]$ – средняя длительность цикла обновления. Предполагается, что $U \gg E[E_C]$ и при больших t распределение потребляемой энергии близко к нормальному. Пренебрегая временем возвращения в последнем цикле обновления, получаем следующие формулы для математического

ожидания и дисперсии нормального распределения $N(\mu_E(t), \sigma_E(t))$ с.в. $E(t)$ количества затрачиваемой энергии:

$$\mu_E(t) = \frac{E[E_C]}{E[T_C]} t, \quad \sigma_E^2(t) = \frac{\sigma^2[T_C]t}{(E[T_C])^3}. \quad (9.26)$$

Тогда, ФР $F_L(t)$ времени жизни датчика имеет вид

$$\begin{aligned} F_L(t) &= Pr\{T_L < t\} = 1 - Pr\{T_L \geq t\} = 1 - Pr\{U - E(t) > 0\} = \\ &= 1 - Pr\{E(t) < U\} = 1 - \Phi_{E(t)}(U), \end{aligned} \quad (9.27)$$

где $\Phi_{E(t)}(t)$ – функция распределения нормального закона с.в. $E(t)$.

9.5.2. Расширенная система с ретрансляцией

Теперь рассмотрим расширенную систему NB-IoT с M мобильными ретрансляторами, установленными на транспортных средствах. Модель движения RDM предполагает, что каждая подвижная БС выбирает случайное направление α , равномерно на $[0, 2\pi]$, и движется по прямой экспоненциально распределенному времени τ с постоянной скоростью v . Предполагается, что система находится в равновесии, т.е. интенсивность входа в зону обслуживания новых подвижных БС равна интенсивности выхода из зоны. В этих предположениях расположение подвижной БС можно считать равномерно распределенным в пределах круга радиуса R , независимо от начального распределения.

Для анализа процесса доставки сообщений используем модель сети массового обслуживания (СМО) в дискретном времени. СМО изображена на рис. 9.13 и состоит из $M + 1$ узлов. Первые M СМО моделируют процесс обслуживания на подвижных ретрансляционных станциях, где величины p_i , $i = 1, 2, \dots, M + 1$, соответствуют долям общего входящего потока интенсивности λ , $\sum_{i=1}^{M+1} p_i = 1$. Выходящие с этих СМО потоки суммируются с потоками, обслужившимися на $(M + 1)$ -й СМО, моделирующей основную БС, где $p_{M+1}\lambda$ – доля поступающего на нее общего входящего потока. Упрощенный анализ этой сети проводится методом декомпозиции, т.е. считается, что все СМО независимы и рассматриваются независимо друг от друга.

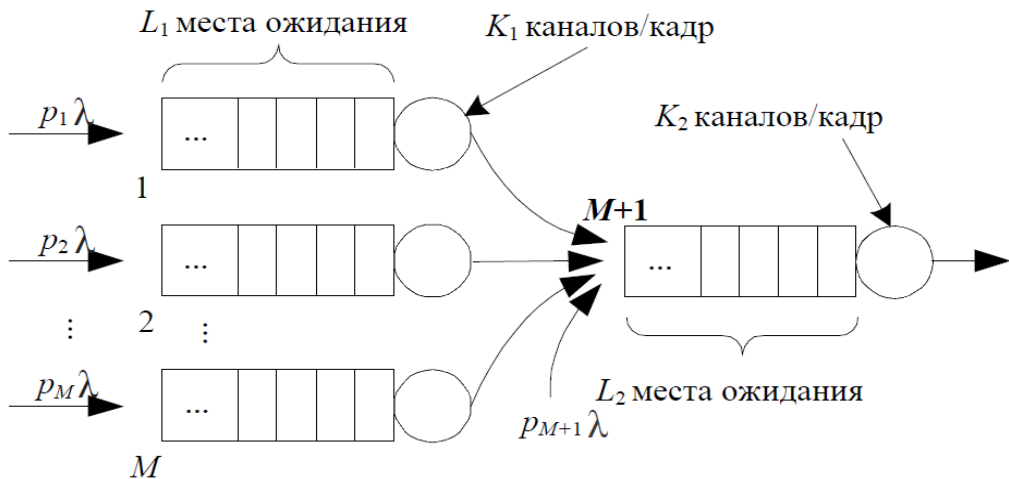


Рис. 9.13. Схема модели СеМО

Задержки и потери сообщений

Рассмотрим случай, когда $p_i = p$, $i = 1, 2, \dots, M$, и напомним, что модель распространения является монотонно убывающей функцией расстояния, а совместное распределение расстояний БС друг от друга является равномерным в круге радиуса R .

В сделанных предположениях плотность распределения расстояния D между двумя точками, равномерно распределенными в круге радиуса R , имеет вид

$$f_D(x, R) = \frac{2x}{R^2} \left(\frac{2}{\pi} \cos^{-1} \left(\frac{x}{2R} \right) - \frac{x}{\pi R} \sqrt{1 - \frac{x^2}{4R^2}} \right), 0 < x < 2R. \quad (9.28)$$

Известно, что ФР минимума M независимых и одинаково распределенных с.в. $Y = \min(X_1, X_2, \dots, X_M)$ определяется по формуле

$$\begin{aligned} F_Y(y) &= 1 - \Pr\{X_1 > y, \dots, X_M > y\} = \\ &= 1 - \prod_{i=1}^M \Pr\{X_i > y\} = 1 - \prod_{i=1}^M [1 - F_i(y)], \end{aligned} \quad (9.29)$$

где $F_i(y)$ ФР с.в. X_i .

Интегрируя (9.28), найдем соответствующую ФР в виде

$$\begin{aligned} F_D(x, R) &= \frac{2}{\pi} \operatorname{csc}^{-1} \left(\frac{2R}{x} \right) - \frac{2R^2 x + x^3}{4\pi R^3} \sqrt{4 - \frac{x^2}{R^2}} + \\ &+ 2Rx^2 \operatorname{sec}^{-1} \left(\frac{2R}{x} \right), 0 < x < 2R. \end{aligned} \quad (9.30)$$

Теперь можно найти ФР расстояния до ближайшей подвижной БС по формуле

$$F_{D_{min}}(x, R) = 1 - \left(1 - \frac{2}{\pi} \operatorname{csc}^{-1}\left(\frac{2R}{x}\right) - \frac{2R^2x + x^3}{4\pi R^3} \sqrt{4 - \frac{x^2}{R^2}} + 2Rx^2 \operatorname{sec}^{-1}\left(\frac{2R}{x}\right)\right)^M, \quad 0 < x < 2R. \quad (9.31)$$

Напомним, что ФР с.в. D_C расстояния от точки, равномерно распределенной в круге, до центра этой окружности имеет вид

$$F_{D_C}(x) = \frac{2x}{R^2}, \quad 0 < x < R. \quad (9.32)$$

Будем считать, что на основной БС обслуживается сообщение от ближайшей к ней подвижной БС. Вероятность этого события определяется формулой

$$p_{M+1} = \Pr\{D_{min} > D_C\} = \Pr\{D_{min} - D_C > 0\}, \quad (9.33)$$

где D_{min} – расстояние до ближайшей подвижной БС. Заметим, что в соответствии с предположениями интенсивности поступающих на подвижные БС пуассоновских потоков имеют вид

$$\lambda_i = \frac{1-p_{M+1}}{M} \lambda, \quad i = 1, 2, \dots, M. \quad (9.34)$$

Теперь каждый узел СеМО, моделирующий подвижную БС может быть описан с использованием модели СМО с дискретным временем из предыдущего раздела.

Учитывая сделанные предположения, можно считать поступающий на основную БС поток пуассоновским, а его интенсивность вычислять по формуле

$$\begin{aligned} \lambda_{M+1} &= M(1 - p_{Loss}) \frac{1 - p_{M+1}}{M} \lambda + p_{M+1} \lambda = \\ &= \lambda[(1 - p_{Loss})(1 - p_{M+1}) + p_{M+1}], \end{aligned} \quad (9.35)$$

где p_{Loss} – вероятность потери сообщения для любого узла подвижных БС. Теперь общая вероятность потери сообщения на всех БС равна

$$\tilde{p}_{Loss} = 1 - (1 - p_{Loss})(1 - p_{Loss, M+1}), \quad (9.36)$$

где $p_{Loss, M+1}$ – вероятность потери сообщений для узла $M + 1$ СеМО.

Потребляемая энергия и время жизни датчика

Как обсуждалось ранее, для расширенной модели с ретрансляцией сообщений могут быть улучшены время жизни датчика и энергия, потребляемая им на одну транзакцию. Преимущество расширенной модели заключается в том, что датчик может использовать подвижные БС для передачи части передаваемых ими сообщений. Введем q_R и q_B – доли времени, в течение которого случайно выбранный датчик отправляет сообщения через ретрансляционные БС и основную БС соответственно. Тогда плотности вероятности с.в. энергии на транзакцию и время жизни датчика могут быть представлены в виде

$$\begin{aligned} f_E(x) &= q_R f_{E_R}(x) + q_B f_{E_B}(x), \\ f_L(x) &= q_R f_{L_R}(x) + q_B f_{L_B}(x), \end{aligned} \quad (9.37)$$

где индексы E и L относятся к обозначениям энергии и времени жизни датчика соответственно. Заметим, что формулы для плотностей $f_{E_B}(x)$ и $f_{L_B}(x)$ были получены ранее и тогда плотности $f_{E_R}(x)$ и $f_{L_R}(x)$ можно найти, заменив плотность расстояния до основной БС на плотность расстояния до ближайшей подвижной БС.

Наконец, заметим, что доли времени q_R и q_B совпадают с вероятностями поступления сообщения на соответствующие БС. Тогда эти вероятности аналогично (9.35) могут быть рассчитаны по формулам

$$q_R = 1 - p_{M+1}, q_B = p_{M+1}. \quad (9.38)$$

9.6. Численный анализ

Рассмотрим результаты численного анализа производительности для системы с ретрансляторами на транспортных средствах. Сначала исследуется базовая модель NB-IoT, а затем добавляем ретрансляторы и сравниваем те же показатели обеих моделей. Для оценки точности аналитической модели результаты сравниваются с компьютерным моделированием, реализованным в системе имитационного моделирования WINTERsim, которая широко использовалась в наших исследованиях для оценки различных сценариев IoT.

9.6.1. Базовый сценарий

В табл. 9.3 приведены исходные данные моделирования, а на рис. 9.14–9.17 показаны результаты вычислений для базовой модели NB-IoT.

Таблица 9.3

Исходные данные моделирования		
Параметр	Обозначения	Значение
Радиус городской зоны обслуживания	R	10 км
Радиус сельской зоны обслуживания	R	30 км
Число подвижных БС в зоне	M	0, ..., 1000
Интенсивность потока сообщений датчика	λ_l	4 сообщения в 1 мин
Плотность датчиков	σ	10, ..., 1000 на км ²
Общее число NB-IoT-каналов	K	72
Число мест ожидания на БС	L	$3 \times K$
Доля каналов	γ_K	0, ..., 0,95
Длительность кадра	Δ	10 мс

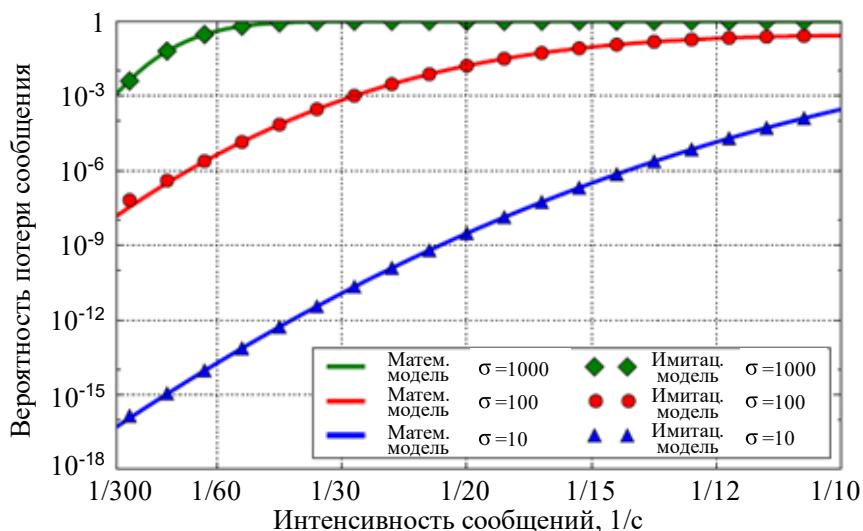


Рис. 9.14. Вероятность потери сообщения для базового сценария, $R = 10$ км

Сначала рассмотрим вероятность потери сообщений в зависимости от интенсивности поступления сообщений от датчиков на рис. 9.14. Интенсивность варьируется от случая, когда сообщение генерируется каждые 5 мин, и до случая, когда сообщение генерируется каждые 10 с. Это соответствует сценариям с регулярно посылаемыми сообщениями (например, мониторинг температуры, давления

воздуха и т.д.), а также для нерегулярного трафика (в частности тревожные сигналы обнаружения вторжений). На рис. 9.14 показана вероятность потери для городского развертывания с радиусом соты 10 км, а на рис. 9.15 для сценария соты в сельской местности с радиусом покрытия 30 км.

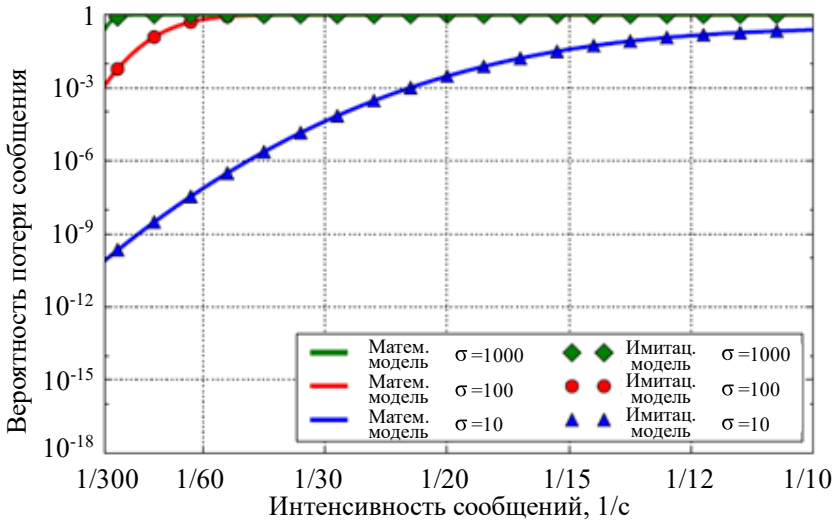


Рис. 9.15. Вероятность потери сообщения для базового сценария, $R = 30$ км

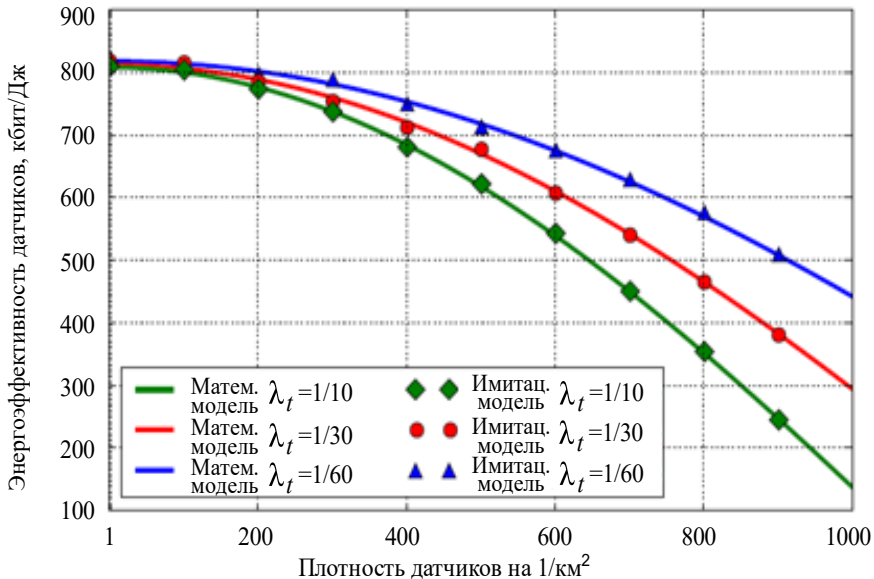


Рис. 9.16. Энергоэффективность датчиков для базового сценария, $R = 10$ км

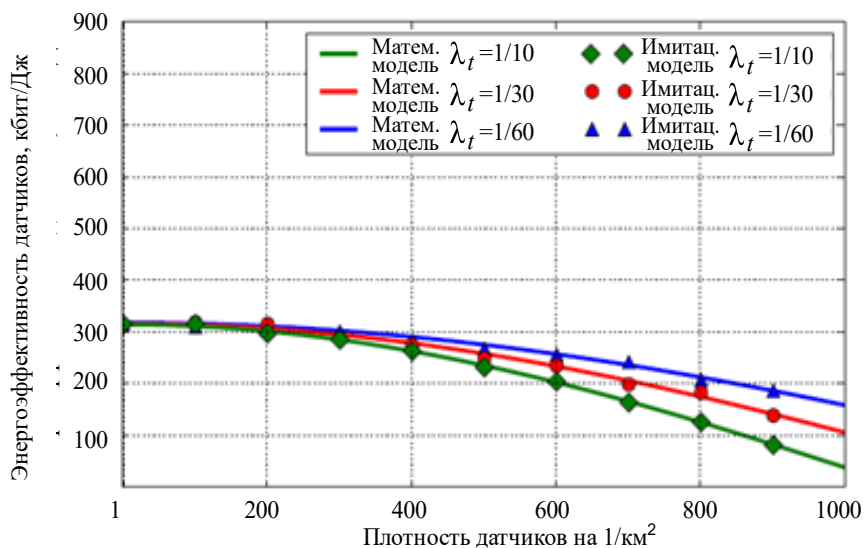


Рис. 9.17. Энергоэффективность датчиков для базового сценария, $R = 30$ км

Из поведения системы в городском сценарии на рис. 9.14 показано, что БС может обеспечить надежную передачу (вероятность потери пакета менее 10^{-6}) для $\sigma = 10$ подключенных датчиков на квадратный километр, которые генерируют приблизительно $\lambda_t = 4$ пакета в минуту от одного датчика. Для более высокой плотности подключенных датчиков ($\sigma = 100$) порог потери 10^{-6} достигается при $\lambda_t = 1$ пакет в минуту. Для $\sigma = 1000$ надежность передачи пакетов не может быть гарантирована, вероятность потери пакета более 10^{-3} . В этих условиях базовая система не может справиться с доставкой чувствительных к задержке данных при больших нагрузках, хотя для меньшего числа датчиков может быть обеспечена надежная связь. Ситуация становится еще сложнее в сельской местности, где плотность $\sigma = 100$ датчиков на квадратный километр приводит к вероятности потери 10^{-3} и более.

Среднее значение энергоэффективности для базовой модели показано на рис. 9.16 для городского сценария и на рис. 9.17 для сельского сценария развертывания. Как можно заметить, передача коротких пакетов по NB-IoT по меньшей мере на порядок превосходит по производительности аналогичные значения для LTE. Тем не менее желаемый срок службы батареи 10 лет и более вряд ли может быть достигнут в рассматриваемом сценарии для типичного недорогого датчика, оснащенного батареей емкостью 50 мАч с напряжением

3 В и сообщения длиной 10 байт. Такие значения приводят к максимальному времени автономной работы от 1 до 8 лет, хотя предполагается, что емкость аккумулятора предназначена исключительно для радиосвязи. Вследствие, как правило, больших расстояний между БС и датчиками в сельском сценарии соответствующие значения энергоэффективности примерно в 2,5 раза ниже. Несмотря на то, что радиус соты в 3 раза больше, коэффициент падения энергоэффективности не равен 3, так как зависимость между количествами потребляемой энергии и радиусом соты не является линейной.

9.6.2. Расширенная система с ретрансляцией

Проведя анализ показателей базового сценария, оценим производительность расширенной системы с БС на транспортных средствах. Во-первых, напомним, что пул доступных каналов NB-IoT в кадре размера K должен быть разделен между ретрансляционными БС, K_1 , стационарной БС, K_2 , $K_1 + K_2 = K$ так, чтобы не было чрезмерных помех, вызванных их одновременным использованием. Введем отношение $\gamma_K = K_1/K$, определяющее долю каналов, выделенных для подвижных ретрансляторов. Если $\gamma_K = 0$, то расширенный сценарий сводится к базовому случаю, а в случае $\gamma_K = 1$ все каналы назначаются подвижным БС.

Важность оптимального распределения радиоресурсов показана на рис. 9.18, иллюстрирующем вероятность потерь в зависимости от γ_K для сельской местности с интенсивностью потока от одного датчика $\lambda_l = 1$ пакет в минуту. Можно видеть, что расширенная система ведет себя аналогично базовому сценарию при $\gamma_K = 0$. С увеличением γ_K надежность доставки пакетов увеличивается до тех пор, пока хватает ресурсов для поддержки всех датчиков без вспомогательных ретрансляторов в непосредственной близости. С этого момента вероятность потерь начинает расти и превышает ее значения для базовой модели.

Анализируя графики на рис. 9.18, заметим, что для любого набора входных параметров существует оптимальное значение γ_K , которое дает наименьшую вероятность потерь сообщений. Когда γ_K становится больше оптимальных значений, производительность уменьшается и в итоге вероятность потерь приближается к 1 для всех

значений исходных данных. Это связано с тем, что все каналы назначаются подвижным БС и нет каналов для передачи генерируемого трафика на стационарную БС. Например, для $M = 1000$ оптимальное значение γ_K составляет около 0,55, что приводит к высокой надежности передачи сообщений.

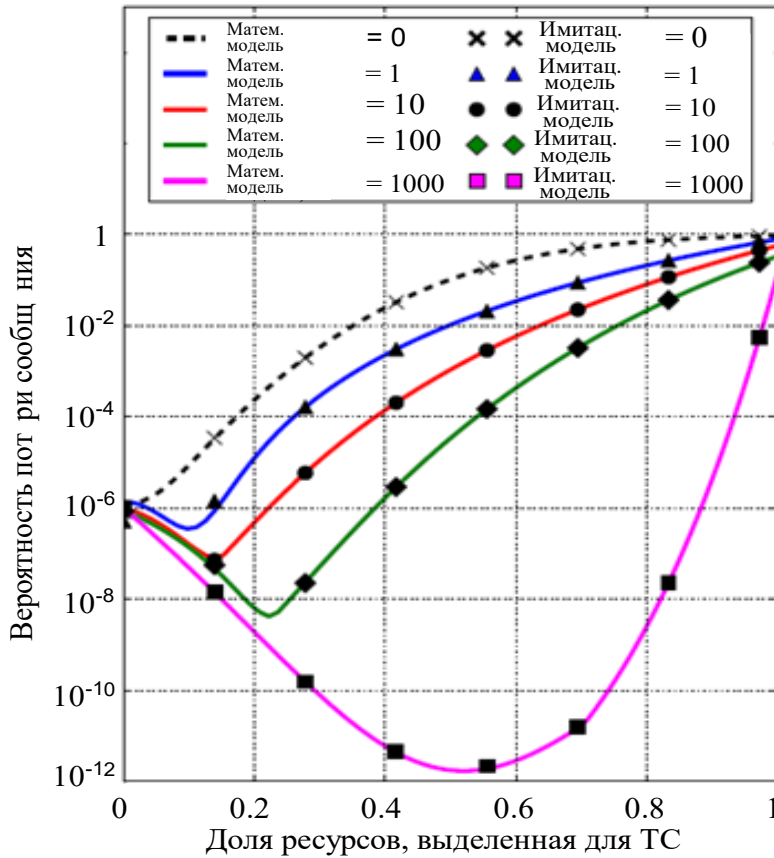


Рис. 9.18. Результат разделения ресурсов между подвижными и неподвижной БС (ТС – транспортное средство, на которое устанавливается подвижная БС)

Подводя итог, отмечаем, что использование оптимального значения γ_K имеет решающее значение для повышения производительности системы. Зная количество доступных вспомогательных транспортных средств в пределах БС, этот параметр можно настроить динамически, чтобы ресурсы распределялись в зависимости от ситуации.

Рассмотрим показатели производительности на рис. 9.18 – вероятность потерь сообщений, энергоэффективность и время жизни датчиков. Мы концентрируемся на сельском сценарии развертывания с радиусом соты $R = 30$. На рис. 9.19 показано, что вероятность потерь быстро уменьшается по мере увеличения числа подвижных БС, что, как говорилось ранее, не только повышает надежность при подключении до сотни датчиков на квадратный километр, но также обеспечивает этот показатель в условиях высокой нагрузки для тысячи подключенных устройств на квадратный километр, что не может быть обеспечено в базовом сценарии.

Наряду со значительным повышением надежности, как показано на рис. 9.20, расширенная система характеризуется повышением энергоэффективности. Начиная с нескольких сотен кбит/Дж для небольшого числа подвижных БС ($M = 300$), расширенная система превосходит базовый сценарий по энергоэффективности в 6 раз при $M = 1000$, что, в свою очередь, позволяет значительно повысить время работы датчика от батареи, как показано на рис. 9.21.

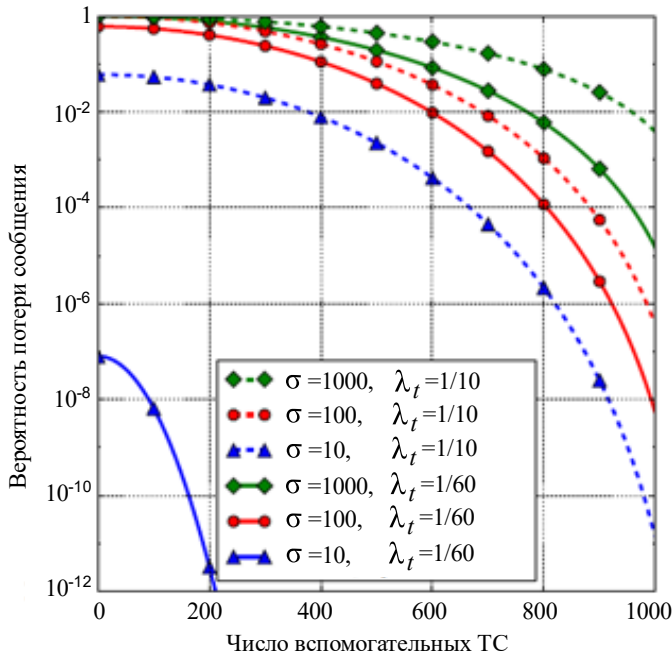


Рис. 9.19. Зависимость вероятности потери сообщения от числа подвижных БС

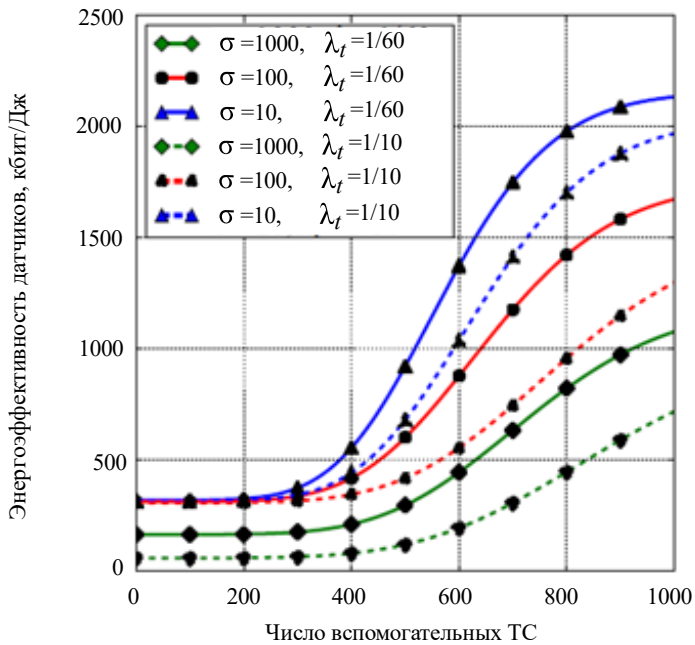


Рис. 9.20. Зависимость энергоэффективности от числа подвижных БС

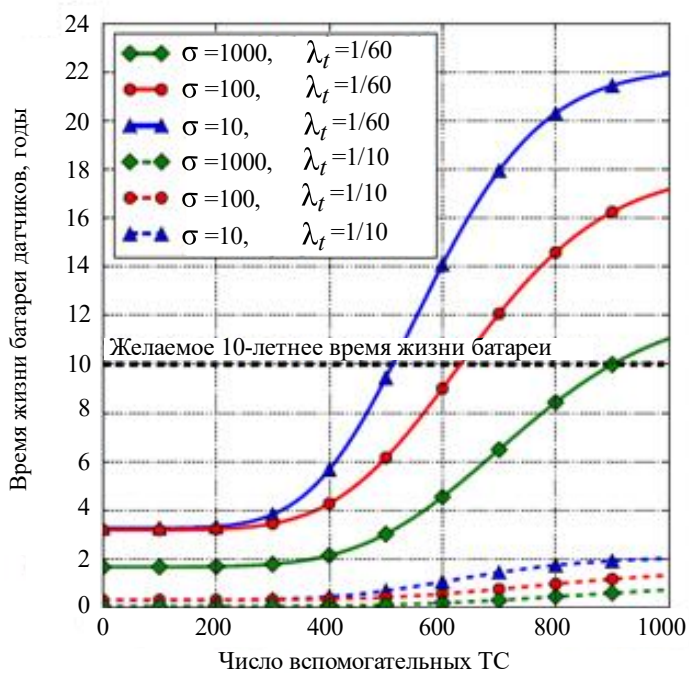


Рис. 9.21. Зависимость времени жизни батареи от числа подвижных БС

Таким образом, предлагаемый подход – это возможная стратегия, которая не только удовлетворяет требованиям к сроку службы батареи до 10 лет, но и повышает ключевой показатель производительности при различных значениях плотности датчиков. Данный подход включает в себя случай для $\sigma = 1000$ подключенных датчиков, что соответствует 2,8 миллиона датчиков в пределах зоны обслуживания БС. Как показано на рис. 9.19 и 9.21, оптимальное использование радиоресурсов, предоставляемое расширенной схемой, дает возможность надежного массового размещения датчиков и существенного увеличения их времени жизни. В то же время, независимо от плотности датчиков, целевой 10-летний срок службы батареи вряд ли может быть достигнут при использовании батареи 50 мАч и 3 В с любым числом вспомогательных транспортных средств, если этот датчик передает сообщения чаще чем раз в минуту.

ГЛАВА 10

РАЗДЕЛЕНИЕ РЕСУРСОВ

ПРИ ИСПОЛЬЗОВАНИИ ТЕХНОЛОГИИ D2D

Ожидается, что технологии прямого взаимодействия устройств D2D будут являться неотъемлемой частью сетей связи 5G/6G, находя активное применение при обслуживании абонентов. В этой главе рассматриваются сценарии использования технологии D2D в будущих сотовых сетях.

Глава начинается с иллюстрации возможностей технологии D2D в параграфе 10.1. Аспекты архитектуры технологии D2D в сети связи 5G/6G рассмотрены в 10.2, а оценка предложенной архитектуры – в параграфах 10.3 и 10.4. Вопросы практической реализации D2D-технологии рассмотрены в параграфе 10.5, а дальнейшее развитие способов ее внедрения обсуждается в 10.6. Пример использования D2D для разгрузки сетей связи обсуждается в параграфе 10.7.

10.1. Возможности D2D в сотовых сетях

В начале 2010-х годов вопрос D2D-взаимодействия не рассматривался в контексте сотовых систем связи. В те годы основная идея состояла в том, что покрытие в сотовых сетях может быть улучшено путем применения систем клиентской ретрансляции, но далее эти идеи были расширены на передачу данных между пользовательскими устройствами. Появление технологии FlashLinQ уже предполагало, что сотовые сети могут использовать одноранговые соединения, что теперь хорошо известно как внутриволновое взаимодействие D2D. Прототип технологии FlashLinQ представляет собой распределенную технологию взаимодействия D2D в лицензируемом спектре. Однако предложенная как решение технология вызвала множество вопросов, в основном связанных с негативным влиянием помех и контролем мощности передачи, которые до сих пор еще полностью не решены и для D2D.

Лицензируемый спектр остается очень ограниченным и дорогим для операторов связи. Как говорилось ранее, есть попытки лицензировать еще больше спектра, поскольку того, который есть, недостаточно для удовлетворения существующего спроса. Поэтому использование любого типа внутриволнового D2D для целей передачи данных является исключительно важной и сложной задачей, поскольку

ресурсы не могут быть окончательно зарезервированы только для D2D и, следовательно, должны динамически использоваться совместно с классическим сотовым радиоинтерфейсом.

Для поддержки нескольких соединений D2D в одной и той же области требуются механизмы управления интерференцией, контролем доступа и мощности. Многие исследовательские работы предлагают способ устранения интерференции, используя информацию о состоянии канала в реальном времени (CSI), местоположении пользователя и пр. В том числе схема управления интерференцией должна решать следующие вопросы:

- распределять ресурсы лицензированной полосы, учитывая интересы стандартного сотового доступа, между пользователями и БС;
- эффективно динамически использовать выделенные ресурсы, освобождая их по мере надобности.

Управление интерференцией является комплексным процессом, известным также как выбор режима D2D. С началом стандартизации D2D-технологий этот вопрос привлекает большое внимание исследователей, сфокусировавших свои исследования на широком круге оптимизационных задач для таких показателей, как ОСШ, пропускная способность, энергоэффективность, задержка передачи данных, справедливое распределение ресурсов и вероятность отказа в обслуживании. Это позволило консорциуму 3GPP всесторонне изучить возможности технологии и включить ее в технологию LTE-A. В конечном счете нехватка ресурсов влияет на лицензионный диапазон D2D так же, как и на сами сотовые сети, которые эта технология пытается улучшить. Текущие исследования включают в себя использование технологий MIMO для D2D, сетевое кодирование, управление интерференцией и использование кэширования.

Конечно, для D2D доступны не только лицензированные диапазоны частот. Действительно, существенная часть спектра по-прежнему не лицензирована, в том числе диапазоны 2,4 и 5 ГГц, используемые Wi-Fi и другими технологиями, такими как Zigbee, Bluetooth и т.д. Использование D2D в нелицензируемом спектре подтверждено исследованиями 3GPP, проведенными в этом направлении, которые известны как LTE-U (LTE Unlicensed, т.е. LTE для нелицензируемого спектра) и LAA (LTE Licensed Assisted Access, т.е. поддержка доступа LTE в лицензируемом спектре). Вместе с тем оператор не может требовать эксклюзивного использования какого-либо нелицензируемого спектра. В результате интерференция становится

неконтролируемой, что требует надежного контроля доступа к среде, который может помочь справиться с этой проблемой.

Поскольку Wi-Fi не имеет встроенного контроля доступа и позволяет каждому устройству занимать канал на основе конкурентного доступа более или менее одинаково, технологию часто критикуют за то, что она не удовлетворяет требованиям по качеству обслуживания QoS. Также Wi-Fi обеспечивает более высокие скорости передачи данных и энергоэффективность, чем любая из сотовых технологий. В принципе при наличии соответствующего интеллекта во всех устройствах Wi-Fi можно добиться и стабильных результатов работы технологии D2D в нелицензируемом спектре частот. Однако, увеличивая число соединений D2D, в том числе с использованием технологии Wi-Fi Direct (WFD), надо отдавать отчет в том, что для этого потребуется дополнительный ресурс управления.

Сотовые сети могут обеспечить ряд важных услуг для систем нелицензируемого спектра, например, интеллектуальное управление радиоресурсами, а также столь необходимую конфиденциальность и безопасность в процессе обслуживания. Если пользовательские устройства постоянно связаны с сотовой сетью, то она может предложить резервное соединение в случае, когда для нелицензируемой полосы качество обслуживания оказывается неудовлетворительным. Кроме того, при поддержке со стороны сотовой сети аутентификация устройств и безопасность соединений D2D могут быть значительно улучшены, что, например, делает любой вид атаки «человек посередине» практически невозможным. Другими словами, степень поддержки сотовой сети для соединений D2D в нелицензируемых полосах спектра представляется очень перспективным подходом.

Области применения технологии D2D в сотовых сетях активно обсуждаются. В том числе рассматриваются и услуги по передаче речи (вызовы между расположенными рядом пользователями) и иные приложения для близко находящихся людей и/или устройств, например, поиск данных, игры и задачи, относящиеся к вычислениям, управление домашними роботами, общественная безопасность. В терминах устройств, помимо смартфонов и планшетов, это, конечно, приложения для автономных машин/роботов, так называемые машинно-ориентированные приложения (MTC/M2M). В целом в технологии D2D на равных используются как ориентированные на человека приложения, так и приложения услуг MTC/M2M.

10.2. Архитектурные аспекты построения D2D

С концептуальных позиций должно быть исследовано взаимодействие двух независимых сетей: существующей сотовой сети и сети D2D, которая не использует сотовый принцип построения сети, а также функционирует в нелицензируемом диапазоне частот. При этом некоторые дополнительные услуги станут доступными для сотовой сети такие, как одноранговое взаимодействие для передачи данных (P2P, Peer-to-Peer), управление соединениями и т.д. При этом сама сеть D2D исходно отвечает за передачу данных P2P и размещение ресурсов, и такое разделение ответственности достаточно обоснованно. Далее будут рассмотрены вопросы выгрузки трафика на уровне приложений, а также решения для обеспечения безопасности соединений D2D.

Далее исследуется технология D2D на примере ее разработки, соответствующей той, что была проведена 3GPP в процессе интеграции D2D в стандарты LTE-A. Далее сосредоточимся на концепции D2D с сетевой поддержкой и примере трансформации этой концепции в полнофункциональное решение, позволяющее обеспечить динамическое управление локальными радиосоединениями в дополнение к соединениям сотовой сети. Стремясь обеспечить совместимость предложения с иными решениями, рассматриваются в основном стандартные интерфейсы и протоколы, облегчающие интеграцию.

Подход заключается в том, чтобы избегать модификации любых технологий канального уровня, так как это может потребовать замены оборудования, которое уже находится в эксплуатации, что экономически нецелесообразно в обозримой перспективе. Аналогично следует не нарушать сотовую архитектуру оператора сети, а модернизировать ее только путем эволюции. Кроме того, ограничим область разработки тем, что новая технология будет оставаться в основном сотовой с добавлением только нескольких функциональных компонентов. Помимо соответствия требованиям 3GPP, такой подход облегчает процесс планирования и эволюции, поскольку в сотовых системах уже существует большой набор специальных функций, таких как хэндовер, аутентификация, безопасность, тарификация, позиционирование и многие другие. Предположим также, что все приложения, которые будут использовать новые функции, должны быть совместимы с существующей экосистемой. Для обеспечения совместимости с планируемыми сетями сфокусируем внимание на 3GPP Release 10 и выше для определения требований к базовой сотовой

сети. Где это возможно, будем рассматривать проблемы с точки зрения восходящих соединений, так как это больше всего влияет на производительность системы. Кроме того, в предлагаемых решениях будет использоваться только оборудование, соответствующее современным пользовательским устройствам.

Далее также рассмотрим набор технологических средств для конвергентной сотовой/D2D-сети, где мультирадиоустройства имеют возможность использовать все свои интерфейсы для одноранговой связи, например, обмен файлами, выгрузка сотового трафика и т.п. Рассмотренный в этой главе набор решений достаточно широк, чтобы охватить сценарии выгрузки, алгоритмы управления и любые потенциальные существующие и будущие беспроводные технологии, в частности IEEE802.11ad/ay, LTE-U и т.д.

Как было указано ранее, для операторов сотовой связи весьма привлекательна возможность выгрузки трафика в сеть D2D нелицензируемого спектра частот, но ни Wi-Fi, ни Wi-Fi Direct не имеют для этого эффективного способа обнаружения устройств и услуг. Анонимность P2P-соединений также остается открытым вопросом для решений D2D в нелицензируемом спектре частот. Ниже рассмотрена архитектура для внедрения поддерживаемых сетью D2D решений для преодоления этих ограничений.

Архитектура уровня обслуживания

Сначала определим, какие именно услуги должны быть обеспечены конечному пользователю. Естественно, очень сложно гарантировать особенные условия для D2D-соединений, а качество обслуживания для таких соединений может существенно отличаться в различные моменты времени и в связи с мобильностью пользователей. Поэтому услуги, нечувствительные к задержкам, такие как распределенное кэширование данных и кооперативные загрузки данных (например, услуга мультивещания) могут рассматриваться как первые кандидаты на выгрузку трафика с использованием D2D. Однако эти услуги достаточно редко необходимы мобильным пользователям. Наиболее интересными для пользователей представляются услуги по запросу, требующие стабильных соединений, предоставляющих высокое качество обслуживания. Но, если устройства находятся близко друг к другу и, следовательно, потенциально могут обладать высоким качеством канала связи, многие стандартные услуги сотовых сетей могут стать возможными для использования в сетях D2D.

Например, на коротком промежутке времени приложения, подобные потоковому видео, многопользовательские игры и многие другие аналогичные услуги могут быть предоставлены посредством D2D-соединений. Для того чтобы сделать эти услуги коммерчески привлекательными, необходимо обеспечить определенные гарантии, а также способы определения того, как выполняются условия использования прямых соединений.

Учитывая изложенное, критически важные требования к разработке архитектуры сети D2D состоят в том, чтобы предоставить клиентам информацию о возможности установления D2D-соединения – с кем, какой контент доступен, и что делать, если по какой-либо причине нет возможности обеспечить требуемое качество обслуживания. Для этого должна быть предусмотрена поддержка таких функций, как:

- идентификация цели взаимодействия (хранение и поиск файловых и потоковых метаданных, идентификация участника игры и т.д.);
- поиск ближайших партнеров, проявляющих интерес к информации, не подверженной риску доступности контента для широкой публики;
- аутентификация и авторизация по отношению к D2D-соединению для передачи контента;
- поддержка надежного возврата к сотовому соединению при отказе D2D-соединения.

Технически нахождение пары для обмена подходящим контентом не является сложной задачей вследствие широкого распространения сегодня облачных систем и социальных сетей. Важнейшая проблема состоит в том, что соединение, которое еще не установлено, не может быть предоставлено и управляться существующими методами, поскольку отсутствуют ассоциированные с ним интерфейсы, отсутствуют IP-адреса и пр. Естественным выходом в такой ситуации является создание новых методов, специально разработанных для управления D2D-соединениями и принятия решений в реальном масштабе времени, основанных на позиционировании пользователей относительно друг друга и доступности радиоресурсов. Поскольку такая информация имеется только в системе управления сетью, то эти методы должны обеспечиваться на сети оператора владельца системой управления.

Вместе с тем процедуры отслеживания контента и нужды систем безопасности разнообразны, а программные прикладные

интерфейсы (API, Application Programming Interface) доступны только поставщику услуг. В этом случае компоненты управления связностью и компоненты социального взаимодействия должны вступить в кооперацию между собой для предоставления комплексных услуг пользователю. Это наиболее естественный интерфейс между двумя, казалось бы, несовместимыми технологиями – специфическими возможностями управления D2D-соединением (управляемого по запросу сервером D2D, располагающимся в транспортной сети оператора связи) и специфическими возможностями отслеживания контента (управляемого по запросу сервера приложений). Относящиеся к безопасности проблемы в этом случае дублируются для обеих систем. Изложенное решение проиллюстрировано на рис. 10.1 и работает следующим образом.

1. Каждое устройство использует возможности уровня приложений для аутентификации сервером приложений. Это позволяет обеспечить взаимодействие с контентом, что авторизуется третьей стороной, а также устраняет требования по обработке метаданных и предоставляет возможности авторизации со стороны социальных сетей.

2. Пользовательское оборудование разрешает серверу D2D-оператора связи представлять информацию о себе на сервере приложений. При этом сервер D2D не получает доступ непосредственно к контенту или его метаданным, а осуществляет идентификацию пользователя, чтобы убедиться, что устройство действительно принадлежит владельцу заявленных учетных данных. Таким образом, сервер D2D позволяет присваивать имена прикладного уровня физическим устройствам, в том числе и в случае, когда имеется несколько устройств, принадлежащих одному и тому же лицу.

3. Пользовательское устройство может публиковать или искать D2D-соединения для контента на сервере приложений, и эти соединения будут ссылаться на конкретный контент конкретного пользователя, но не устройства. D2D-соединения могут быть постоянными, периодическими или одноразовыми в зависимости от приложения.

4. Пользователь запрашивает сервер D2D для упрощения создания соединения таким образом, чтобы обеспечить соединение на уровне приложений, а сервер приложений преобразует этот запрос в сеанс связи на уровне соединения и для этого предоставляет IP-адрес, к которому в дальнейшем могут быть привязаны сокеты приложений.

5. Наконец, на данном этапе может начаться обмен данными. Заметим, что сервер приложений не участвует в этом процессе и не отслеживает обмен данными по D2D-соединению, будучи только уверенным в том, что соединения следуют модели безопасности. В то же время сервер D2D, при необходимости, может осуществлять мониторинг и управлять свойствами соединений.

Предложенная на рис. 10.1 схема позволяет сохранить модели безопасности и разрешений, используемых услугами приложений. Например, если доступ к контенту ограничен для определенной группы пользователей, то только пользователи, которые имеют соответствующие права, получают доступ к этому контенту. Дополнительной целью является гарантия того, что пользователи не будут отслеживаться после того, как их D2D-сессия будет завершена. Предложенная схема поддерживает хорошо известные решения для обеспечения целей безопасности.

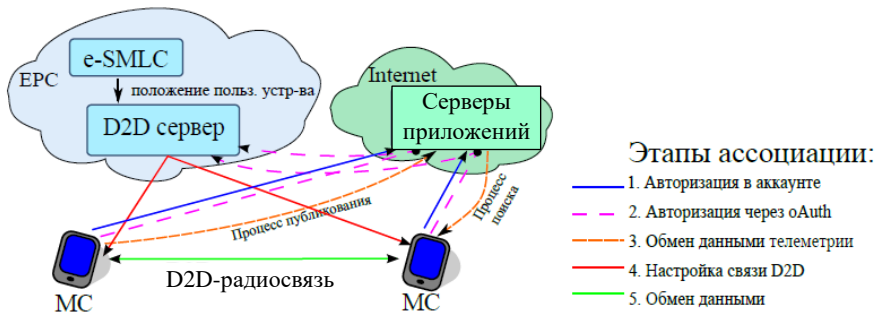


Рис. 10.1. Иллюстрация установления D2D-соединения

Далее кратко перечислим особенности архитектуры, которые делают ее приемлемым решением с точки зрения безопасности в поддерживаемой сотовой сетью D2D-технологии.

1. Каждое соединение начинается с проверки контента соединения. Потенциальным злоумышленникам необходимо будет найти компромисс с сервером приложений для получения прав на установление такого соединения. Поэтому любая атака на систему, по существу, блокируется паролем, распространяемым сервером приложений через существующие защищенные каналы сотовой системы связи.

2. Ни одно из пользовательских устройств не должно широковещательным образом транслировать информацию об обнаружении

потенциальных партнеров для соединения или прослушивать запросы на обнаружение. Поскольку устройства поддерживают радиointерфейсы D2D выключенными до тех пор, пока не будет получена соответствующая команда активации, они не могут быть атакованы или отслежены, прежде чем перейдут в активное состояние.

Предложенные механизмы в значительной степени опираются на существующие протоколы, такие как OAuth и SSL, для авторизации на сервере D2D для управления соединениями и передачи команд соответственно. Кроме того, предполагается, что серверу D2D всегда можно доверять только потому, что он находится в сети оператора. Поэтому необходимо оценить потенциальные риски безопасности в конкретных случаях, когда применяется D2D с сетевой поддержкой со стороны сотовой сети.

Интеграция с архитектурой 3GPP

Теперь, когда ясно, как должна работать система, можно сосредоточиться на вопросах ее интеграции с существующими сетевыми архитектурами. Благодаря гибкости ядра сети LTE относительно легко выполнить интеграцию сервера D2D в существующую архитектуру LTE. На рис. 10.1 показано, как эта интеграция может быть осуществлена на практике. Действительно, рассмотренная в этой главе схема D2D может быть сопоставлена с объектами архитектуры 3GPP практически однозначным образом, с единственным дополнительным объектом – сервером D2D, находящимся в транспортной сети оператора связи. Такое размещение позволяет серверу D2D связываться с центром местоположения (SMLC, Serving Mobile Location Center), чтобы позиционировать пользовательское оборудование на местности, а также взаимодействовать с серверами приложений внешних провайдеров услуг.

Сигнализация для выгрузки трафика в D2D

Протокол выгрузки трафика для D2D сегодня, по сути, является частью стандартов 3GPP. При этом некоторые детали стандартизованного решения 3GPP отличаются от первоначально предложенных. В том числе система сигнализации, предложенная в документах 3GPP, не учитывает проблемы взаимодействия с серверами приложений, поскольку это выходит за рамки исследований консорциума. В дальнейшем предполагается, что используется такая

система сигнализации, которая адекватно реализует описанную выше функциональность.

Что касается состояния исследований, то технология D2D с сетевой поддержкой по-прежнему в значительной степени не изучена. Например, если сеть давала бы возможность контролировать, какие D2D-соединения используются, возможно было бы избежать выгрузки трафика на те соединения, которые могут ухудшить производительность сети и/или качество обслуживания пользователя. Аналогично, если бы сеть могла управлять процессом передачи по соединению D2D, возможно было бы создать расписание, учитывающее, когда группы неконкурирующих соединений могут взаимодействовать, что существенно уменьшило бы интерференцию и, соответственно, увеличило бы пропускную способность и энергоэффективность. Необходимо подчеркнуть, что становятся доступными улучшенные показатели управления мощностью, когда сеть участвует в установлении D2D-соединений. В любом случае рассмотренная архитектура достаточно хорошо масштабируема, учитывает почти все известные модели управления со стороны и, таким образом, обеспечивает основу для дальнейшего развития технологии D2D, управляемой сотовой сетью оператора связи.

Следует отметить, что модели взаимодействия пользователей в пределах технологии D2D достаточно разнообразны. В соответствии с эволюцией технологий новые возможности интегрируются в повседневную жизнь и оказывают на нее существенное влияние. Рассмотренная архитектура содержит начальный набор функциональных возможностей, необходимых для интеграции технологии D2D, управляемой сетью оператора связи и взаимодействия с остальной частью всемирной сети, в том числе, с поставщиками услуг. При этом можно достаточно легко вернуться к стандартному неуправляемому функционалу D2D, когда по какой-либо причине сетевая поддержка становится недоступной или даже контрпродуктивной, или предоставить возможность устанавливать соединения с приложениями каким-либо другим способом.

Далее сфокусируем внимание на анализе производительности предложенных решений по выгрузке трафика и обсудим их применимость на практике.

D2D в сетях 5G/6G

Важно отметить, что попытки отслеживать пользовательскую активность в целях выгрузки сессий по альтернативному пути,

например, посредством прямых соединений, скорее всего, не дадут конкретных практических решений и, таким образом, пользователи не смогут использовать весь потенциал сетей D2D. Вместе с тем разрешение устанавливать прямые соединения отражает скорее предпочтения пользователя, а не стремление сети к выгрузке сессии, существенно улучшая качество обслуживания для этой сессии в перспективе. Отметим, что не только потенциальная производительность соединений D2D относительно сотовых каналов делает их привлекательными для пользователей. Технология D2D позволяет еще и организовать взаимодействие между пользователями, расположенными в некоторой географической области. Это важное положение не только лежит в основе исследований, но и подчеркивает также изменения представления о том, как пользователи могут взаимодействовать с сотовыми сетями связи в ближайшем будущем.

Ожидается, что мобильные сети 5G/6G позволят принципиально увеличить пиковые скорости передачи данных, спектральную эффективность и энергоэффективность, а также обеспечить обслуживание пользователей в любое время, в любом месте. Ясно, что для сетей 5G/6G одна технология не сможет удовлетворить все требования, в том числе по производительности, показателям качества обслуживания и качества восприятия. Например, сверхплотные гетерогенные сети HetNets (см. главу 11), скорее всего, станут одним из методов увеличения емкости сотовых сетей. Этот подход видится весьма дорогостоящим и, следовательно, скорее всего, не будет полностью решать проблемы 5G/6G во всех предполагаемых сценариях развертывания.

Технология D2D, скорее всего, станет частью сетевых сервисов. С точки зрения оператора, взаимодействие по технологии D2D, вероятно, может быть бесплатным как для обычных, так и новых «контекстных» приложений. D2D можно использовать для реализации недорогих, с ультрамалыми задержками, или высокоскоростных соединений в приложениях, которые нуждаются в любой из этих особенностей, в то время как обычные услуги будут предоставляться сотовой сетью.

Поддержка услуг на базе технологии D2D, скорее всего, станет частью будущего портфеля оператора вместе с взаимодействием *машина–машина*, услугами с ультрамалой задержкой и множеством других услуг, которые становятся возможными в сетях 5G/6G. Сегодня сложно предсказать, будет ли D2D завтрашнего дня выглядеть точно, как сейчас. Технология может мигрировать в лицензируемые

диапазоны, несмотря на проблемы управления спектром частот. Так, уже стандартизирована технология D2D для сетей 5G NR в миллиметровом диапазоне частот. Остается надеяться, что в сетях 5G/6G отпадет необходимость общаться с абонентом, находящимся в непосредственной близости на расстоянии нескольких метров, при этом синхронизируя данные через облачные технологии.

10.3. Оценка возможностей архитектуры D2D

Оценка производительности всегда являлась достаточно сложной задачей. В этом разделе обсудим, как именно производится оценка производительности технологии D2D и почему именно такая методология правильна. Далее представим подходы к оценке систем D2D путем имитационного моделирования и, наконец, используем результаты некоторых исследовательских работ.

10.3.1. Методология оценки

В этом разделе определены сетевые объекты и соответствующие механизмы, которые требуются для поддерживаемой со стороны сетевой сети технологии Wi-Fi Direct. Рассматривается методология оценки, которая может быть использована для широкого спектра перспективных технологий D2D. Эта же методология применяется во второй части этого раздела.

Сетевые объекты

Беспроводная сеть представляется как совокупность объектов, которые имеют возможности поддерживать различные радиотехнологии доступа. Это, во-первых, наложенная сеть 3GPP LTE, представленная БС eNodeB. Каждая БС связана с ядром сети, обеспечивающим сотовую связь для всех пользовательских устройств, ассоциированных с сетью. Каждая БС eNodeB доступна ряду пользовательских устройств, каждое из которых имеет возможность взаимодействовать как с LTE, так и с Wi-Fi. Каждому клиенту доступны услуги, которые поддерживаются координатором, по сути являющимся коммутатором. В зависимости от рекомендаций координатора пользователь может задействовать для передачи данных интерфейсы LTE или Wi-Fi. Следует подчеркнуть, что все сказанное далее относится и к технологии 5G NR.

Необходимо также принимать во внимание интерференцию в нелицензируемом спектре от устройств WLAN, которые могут быть расположены в непосредственной близости от точек доступа Wi-Fi. Эти устройства имеют равные права с рассматриваемыми пользовательскими устройствами в контексте доступа к ресурсам сети в соответствии с IEEE 802.11. В то же время они могут быть не ассоциированы с сотовой сетью оператора связи и их работа в нелицензируемом спектре не может быть контролируема сетью LTE.

Трафик и загрузка сети

В соответствии с методологией исследований, отвечающей рекомендациям 3GPP, предположим, что N клиентов LTE, которые имеют также возможность использования D2D, расположены равномерно в зоне развертывания сети. Плотность устройств считается достаточно высокой так, что каждый клиент в зоне радиуса D2D-взаимодействия имеет возможность установить D2D-соединение по крайней мере с еще одним клиентом. Однако только половина клиентов могут быть источниками данных и передавать их, поскольку P2P-взаимодействия являются однонаправленными, и при этом существует только один источник и один получатель данных в каждой сессии. Такой трафик описывается моделью «полный буфер» (Full Buffer в терминологии 3GPP) с пакетами длиной 1500 байт. Предполагается, что некоторый процент x клиентов-источников в радиусе действия D2D имеет своих клиентов – получателей данных.

Устройства, которые не контролируются сетью LTE, как правило, генерируют «тяжелый» трафик передачи данных. Трафик таких клиентов представляется моделью «полный буфер» с пакетами длиной 1500 байт каждый, направляемыми к своей точке доступа Wi-Fi. Для упрощения оценки не будем использовать модель точки доступа Wi-Fi для нисходящего трафика, поскольку это, по существу, другой пользователь с точки зрения протоколов множественного доступа к среде семейства IEEE 802.11.

10.3.2. Имитационное моделирование D2D

Для проведения исследований в рассматриваемом разделе использован имитатор WINTERsim (<http://winter-group.net/download/>), базирующийся на методологии моделирования LTE, предложенной в TR 36.814 и в других документах 3GPP, а также спецификациях

для оценки производительности беспроводных сетей в стандартах IEEE 802.11. Имитатор может поддерживать различные стратегии построения сетей, модели трафика, характеристики каналов связи и протоколы различных уровней. Он позволяет моделировать практически всю существующую инфраструктуру LTE-A и все известные способы размещения пользователей и БС (гексагональные и квадратные соты, внешние условия, равномерное и кластерное распределения клиентов и т.д.).

Для каждого пользователя предусмотрен собственный генератор трафика. Каналы моделируются с учетом соответствующих характеристик источника, получателя и среды распространения. Каждый узел способен поддерживать несколько радиointерфейсов, которые могут использоваться для восходящих и нисходящих соединений и моделируются в пределах единой временной шкалы.

Ниже сформулированы основные предположения для имитационного моделирования, которые легли в основу сценариев для оценки эффективности развертывания технологии D2D.

Сотовая сеть

Для оценки характеристик технологии D2D, основанной на использовании Wi-Fi Direct с поддержкой сотовой сети связи, рассмотрим пример рекомендованного 3GPP сценария, для развертывания сетей в городских условиях. Сеть LTE включает в себя 19 гексагональных сот с тремя секторами в каждой (рис. 10.2). БС eNodeB поддерживают технологические решения как минимум 3GPP Release 11 LTE, расстояние между близлежащими БС eNodeB составляет 200 метров, а радиус соты равен примерно 110 метрам.

Все соты работают в одном 60 МГц лицензированном спектре частот, который разделен на три пары полос по 10 МГц для функционирования в режиме FDD. Каждая сота, в свою очередь, разделена на три сектора, и каждому сектору выделяется пара полос по 10 МГц, что приводит к шаблону повторного использования частот $1 \times 3 \times 3$. Клиенты 3GPP LTE ассоциируются с БС eNodeB на основе лучшего значения ОСШ с порогом хэндовера 1 дБ. Для дополнительной информации о стандартной конфигурации LTE можно обратиться к табл. 10.1 и соответствующим документам по стандартизации (например, 3GPP TR 36.814-900 и МСЭ-R М.2135-1). Для целей проверки производительности реализован сценарий калибровки в соответствии с документом 3GPP TR 36.814-900 и проведены

соответствующие тесты. Результаты моделирования хорошо укладываются в ограничения по спектральной эффективности как для центра соты, так и для ее границ (рис. 10.3).

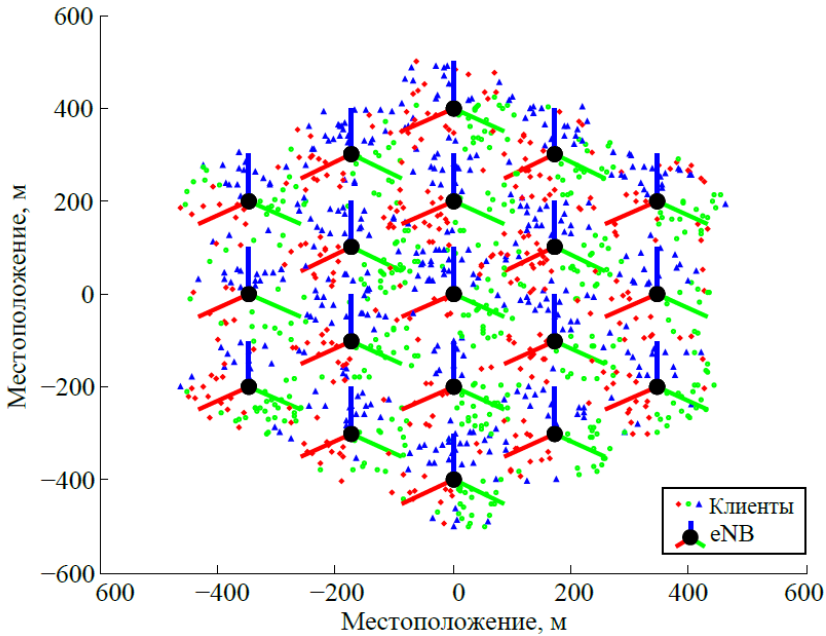


Рис. 10.2. Схема сети с 19 сотами

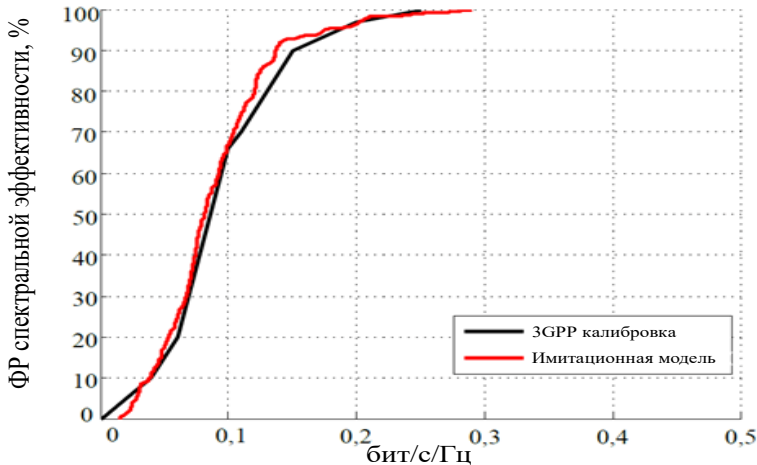


Рис. 10.3. Калибровка LTE: спектральная эффективность

Развертывание сети с поддержкой Wi-Fi Direct

Будем полагать, что все точки доступа Wi-Fi и устройства, не управляемые сетью LTE по одной и той же технологии, IEEE 802.11n.

Для того чтобы имитировать сетевое взаимодействие, эти устройства расположены вокруг их точек доступа. Точки доступа Wi-Fi могут быть расположены произвольно внутри моделируемой сети: в кафе, на транспорте, в магазинах и торговых центрах и т.д. Расстояние от устройств, не управляемых сетью LTE, до точек доступа ограничено потерями распространения и наихудшей допустимой модуляционно-кодированной схемой. Будем полагать, что эти устройства не изменяют положения, поэтому процедура хэндовера не рассматривается. Далее будем рассматривать случай, когда число устройств, не управляемых сетью LTE, равно максимально возможному числу активных D2D-соединений, поэтому эти устройства могут задействовать около половины эффективных ресурсов в условиях максимальной выгрузки от систем D2D.

Предполагается, что все Wi-Fi-соединения используют один и тот же частотный диапазон. Дополнительная информация о конфигурации сетей Wi-Fi содержится в табл. 10.1 и в документации драйвера с открытым исходным кодом Atheros. Для целей калибровки были использованы результаты публикаций по одноранговым сетям WLAN. При калибровке производительности Wi-Fi достигается почти идеальное соответствие (рис. 10.4).

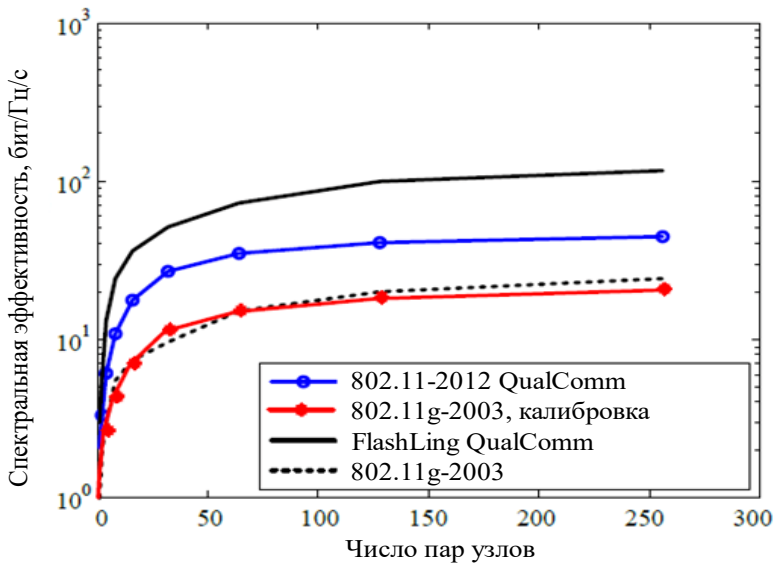


Рис. 10.4. Калибровка Wi-Fi: спектральная эффективность на квадратный километр

Будем рассматривать стандартный канал шириной 20 МГц для сетей Wi-Fi. Естественно, это можно легко экстраполировать на более общий случай, когда несколько каналов доступны в одно и то же время или агрегированный канал.

Таблица 10.1

Данные для имитационного моделирования

Параметр	Значение
Параметры ядра сети	
Мощность передачи	23 дБм IRP на интерфейс
Модель мобильности	RDM, 3 км/ч скорость
LTE радиointерфейс	
Модель затухания	ITU-R M.2135-1
Модель замираний	ITU-R M.2135-1
Доступ к среде	Round-Robin расписание
Контроль скорости	Целевое значение 15 дБ
Частотные ресурсы	10 + 10 МГц FDD
Режим сигнализации	2 из 20 кадров, 10 мс кадр
Радиооборудование	ITU-R M.2135-1, Tbl. 8-4
Конфигурация антенны	1×2 eNodeB
Wi-Fi радиointерфейс	
Модель затухания	Эмпирическая
Модель замираний	Только корреляция
Доступ к среде	CSMA/CA, -76 дБм
Управление скоростью	Целевое значение SINR 25 дБ
Частотные ресурсы	20 МГц TDMA
Режим сигнализации	RTS/CTS
Радиооборудование	Чувствительность -95 дБм
Конфигурация антенны	1×1 (простая антенна)

10.3.3. Сценарий применения гетерогенных сетей

Рассматривается типичный сценарий использования гетерогенных сетей (HetNet, Heterogeneous Network), представленный на рис. 10.5. В основе покрытия беспроводной сети лежит макро-БС с радиусом покрытия r_M . С целью повышения пространственного переиспользования ресурсов сети также используются M микро-БС, расположенные вокруг макро-БС, каждая из которых имеет радиус покрытия, равный r_m . Также в сценарии присутствуют мобильные пользователи с устройствами, поддерживающими технологию D2D, при этом некоторые из них могут выступать в роли активных D2D-партнеров, тем самым предоставляя возможности для подключений. Пользователи, желающие получить запрошенные данные,

могут быть подключены только к одному уровню рассматриваемой гетерогенной сети, т.е. только к макро-БС, микро-БС или же к активному D2D-партнеру, если последний находится в непосредственной близости, не далее чем на расстоянии r_D .

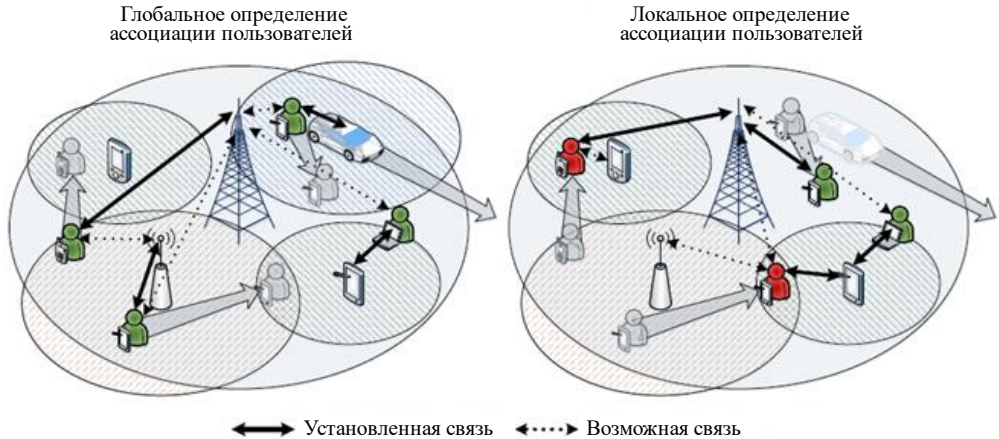


Рис. 10.5. Сценарий организации D2D-взаимодействий

В моменты оптимизации скорости пользователей макро-БС следует алгоритму, ориентированному на перераспределение беспроводных радиоресурсов всей сети с целью обеспечения каждого пользователя максимально возможным качеством обслуживания, т.е. применяется алгоритм, нацеленный на максимальную пропускную способность. Так как целью является анализ временных зависимостей поведения системы, в случае любых локальных изменений подключений пользователям дается возможность самим решать, к какому уровню сети они подключатся, максимизируя собственную пропускную способность. При этом могут наблюдаться случаи, когда выбранное подключение не будет оптимальным по причине отсутствия достаточного количества радиоресурсов, повышенного уровня интерференции или же отсутствия прямого соединения. Поскольку пользователи в первую очередь пытаются получить максимальное качество обслуживания, D2D-соединения, обладающие наибольшей пропускной способностью, получают больший приоритет, а макро-БС – меньший приоритет. Следует отметить, что возможны случаи, когда применение алгоритма, гарантирующего максимальную пропускную способность для пользователей, приведет и к улучшению энергоэффективности.

Будем предполагать, что пользователи запрашивают какой-либо популярный медиаресурс (к примеру, обмен медиа на массовом мероприятии, таком как футбольный матч или концерт). Подобного рода информация обычно кэшируется макро- и микро-БС, расположенными вблизи точки концентрации запросов. Также предполагается, что этот ресурс доступен $N - K$ из N пользователей. Таким образом, остается K пользователей, генерирующих запросы на получение данного ресурса. Все пользователи мобильны и движутся согласно модели движения с выбором случайного направления (RDM, Random Direction Model). Во многих работах показано, что данная модель движения дает качественно близкие результаты к более сложным, близким к реальным моделям движения. Так как макро-БС представляет собой центральную сущность, координирующую процесс подключения пользователей, ей доступна дополнительная информация о местоположении пользователей, карта интерференции и возможности подключения пользователей. Эта информация используется макро-БС для расчета оптимальных подключений и распределения радиоресурса для пользователей, которые сообщаются им с целью управления выбором оптимального подключения в рассматриваемой сети.

10.4. Результаты имитационного моделирования

Численный анализ системы проведен с помощью системы имитационного моделирования WINTERsim. Построен сценарий, в котором пользователи равномерно распределены в покрытии сети 5G с заданным набором технологий доступа и моделью движения RDM. Учтены три типа уровней HetNet сети: а) макросоты LTE eNodeB, б) микросоты 3GPP LTE и в) пользовательские устройства с возможностью взаимодействия D2D на основе Wi-Fi Direct. При этом 20% пользователей являются источниками данных. Макросоты LTE eNodeB с радиусом покрытия 500 м обеспечивают 100 ресурсных блоков, из которых 20 доступны непосредственно на БС, а остальные отданы микросотам.

Состояния беспроводных каналов оцениваются по уровню ОСИП на каждой из поднесущих с учетом затухания и замирания сигнала. Обнаружение и установление D2D-соединений управляется eNodeB, а передача данных осуществляется при помощи протокола Wi-Fi Direct.

Передаваемый поток данных моделируется для сервиса потокового вещания видео с качеством 720p и максимальной скоростью передачи данных в 4 Мбит/с. Оцениваются системные метрики: а) общая пропускная способность сети; б) энергоэффективность системы в зависимости от времени. Данные, использованные при моделировании, приведены в табл. 10.2.

Таблица 10.2

Исходные данные моделирования

Системные параметры	Значение
Радиус макро-БС	500 м
Радиус D2D-канала	30 м
Радиус фемтосот	50 м
Частота несущей LTE	2.6 ГГц
Частота несущей Wi-Fi Direct	2.4 ГГц
Мощность передачи абонента	23 дБм
Мощность передачи макро-БС	46 дБм
Мощность передачи фемтосоты	20 дБм
Время установления D2D-канала	1 с
Макс. скорость D2D-канала	40 Мбит/с
Количество запусков	500
Параметры приложения	Значение
Разрешение видео	720p, 30 кадров в секунду
Интервал ключевого кадра	Каждые 2 с
Макс. скорость	4 Мбит/с
Управление скоростью	CBR
Частота аудио	44.1 кГц
Скорость аудиопотока	128 кбит/с

Модель предполагает управление выбором наиболее подходящих подключений для пользователей с помощью специальной функции, которая учитывает пропускную способность пользователя и мощность передатчика по отношению к текущей загрузке сети.

На рис. 10.6 и 10.7 представлены результаты моделирования для трех параметров мобильности пользователей: 5,5 км/ч (малая мобильность), 45 км/ч (средняя) и 100 км/ч (высокая). Из графиков на рисунках видно, что общая пропускная способность сети и ее энергоэффективность быстро ухудшаются со временем с момента глобальной оптимизации системы вне зависимости от мобильности пользователей. В модели время перехода сети из оптимального состояния в состояние, управляемое пользователями, определяется

путем применения экспоненциально-взвешенного скользящего среднего (EWMA, Exponentially-Weighted Moving Average) с коэффициентом, равным $\gamma = 0,1$.

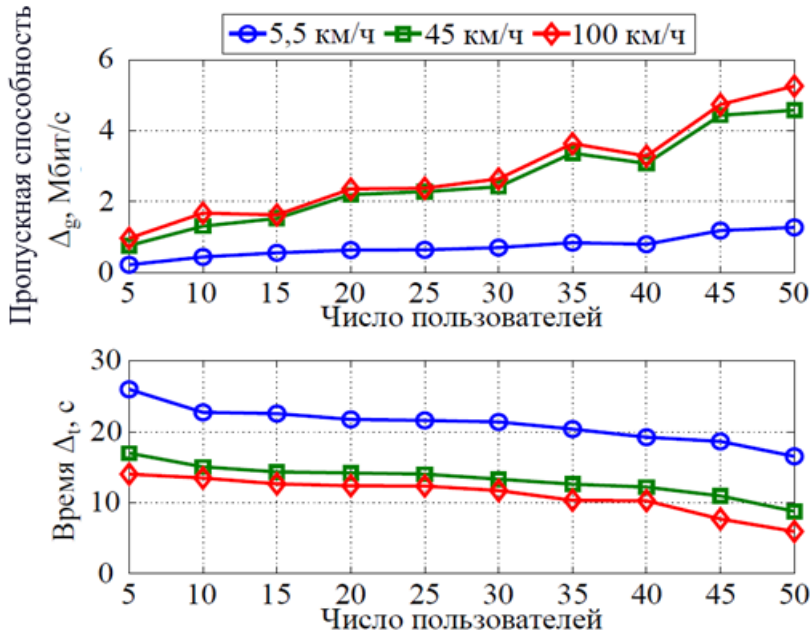


Рис. 10.6. Пропускная способность сети в зависимости от числа пользователей

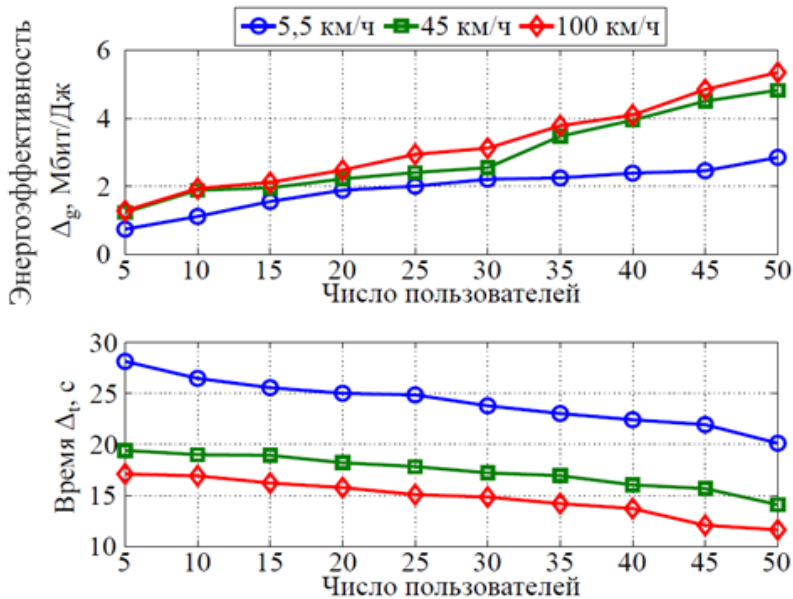


Рис. 10.7. Энергоэффективность в зависимости от числа пользователей

Заметим, что снижение уровня мобильности пользователей влечет увеличение времени оптимальной работы системы, так как сама система в целом начинает гораздо медленнее изменять свое состояние – высокоскоростные пользователи быстрее покидают свою зону обслуживания и у них появляется больше шансов встретить альтернативный (неоптимальный) источник данных. Не менее важно отметить, что разница между оптимальным и неоптимальным состояниями значительна. В частности, энергоэффективность может ухудшиться на порядок.

На рис. 10.8 представлена пропускная способность одного пользователя как функция времени между оптимизацией сети в зависимости от скорости перемещения пользователей, запрашивающих потоковые данные для $N = 50$ пользователей. При более высокой мобильности пользователей наблюдается быстрое ухудшение пропускной способности, и поэтому приходится гораздо чаще проводить глобальную оптимизацию сети. Аналогичный эффект наблюдается и для характеристик энергоэффективности сети (рис. 10.9). Заметим, что при определенных значениях скорости пользователей увеличение числа пользователей несущественно влияет на характеристики системы.

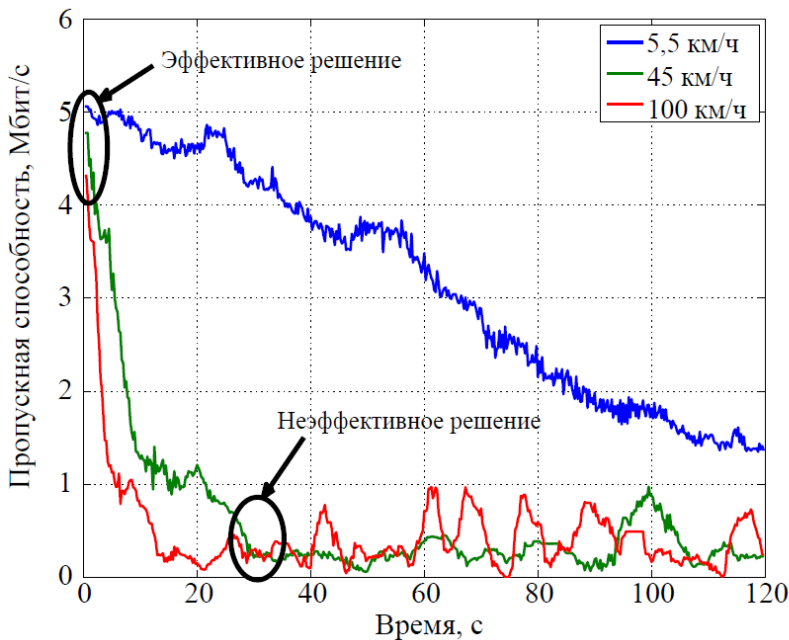


Рис. 10.8. Общая пропускная способность сети

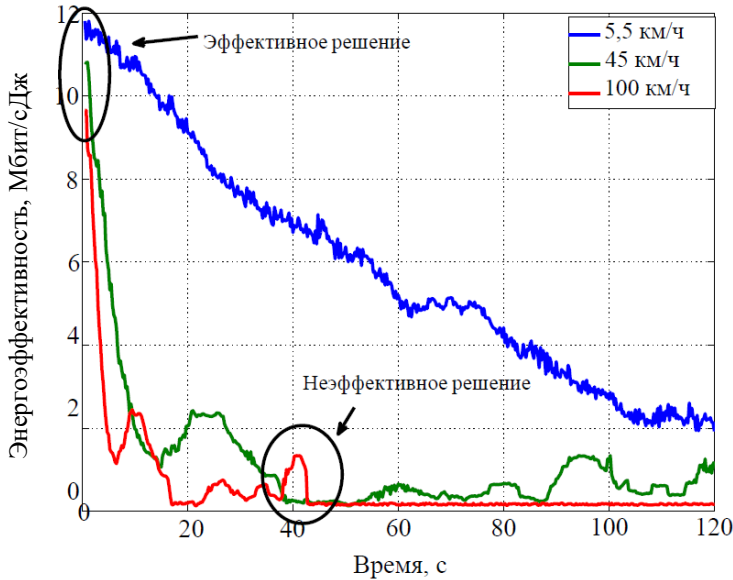


Рис. 10.9. Энергоэффективность сети ($N = 50$)

Подводя итог, можно сказать, что мобильность пользователей является ключевым параметром, оказывающим влияние на выбор времени между точками оптимизации системы. При малой мобильности (около 5 км/ч, что соответствует скорости пешехода) время ухудшения состояния сети достаточно велико и остается на уровне нескольких минут, а для высокой мобильности это время значительно сокращается. В результате этого выбор моментов времени, когда необходимо оптимизировать беспроводную сеть, становится основной проблемой для операторов сотовых сетей.

Значения константы отклонения состояния сети от оптимизированного состояния λ (безразмерная величина) могут быть получены путем решения уравнения $y(t) = Ce^{-\lambda t}$ относительно C и λ в моменты времени $t = 0$ и $t = \Delta t$. Решая уравнение, получаем $C = G_2$, $\lambda = (1/\Delta t) \ln(G_2/G_1)$. Результаты вычислений приведены в табл. 10.3.

Таблица 10.3

Константа отклонения от оптимизированного состояния сети

Скорость пользователей сети		
5,5 км/ч	45 км/ч	100 км/ч
0,01022	0,09001	0,17585
0,01233	0,08188	0,22426
0,01625	0,12330	0,18381
0,01235	0,13359	0,23081

Скорость пользователей сети		
5,5 км/ч	45 км/ч	100 км/ч
0,01207	0,13837	0,18455
0,01279	0,13927	0,23914
0,01172	0,15923	0,28507
0,01255	0,16279	0,28836
0,01400	0,21258	0,45732
0,01600	0,24040	0,59603

10.5. Аспекты практической реализации

В этом параграфе рассмотрим некоторые практические аспекты, позволяющие внедрять рассмотренную систему выгрузки трафика на D2D-соединения при поддержке технологии оператором сети. На сегодняшний день ключевые аспекты протокола, а также сервисная архитектура уже являются частью версий 3GPP Release 12 и 13. Однако, как правило, некоторые аспекты реализации не являются частью каких-либо спецификаций или стандартов, поскольку достаточно много требований всегда отдается на усмотрение разработчиков. Хотя большинство требований не очень сложные, но необходимо обеспечить их комплексное взаимодействие при реализации технологии D2D с поддержкой со стороны оператора сети, что может вызвать определенные затруднения.

10.5.1. Программная платформа

Рассмотрим программные компоненты для реализации технологии D2D с сетевой поддержкой, а также необходимые изменения в них.

Операционная система Android и IP-сети

Android, как любая Linux-базирующаяся система, позволяет устанавливать одновременные соединения с несколькими сетями более чем через один интерфейс. Однако даже если оба интерфейса 3GPP LTE и Wi-Fi активны, АУ обычно использует только один из них в качестве шлюза по умолчанию для IP-трафика. При этом можно использовать D2D-соединение только тогда, когда адресом назначения IP-пакета является Wi-Fi Direct адрес вызываемого абонента. Кроме того, из-за отсутствия резервных IP-адресов для

соединения Wi-Fi Direct необходимо использовать некоторый известный набор адресов. Желательно, чтобы можно было использовать общедоступный IP-адрес партнера в системе 3GPP LTE в Wi-Fi Direct.

Таким образом, одна из задач, которую необходимо решить для практического использования, заключается в создании решения, которое будет прозрачным для существующих приложений и, таким образом, облегчит совместимость разработанной технологии выгрузки трафика на D2D-соединения. По этой причине изменения физического уровня не могут принести желаемого результата, поскольку он в большой степени зависит от производителя оборудования. Аналогичная ситуация складывается и с канальным уровнем – для изменения логики маршрутизации в существующем оборудовании потребуется поддержка производителей оборудования, так как создание виртуальных интерфейсов приводит к значительным накладным расходам. Поскольку приложения существенно зависят от стандартных принципов их использования в Интернете и сетях беспроводного доступа, модификации на этом уровне также представляются затруднительными. Далее на сетевом уровне IP-адреса привязаны к физическим интерфейсам, но решения о переадресации принимаются независимо от интерфейсов, что позволяет создать независимое от интерфейса решение без необходимости внесения изменений в верхние уровни стека TCP/IP.

Конфигурация по умолчанию в системе Android позволяет использовать несколько маршрутов для шлюзов, но при этом шлюзы включаются в таблицу маршрутизации с различными метриками. Поэтому балансировка нагрузки не может быть осуществлена и только один маршрут для шлюза может быть использован в конкретный момент времени. В случае LTE (или любой другой сотовой сети) и Wi-Fi выбор маршрута шлюзом LTE предпочтительнее, когда ожидается соединение с сетью Интернет. Следовательно, изменение стоимости маршрута шлюза может привести к тому, что Интернет будет недоступен для всех приложений в мобильном устройстве, поскольку весь трафик будет перенаправлен на прямое соединение D2D.

Можно рассмотреть решение о введении маршрута, основанное на разрешении мобильному устройству маршрутизировать IP-трафик, как обычно, и вводить в таблицу маршрут для связи с пользователем для определенного пути D2D. Благодаря ядру Linux в системе Android это возможно путем изменения конфигурации

net.ipv4.conf.all.forwarding. После этого изменения мобильное устройство Android имеет возможности обычного маршрутизатора и, таким образом, может передавать пакеты от одного интерфейса к другому. Это позволяет отправлять IP-пакеты с помощью адреса источника, эквивалентного общему IP-адресу интерфейса LTE и месту назначения, а также адресу, совпадающему с используемым адресом D2D соединения.

Поскольку следует осуществлять выгрузку трафика только для одного партнера, можно вставить маршрут в таблицу маршрутизации таким образом, что общедоступный IP-адрес LTE будет возможно использовать через Wi-Fi Direct как частный IP-адрес. Данная вставка выполняется командой

```
ip route add PEER_LTE_IP/32 via PEER_WD_IP.
```

После этого весь трафик с IP-адресом назначения PEER_LTE_IP будет перенаправлен напрямую на Wi-Fi Direct интерфейс с кратчайшим путем, проходящим через соединение Wi-Fi Direct, как это и предполагалось. Когда такая связь больше не нужна, маршрут можно удалить следующим образом:

```
ip route del PEER_LTE_IP/32 via PEER_WD_IP.
```

Вставка и/или удаление выполняются приложением на стороне клиента, которое работает как системная услуга. Обе операции могут быть выполнены под управлением сервера D2D, но удаление может быть также осуществлено и на основе качества соединения, например, при неудовлетворительном значении ОСШ. В том числе при достижении определенного порогового значения уровня сигнала маршрут можно вставлять или удалять по мере необходимости. Пример таблицы маршрутизации во время введения маршрута в случае Wi-Fi Direct представлен на рис. 10.10. Показано, что трафиком можно управлять не только через Wi-Fi Direct, но и через другие методы не 3GPP доступа, например частные сети.

Рассмотренный метод введения маршрута в таблицу маршрутизации не требуется на обоих устройствах, участвующих в выгрузке трафика в D2D. Так, например, обратный трафик может быть обслужен через сотовый интерфейс, что может быть полезно в некоторых сценариях применения. Также введение маршрута не ограничено единственным одноранговым узлом: множество партнеров могут иметь свои собственные маршруты при условии, что им выделены различные частные IP-адреса, а сервер D2D вполне способен это обеспечить.

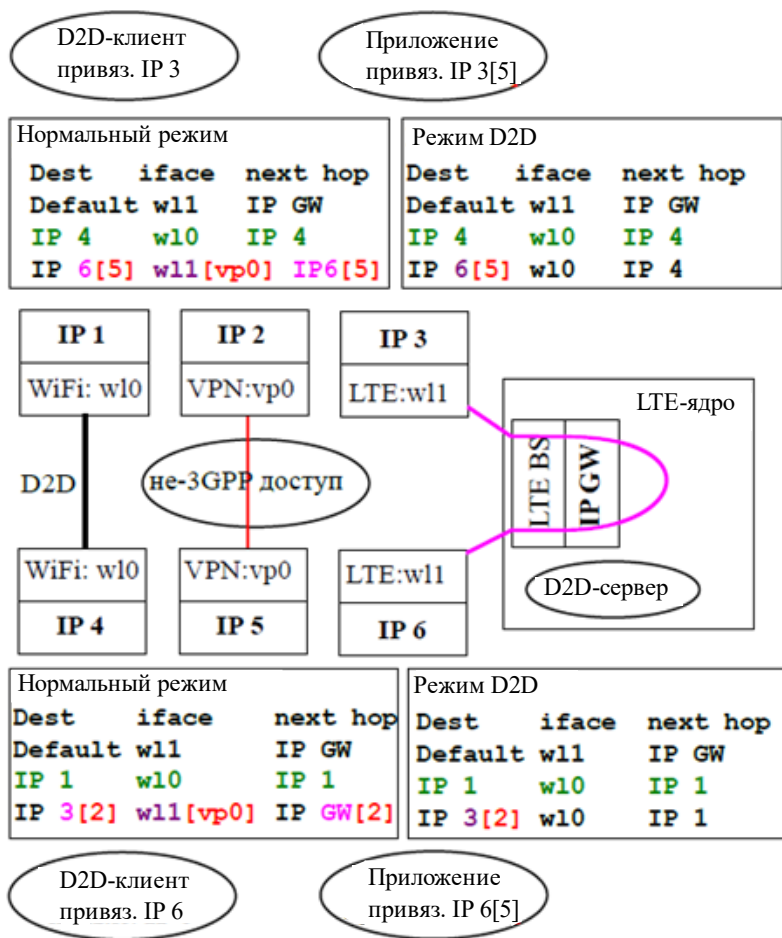


Рис. 10.10. Пример маршрута D2D в таблице маршрутизации

Инфраструктура и поддержка платформы

Был разработан прототип технологического решения для выгрузки трафика при поддержке сотовой сети, включающий в себя три основные функции.

1. Обслуживание на клиентской стороне работает в фоновом режиме на пользовательских устройствах.

2. База данных контента, которая поддерживает D2D-соединение и управление доступом.

3. Сервер D2D, который также выполняет функции прокси-сервера (ProSe, proximity-based services), управляет соединениями и координирует их.

База данных контента – это фактически веб-платформа, предоставляющая услуги обмена данными для пользователей, примерами являются социальные сети, такие как LinkedIn, YouTube. Сервер ProSe предназначен для оператора сотовой сети и/или поставщика услуг (ISP, Internet Service Provider). Следует отметить, что вместо предложенной архитектуры аналогичная услуга может быть предоставлена также, когда оба объекта управляются напрямую оператором связи. Но при этом перенос пользователей из известных социальных сетей на новую платформу может быть более сложной задачей. Примеры реализации приведены на рис. 10.11.

В качестве обоснования предложенной концепции приведем то, что все функции на стороне сервера были развернуты внутри разработанной облачной инфраструктуры. Две виртуальные машины были настроены так, чтобы выполнять функции регистра контента и сервера D2D соответственно. Контентный регистр был реализован как PHP-скрипт, обслуживаемый веб-сервером Apache, используя при этом API Google+. В предложенной реализации приложение является обычным веб-сайтом, что дает зарегистрированному пользователю возможность опубликовать свое намерение обменяться некоторым контентом или искать общие записи данных (например, торрент-трекер). Пользователь вводит только информацию, которая необходима для доступа к данным, а не к их местоположению, т.е. протокол совместного доступа и номер порта, в то время как настройке IP-соединения помогает сервер D2D ProSe. При этом внедрена новая схема адресации для D2D-соединений, а также новый идентификатор протокола в URI: d2d://. Устройства конечного пользователя могут быть легко сконфигурированы для того, чтобы интерпретировать этот тип протокола как запрос на обслуживание клиента, который, в свою очередь, способен связываться с сервером ProSe для того, чтобы разрешить использовать имя пользователя на уровне приложения для совместного доступа и соединения с одним из его устройств, на котором размещается контент.

Сервер ProSe реализован на Python как отдельное приложение с использованием https в качестве транспорта для управляющих сообщений. Система предполагает, что мобильное соединение для передачи данных пользовательского оборудования должно быть включено во время предоставления услуги. Предложенное решение показывает, что регистр контента и сервер(ы) ProSe легко интегрируются в беспроводной форме в существующую инфраструктуру веб-обслуживания и не требуют каких-либо решений, которые не вписывались

бы в строго определенную веб-парадигму. Протокол управления соединением D2D показан на рис. 10.12.

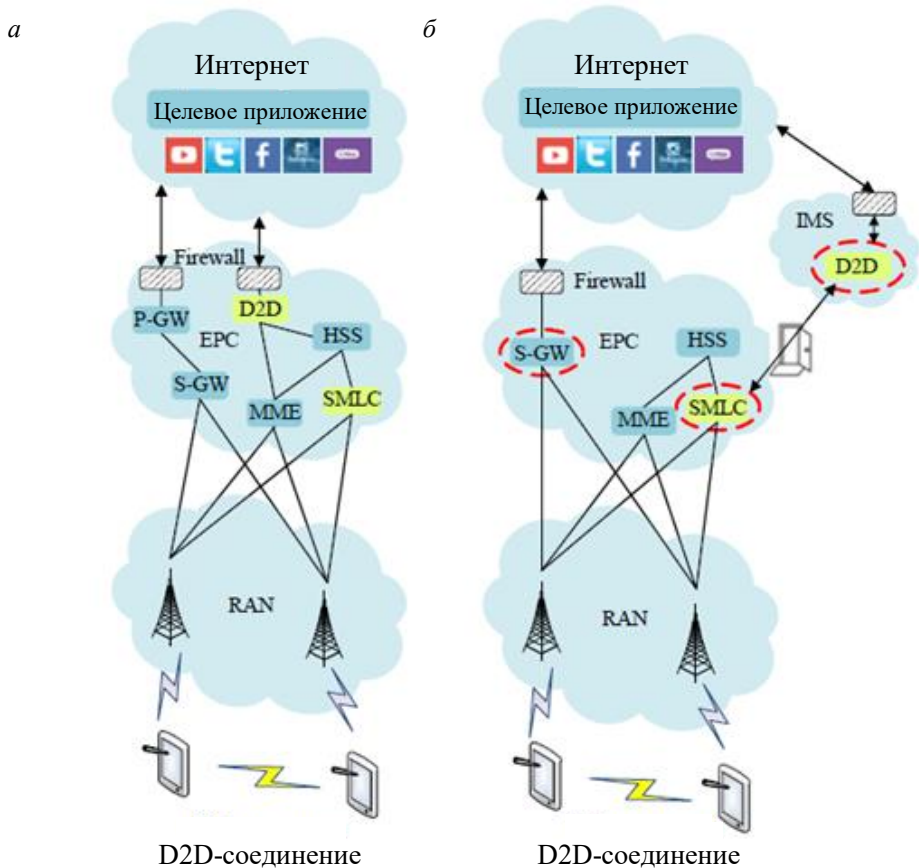


Рис. 10.11. Сравнение (а) стандартизированной и (б) предложенной архитектуры

В демоверсии использовались смартфоны Sony Xperia ZL с Android OS и Samsung Galaxy S8. Одной из основных функций, необходимых для пользовательского оборудования, была поддержка сотовых и Wi-Fi-соединений одновременно. Из-за ограничения потребления энергии большинство систем не используют такую функцию, поэтому необходимо было обойти систему управления услугами на базе Android Wi-Fi и напрямую взаимодействовать с драйвером Wi-Fi. Аналогично GNU Linux Android предоставляет необходимые инструменты: `wpa_supplicant` – интерфейс, управляемый с помощью утилиты `wpa_cli`. К сожалению, никакая прошивка не позволяет получить доступ к этим утилитам, даже для

разработчиков, поэтому использовалась послепродажная прошивка Cyanogenmod, которая может быть применена любым пользователем, кто владеет соответствующим устройством.



Рис. 10.12. Протокол управления соединением D2D

Одним из ключевых требований для мобильной платформы конечного пользователя является способность принимать входящие P2P-соединения через сотовый канал передачи данных. В связи с тем что большинство операторов используют частные пулы адресов IPv4 для назначения пользовательским устройствам и предоставления интернет-подключения через NAT и брандмауэры, обеспечение доступа к услугам на устройстве пользователя является нетривиальной задачей. Одним из возможных решений этой проблемы может быть использование IPv6, но мобильные операторы еще не полностью перешли на IPv6. И даже тогда маловероятно, что операторы удалят брандмауэры, поскольку это – эффективное средство борьбы с нелегальным P2P-трафиком. С учетом изложенного простейшим

протестированным вариантом для демонстрации технологии явилась инкапсуляция мобильных соединений обоих коммуникационных устройств внутрь туннеля VPN с общим VPN-сервером, тем самым располагая оба устройства в одной и той же IP-подсети. Однако из-за чрезмерного усложнения решения и больших накладных расходов на заголовки этот вариант мало подходит для практической реализации.

При фактическом развертывании очень важно разработать способ динамического конфигурирования брандмауэра оператора, поскольку нет достаточного числа адресов IPv4 для размещения на каждом пользовательском устройстве. Наиболее разумное решение заключается в том, чтобы позволить серверу D2D выбирать политику межсетевое экранирования непосредственно перед фактическим подключением с целью обхода функций NAT. Такое решение позволит устройствам обеспечивать прямое соединение, не прибегая к возможностям VPN, и передавать при этом данные без заголовка с очень небольшим влиянием на безопасность.

10.5.2. Результаты экспериментов

Рассмотренная реализация архитектуры системы D2D столкнулась с рядом практических проблем, не предусмотренных в стандартизирующих документах. Текущие архитектурные представления о сетях LTE не позволяют создавать функции сетевой поддержки так, как было бы естественно с технической точки зрения, и вместо этого приходится искать обходные пути для адаптации существующих систем и технологий. Далее рассмотрим ключевые решения, принятые на этапах разработки и испытаний, и опишем процесс создания экспериментальной сотовой сети на базе департамента телекоммуникаций Технологического университета г. Брно (BUT, Brno University of Technology) в Чехии.

BUT LTE-зона – полностью коммерчески пригодная версия для внедрения всех важнейших подсистем, входящих в современные сети 4G (рис. 10.13). Тестовая версия была сконфигурирована для обеспечения необходимых услуг при доступе к сети с коммутацией пакетов, таких как VoIP поверх инфраструктуры радиодоступа LTE и Wi-Fi. Ядро сети EPC обеспечивает высокую скорость передачи данных с соответствующим уровнем QoS, имея при этом возможность поддерживать до 100 тысяч одновременно обслуживаемых

пользователей. Для голосовых и видеослужб используется высокопроизводительная IP-мультимедийная подсистема (IMS, IP Multimedia Subsystem), поддерживающая как мобильный, так и фиксированный доступ, а также соединения с существующими сетями телефонной связи и системами видеоконференции. Упрощенный набор фаз работы D2D-соединений приведен на рис. 10.14.

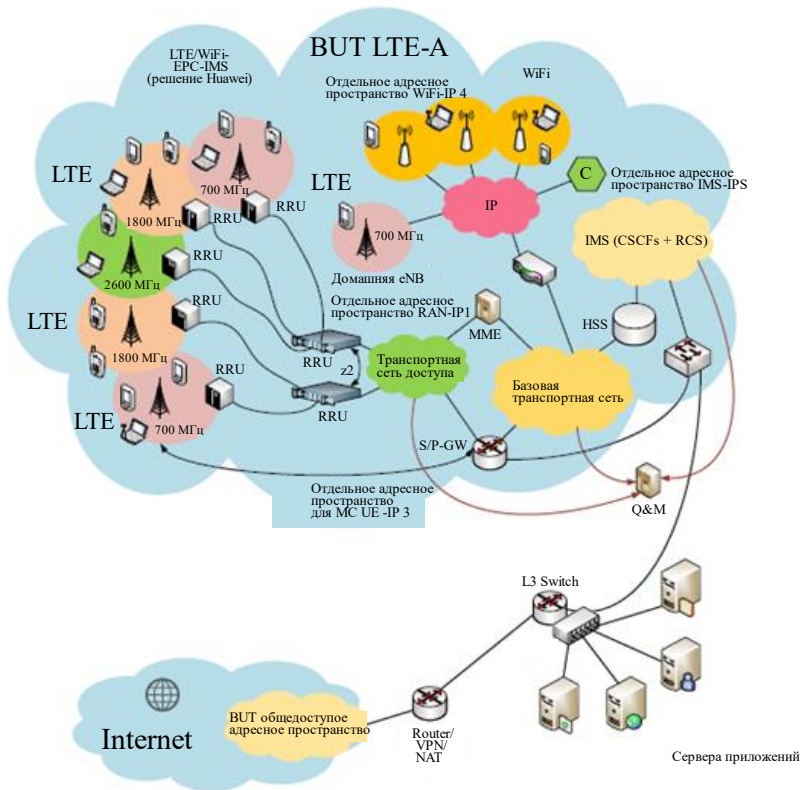


Рис. 10.13. Сеть LTE Технологического университета г. Брно, Чехия

Глобальная цель построения опытной зоны заключается в доказательстве положительного эффекта при совместном использовании мобильной сети и сети D2D, что подробно исследуется в данном параграфе. К сожалению, не все механизмы поддержки со стороны LTE для D2D были доступны в оборудовании опытной зоны. В частности, отсутствовала возможность использования сервера сервисного центра мобильной связи (e-SMLC), поэтому в качестве замены был разработан альтернативный интерфейс для получения информации

о местоположении устройства с использованием узла управления мобильностью сети сотовой связи (MME, Mobile Management Entity). Кроме того, сервер D2D был реализован как виртуальное устройство, тогда как окончательная реализация для коммерческих сетей связи оператора должна использовать более надежную технологию.

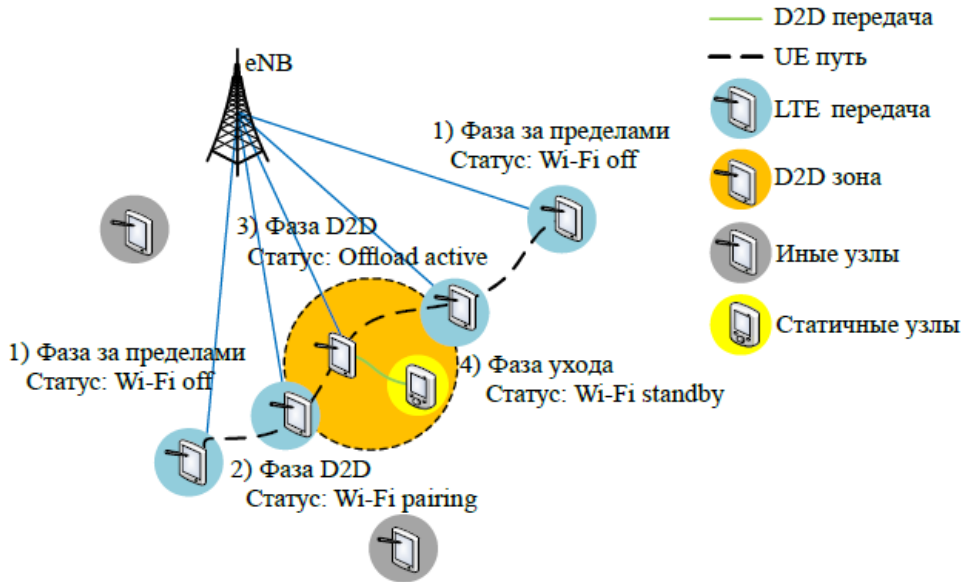


Рис. 10.14. Пример переключения между интерфейсами D2D

Методология оценки эффективности

Достаточно сложно определить методологию тестирования в модельной сети так, чтобы в дальнейшем все корректно функционировало в реальных сетях D2D. Для этого, в первую очередь, требуется организовать управление нагрузкой сотен пар соединений D2D, когда сотни клиентов используют в работающих смартфонах пользовательское программное обеспечение. Создание таких условий представляет очевидную сложность. Поэтому основные цели оценки эффективности решений были ограничены таким образом, чтобы дополнить результаты, которые были получены при аналитическом и имитационном моделировании. Поскольку пропускная способность соединения легко тестируется, наибольший интерес вызывали те ситуации, когда возникали локальные перегрузки, что потенциально могло бы уменьшить связность для D2D-технологий в реальных сетях. Установление соответствующих пределов производительности

и ограничений для технологии D2D и определение того, какие услуги могут быть наиболее подходящими для такого взаимодействия в обозримом будущем, является одной из важнейших целей создания опытной зоны. Для достижения этих целей предложены процедуры измерений, изложенные ниже в данном параграфе.

Одним из важнейших параметров при установлении соединений D2D является длительность интервала времени занятия соединения. Этот параметр имеет решающее значение, так как его низкое значение существенно снижает нагрузку в инфраструктуре LTE RAN, перемещая поток на альтернативное соединение, а также улучшает другие параметры, такие как энергоэффективность и качество восприятия. Время ответа при этом может быть разделено на несколько компонентов. Из-за небольшого размера сообщения, всего несколько десятков байтов, другие требования к QoS для сигнализации D2D легко достижимы и являются такими же, как и для любой сети доступа и, следовательно, не рассматриваются как ключевые показатели эффективности. Иные аспекты производительности соединения D2D, которые напрямую не связаны с инфраструктурой мобильной сети, выходят за рамки рассматриваемых оценок.

Изучая протокол для взаимодействия D2D, можно установить, что ключевым показателем является время отклика сервера, которое существенно зависит от нагрузки на соту. Численные результаты представлены на рис. 10.15, где можно видеть, что задержка не достигает предельных значений, если сота не загружена на 100%, но, когда это происходит, дисперсия задержки растет до недопустимых значений. На практике система работает неудовлетворительно, если задержка превышает 100 мс, когда сигнальный трафик D2D должен получить приоритет в планировщике LTE, если это возможно. Результаты измерения задержек при различных значениях нагрузки на соту приведены в табл. 10.4.

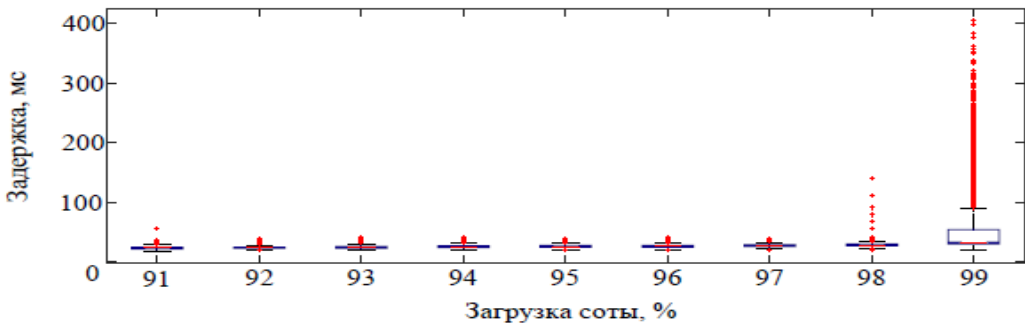


Рис. 10.15. Измерение задержки ответа сервера

Исходные данные имитационного моделирования

Нагрузка, %	10	50	90	99
RTT, мс	18	25	27	60
Фаза 2, мс	36	50	54	120
Фаза 3, мс	786	900	121	121
Фаза 4, мс	9	13	13	30
Время, мс	849	988	1.148	1.480

Результаты оценки качества восприятия QoE

С учетом приведенных выше показателей эффективности можно сделать следующие выводы по вопросам реализации Wi-Fi Direct с сетевой поддержкой LTE. Согласование соединения D2D является основным узким местом в системе, поскольку это может включать в себя множество циклов задержки из «конца-в-конец» (RTT, Round-Trip Time). В загруженной сети любая связь с сервером D2D может потребовать значительного времени, когда сигнальный трафик выбирает такой же канал, как канал с трафиком данных при отсутствии поддержки в мобильном API. Если пользователи высокомобильны, отслеживание их местоположения становится неточным, что приводит к ложному уведомлению о близости пользователя или, наоборот, позднему установлению соединения, из-за чего оно может быть потеряно. Несмотря на это, можно с уверенностью утверждать, что предложенный механизм D2D с поддержкой со стороны сети вполне пригоден для пользовательских устройств, перемещающихся со скоростью пешехода.

Принимая во внимание, что может существовать и некоторая дополнительная задержка, своевременное установление соединения даст пользователям уверенность в их контакте, в том числе и в случае, когда они перемещаются относительно друг друга. Можно сделать вывод о том, что в зависимости от качества сотовой связи полезность соединений D2D может отличаться. Однако даже с использованием современных технологий LTE задержки передачи информации могут быть до 10 раз больше, чем при использовании Wi-Fi, что означает, что технология D2D поверх Wi-Fi Direct является весьма конкурентоспособным решением.

10.5.3. Перспективы внедрения

Для операторов мобильной связи взаимодействие D2D может стать весьма эффективным механизмом для выгрузки трафика с БС, а также обеспечения поддержки социальных сетей посредством локализации предоставления услуг. Выше была рассмотрена реализация прототипа архитектуры технологии D2D с поддержкой со стороны сотовой сети, определены основные проблемы и потенциальные выгоды от прямого взаимодействия между пользовательскими устройствами. Далее рассмотрим наиболее важные выводы и уроки из проведенных экспериментальных исследований.

1. Успешная интеграция D2D и обеспечение совместимости с существующей архитектурой 3GPP LTE доказывают, что нет технических проблем, которые препятствовали бы операторам и поставщикам услуг предлагать приложения D2D для своих клиентов. Более того, некоторые из них могли бы существенно улучшить существующую инфраструктуру и предоставлять такие новые услуги, которые были бы невозможны без необходимости непрерывного отслеживания местоположения с помощью глобального позиционирования.

2. Успешное создание системы D2D с поддержкой со стороны сотовой сети с использованием ОС Android доказывает, что системы управления, эксплуатации и обслуживания можно будет достаточно просто реализовать для требуемых протоколов управления. Только некоторые платформы, которые не реализуют стандартизованные стеки, могут столкнуться с определенными трудностями с процедурами введения дополнительных маршрутов, необходимых для управления трафиком.

3. Качество восприятия для D2D-соединений будет превосходить лучшие сотовые соединения, по крайней мере в сетях 4G LTE, в пределах разумных расстояний между пользователями D2D (менее 50 метров в городских условиях). При этом Wi-Fi Direct позволяет передавать потоковое видео высокой четкости, а также приложения реального времени с высокой энергоэффективностью.

4. Отсутствие возможностей внешнего подключения и политик экранирования, имеющих в распоряжении операторов мобильной связи, в существенной степени ограничивает доступность соединений D2D там, где еще пока используется IPv4, поскольку число пользовательских устройств больше, чем доступных IPv4-адресов. Однако эти ограничения могут быть достаточно эффективно обойдены соответствующим конфигурированием сервера D2D и брандмауэров.

5. Успех всей концепции D2D во многом зависит от поддержки оператором позиционирования в сети, а также от готовности к межоператорскому взаимодействию. Это, как обычно, зависит от соответствующих стандартов консорциума 3GPP.

В целом можно сказать, что проблемы, выявленные в ходе реализации прототипа выгрузки трафика в D2D в условиях сетевой поддержки, можно решить оперативно на этапе внедрения на сетях операторов связи. Поскольку требования пользователей по новым услугам возрастают, можно утверждать, что рассмотренная архитектура D2D с сетевой поддержкой найдет в той или иной форме широкое применение на сетях связи при внедрении сетей четвертого и, в особенности, пятого поколения.

10.6. Дальнейшее развитие D2D

Одним из ключевых элементов в диверсификации топологии в сотовых сетях связи при использовании технологии D2D является возможность непосредственного взаимодействия пользователей друг с другом, когда в то же время существует управление со стороны сотовой сети. Сочетание высокой производительности, энергоэффективности и жесткой системной интеграции в качестве целей разработки системы D2D представляет собой исключительно большую область исследований и разработок, где пока, к сожалению, достаточно мало успешных работ.

Это исследование подчеркивает необходимость разработки нового набора технологий и протоколов, которые сделают сеть D2D неотъемлемой частью современной инфраструктуры сотовых сетей. Во введении были рассмотрены несколько аспектов проблемы, что вместе с исследованием особенностей системы продемонстрировало существенный потенциал системы *сотовая сеть – D2D Wi-Fi Direct* для обеспечения увеличения производительности существующих и будущих беспроводных сотовых сетей.

Предложения по архитектуре для решения D2D на базе 3GPP LTE и технологии IEEE 802.11 следуют лучшим известным практикам. При этом разработанная архитектура предоставляет различные возможности для интеграции с существующими социальными сетями и платформами распределения контента, обеспечивая требуемое масштабирование. Характеристики предложенной архитектуры были подтверждены при комплексном использовании аналитического и имитационного моделирования для оценки эффективности

на основе моделирования. Это дает глубокое понимание пределов функционирования сети D2D для нелицензируемого диапазона. С практической точки зрения представленные исследования сыграли ключевую роль в нескольких демонстрациях коммерческого уровня, а также в создании экспериментальной сети LTE. Были продемонстрированы возможности совместимости с существующим программным и аппаратным обеспечением, открыв тем самым перспективы для будущих коммерческих проектов.

10.7. Mesh-сети с использованием технологии блокчейн

Логическим развитием концепции D2D является создание многошаговых Mesh (multi-hop mesh networks) сетей, которые могут существовать независимо от сетей оператора связи или же использоваться оператором для разгрузки сотовой сети. В этом параграфе рассмотрены аспекты, связанные с использованием таких сетей для выгрузки трафика из сотовых сетей.

Реализация и внедрение Mesh-сетей в значительной степени зависит от готовности пользователей делиться своими ресурсами, исходя из соображений справедливости, анонимности и безопасности. Если участвующие пользователи решат покинуть сотовую сеть, это может серьезно ухудшить возможности Mesh-сетей на основе D2D, так как количество возможных путей для маршрутизации трафика, общая внутренняя и емкость сети сократятся. Поэтому одним из решающих факторов для повышения эффективности разгрузки сотовой сети при помощи локальных беспроводных Mesh-сетей является стимулирование пользователей делиться своими ресурсами в Mesh-сетях на основе D2D.

Одной из движущих сил для пользователей является то, что они могут сэкономить деньги из-за внедрения дополнительных скидок, так что фактическая цена, которую необходимо заплатить за использование беспроводного (или же какого-либо другого – вычислительного ресурса или же ресурса хранения данных), может соответствовать фактической ценности и интересу, который они придают услуге с точки зрения затрат энергии, пропускной способности или времени загрузки по сравнению с системами без стимулирующих факторов. Многие исследователи изучали потенциал стимулирования в распределенных сетях связи в контексте сетей 4G/5G/6G, стремясь улучшить общее качество работы системы и пользователей за счет более эффективного управления доступными ресурсами.

Для сетей на основе D2D наложенная сеть (оверлей) с поддержкой технологии блокчейн может потенциально уменьшить опасения пользователей относительно справедливости распределения ресурсов, привлекая их к участию в совместной работе. Такие стимулированные блокчейном Mesh-сети для локальной выгрузки трафика взаимовыгодны как для сотовых операторов, так и пользователей. Последние получают более высокое качество обслуживания за ту же цену и при этом вносят свой вклад в покрытие сети оператора. В некоторых сценариях пользователи могут получать дополнительное вознаграждение от оператора за использование своих ресурсов. Эти вознаграждения могут рассматриваться как инвестиции оператора в инфраструктуру сети. Действительно, с точки зрения оператора, Mesh-топологии по требованию временно расширяют доступность сетевых услуг, а также улучшают их качество без дополнительных капитальных затрат.

Технологии блокчейн

Появление первой технологии на основе блокчейн, криптовалюта Bitcoin, оказало заметное влияние на общество. Хотя криптовалюты стали инновационной альтернативой для финансового сектора, за ними всеми стоит концепция блокчейна. Благодаря надежно распределенным учетным записям, составляющим основу блокчейна, приложения, которые раньше могли работать только через доверенный центральный узел (например, банк, в контексте криптовалюты), теперь могут работать без него. Эта уникальная особенность расширяет сферу применения блокчейна далеко за пределы финансового сектора. В частности, недавно появились различные варианты использования блокчейн в беспроводных сетях. Ключевые варианты использования могут быть разделены на следующие категории: а) приложения машинного типа; б) функции безопасности и конфиденциальности; в) усовершенствование коммуникационных технологий; г) стимулирование пользователей. Далее мы кратко рассмотрим каждую из перечисленных категорий.

10.7.1. Использование блокчейн в Mesh-сетях

Основная проблема, вызывающая необходимость стимулирования пользователей, связана с их опасениями и потенциальными рисками, связанными с совместным использованием ресурсов —

несправедливое распределение ресурсов, проблемы конфиденциальности и эгоистичное поведение. В результате пользователи могут проявлять ограниченную готовность к совместному использованию ресурсов. Наложение сетей на основе блокчейна, развернутые поверх сетевой инфраструктуры, могут снять эти проблемы, предлагая криптографически надежную функциональность авторизации и аутентификации пользователей. Однако чтобы эффективно использовать эту технологию в беспроводных системах связи, она должна учитывать неоднородные вычислительные возможности различных устройств. По сравнению с традиционными технологиями (например, Bitcoin или Ethereum), недавно предложенные блокчейн-решения значительно более эффективны с точки зрения потребления ресурсов конечных систем и нацелены на обработку миллионов транзакций в секунду с использованием легких алгоритмов консенсуса и многоуровневой иерархической архитектуры.

Стандартные алгоритмы консенсуса в блокчейне опирались на методы доказательства выполнения работы (PoW, Proof-of-Work), требующие больших вычислительных затрат. Современные блокчейн-решения могут использовать значительно менее трудоемкий метод доказательства ставки (PoS, Proof-of-Stake). При обработке PoS узлы вносят ставки, чтобы гарантировать свою надежность, и достигают консенсуса с помощью византийского отказоустойчивого протокола. Такой подход является чрезвычайно легким и может выполняться на устройствах с ограниченными вычислительными ресурсами, требуя от узлов блокчейна только обработки транзакций и логики смарт-контрактов, что увеличивает скорость и энергоэффективность всей сети.

Другим недостатком классических блокчейн систем является значительный объем памяти, необходимый для постоянно растущих связанных списков (например, размер блокчейна Bitcoin на I квартал 2019 года составлял около 210 ГБ), что препятствует использованию технологии на устройствах с ограниченными ресурсами, таких как смартфоны и бытовая электроника. Эта проблема может быть решена с помощью концепции многоуровневого блокчейна, также называемого «блокчейн блокчейнов». Такая система управляется главным блокчейном, расположенным на вершине, который содержит общую информацию о протоколе, набор валидаторов и их ставки, набор активных в настоящее время блокчейнов нижнего уровня и набор хэшей их последних блоков. За главным блокчейном может быть несколько уровней вложенных блокчейнов. Хотя нижележащие

блокчейны могут иметь различные форматы адресов счетов и транзакций, все они удовлетворяют набору критериев совместимости для обеспечения взаимодействия между различными блокчейнами.

Многоуровневая архитектура блокчейна может быть использована для рассматриваемых динамических Mesh-сетей. Валидация блоков в главном блокчейне осуществляется оператором связи. Затем меньшее подмножество валидаторов назначается в каждый блокчейн более низкого уровня. Каждый валидатор может участвовать в нескольких блокчейнах одного или разных уровней. Таким образом, все алгоритмы валидации и консенсуса могут работать параллельно. Когда нижний блокчейн прекращает свое существование, баланс узлов сохраняется в блокчейнах более высоких уровней, обеспечивая сохранение баланса пользователей во времени. Например, временный блокчейн, используемый для поддержки работы Mesh-сети в определенном месте, может рассматриваться как блок блокчейна более высокого уровня.

10.7.2. Пример сценария стимулирования пользователей

В этом разделе рассмотрен и численно исследован сценарий, в котором использование блокчейна обеспечивает стимулирующую основу для разгрузки трафика в Mesh-сеть на базе технологии D2D. Предложенная система безопасности может позволить устанавливать доверенные маршруты в Mesh-сети через пользователей, находящихся поблизости. Однако это ставит перед исследователями важную задачу: каким образом стимулировать пользователей участвовать в поддержке Mesh-сети и процедуры выгрузки трафика?

В качестве одного из потенциальных решений для стимулирования пользователей мы предлагаем использовать специальные жетоны, которые могут быть потрачены на прямые действия (например, оплата фактических данных, которые будут переданы) или использованы для других услуг, предоставляемых оператором связи. В целом можно рассматривать первоначальное распределение жетонов как часть контракта с сотовым оператором. Пользователи, предлагающие свои ресурсы, выполняют функцию временных БС в инфраструктуре сети оператора связи, а также ретрансляторов для прямого подключения других пользователей. Поэтому такой обмен может напрямую поддерживаться оператором: пользователь, получающий выгоду от ретрансляции, тратит x жетонов, которые распределяются между участвующими k пользователями. При этом каждый из этих k

узлов получает $x/k + \sigma$ жетонов, где σ – дополнительный стимул оператора. Последнее можно рассматривать как инвестиции оператора в инфраструктуру сети.

Для практического обоснования рассматриваемой методологии стимулирования были опрошены 120 студентов двух университетов на предмет принятия данной концепции ретрансляции с помощью опроса по пятибалльной шкале Лайкерта (не согласен: 1, согласен: 5). Опрашиваемым был предложен следующий контекст. Пользователь подключен к сотовой сети, но имеет крайне низкое качество обслуживания. Однако он/она хочет поделиться данными со студентом того же университета. В данный момент нет возможности связаться с адресатом ни через Wi-Fi Direct, ни через сотовую сеть из-за ограниченного покрытия. Однако между этими двумя университетами находится еще один пользователь, который в принципе может передать данные. Предположим, что передача данных полностью безопасна и конфиденциальна. Обратите внимание, что передача данных разряжает батарею.

Анкета содержала следующие вопросы:

1. Хотели бы вы выступать в роли такого ретранслятора в случае низкого качества соединения инфраструктуры?

2. Хотели бы вы это делать, если вы можете использовать кого-то в качестве ретранслятора для своей передачи?

3. Хотели бы вы это делать, если вы будете вознаграждены каким-либо жетоном от пользователя, который просит о помощи?

Большинство респондентов положительно относятся к передаче своих данных через другие смартфоны в случае плохого сотового соединения (медиана: 3,95, стандартное отклонение: 1,22). В целом опрошенные нейтрально относятся к использованию своего смартфона в качестве ретранслятора для передачи таких данных (медиана: 3,04, стандартное отклонение: 1,13). Однако мнение большинства меняется на положительное, если им предлагают вознаграждение от пользователя, который запрашивает услугу ретрансляции (медиана: 4,21, стандартное отклонение: 1,08). Эти результаты подтверждают, что использование дополнительных операционных расходов оператора действительно может изменить отношение мобильных пользователей к совместному использованию ресурсов сети.

Для оценки влияния процесса стимулирования пользователей на характеристики выгрузки трафика в Mesh-сети была проведена оценка производительности на уровне системы с помощью среды имитационного моделирования. Исходные данные моделирования

приведены в табл. 10.5. Модель воссоздает плотный городской район, полностью покрытый сотовой связью LTE. Общая емкость сети мобильного доступа ниже, чем совокупный пользовательский трафик, генерируемый пользователями. Пользователи свободно перемещаются по рассматриваемой территории в соответствии с моделью движения RDM. Предполагается, что пользовательский трафик создается многопользовательской игрой AR/VR.

Сигнализация сети на основе D2D (функции обнаружения и установки соединения) осуществляется благодаря покрытию LTE, в то время как передача данных между узлами сети осуществляется по каналам связи в миллиметровом диапазоне частот, используя D2D-соединения технологии 5G NR. Напомним, что такие передачи могут быть заблокированы из-за движущихся препятствий, при этом учитывается динамическая блокировка каждого канала связи. В сценарии пользователи могут демонстрировать различное поведение в зависимости от уровня их мотивации: а) готовы действовать как ретранслятор все время, б) готовы действовать только с определенной вероятностью, в) не хотят ретранслировать, но стремятся потреблять ресурсы сотовой связи. Сравниваются следующие альтернативные варианты.

- *Базовое сотовое решение.* В этом случае подключение к сети доступно только через сотовые каналы связи, без поддержки сети на основе D2D.

- *Стандартная выгрузка в Mesh-сеть.* Пользователи имеют возможность установить путь в Mesh-сети, если выбранный ими партнер не может подключиться напрямую. Если какой-либо канал на пути перегружен, текущая сессия направляется через сотовую сеть, если ее радиоресурсов достаточно. В противном случае сессия прерывается и немедленно запускается процедура восстановления соединения.

- *Стимулированная выгрузка на Mesh-сеть.* В этом сценарии оператор предлагает дополнительное вознаграждение каждому ретранслирующему узлу, который участвует в разгрузке сотовой сети. В то же время пользователи, использующие услугу ретрансляции, платят жетоны своим помощникам. Чтобы подчеркнуть влияние слоя стимулирования, мы рассматриваем различные значения $\alpha = 0,5$ и $0,8$.

Исходные данные имитационного моделирования

Системный параметр	Значение
Рабочая частота	28 ГГц
Полоса пропускания	2,0 ГГц
Мощность точки доступа	2,0 Вт
Мощность передачи АУ	0,2 Вт
Антенная решетка АУ	16×16 элементов
Антенная решетка точки доступа	128×128 элементов
Модель распространения	3D кластерная модель 3GPP
Ослабление сигнала блокировкой	20 дБ
Скорость	400 Мбит/с
Размер соты	100×100 м
Максимальное количество хопов	3
Модель мобильности	Процесс Леви (параметром 1,5)
Количество пользователей	[10–100]
Параметр стимулирования, α	[0; 0,3; 0,5; 0,8]

Исследуемыми показателями эффективности выгрузки являются: а) совокупная пропускная способность сети, которая представляет собой скорость передачи данных, доступную при использовании как сотовых, так и Mesh-сетей, б) вероятность потери активной сессии и в) операционные расходы оператора, такие как количество жетонов, предоставленных оператором для стимулирования пользователей.

Ниже предложена визуализация времени жизни системы от инициализации сети до фазы стабильной работы. На рис. 10.16 представлен обзор случайным образом распределенных и пронумерованных узлов в пределах интересующей области. В сценарии зафиксировано положение узлов для лучшего представления работы сети. Связи между узлами представляют собой установленные D2D-соединения, а узлы без связей используют сотовую связь или не используют ее вообще в текущий момент времени. На рис. 10.16а показан момент установления D2D-соединений, где каждый узел имеет равный радиус, так как стимуляция еще не произошла. Далее, во время работы системы, некоторые узлы активно участвуют в работе Mesh-сети, предлагая свои ресурсы с большей интенсивностью. На рис. 10.16б представлен тот же самый сценарий, но некоторые узлы уже достигли большего вознаграждения, что отражено окружностями большего радиуса. Нагрузка на канал связи показана шириной линий, соединяющих узлы. На рис. 10.16в показано состояние системы после относительно длительного времени работы. Следует

отметить, что некоторые узлы могут иметь небольшую собственную нагрузку, но их каналы интенсивно используются.

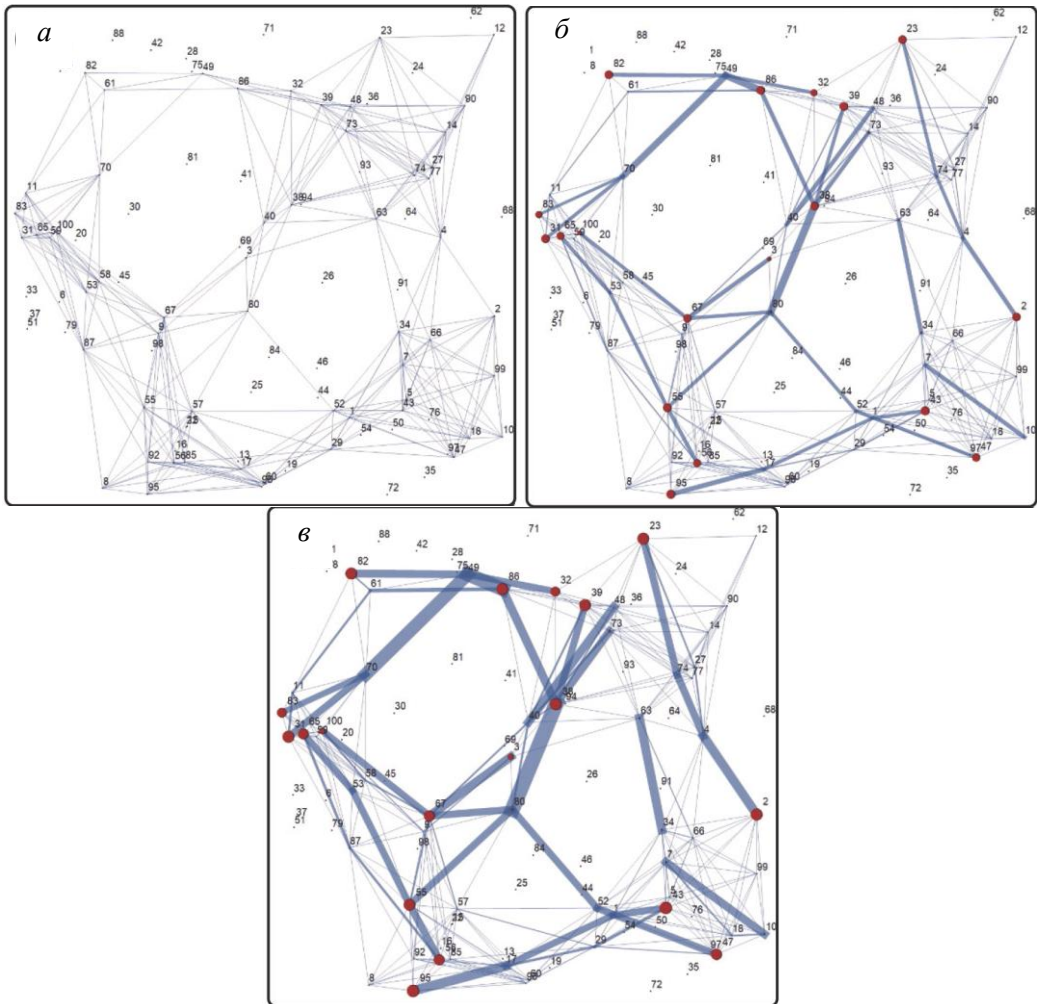


Рис. 10.16. Визуализация функционирования системы D2D
(пояснения в тексте)

Для того чтобы проиллюстрировать стимулирующий эффект технологии блокчейн, сначала продемонстрируем среднее количество узлов, участвующих в ретрансляции (рис. 10.17). Эта оценка основана на результатах опроса, представленных ранее. Как можно заметить, при увеличении значения параметра α , определяющего вовлеченность пользователей, число узлов, участвующих в ретрансляции, сначала увеличивается линейно с числом узлов в сети. Причина заключается в том, что большее количество узлов увеличивает разнообразие путей в сети, как можно видеть на рис. 10.18,

иллюстрирующем количество соединений в Mesh-сети. Таким образом, можно ожидать, что более высокие значения α будут положительно влиять на пропускную способность сети.

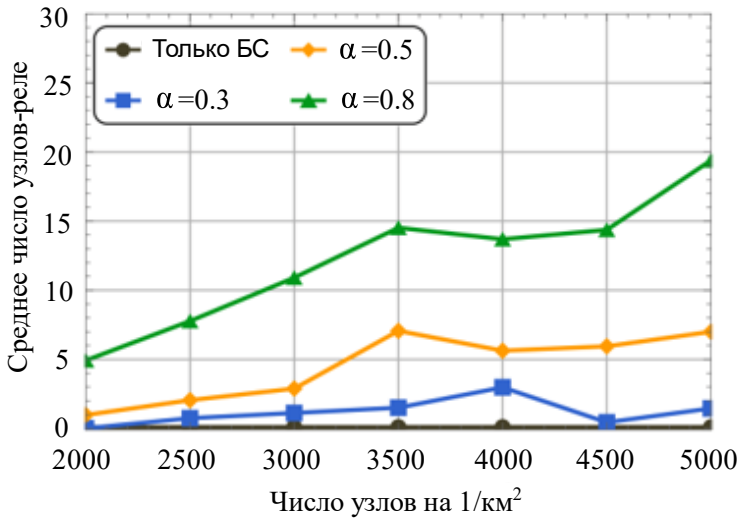


Рис. 10.17. Среднее число узлов, участвующих в ретрансляции

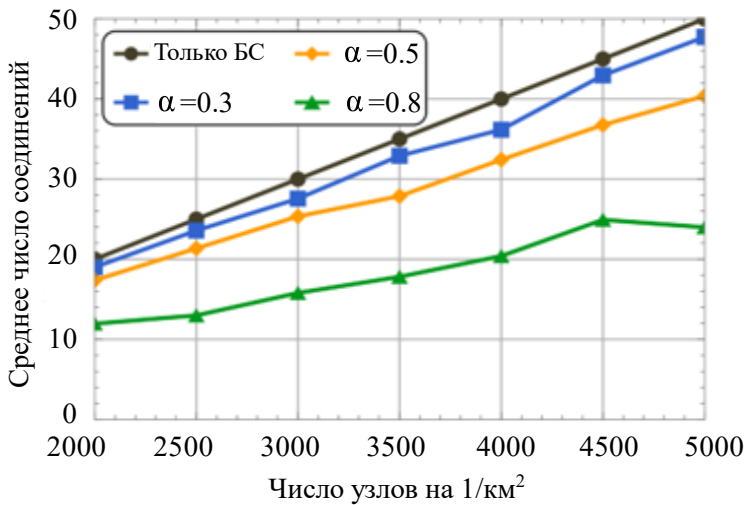


Рис. 10.18. Среднее число соединений без маршрута через Mesh-сеть

Увеличение числа вовлеченных в Mesh-сети пользователей, стимулируемых оператором через технологию блокчейн, способствует появлению новых маршрутов, как показано на рис. 10.18, что приводит к более эффективной разгрузке трафика и, соответственно,

увеличивает суммарную пропускную способность системы (сотовой и Mesh-сети), как показано на рис. 10.19. В отличие от микроволновых технологий D2D, в рассматриваемой Mesh-сети на основе миллиметровых волн пользователи также выигрывают от значительного снижения интерференции. В том числе наличие 16 антенных элементов приводит к ширине луча половинной мощности приблизительно $6,5^\circ$ как на передающем, так и приемном устройствах. Как результат, увеличение плотности соединений сети не оказывает существенного влияния на скорость соединения для такого класса сценариев.

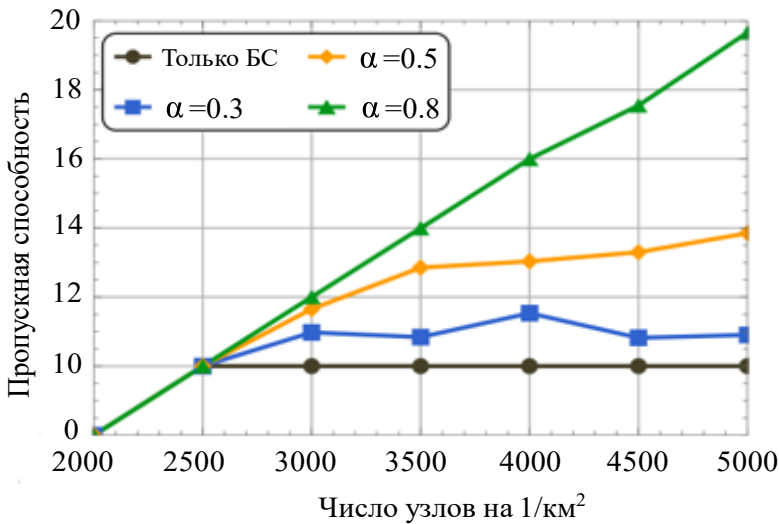


Рис. 10.19. Влияние стимулирования пользователей на пропускную способность системы

Как можно заметить, в базовом сценарии все пользователи полагаются на инфраструктуру сотовой сети, которая быстро насыщается по мере увеличения числа пользователей. При таком уровне загрузки вероятность потери активной сессии начинает расти, как показано на рис. 10.20. Увеличение параметра стимулирования привлекает больше пользователей в Mesh-сеть, увеличивая разнообразие путей, и, таким образом, эффективность выгрузки трафика увеличивается. Это разгружает инфраструктуру сотовой сети, а также снижает вероятность потери активной сессии. В абсолютных значениях вероятность потери сессии снижается примерно на 0,45 при использовании $\alpha = 0,8$ по сравнению с сотовой инфраструктурой с $\alpha = 0$. При этом увеличение пропускной способности сети достигает 200%.

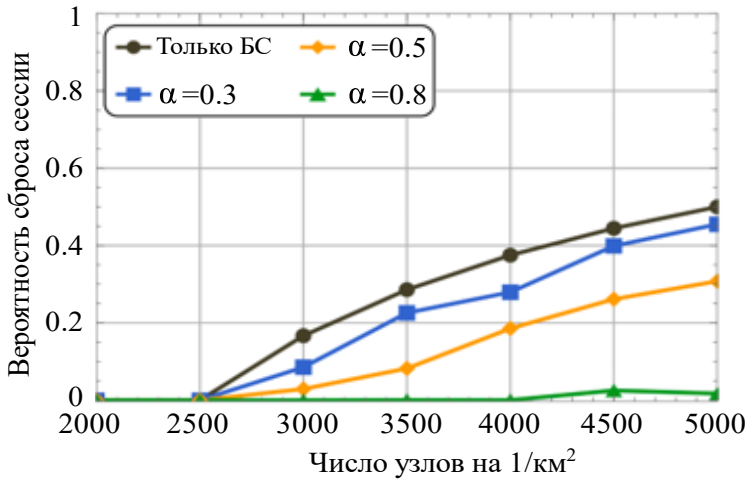


Рис. 10.20. Вероятность сброса активной сессии

Далее, чтобы охарактеризовать расходы оператора, мы оцениваем среднее количество накопленных жетонов пользователями, вовлеченными в Mesh-сеть, как показано на рис. 10.21. Можно заметить, что для выбранных параметров системы средние операционные расходы системы изменяются аналогично среднему объему совокупного трафика, передаваемого в Mesh-сети. Поставщик услуг может рассматривать эти расходы как часть операционных расходов на функционирование сети, заменяя капитальные затраты на инфраструктуру сети в тех местах, где скачки пользовательского спроса нечасты, но значительны.

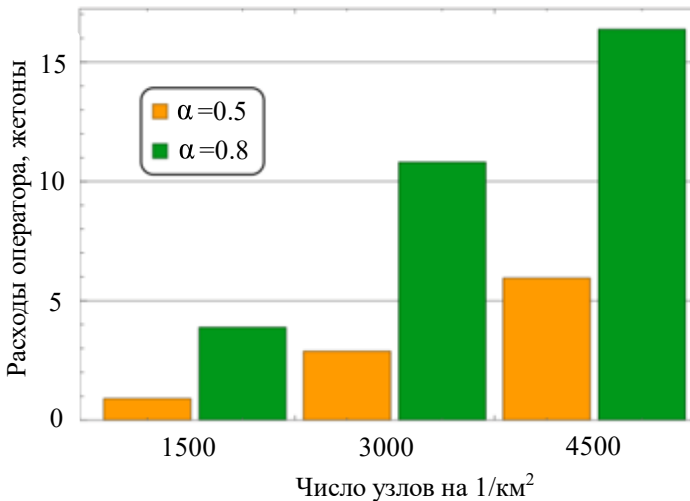


Рис. 10.21. Расходы оператора

10.7.3. Открытые исследовательские задачи

Использование технологии блокчейн для целей стимулирования пользователей – это перспективная концепция с большим потенциалом. Однако свойства, присущие распределенным системам, могут существенно повлиять на эффективность практических реализаций. Далее мы кратко охарактеризуем основные проблемы и трудности.

Тип блокчейна

Можно представить себе два концептуально различных варианта технологии блокчейн для целей стимулирования. Решения на основе частного блокчейна предполагают наличие единой инстанции для присоединения к сети и, следовательно, должны быть созданы оператором. Такой подход требует дополнительных межоператорских интерфейсов для согласования их действий и стратегий. В качестве альтернативы использование схем публичного блокчейна может потенциально обеспечить независимое от оператора функционирование системы и при этом гарантировать, что оператор участвует в построении сети как обычный узел.

Гетерогенность устройств

В мобильных блокчейн-системах узлы могут иметь сильно различающиеся ресурсы, включая вычислительные возможности, емкость батарей, памяти и пр. Здесь использование обычных алгоритмов консенсуса PoW может привести к значительным задержкам при подтверждении транзакций, поскольку некоторые узлы могут сильно отставать при выполнении необходимых вычислений. Кроме того, алгоритмы с интенсивными вычислениями могут негативно влиять на время автономной работы узлов. Поэтому для эффективной реализации технологии блокчейн перспективные системы могут либо потребовать новых облегченных схем консенсуса, либо использовать различные алгоритмы консенсуса, которые разумно распределяют задания на вычисления между подмножеством тщательно отобранных узлов сети.

Устойчивость к атакам

В обычных случаях использования устойчивость блокчейна к различным атакам обеспечивается либо размером самой системы

(публичный блокчейн), либо одним контролирующим органом (частный блокчейн), либо, что предпочтительнее, группой контролирующих органов (федеративный блокчейн). Широко известной атакой на публичные блокчейны, работающие по схеме консенсуса PoW, является атака большинства, когда пользователь, обладающий более 50% вычислительной мощности сети, может создать вредоносную копию блокчейна. Поскольку в мобильной системе количество участвующих узлов может изменяться со временем, необходимо либо предоставить эффективные средства для смягчения потенциальных последствий этих атак (например, реализовать алгоритмы консенсуса, обладающие свойством стойкости), либо ввести определенный уровень централизованной координации (например, использовать частный блокчейн).

Стратегия стимулирования

Развивающаяся технология блокчейн, применяемая для стимулирования пользователей, может как опираться на контроль оператора, так и нет. В первом случае управление стимулирующими выплатами (жетонами) имеет решающее значение для эффективной стимуляции пользователей. Когда оператор вводит жетоны в Mesh-сеть, общий эффект стимулирования принимает лавинообразную картину за счет возможности обмена заработанных жетонов на дополнительные услуги. Выбор типа стимулирующих выплат и его объем, вводимый оператором, сильно влияет на производительность Mesh-сетей и является важным вопросом для будущих исследований.

Интеграция в 5G/6G

Предложенный подход может быть использован в качестве инструмента для адаптивного управления сетью в дополнение к воздушным точкам доступа. Однако использование всего потенциала технологии блокчейн в сетях 5G/6G требует дополнительных усилий по стандартизации, связанных с новыми интерфейсами, включая сигнальные протоколы и разработку новых методов учета трафика с внешними системами стимулирования. В первую очередь необходимы усилия по стандартизации для определения терминологии, протоколов передачи сигналов, методологии учета трафика, а также специфических аспектов технологии блокчейн, например политики консенсуса, процедур майнинга, взаимодействия между блокчейнами разных уровней и т.д.

ГЛАВА 11

ЗАДАЧА СПРАВЕДЛИВОГО РАЗДЕЛЕНИЯ РЕСУРСОВ ГЕТЕРОГЕННОЙ СЕТИ РАДИОДОСТУПА

Системы сотовой связи 5G/6G, приходя на смену технологии 4G, призваны не полностью заменить последнюю, а дополнить ее как новыми функциональными возможностями, так и новыми типами радиоинтерфейсов. Таким образом, сети связи 5G/6G становятся гетерогенными, сочетая в себе несколько технологий доступа, каждая из которых должна использоваться в подходящем именно для нее случае. Так, например, технология 5G NR в миллиметровом диапазоне частот, а также будущий радиоинтерфейс 6G, работающий в терагерцовом диапазоне частот, будут удовлетворять потребности пользователей в сверхвысокой скорости на абонентском участке доступа, а технология LTE – обеспечивать непрерывность связи в условиях высокой мобильности. Слаженная работа интерфейсов должна обеспечиваться динамической системой контроля на стороне оператора и удовлетворять заданным параметрам качества обслуживания. В данной главе мы рассмотрим принципы одновременного использования нескольких интерфейсов на пользовательских устройствах и предложим методы оптимального использования ресурсов гетерогенных сетей.

3GPP-архитектура гетерогенных сетей рассматривается в параграфе 11.1. Критерии справедливого разделения ресурсов на беспроводном участке доступа обсуждаются в 11.2. Параграф 11.3 посвящен унификации известных критериев и выработке нового критерия, который учитывает специфику беспроводных каналов связи и позволяет динамически изменять требуемый баланс между справедливостью разделения ресурсов и общей пропускной способностью сети радиодоступа. Результаты, иллюстрирующие предложенную методику, представлены в 11.4. Параграфы 11.5–11.7 посвящены анализу гетерогенной архитектуры N-CRAN, стандартизированной 3GPP. Практическая реализация гетерогенных архитектур рассмотрена в параграфе 11.8.

11.1. Архитектура гетерогенной сети доступа

В данном параграфе обсуждаются концептуальные понятия для гетерогенных сетей и детально рассматриваются их компоненты. Кроме того, дан краткий анализ состояния исследований и их

развития в этой предметной области. Следует подчеркнуть, что концепция гетерогенных сетей начала формироваться в спецификациях второго поколения (2G) сотовых сетей в виде понятия о малых сотах. Впоследствии в спецификациях 3G эта концепция не получила развития, поскольку основные усилия тогда были сконцентрированы на оптимизации отдельных радиотехнологий (сотовых, беспроводных локальных сетей и т.д.).

Политика централизации управления для мультирадиосетей (например, WLAN и 3GPP систем) была введена в спецификациях четвертого поколения (4G) путем использования механизма центрального сервера, известного как функция обнаружения и выбора сети доступа (ANDSF, Access Network Discovery and Selection Function). Эта стандартизованная схема позволяет сотовым сетям включать в себя политики управления обнаружением и доступом в сеть для других, то есть не 3GPP-сетей. В частности, функция ANDSF может поддерживать пользовательское оборудование (UE, user equipment) для близлежащих не-3GPP сетей доступа (функция обнаружения). Также отметим, что UE определяет ограничения и предпочтения для близлежащих сетей доступа посредством функции межсистемной политики мобильности. Если пользовательское оборудование позволяет использовать более чем один радиointерфейс в конкурентной среде, выгрузка трафика может быть специфицирована посредством выработки межсистемной политики маршрутизации. В целом функция ANDSF реализует политики обнаружения и доступа, которые используются для длительных периодов функционирования, но не поддерживают динамическое управление и расписание использования радиоресурсов для мультирадиотехнологий. Ниже мы концентрируем внимание на концептуальных вопросах создания гетерогенных сетей, когда имеется больше возможностей для управления радиоресурсами, чем это возможно только с функцией ANDSF, когда пользовательское оборудование может одновременно взаимодействовать с любым числом радиотехнологий.

Каждая из радиотехнологий имеет свои собственные архитектуру и принципы передачи данных. Эффективное совместное использование мультирадиотехнологий является одной из ключевых особенностей HetNet, к реализации которой известны два подхода. Первый состоит в создании централизованного координатора, который обеспечивает полное управление всеми ассоциированными сетями радиодоступа. Однако необходимость соответствующих сигнальных процедур делает этот подход осуществимым только тогда,

когда спецификация каждой из радиотехнологий разрабатывается как часть некой взаимосвязанной структуры. Трудности возникают сразу же при попытке удовлетворить разнообразные требования каждой из вовлеченных в координацию радиотехнологий и создать сбалансированные сигнальные процедуры, не ухудшая эффективность индивидуальных сетей. Второй подход к координации сетей и систем гетерогенной сети состоит в обеспечении функциональной совместимости составляющих сеть индивидуальных радиотехнологий. В этом случае оптимальное управление гетерогенной сетью может стать невозможным из-за отсутствия полной информации о доступных сетевых ресурсах со стороны каждой из радиосетей доступа. При этом система координации радиотехнологий может быть реализована посредством межсетевых транзитных каналов, что может увеличить нагрузку на каналы управления. Альтернативный подход заключается в том, чтобы делегировать функцию выбора радиотехнологии пользователю оборудования путем обеспечения его всей необходимой информацией. Однако это решение может быть неоптимальным, поскольку предполагает достаточно большую гибкость как в решениях производителей мобильных устройств, так и в поведении конечных пользователей.

Кроме того, для беспроводных персональных сетей WPAN, беспроводных локальных сетей WLAN, беспроводных сетей масштаба города WMAN и беспроводных глобальных вычислительных сетей WWAN требуется единый подход к планированию покрытия. Базовые станции традиционных сотовых сетей, как правило, классифицируются по эффективности покрытия, например, макро- и микросоты. Подобным образом соты гетерогенных сетей могут варьироваться по размеру зоны покрытия, начиная с WPAN и WLAN (например, Bluetooth и Wi-Fi технологий) и заканчивая крупномасштабными зонами покрытия LTE WWANs, обычно имеющими радиус покрытия в несколько километров. Однако слишком велика разница в покрытии для различных технологий. Типовая макросота WMAN имеет зону обслуживания, которая в существенной степени зависит от условий эксплуатации, в то время как небольшие пико- и фемтосоты WLAN имеют покрытие радиуса от десятков до сотен метров.

В дополнение к архитектурным различиям гетерогенных сетей возможна также дифференциация по отношению к ожидаемому использованию отдельных радиотехнологий. Эта классификация относится к уровню приложений, поскольку сети 2G/3G ориентируются на передачу речи, а, например, WLAN предназначается в основном

для передачи данных. При таком подходе обеспечение пользователей услугами становится основным в системе оптимизации. Теоретически гетерогенные сети могут обеспечить поддержку для смартфонов LTE, ноутбуков и планшетов с Wi-Fi, мультирадиоустройств LTE/Wi-Fi, прямого взаимодействия *устройство–устройство* (device-to-device, D2D) через Wi-Fi и LTE-интерфейсы, а также масштабные машинно-ориентированные приложения.

Определим важнейшие характеристики модели построения гетерогенной сети, которая включает в себя две радиотехнологии (LTE и Wi-Fi), два типа покрытия (макро- и пикосоты для LTE и Wi-Fi) и простейшее приложение – передачу данных с использованием модели эластичного трафика. Сфокусируем внимание на случае, когда гетерогенная сеть управляется с помощью координирующего узла. Этот узел собирает всю текущую информацию о трафике пользователей и о доступности покрытия конкретной радиотехнологией в зоне обслуживания. Как только такая информация становится известна, координатор делает выбор наилучшей радиотехнологии для всех ассоциированных пользователей, а также сообщает им, какие именно ресурсы доступных радиотехнологий могут быть использованы. Кроме того, предполагается, что координирующий узел имеет полное знание о доступности услуг в каждой сети радиодоступа вместе с информацией о текущих характеристиках каналов для каждого пользователя.

Координатор может физически располагаться на стороне макро-базовой станции LTE, куда текущая информация с пикосот и точек доступа Wi-Fi может поставляться по прямым проводным и беспроводным каналам связи. В этом случае выделенный шлюз между точками доступа Wi-Fi и макро-БС LTE должен быть установлен на этапе ввода сети в эксплуатацию. Предложенная система управления функционирует на пакетном уровне, что дает далее возможность для пакетов, принадлежащих одному и тому же потоку, быть переданными и по конкурентному радиоинтерфейсу. При этом координатор гарантирует, что пакеты одного и того же интерфейса будут впоследствии собраны корректным образом. На практике это может быть реализовано с помощью механизма умного туннелирования между пользовательским оборудованием и узлом координатором. В доказательство этого в нескольких исследованиях уже продемонстрированы эти возможности, когда пользовательское оборудование использует такое умное туннелирование (основанное на архитектуре

OpenFlow) для разделения своего трафика динамически между LTE и Wi-Fi радиointерфейсами.

Отметим, что решение проблемы централизованного размещения радиоресурсов в мультирадиогетерогенных сетях может быть найдено благодаря использованию методов оптимизации. В этом случае задача централизации координирующего узла должна быть решена исходя из разделения общего ресурса системы, который необходимо оптимизировать с учетом возможных интересов всех подсистем. Следовательно, важнейшим подходом к оптимизации функционирования систем с множественными радиотехнологиями является выбор критерия оптимизации, который мог бы удовлетворять ожиданиям и операторов, и конечных пользователей.

11.2. Выбор критерия разделения ресурсов

При разделении ресурсов сети, как правило, достигается компромисс между двумя основополагающими принципами: а) равноправный доступ пользователей к ресурсам; б) максимизация общей пропускной способности системы. Возможность компромисса зависит от эффективности политики оператора сети при разделении ресурсов между пользователями. Для решения проблемы известны два подхода – стратегия *max-min* и стратегия пропорциональной справедливости (PF, *proportional fairness*).

Критерий max-min

Стратегия *max-min* является одним из наиболее распространенных подходов решения проблемы справедливого разделения ресурсов в сети с ограниченной пропускной способностью с множеством звеньев (каналов) $e = 1, \dots, E$ и N пользователями с множеством запросов $d = 1, 2, \dots, N$ на передачу эластичного трафика по доступным для каждого запроса P_d путям. Обозначив x_{dp} поток (бит/с) запроса d по пути p и $x_d = \sum_{p=1}^{P_d} x_{dp}$ как общий поток запроса d , приходим к известной формулировке *max-min* задачи – найти вектор разделения ресурсов сети $\mathbf{x} = (x_1, x_2, \dots, x_N)$, который при сортировке в неубывающем порядке является лексикографически максимальным среди

всех векторов разделения ресурсов, отсортированных в неубывающем порядке. При этом должны выполняться ограничения на потоки по звеньям сети $\sum_{d=1}^N \delta_{ed} x_d \leq c_e$, где δ_{ed} – индикаторная функция.

Напомним, что удовлетворяющий ограничениям вектор разделения ресурсов является *max-min* справедливым, если для каждого запроса d существует принадлежащее соответствующему пути насыщенное ($\sum_{d=1}^N \delta_{ed} x_d = c_e$) звено e , такое, что поток x_d на нем является максимальным, т.е. $x_d = \max \{x_d: d = 1, 2, \dots, N, \delta_{ed} = 1\}$. Для решения задачи необходимо найти вектор x , в котором все элементы расположены в порядке возрастания и который является лексикографически максимальным среди всех возможных векторов разделения ресурсов сети. Вектор x лексикографически больше, чем z , если существует индекс k – такой, что $x_i = z_i, i = 1, 2, \dots, k$, и $x_k > z_k$.

Таким образом, разделение ресурсов является *max-min* оптимальным, если невозможно увеличить ресурс, выделяемый конкретному пользователю без уменьшения ресурса, выделяемого любому другому пользователю. Заметим, что лексикографическая максимизация относится к задачам линейного программирования. Например, возможно использовать для решения задачи симплекс-метод. Простота решения оптимизационной задачи является привлекательным фактором, поскольку дает возможность использовать метод в реальном времени для управления ресурсами сети.

Пропорциональное разделение ресурсов

Традиционно во многих случаях для решения задачи справедливого разделения ресурсов применяется критерий

$$\sum_{d=1}^N \log x_d \rightarrow \max, \quad (11.1)$$

который, благодаря свойствам функции $f(x) = \log(x)$, имеет ряд преимуществ. Во-первых, когда необходимая для определенного запроса скорость передачи мала, вклад в (11.1) является отрицательным. Во-вторых, хотя функция и является монотонно возрастающей, ее производная $f'(x) = (\log x)' = 1/x \ln(x)$ показывает, что с увеличением x скорость роста уменьшается, что практически исключает вариант выделения ресурсов в больших размерах. Заметим, что основание логарифмической функции в (11.1) для решения задачи практически не имеет значения. По сравнению с *max-min*-критерием пропорциональное разделение обеспечивает лучшие результаты для

общей пропускной способности за счет предотвращения выделения очень больших или малых объемов ресурсов, что, однако, приводит к некоторой потере уровня справедливости.

Важно, что в классической формулировке, аналогично критерию $\max\text{-min}$, пропорциональное распределение ресурсов не обеспечивает гибкости в выборе точки баланса справедливого распределения пропускной способности. В частности, невозможно в явном виде регулировать баланс между справедливостью распределения ресурсов и общей пропускной способностью системы. Этот аспект существенен для поиска других целевых функций со свойствами, аналогичными логарифмической функции. При этом необходимо помнить, что указанная задача относится к классу задач выпуклого программирования, которые существенно сложнее по сравнению с задачами линейного программирования, что влечет за собой проблему выбора метрики справедливого пропорционального распределения для алгоритмов, работающих в реальном времени, особенно в сверхплотных гетерогенных сетях.

Как следует из весьма краткого анализа, сложная топология гетерогенной сети с уникальными свойствами отдельных компонентов радиотехнологий требует некоторого пересмотра выбора критерия справедливого распределения ресурсов. Например, в проводных сетях рассмотренные критерии штрафуют более длинные (по числу переприемов) потоки больше, чем короткие, в условиях конкуренции за ресурсы в пределах множеств путей. Поскольку пользователи с более высоким качеством канала могут эффективнее использовать беспроводные ресурсы, эквивалент критерия пропорционального справедливого распределения ресурсов для гетерогенных сетей должен быть основан на некоторой метрике, учитывающей текущее состояние канала. Критерий $\max\text{-min}$ пытается разделить ресурсы справедливо настолько, насколько это возможно, но это свойство, полезное для проводных сетей, может быть неподходящим для гетерогенных сетей. В самом деле, текущее расстояние между пользователем и БС в гетерогенной сети может существенным образом влиять на выбор схемы модуляции и кодирования (MCS, Modulation and Coding Scheme) и, следовательно, на эффективную скорость передачи, обеспечиваемую применяемой радиотехнологией. При использовании критерия $\max\text{-min}$ для пользовательского оборудования с худшими условиями канала компенсация должна осуществляться за счет

большого объема ресурсов, выделенного для передачи. Последнее, в свою очередь, предполагает, что попытка обеспечить полностью справедливое разделение приведет к резкому уменьшению пропускной способности всей сети, что может сказаться негативно.

Как альтернативу для обеспечения баланса между сетевыми возможностями и разделением пропускной способности, можно использовать основанный на полосе пропускания max-min-критерий справедливости. В этом случае, вместо того чтобы сразу справедливо разделить ресурсы между пользователями, предлагается сначала как можно более справедливо разделить совокупную полосу пропускания доступных радиотехнологий доступа. Очевидно, что пользователи, которые находятся ближе к обслуживающей БС, получают возможность использовать более высокую скорость передачи, чем те, которые расположены дальше с точки зрения использования MCS. Однако предложенный критерий не допускает малых скоростей для пользователей с плохими условиями канала, тогда как высокие скорости пользователей с хорошими условиями ограничены возможностями MCS, которые определяются конкретной радиотехнологией.

Таким образом, предложенная метрика выражается в компромиссе между пропускной способностью системы (выгода от динамического выбора MCS) и справедливым разделением скорости передачи данных. Далее, для обеспечения «обмена» пропускной способности на справедливость и наоборот, мы дополним базовую целевую функцию весовыми коэффициентами. Эти коэффициенты определяются подбором функций, использующих мгновенные значения спектральной эффективности пользовательского оборудования и, следовательно, позволяют гибко корректировать текущий режим функционирования гетерогенной сети. Ввиду сказанного предложенный критерий может рассматриваться как разновидность пропорциональной справедливой метрики для применения в гетерогенных сетях, где пользователи, расположенные на больших расстояниях от БС, получают больше или меньше ресурсов, в зависимости от выбранных весовых коэффициентов. Проблема формализована в рамках задачи линейного программирования, что важно для разработки алгоритмов, сложность которых соответствует размерности задачи, т.е. размеру гетерогенной сети и числу пользователей.

11.3. Критерий для гетерогенной сети радиодоступа

Трафик и модель сети

Как обсуждалось выше, тип трафика влияет на выбор целевой функции и, следовательно, на алгоритм решения задачи, и в задаче предполагается, что трафик является «жадным» и эластичным. Последний в спецификациях 3GPP также известен как трафик типа «полный буфер» или «насыщенный». Напомним, что «жадный» трафик занимает все доступные радиоресурсы, в то время как эластичность позволяет использовать ресурсы в доступном на момент применения объеме. Заметим, что эластичный трафик является типичным не только для сессий протокола ТСП, но и современных мультимедийных видео и голосовых приложений, реализованных посредством алгоритмов с динамической адаптацией скорости на уровне приложения.

Далее рассмотрим сценарий развертывания одной соты в трехуровневой гетерогенной сети HetNet, схематично изображенной на рис. 11.1. Сеть включает N пользователей, распределенных по трем уровням: $M_1 = 1$ – БС первого уровня (макросота LTE); M_2 – БС второго уровня (пикосоты LTE); M_3 – БС третьего уровня (точки доступа Wi-Fi). Всего в сети $M = M_1 + M_2 + M_3$ БС, а также один координирующий узел (например, агрегатор, суммирующий трафик после его разделения в различных радиотехнологиях). Предположим, что пользователь имеет доступ не менее чем к одной БС на каждом уровне HetNet; жадный и эластичный трафик; местоположение пользователя известно на координаторе; достижимая скорость передачи в сети конкретной радиотехнологии зависит от текущего значения MCS.

Считается, что координатор, выполняющий функции централизованного разделения ресурсов, интегрирован в оборудование БС первого уровня. Следовательно, решения о разделении ресурсов принимаются в момент входа пользователей в систему. Пример топологии гетерогенной сети с четырьмя пользователями показан в части В на рис. 11.1. Далее рассмотрим так называемое разветвленное разделение ресурсов, которое предполагает одновременное использование более чем одного радиоинтерфейса и разделение трафика между ними самим пользователем.

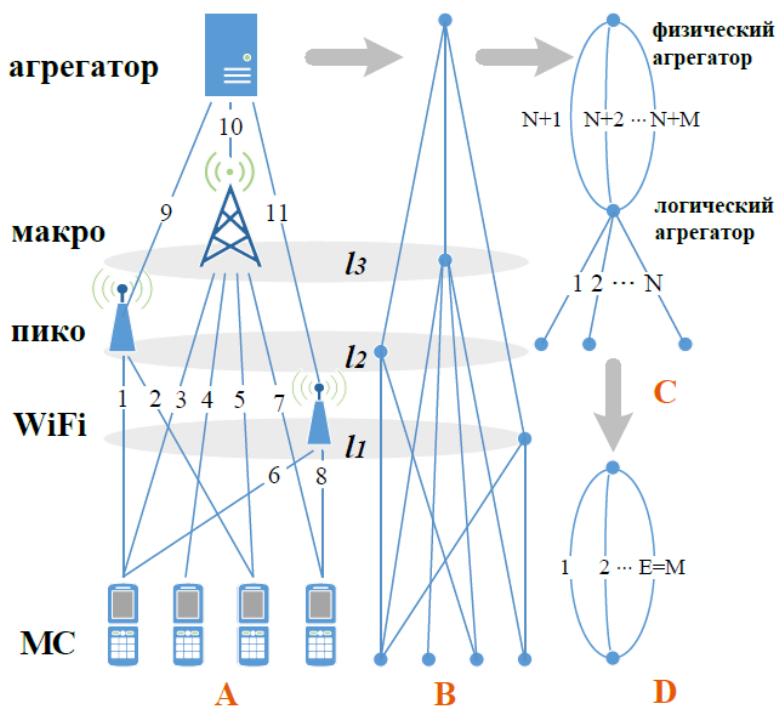


Рис. 11.1. Формализация системной топологии HetNet:

A – сеть; B – физические соединения в сети; C – логические соединения в сети;
 D – упрощенная модель, учитывающая участок доступа

Рассмотрим два дополнительных элемента системы – логический и физический агрегаторы. Физический агрегатор является элементом сети оператора связи. Логический агрегатор введен в рассмотрение для упрощения модели и не учитывает передачу трафика по каналам связи от БС до физического агрегатора. Отметим, что топология на рис 11.1(C) является избыточной, поскольку для соединения БС с физическим агрегатором желательно иметь равную или большую пропускную способность, чем это обеспечивается БС по радиоинтерфейсу. Например, пусть $e = 1, 2, \dots, E$ – набор каналов и c_e – их пропускная способности на физическом уровне. Из данных рис. 11.1 видно, что $c_9 \geq c_1 + c_2$, $c_{10} \geq c_3 + c_4 + c_5 + c_7$ и $c_{11} \geq c_6 + c_8$, причем каналы c_e , $e = 9, 10, 11$, не создают дополнительных ограничений по скорости передачи. На рис. 11.1(D) показаны топологии, где для простоты такие «лишние» каналы удалены.

В итоге наша окончательная модель топологии сети HetNet представлена на рис. 11.1(D), где все требования сконцентрированы между двумя узлами: логическим и физическим агрегаторами.

Число каналов в такой системе равно числу базовых станций $M = M_1 + M_2 + M_3$, в то же время их пропускные способности равны эффективным пропускным способностям соответствующих базовых станций. Некоторые из этих каналов (однако не все) являются общими для требуемых пар. В конечном счете пути для каждого требования являются известными, и они однозначно идентифицируют потоки данных, которые должны быть реализованы в системе для удовлетворения текущих потребностей пользователей.

Формулировка задачи

Рассмотрим разделение ресурсов на основе скоростей передачи данных. В этом случае спектральная эффективность (Spectral Efficiency, SE) пользователя не принимается в расчет, а все пользовательские узлы равны с точки зрения разделения полосы пропускания. Далее расширим формулировку, учитывая SE как часть задачи. В последнем случае ожидается более справедливое разделение ресурсов, при этом должен быть достигнут компромисс между справедливым разделением и общей пропускной способностью системы. Таким образом, вводим управляемое справедливое разделение путем введения весовых коэффициентов как функций спектральной эффективности. Напомним, что интерес представляют «разветвленные» решения, т.е. схемы, когда пользователем могут быть использованы одновременно все доступные радиотехнологии. Далее формулируем задачу разделения по критерию *max-min* для полосы пропускания и расширенному – для случая использования спектральной эффективности SE.

Обозначим N число пользователей в системе, которые определяют требования h_d (бит/с) к пропускной способности $d = 1, 2, \dots, N$ в соответствии с предположением о «жадном» эластичном трафике так, чтобы соответствовать критерию справедливого разделения ресурсов. Для каждого запроса введем набор путей

$$P_d = \{P_{d1}, P_{d2}, \dots, P_{dP_d}\}, d = 1, 2, \dots, N, \quad (11.2)$$

где P_d – общее число путей, доступных запросу d . На практике, P_d – это число БС на всех уровнях гетерогенной сети, с которыми ассоциирован соответствующий пользователь. В рассматриваемой топологии эти пути очевидно определяются множествами P_{dp} , $p = 1, 2, \dots, P_d$,

$d = 1, 2, \dots, N$, содержащими по крайней мере один элемент, связанный с агрегаторами.

Обозначим x_{dp} (переменную потока) часть запроса d по пути p , определим ограничения на переменные потока в виде

$$\sum_{p=1}^{P_d} x_{dp} = h_d, \quad d = 1, 2, \dots, N \quad (11.3)$$

и введем индикаторную переменную

$$\delta_{edp} = \begin{cases} 1, & e \cap P_{dp} = e, \\ 0, & \text{в противном случае.} \end{cases}$$

Тогда ограничения по пропускной способности можно записать в виде

$$\sum_{d=1}^N \sum_{p=1}^{P_d} \delta_{edp} x_{dp} = R_e, \quad e = 1, 2, \dots, E, \quad (11.4)$$

где R_e – суммарная скорость каналов связи между пользователями и БС. Таким образом, учтено, что пропускная способность ограничена, что гарантирует отсутствие перегрузок каналов связи.

Вводя вектор разделения ресурсов

$$\mathbf{h} = (h_1, h_2, \dots, h_N), h_d = \sum_{p=1}^{P_d} x_{dp}, \quad (11.5)$$

сводим задачу к лексикографической максимизации вектора \mathbf{h} с заданными ограничениями (11.3) – (11.7). Решение дает max-min-разделение ресурсов.

Теперь рассмотрим второй случай, когда учитывается спектральная эффективность SE пользователей, что требует модификации ограничений и для запросов по пропускной способности следующим образом. Пусть s_{dp} есть SE для потока с переменной x_{dp} для части запроса d по пути p , измеряемая в бит/с на один Гц (бит/с/Гц). Напомним, что для случая с SE справедливое разделение ресурсов для потока с переменной x_{dp} зависит от MCS, используемой в текущий момент времени. В соответствии с обеспечением max-min-разделения ресурсов различных MCS между пользователями необходимо распределить большую полосу пропускания для оборудования с MCS с низкой скоростью передачи данных. В этом случае ограничения для переменных потоков имеют вид

$$h_d = \sum_{p=1}^{P_d} s_{dp} x_{dp}, \quad d = 1, 2, \dots, N, \quad (11.6)$$

где произведение $x_{dp} s_{dp}$ отвечает за разделение ресурсов.

Ограничения по пропускной способности остаются такими же, как в (11.4), со скоростями R_e в полосе пропускания B_e , измеренной в герцах. Задача также состоит в лексикографической максимизации вектора \mathbf{h} . Задача с ограничениями (11.6) обеспечивает max-min-разделение скоростей передачи для пользователей. В этом случае выделяемая им часть полосы пропускания изменяется пропорционально весам соответствующих MCS-пользователей.

Поскольку проблема заключается в справедливом разделении ресурсов при переменной пропускной способности, можно в классической постановке заменить неравенства на равенства с ограничениями на пропускную способность. Заметим также, что для ресурсов, выделяемых по расписанию в базовых станциях LTE и основанных на дискретном разделении ресурсов по ресурсным блокам, полученное в решении задачи разделение ресурсов является приближенным с точностью до целых значений числа РБ.

Гибкое управление ресурсами

Для управления балансом между справедливостью разделения ресурсов и суммарной пропускной способностью сети рассмотрим возможные модификации сформулированной задачи оптимизации. Для достижения такого эффекта в простейшем случае достаточно модифицировать значения спектральной эффективности пользователей. Например, искусственно уменьшая спектральную эффективность пользователей, можно добиться эффекта увеличения выделяемых ресурсов для этих пользователей, что приведет к повышению суммарной пропускной способности системы, так как эти пользователи заведомо находятся в более выгодных радиоусловиях. Напротив, искусственно увеличивая спектральную эффективность пользователей, которые находятся в хороших радиоусловиях, можно увеличить справедливость.

Таким образом, основной задачей является выбор функции управления, для которой рассмотрим две альтернативы: $f(x) = x^\alpha$ и $f(x) = \alpha^x$, где x – спектральная эффективность (рис. 11.2). В первом варианте полностью справедливое разделение ресурсов достигается при значении $\alpha = 1$, а максимальная пропускная способность системы при $\alpha \rightarrow 0$. Выбор второй функции, которая, например, применяется при сжатии диапазона приема сигнала в радиочастотных усилителях, может быть предпочтительным ввиду наличия точки баланса при $\alpha = 1$ (см. рис. 11.2). Здесь значения $\alpha < 1$ позволяют

увеличить общую пропускную способность системы. Значение $\alpha = 1$ обеспечивает разделение ресурсов по скорости передачи, что также способствует увеличению пропускной способности системы. Для $\alpha > 1$ схема имеет тенденцию к улучшению справедливого разделения ресурсов по полосе пропускания. Отметим, что рассмотренные модификации оставляют решение в том же классе оптимизационных задач (линейное программирование), поскольку уточняют лишь весовые коэффициенты.

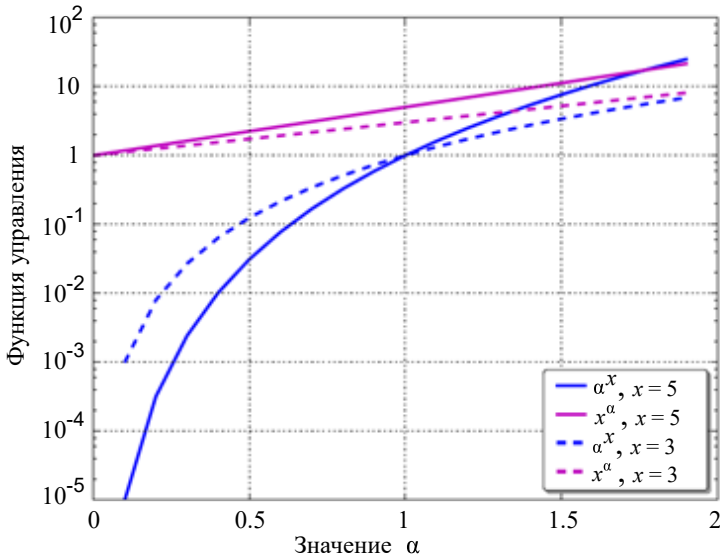


Рис. 11.2. Сравнение функций управления спектральной эффективностью

Алгоритм решения задачи

Решение задачи max-min является расширением алгоритма для задачи нахождения единственного пути. Для разработки могут быть использованы встроенные средства MATLAB, поскольку задача классифицируется как задача линейного программирования.

Такой алгоритм используется на БС макросоты LTE каждый раз, когда новый пользователь поступает в систему или существующий пользователь изменяет состояние своей связности с радиотехнологией гетерогенной сети. Причина в том, что max-min-разделение ресурсов для всех пользователей зависит от доступности беспроводных интерфейсов. При этом предполагается, что текущие сессии не прерываются вследствие прихода новых пользователей, поскольку

доступные для разделения ресурсы пересчитываются. По этой же причине нет необходимости решать задачи разделения ресурсов, когда не наблюдается изменений в связности пользователей. В заключение следует подчеркнуть, что, поскольку задачи относятся к классу задач линейного программирования, вычислительная сложность для БС макросоты LTE или 5G NR вполне приемлема даже для сценариев сетей со сверхвысокой плотностью.

Эвристические подходы

Рассмотрим дополнительно две эвристические стратегии разделения ресурсов, которые могут служить в качестве рейтинговых тестов для предложенных выше критериев разделения радиоресурсов гетерогенной сети.

Первый подход – централизованный, когда пользовательское оборудование независимо решает, сколько ресурсов ему требуется на каждом сетевом уровне. Важно отметить, что в этом случае пользователь не знает, сколько ресурсов доступно на каждом уровне сети и как можно конкурировать за каждый доступный радиоинтерфейс. Второй подход – распределенный с сетевой поддержкой, когда беспроводная сеть помогает пользовательскому оборудованию сделать выбор интеллектуальным способом, обладая информацией о текущих условиях радиоканала, а также об общем количестве пользователей в системе. Такие алгоритмы известны начиная с 3GPP Release 12 и последующих для LTE и 5G NR. Ниже рассматривается упрощенный вариант алгоритма, в котором пользователям предоставляется информация только о предпочтительном пороге ассоциации с каждым уровнем HetNet. С помощью соответствующей регулировки этого порога сеть может уменьшать/увеличивать эффективную дальность покрытия пикосот LTE/NR и точек доступа Wi-Fi для выгрузки трафика от/к макросоте LTE/NR.

Предполагается, что пользовательское оборудование может одновременно подключиться только к одной точке обслуживания, например Wi-Fi или пикосоте LTE/NR. Зная об этом ограничении, сначала пользователь пытается подключиться к точке доступа Wi-Fi, что является типовым решением для высокой скорости передачи данных. Если обеспечиваемый сетью порог ассоциации не достигается на Wi-Fi, БС пикосоты LTE/NR является следующей в очереди попыток ассоциации с сетью. В итоге, если ни одна из попыток не

заканчивается успешно, пользователю остается установить соединение с БС макросоты LTE/NR. Как правило, назначенные пороги могут быть использованы индивидуально для каждой БС. Вначале выбираем порог, основанный на средних значениях SE (s_{dp}) пользователя на каждом уровне сети. Далее рассматриваем различные типы порогов, выбирая компромисс между справедливым разделением ресурсов и общей пропускной способностью системы. В дальнейшем будем называть этот эвристический алгоритм «Wi-Fi-предпочтительным».

11.4. Численный анализ

Параметры моделирования

Типовой сценарий HetNet, который мы рассматриваем, включает следующие системные составляющие: БС макросоты, называемая eNodeB, несколько пикосотовых БС eNodeB, точки доступа Wi-Fi, а также пользовательские узлы с мультирадиотехнологиями (Wi-Fi и LTE/NR). Пико-соты, точки доступа Wi-Fi и пользователи размещены в зоне покрытия макросотовой БС. Предполагается, что каждый пользователь имеет несколько активных соединений, что зависит от доступных радиотехнологий и уровня принимаемого сигнала.

В соответствии со спецификациями 3GPP для пользователей макросот зависящие от расстояния потери определяются моделью 3GPP для каналов в городской среде. Для точки доступа Wi-Fi и пользователей пикосот LTE/NR потери определяются с помощью модели 3GPP для микрогородских каналов со средней высотой зданий $h = 20$ м и средней шириной улицы $W = 20$ м. Для детального моделирования потерь распространения сигнала используются модели медленного и быстрого замирания сигналов (см. главу 2).

Основные параметры моделирования приведены в табл 11.1.

Таблица 11.1

Значения параметров модели

Параметр	Значение
Частоты LTE/Wi-Fi	10 МГц FDD/20 МГц
Конфигурация сценария	1 макросота, несколько LPN
Модель канала макро/LPN-UE	ITU UMa/UMi
Усиление антенны макро/LPN	17/6 дБ
Мощности макро/пико/Wi-Fi	43/23/20 дБм

Параметр	Значение
Максимальная мощность UE	23/20 (LTE/Wi-Fi) дБм
Контроль мощности LTE/Wi-Fi	макс. мощность
Высота антенны UE/макро/LPN	1,5/25/10 м
Потери кабеля антенны	5 дБ / 0 дБ
Модель трафика	эластичный
Тип разделения LPN/UE	равномерный
Мин. дистанция LPN/UE-макро	< 75/35 м
Мин. дистанция LPN/UE-UE	< 40/10 м

Проверка аналитических условий

Для верификации предположений об интересующих метриках (включая пропускную способность и справедливость разделения ресурсов) необходимо протестировать модель в близких к реальным условиям гетерогенной сети. С этой целью использована система имитационного моделирования WINTERsim, которая учитывает все существенные детали функционирования HetNet. На рис. 11.3 сравниваются результаты, полученные с использованием известных аналитических методов решения оптимизационных задач, с результатами, полученными с помощью имитационной модели.

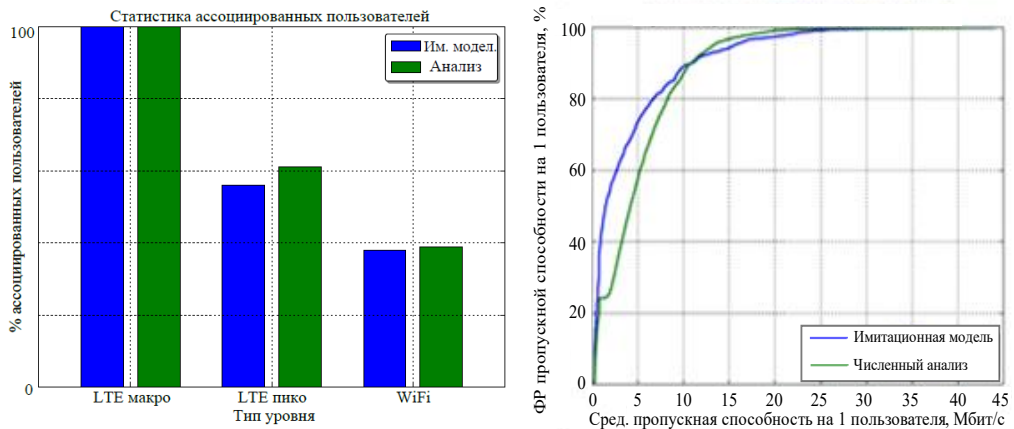


Рис. 11.3. Калибровка аналитической модели:
слева – процент пользователей, использующих определенную технологию в имитационном моделировании,
справа – ФР пропускной способности одного пользователя

Исходными для тестирования параметрами являются число пользователей – 60 пользователей равномерно распределены в зоне макросоты, число пикосотовых БС LTE/NR и точек доступа Wi-Fi – по

5 узлов каждого типа с равномерным размещением в макросоте. Критерием функционирования выбрано справедливое разделение ресурсов. Для оценки результата справедливого разделения применяется индекс Джейна

$$J = (\sum_{d=1}^N x_d)^2 \frac{1}{N \sum_{d=1}^N x_d^2}, \quad (11.7)$$

где x_d – пропускная способность пользователя и N – число пользователей.

Результаты моделирования, усредненные для 100 экспериментов (реальное время составляло 4 с для каждой реализации в имитаторе), представлены на рис. 11.3, где левый график показывает статистику ассоциированных с радиотехнологией пользователей, а на правом графике приведена средняя пропускная способность пользовательского устройства. Подтверждена разница в 5% в значениях метрик для двух подходов, которую можно считать достаточно малой. Наблюдаемые отклонения в основном связаны с отсутствием учета межсотовой интерференции в аналитической модели.

Несмотря на то, что для моделирования пикосотовой инфраструктуры LTE/NR осуществлено в соответствии со стандартами 3GPP, которые обеспечивают минимальное расстояние между БС, в имитационной модели учтена интерференция между близлежащими сотами. Наблюдая на рис. 11.3 ФР пропускной способности пользователя, отмечаем небольшое отклонение начиная с 25% ФР. Такое поведение объясняется разницей в пропускной способности пользователей, использующих для связи только макросоту LTE/NR с теми, кто имеют дополнительные радиотехнологии. Заметим, что в имитаторе этот эффект сглажен вследствие учета интерференции. Подобный эффект наблюдается в случаях, когда пользовательское оборудование с лучшей связностью (например, ассоциированное одновременно со всеми тремя уровнями) не обладает преимуществами в пропускной способности по сравнению с пользовательским оборудованием, имеющим связь только с макросотой, вследствие большой интерференции на уровне Wi-Fi и/или пикоуровнях. В заключении продемонстрируем результаты, которые позволяют сделать вывод о том, что разработанный инструмент аналитического моделирования является достаточным для оценки влияния предлагаемых стратегий для разделения ресурсов HetNet.

Эффективность разделения ресурсов

Сравним теперь эффективность аналитической схемы разделения ресурсов, основанной на полосе пропускания, ($f(x) = s_{dp}^\alpha$ при $\alpha = 1$) с эвристическими методами. Результаты сравнения проиллюстрированы на рис. 11.4 и 11.5. На рис. 11.4 показаны графики ФР пропускной способности пользователя и 5%-процентиля для всех рассмотренных стратегий. Можно видеть, что средние значения пропускной способности практически совпадают, но разница в справедливом разделении ресурсов с точки зрения 5%-процентиля существенна.

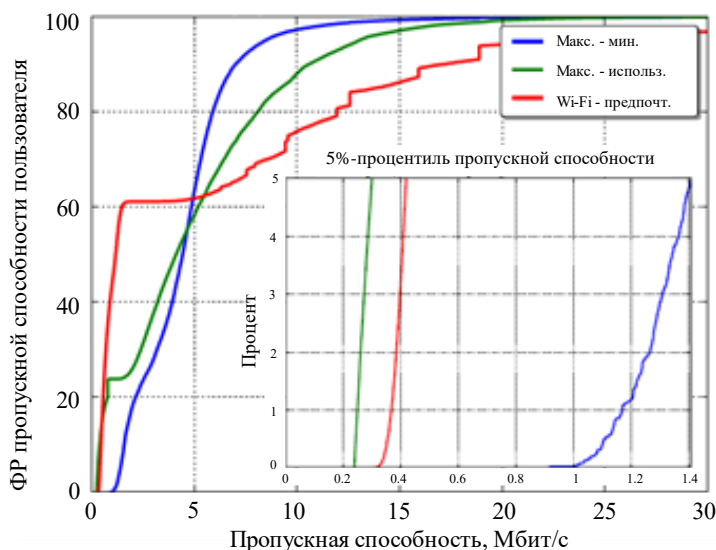


Рис. 11.4. ФР пропускной способности пользователя

Для Wi-Fi предпочтительной стратегии наибольшая пропускная способность достигается устройствами, имеющими связность с Wi-Fi, в то время как пользователи пикосот и макросот LTE/NR имеют существенно более низкую пропускную способность (разница достигает 60%). Кроме того, для такого разделения ресурсов неравномерные изменения на графике ФР отмечаются только для Wi-Fi-пользователей. Этот эффект вызван отсутствием разветвления потока для пользователей, использующих одну радиотехнологию в каждый момент времени. Предложенная схема обеспечивает лучший 5%-процентиль пропускной способности пользователя, что достигается за счет некоторого снижения среднего значения.

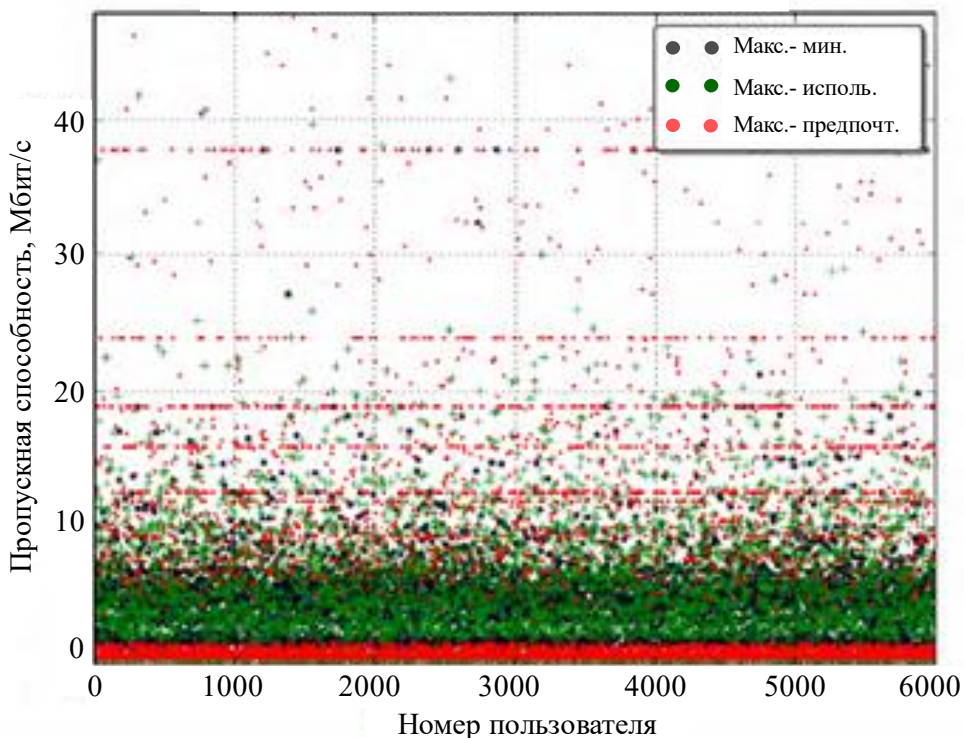


Рис. 11.5. Пропускная способность пользователей

Данные рис. 11.5 отражают пропускную способность на одного пользователя в виде диаграммы скоростей. На рисунке красные точки (Wi-Fi предпочтительная схема) сконцентрированы в верхней и нижней частях диаграммы вследствие неразветвленного механизма разделения ресурсов. Зеленые точки, соответствующие схеме максимального использования, распределены по всей диаграмме, а синие точки, соответствующие max-min-схеме, группируются в центре диаграммы, показывая наилучшее справедливое разделение ресурсов полосы пропускания.

Для лучшего понимания влияния пропускной способности и связности на эффективность предложенных алгоритмов оценим рассматриваемые схемы разделения ресурсов для различных значений плотности узлов на уровнях гетерогенной сети. Результирующее справедливое разделение ресурсов и пропускная способность на одного пользователя для всех трех схем приведены на рис. 11.6, где варьируется число узлов LTE/NR и Wi-Fi. Анализируя результаты, можно отметить, что при малых значениях связности пользователей и пропускной способности сети (например, почти все пользователи

задействуют макросоты LTE/NR) уровень справедливости распределения ресурсов мал. Можно заметить, что среди рассмотренных схем отсутствует стратегия, которая может улучшить эту ситуацию.

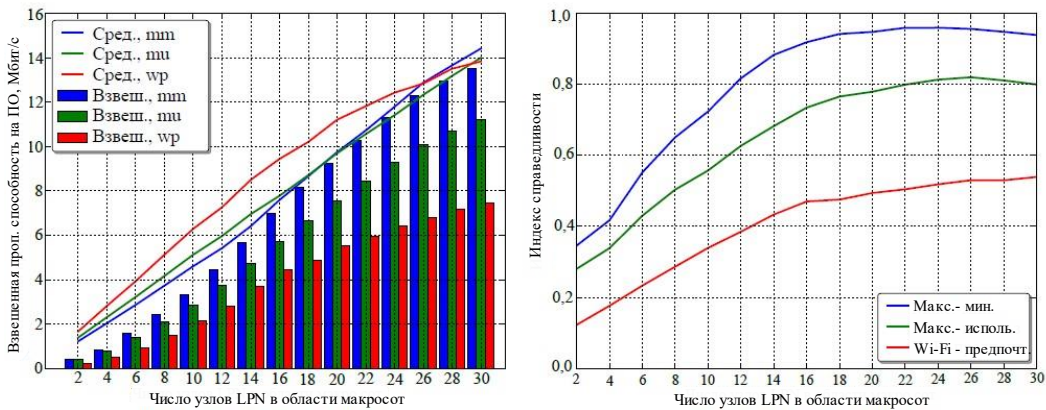


Рис. 11.6. Взвешенная справедливая и средняя пропускная способность (слева), индекс Джейна (справа) для различных стратегий (wp – Wi-Fi-предпочтительная стратегия, mu – стратегия максимального использования, mm – max-min-стратегия)

С увеличением связности пользователей средняя пропускная способность при Wi-Fi предпочтительной схеме становится значительно выше по сравнению с другими стратегиями. Однако этот эффект существенен только для части пользователей, использующих Wi-Fi. В то же время доля пропускной способности Wi-Fi предпочтительной схемы мала по причине неоптимального распределения ресурсов. Однако, если связность пользователей увеличивается, max-min-стратегия позволяет формировать лучшие условия с точки зрения и средней пропускной способности и справедливого распределения ресурсов. Другой интересный вывод может быть сделан при анализе поведения индекса Джейна (правый график на рис. 11.6). Как уже было отмечено, однажды уже использовав степень связности, получить дальнейшее улучшение справедливого распределения ресурсов не представляется возможным. Мы называем этот эффект «насыщением справедливостью» и отмечаем, что он наблюдается для всех рассмотренных схем. Это явление является следствием близкой к идеальной связности (когда почти каждый пользователь может получить доступ ко всем трем уровням сети) и равномерного распределения расположения пользователей в зоне покрытия макросоты LTE/NR. Абсолютная

величина насыщения справедливостью также зависит от деталей конкретной стратегии распределения ресурсов.

Аспекты управления распределением ресурсов

В этом параграфе обсуждается эффективность схем max-min и максимального использования в аспекте определения управления достижением баланса между суммарной пропускной способностью системы и справедливым разделением ресурсов. Напомним, что в случае max-min-схемы распределения ресурсов могут быть использованы две функции управления, принимающие во внимание спектральную эффективность SE-оборудования пользователей. На рис. 11.7 показан баланс между справедливым разделением ресурсов и пропускной способностью одного пользователя для max-min-стратегии распределения ресурсов и функций управления α^x и x^α . Как и ожидалось, в определенном диапазоне значений α эффективность max-min-схемы для обеих функций почти идентична. Однако функция управления $f(x) = \alpha^x$ имеет больший диапазон изменения значений (рис. 11.7), потенциально обеспечивая большую гибкость управления ресурсами сети.

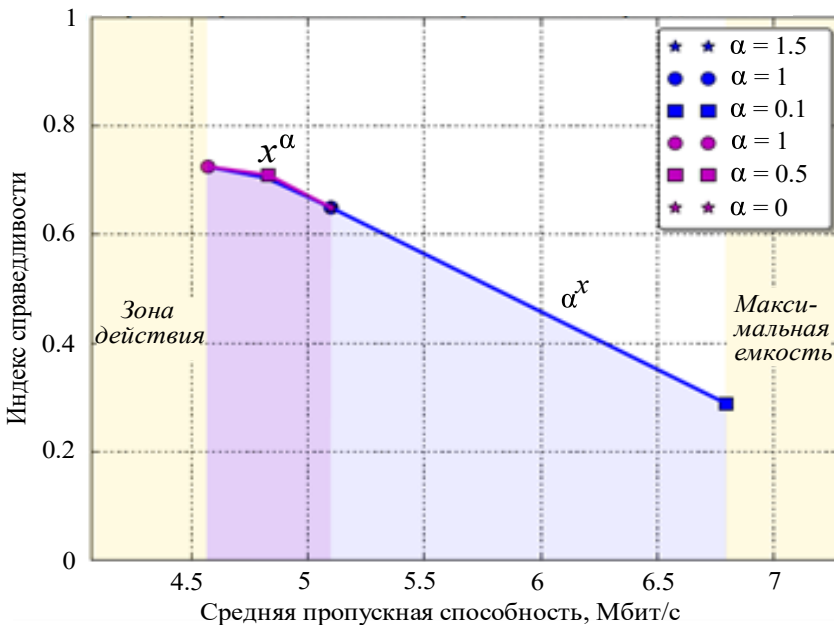


Рис. 11.7. Индекс справедливости Джеейна в зависимости от средней пропускной способности пользователя для двух функций управления

Завершая параграф, сравним эффективность управления разделением ресурсов для схемы max-min и эвристического подхода с максимальным использованием для двух условий покрытия и значений пропускной способности. В этих экспериментах мы рассматриваем гетерогенную сеть с целью идентификации доступных режимов ее эксплуатации. Результаты численного анализа показаны на рис. 11.8, где функция для взвешенного max-min-управления имеет вид $f(x) = \alpha^x$. В первом сценарии пять точек доступа Wi-Fi (40% покрытия) и пять пикосотовых БС LTE/NR (60% покрытия) расположены в зоне покрытия макросоты БС LTE/NR. Второй сценарий предусматривает развертывание сети сверхвысокой плотности, где 15 узлов каждого типа расположены в той же зоне, покрывая почти 100% уровней и пикосот БС LTE/NR и точек доступа Wi-Fi.

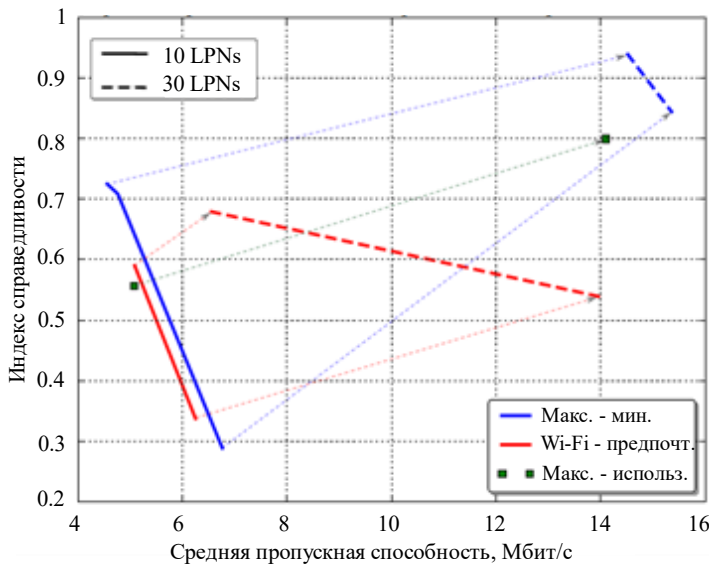


Рис. 11.8. Диапазоны применения схем разделения ресурсов

В рассматриваемых сценариях в дополнение к централизованному max-min-решению возможно использовать справедливое разделение ресурсов за счет Wi-Fi предпочтительной схемы путем установки порога ассоциации очень низким (высокая степень справедливости) или очень высоким (большая пропускная способность). Однако, как видно из графиков на рис. 11.8, применение эвристической стратегии дает худший результаты, чем взвешенное max-min-разделение. При этом справедливость эвристического подхода ниже справедливости взвешенного max-min-решения. Другой интересный

эффект наблюдается, когда для пользователей централизованно применяется схема максимального использования. В условиях сверхвысокой плотности размещения узлов сети этот подход оказывается лучше, чем Wi-Fi-предпочтительная стратегия с сетевой поддержкой. Причина состоит в том, что такой сценарий даже в простейшей разветвленной схеме превосходит большинство неразветвленных решений.

11.5. Гетерогенная облачная сеть радиодоступа H-CRAN

В зависимости от ограничений транзитных соединений могут быть реализованы различные уровни координации в архитектуре 5G сетей HetNet. Например, если доступна магистральная сеть только мобильных операторов, то может быть использована координация посредством архитектуры *якорь–бустер* (anchor-booster architecture), где в качестве «якоря» может выступать БС макросоты, обеспечивая управление сетью и малыми сотами («бустерами») для передачи данных пользователей путем выгрузки трафика. Как альтернатива, в случае магистральной сети с высокой пропускной способностью и малыми задержками (например, оптоволоконная сеть), сигналы с небольших сот могут быть приняты и обработаны на удаленной централизованной платформе. Эта архитектура, получившая название облачной сети радиодоступа Cloud RAN, становится основной для сетевых операторов, которые все чаще отдают предпочтение сетям с преобладанием оптоволоконных систем, в первую очередь для сверхплотных гетерогенных сетей, покрывающих зоны с высокой интенсивностью трафика. Сегодня, когда до 80% капитальных затрат операторов сотовых сетей приходится на сети радиодоступа, концепция облачной сети радиодоступа Cloud RAN позволяет существенно снизить капитальные/эксплуатационные расходы, а также значительно сократить потребление энергии беспроводной сети.

В Cloud RAN удаленная радиоголовка (RRH, Remote Radio Head), которая является достаточно простым малопотребляющим энергию устройством, использует высокоскоростные каналы прямого доступа (fronthaul), сжимает в реальном масштабе времени и направляет необработанные модулированные сигналы от мобильного пользовательского оборудования на централизованный узел (BBU, Baseband Unit). Ограничения на пропускную способность каналов прямого доступа накладывают фундаментальное ограничение на функционирование системы, когда совместно с динамическим

управлением радиоресурсами требуется поддерживать приемлемый уровень пропускной способности Cloud RAN.

Концепция радиосетей доступа гетерогенной облачной сети Heterogeneous Cloud RAN (H-CRAN) была предложена как экономически эффективное решение для повышения степени переиспользования ресурсов (в частности сигнальных процессоров) в гетерогенных сетях посредством их комбинирования с облачной обработкой сигналов. Располагаясь между гетерогенной сетью и облачными платформами, H-CRAN наследует их преимущества, способствуя уменьшению интерференции, увеличению масштабируемости и возможностей по управлению ресурсами радиосвязи за счет совместного применения методов обработки сигналов и сетевых технологий. Управление ресурсами малопотребляющих узлов продвигается в направлении создания виртуальных БС, которые являются частью процессорной мощности, расположенной в пуле физических BBU облачного сервера. Резюмируя, заметим, что технологические возможности и принципы построения ядра H-CRAN формируются в направлении создания коммерческих H-CRAN, но для систем 5G/6G основные задачи состоят в анализе эффективности и оптимального разделения ресурсов для понимания потенциала их использования.

Основываясь на знании этой технологии и последних разработок 3GPP, обращаем внимание на проблему совместного управления радиоресурсами в 5G H-CRANs посредством создания комплексного метода оптимизации для пропускной способности в реальном масштабе времени. Возможно такое решение позволит динамически управлять ресурсами, назначаемыми конечным пользователям, при этом осуществляя баланс между справедливым разделением ресурсов всех доступных сетей радиодоступа и общей пропускной способностью системы. Рассмотрим динамическое управление системами H-CRAN для достижения компромисса между этими метриками и сравним эффективность механизмов для централизованного и децентрализованного механизмов назначения ресурсов в характерных для H-CRAN условиях с различными уровнями доступа к LTE/NR/Wi-Fi, а также проведем анализ концепции для доказательства ее обоснованности.

В модели управления сетью 3GPP Release 8 вводится концепция открытой сети доступа и функции обнаружения и выбора (ANDSF, Access Network Discovery and Selection Function), которая используется для управления взаимодействием между беспроводными локальными сетями (WLAN, Wireless Local Area Network) и сетями

3GPP с помощью сервера политик ANDSF в ядре сети (core network). При этом оператор может специфицировать политики использования ресурсов WLAN, оставляя задачу выбора сети пользователю, который в своих решениях использует факторы локального окружения и динамически изменяющиеся условия радиоканалов. Архитектура, основанная на функции ANDSF, показана на рис. 11.9а. Заметим, что децентрализованные решения для пользовательского оборудования являются неоптимальными, поскольку пользователь не знает об условиях, в которых находятся каналы, и о требованиях к радиоресурсам других пользователей. Кроме того, узлы сети обладают весьма ограниченными знаниями о радиоресурсах друг друга, что еще больше снижает эффективность их использования. Более того, управление мобильностью между WLAN и 3GPP-каналами связи осуществляется в ядре сети, делая обмен трафика между WLAN и 3GPP-каналами очень дорогим при адаптации к динамическим изменениям.

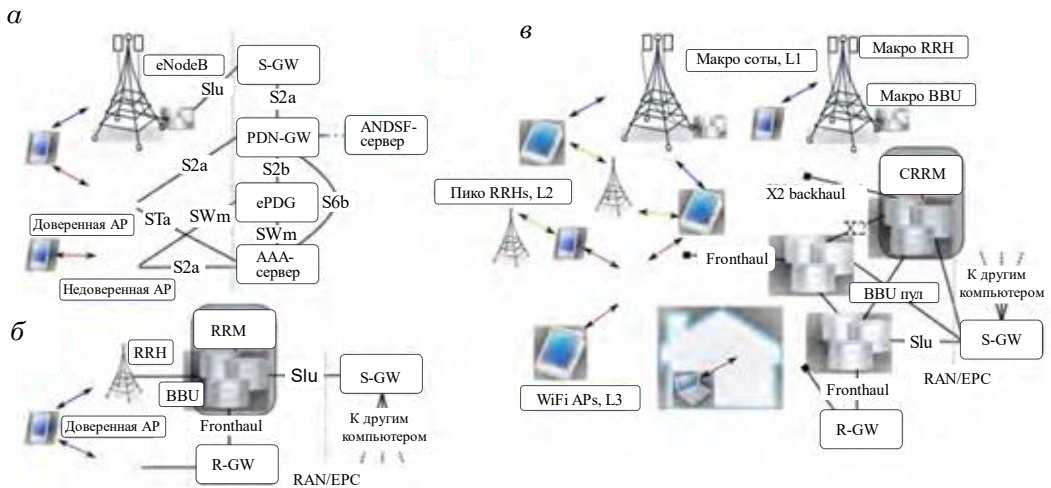


Рис. 11.9. Примеры развертывания сети N-CRAN:

- a* – архитектура, основанная на функции ANDSF;
- б* – упрощенная архитектура управления ресурсами исключительно для выгрузки трафика;
- в* – структура и системная модель подключения к CRRM-серверу

В 3GPP Release 12 было найдено изложенное выше решение проблемы взаимодействия 3GPP/WLAN. Для координации с сетью радиодоступа для БС eNodeB стал доступен набор порогов, позволяющий соотнести качество канала и решение о выгрузке трафика в

WLAN, что дает возможность улучшить соответствующий обмен трафиком, хотя координация, требуемая для выбора подходящих порогов, до сих пор не специфицирована. Но, поскольку мобильный «якорь» для каналов WLAN/3GPP находится в ядре сети, отсутствует возможность эффективной смены радиointерфейса пользователем.

Разработки 3GPP направлены на интеграцию WLAN в сотовые сети. Существующие предложения стремятся к использованию WLAN в качестве вторичного транспорта, с якорем на БС eNodeB в пределах сети радиодоступа RAN. Такая архитектура расширяет возможности 3GPP Release 12, предусматривая двойное использование системы *якорь–бустер*, введенное для малых сот, так же как и существующая структура агрегации в 3GPP включает в себя не-3GPP радиотехнологии (например, доступ WLAN).

Также ожидается, что 3GPP рассмотрит стандартизацию интерфейса между БС eNodeB и точкой доступа WLAN для несовмещенных WLAN/3GPP. Важно, что использование БС eNodeB в качестве якорного узла WLAN позволяет пользователям использовать LTE/NR сеть для реализации функций управления ресурсами исключительно для выгрузки трафика. Упрощенная архитектура такого подхода показана на рис. 11.9б. Функция управления радиоресурсами показана как RRM-модуль, но поскольку интерфейс Wi-Fi-BBU еще не стандартизован, предполагается, что соединение будет проходить через специальный шлюз (шлюз RAN или R-GW), который выполняет согласование интерфейсов.

Учитывая тот факт, что архитектура *якорь–бустер* позволяет координировать использование радиоресурсов в зоне покрытия якорных сот (как правило, макро-сот LTE/NR), эффективность системы в целом может быть увеличена, если будет использована координация радиоресурсов через якорные соты. На практике для достижения такой координации могут быть использованы различные подходы. В распределенной модели координация БС eNodeB с другой БС может быть осуществлена через интерфейс X2. Как альтернатива, контроллер централизованного разделения радиоресурсов может быть использован для системного администрирования радиоресурсов.

Для сетей, которые используют высокоскоростные соединения по оптоволокну, архитектура облачной сети доступа Cloud RAN становится осуществимой. При этом каналы RRH с простой функциональностью централизуются на уровне пула BBU в облаке. Для таких архитектур в сети можно разместить не-3GPP RRH узлы, что позволяет централизовать координацию для мультирадиосетей. Отметим,

что централизация сети для управления радиоресурсами позволяет принимать во внимание не только проблемы интерференции, но и изменения нагрузки и мобильность пользовательского оборудования (например, пользовательское оборудование с высокой мобильностью может быть выгружено в макросоту).

Далее будем предполагать, что H-CRAN позволяет централизовать процесс управления радиоресурсами с помощью выделенного объекта, называемого менеджером кооперативных радиоресурсов (CRRM, Cooperative Radio Resource Manager), при этом подключение к CRRM-серверу осуществляется с помощью такого же интерфейса X2, как и для магистральной сети. Структура и системная модель такого подхода показаны на рис. 11.9в.

11.6. Разделение ресурсов в H-CRAN

Рассмотрим H-CRAN с некоторым числом мультирадиосетей доступа RANs, называемых далее уровнями. Следуя концепции, изложенной в параграфе 11.5, все узлы радиодоступа на всех уровнях должны иметь возможность взаимодействия с CRRM, который отвечает за централизованное разделение ресурсов сетей доступа. Отдельная радиотехнология, как правило, имеет индивидуальный набор возможностей для адаптации каналов низкого уровня и механизмы обработки сигналов для управления радиоканалами ассоциированных пользователей. Эти параметры доступны CRRM, вследствие чего CRRM становится ответственным за оптимизацию H-CRAN. Однако разработка приемлемых процедур оптимизации и относительно несложных алгоритмов является очень сложной комплексной исследовательской проблемой. Вместе с тем роль CRRM может быть ограничена исключительно оптимизацией трафика в H-CRAN. Ниже рассмотрим именно такой подход кооперативного администрирования радиоресурсов.

Состояние H-CRAN в любой момент времени может быть описано состоянием запросов на передачу трафика в узлах, их возможностями связи с сетью радиодоступа и географическим расположением зоны обслуживания. Информация об изменении этих метрик сообщается пользователем или узлами радиодоступа в CRRM. В моменты изменений состояния CRRM использует величины этих метрик как исходные для оптимизации разделения ресурсов через доступные сети радиодоступа. Оптимальное разделение ресурсов доводится до пользователя посредством интерфейсов управления и используется

до тех пор, пока не произойдет следующее изменение состояния системы. Частота проведения такой оптимизации зависит от динамики исходных параметров.

Далее рассмотрим некоторый момент времени, когда изменение состояния системы запускает функцию оптимизации ресурсов в CRRM. Соответствующая модель разделения ресурсов включает в себя три составляющие: а) модель топологии H-CRAN, б) спецификацию подходящей модели оптимизации и в) эффективный алгоритм решения. Для исходящих каналов H-CRAN интерес представляет так называемое *разветвленное* разделение ресурсов, позволяющее гибко распределить объемы трафика между доступными радиointерфейсами. Это соответствует случаю, когда мультирадиопользовательское оборудование может вести передачу через один или более радиointерфейсов одновременно и результаты проявляются в относительной простоте алгоритма разделения ресурсов, так же как и при оптимизации ядра сети, делая их внедрение реалистичным для CRRM.

Традиционный подход к описанной проблеме разделения ресурсов состоит в максимизации минимального значения ОСШ для всех пользователей, подобно методу *max-min*-разделения ресурсов. Однако это может быть не самым гибким решением вследствие ряда присущих системе ограничений. Первая проблема состоит в том, что потребуются решение задачи выпуклого программирования со значительно более трудоемкими вычислениями, чем для задачи линейного программирования. Второе ограничение связано с тем, что существующие мультирадиосети доступа могут работать в различных полосах пропускания, и тогда потребуются более сложная целевая функция.

Еще одно важное предположение состоит в предположении о «жадном» и эластичном трафике, причем «жадный» трафик занимает все ресурсы, в то время как эластичность подразумевает адаптацию к реально доступному объему ресурсов. С этими двумя предположениями рассматриваемая H-CRAN-система полностью характеризуется возможностями связности пользователей. Как увидим далее, предложенный подход достаточен для оптимизации системы в реальном масштабе времени для крупномасштабных гетерогенных H-CRANs.

Узким местом для H-CRAN являются радиointерфейсы. Здесь, как и ранее, можно абстрагироваться от каналов, связывающих узлы сети радиодоступа с пулом BBU. Далее, не принимая во внимание

набор каналов, связывающих пользовательское оборудование с узлами сети радиодоступа, приходим к представлению топологии в виде графа с двумя вершинами, одна из которых является логическим, а другая физическим агрегатором. Число каналов, соединяющих агрегаторы, равно числу узлов сетей радиодоступа на всех уровнях и в таких предположениях пользователей поддерживается набором доступных путей между двумя узлами (см. параграф 11.3 и рис. 11.1).

Таким образом, ограничения задачи по пропускной способности и по запросам имеют вид

$$\sum_{p=1}^{P_d} \alpha_{dp} x_{dp} = h_d,$$

$$\sum_{d=1}^M \sum_{p=1}^{P_d} \delta_{edp} x_{dp} = B_e, \quad (11.8)$$

где δ_{edp} – индикаторная функция, принимающая значение 1, если канал e используется для обслуживания запроса d по пути p , и 0 – в противном случае; B_e – полоса частот радиодоступа канала e . Как и ранее, задача для max-min-решения состоит в лексикографической максимизации вектора разделения полосы пропускания \vec{h}_d , что сводится к задаче линейного программирования.

Чтобы определить коэффициент α_{dp} в (11.8), используется спектральная эффективность пользователя d на беспроводном интерфейсе p . Выбор этих коэффициентов должен быть в согласии со справедливым разделением ресурсов. Для обеспечения возможности контролируемого баланса между справедливостью разделения ресурсов и пропускной способностью системы предлагается определенная выше функция спектральной эффективности. Определив $\alpha_{dp} = \beta^{s_{dp}}$, можно использовать параметр β в предложенном выше оптимизационном алгоритме. Принимая $\beta = 1$, приходим к классическому max-min-решению в терминах справедливого распределения скоростей передачи данных. При $\beta > 0$ пользователи имеют более высокую спектральную эффективность, что приводит к лучшей пропускной способности системы. При $\beta \rightarrow \infty$ занимается вся полоса частот пользователями с высокой спектральной эффективностью и, следовательно, максимизируется пропускная способность системы.

Предложенная задача оптимизации является базовой и может быть далее расширена путем учета дополнительных метрик. Одно из

потенциальных расширений базируется на введении дополнительных коэффициентов в (11.8). Например, добавляя коэффициенты γ_d к первому уравнению, можно учесть приоритетное разделение ресурсов для некоторого заданного числа классов обслуживания. Далее, вводя дополнительные весовые коэффициенты, связанные с потреблением энергии различными радиоинтерфейсами и/или спектральной эффективностью, можно оптимизировать потребление энергии пользовательским оборудованием при совместной работе с H-CRAN. Кроме того, ограничение запросов трафика может быть введено в модель без каких-либо значимых модификаций посредством указания их в явном виде.

Рассмотренная в этой главе модель может быть модифицирована для различных целей и/или условий внешней среды. Неразветвленное разделение ресурсов (когда пользовательское оборудование не может делить трафик между несколькими радиотехнологиями) переводит задачу в класс задач смешанного целочисленного программирования (MIP, Mixed-Integer Programming), существенно увеличивая сложность алгоритма решения. Такие задачи известны как NP-сложные даже в случае, когда имеется только один источник и несколько радиотехнологий доступа. Однако в случае H-CRAN, где все пути передачи данных имеют фиксированную длину, равную 2, создание эффективного алгоритма оптимизации представляется возможным.

Модифицируя предложенную целевую функцию, можно рассмотреть случаи с различными стоимостями маршрутизации через конкретные сети доступа. Например, более высокие затраты на маршрутизацию через макросоту LTE/NR могут заставить пользовательское оборудование выгружать больше трафика через Wi-Fi или интерфейсы пикосот. Заметим, что эти модификации не оказывают влияния на задачу оптимизации, которая по-прежнему будет являться задачей линейного программирования. Если целью оптимизации для H-CRAN являются задержки, тогда требуется изменить целевую функцию и использовать выпуклое программирование. В заключение отметим, что с приемлемыми модификациями предложенная модель пригодна и для оптимизации нисходящих потоков трафика.

11.7. Численный анализ решения H-CRAN

Для того чтобы оценить рассмотренные выше решения, рассмотрим и численно оценим метрики схем управления радиоресурсами в

близких к реальным условиям внедрения H-CRAN. Имитационное моделирование выполнено, как и ранее, с помощью системы имитационного моделирования WINTERsim. Отметим, что ряд особенностей системного уровня в имитаторе представлен в упрощенном варианте, например, межсотовая интерференция и замирания канала моделируются с помощью констант. Поэтому используемые аналитические решения можно считать достаточно точными для понимания обсуждаемых концепций и технологий. Упрощение моделей межсотовой и межуровневой интерференции дает возможность фокусироваться на типовых характеристиках макросот БС LTE/NR. Предполагается, что БС пикосот и точки доступа Wi-Fi размещаются равномерно в зоне покрытия.

Далее рассмотрим три сценария реализации H-CRAN по выбору оператором сети способа интеграции LTE/NR/Wi-Fi: а) оператор создает макросеть LTE/NR в лицензируемом спектре и собственную Wi-Fi сеть в нелицензируемом спектре; б) только LTE/NR-технология используется оператором с применением как макросот, так и пикосот на отдельных лицензированных частотах; в) все возможные опции доступны оператору, а именно: макро- и пикосоты LTE/NR и Wi-Fi-сети радиодоступа. Естественно, что последний сценарий в потенциале обладает большей пропускной способностью и наибольшей гибкостью для пользователей.

Плотность пикосот и точек доступа Wi-Fi в сценариях может быть различна, но целевое значение четырех доступных радиоузлов на одну макросоту предполагается соответствующим современным городским условиям. Число единиц пользовательского оборудования в соответствии с рекомендациями 3GPP предполагается равным 60, равномерно распределенным в зоне обслуживания и статичным. Как и ранее, предполагается, что каналы магистральной сети имеют большую пропускную способность, чем для радиоканалов, и, поэтому анализ направлен на аспекты, относящиеся только к пропускной способности радиоканалов.

Далее проведем сравнение эффективности двух вариантов реализации H-CRAN, позволяющих обеспечивать в полном объеме управление сетью – централизованно и децентрализованно с ограниченной поддержкой со стороны сети. Предполагается «жадная» схема максимального использования радиоресурсов. Для решения управления ресурсами с сетевой поддержкой рассматривается неразветвленный вариант, который позволяет пользователю выбрать предпочтительную сеть радиодоступа в конкретный момент

времени. Алгоритм выбора состоит в следующем. Если первая попытка сопользователя занять ресурсы Wi-Fi с низким уровнем сигнала оказалась неудачной, так же как и вторая попытка для пикосоты, то пользователь будет обслужен на уровне макросоты. Рассматривается также централизованное управление, основанное на H-CRAN/CRRM, которое позволяет оптимально разделять ресурсы между радиотехнологиями на основе качества радиоканала пользователя.

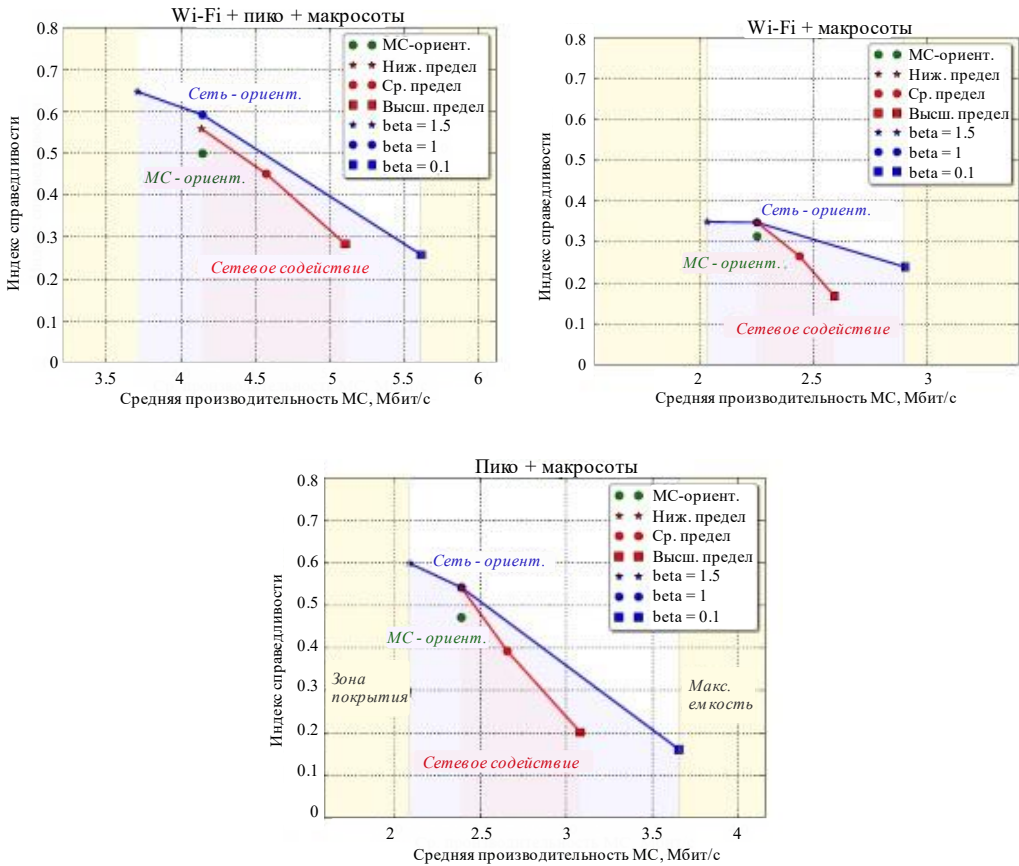


Рис. 11.10. Сравнение эффективности трех сценариев интеграции LTE/NR/Wi-Fi

На рис. 11.10 показано сравнение эффективности разделения ресурсов для трех схем: Wi-Fi и макросоты LTE/NR (левая часть), пикосоты и макросоты LTE/NR (центральная часть) и гетерогенный сценарий с интеграцией макросоты LTE/NR, Wi-Fi и пикосот LTE/NR

(нижняя часть рисунка). Анализируя полученные результаты, следует отметить, что увеличение пропускной способности влечет снижение уровня справедливости, рассчитанного по (11.7) с использованием индекса справедливости Джейна на вертикальной оси графиков. Ограничения доступного компромисса показаны на рис. 11.10 различными цветами. Например, если попытка запроса на увеличение уровня справедливости встречает ограничения по покрытию, то выделение дополнительных ресурсов для пользователей со слабой связностью будет необоснованно, в то время как пользователи с высокой связностью не будет больше иметь дополнительных привилегий в справедливости. Вместе с тем если будут предприняты действия по увеличению пропускной способности сверх установленного предела, то для сильно связанных пользователей это приведет к ограничениям по пропускной способности и, следовательно, не будет выделен объем ресурсов.

Аналогичный эффект наблюдается для схемы с сетевой поддержкой, когда варьируется степень ассоциации пользователей с малыми сотами LTE/NR и Wi-Fi, он основан на значении ОСШ. Практически это означает, что управление ресурсами на базе H-CRAN показывает большую эффективность с точки зрения справедливости разделения ресурсов и пропускной способности сети. В частности, система имеет высокие значения справедливости и пропускной способности с увеличением доступных сетей RAN, но сочетание уровней «пико-плюс-макро» более выгодно, чем «Wi-Fi-плюс-макро», несмотря на то, что использование полосы пропускания в 20 МГц для Wi-Fi должно было бы иметь преимущество по сравнению с 10 МГц пико-сотового канала LTE/NR. Это происходит вследствие исходно лучшего покрытия для уровня пикосетей за счет меньших потерь распространения.

11.8. Аспекты практической реализации

Для оценки практической ценности рассмотренной выше концепции H-CRAN, экспериментальной проверки полученных результатов и расширения возможных опций интеграции соответствующей функциональности в сети мобильных операторов были созданы соответствующие прототипы решений путем реализации тестовой среды, демонстрирующей полную интеграцию Wi-Fi технологии в существующую сеть LTE/NR. Ниже перечислены условия тестовой среды, основные проблемы, с которыми приходится сталкиваться при создании такого рода практических моделей, и предлагаются решения по

расширению предложенной технологической концепции для операторов при развертывании сетей 5G/6G.

Важно отметить, что большинство существующих пользовательских устройств уже сегодня поддерживают и сотовые интерфейсы, и интерфейсы беспроводных локальных сетей. Однако борьба производителей оборудования за энергетическую эффективность требует ограничений по использованию оборудования в части использования только одного радиointерфейса в каждый момент времени. В настоящее время эта ситуация изменяется вследствие внедрения систем 5G/6G, которые будут представлять собой интеграцию различных технологий и решений радиодоступа, а не одну новую технологию. Вследствие этого современные требования по увеличению связности заставляют производителей поддерживать разработку перспективных платформ для доступа и эффективного использования всех доступных радиointерфейсов. В качестве примера можно назвать операционную систему Sailfish.

Расширяя видение развития H-CRAN, следует рассмотреть возможные варианты интеграции сотовых технологий и технологий локальных беспроводных сетей. В идеале интеграция точек доступа Wi-Fi на стороне сетевого провайдера потребует доступа к реализации открытой для разработчиков среды сотовой БС (например, согласно стандарту Open RAN, O-RAN), которая все еще недоступна для исследователей. Поэтому эффективная интеграция Wi-Fi и сотовых радиотехнологий в одной и той же сети может быть выполнена посредством инкапсуляции соответствующих радиоканалов в выделенных OpenVPN-туннелях. Они будут заканчиваться в виртуальной машине, функционирующей в операторском узле агрегации, что может расширить функциональность пакетного шлюза или выделенного объекта в пуле BBU.

Разработанная тестовая среда H-CRAN использует архитектуру программно-конфигурируемой сети (SDN) и протокол OpenFlow для динамического управления связностью пользовательского оборудования. Все радиоканалы, доступные пользовательскому устройству, анализируются в реальном масштабе времени, что эмулируется посредством программного обеспечения Open vSwitch. Реальный выбор направления выполняется программным контроллером, который реализуется на основе правил OpenFlow, в том числе на активных измерениях условий радиоканала. В тестовой конфигурации контроллер работает на самом пользовательском оборудовании, но предложенное решение может быть обобщено. В частности,

конфигурация оборудования может быть передана сети оператора связи под управление единого контроллера для всех обслуживаемых пользователей.

Более конкретно якорные точки для туннелей VPN на операторской стороне функционируют в изоляции внутри соответствующих контейнеров Linux, которые поддерживаются в Docker Engine. Виртуальные магистральные каналы от контейнеров подключены через Open vSwitch daemon к эмулятору маршрутизации. Сотовые каналы инициализируются в системе посредством RmNet или RmNet USB-драйвера. Поэтому добавление радиоканала в виртуальный план маршрутизации требует соответствующей модификации vSwitch daemon. В архитектуре тестовой среды введен уровень абстракции для канала, определенного посредством GRE-туннелей, эмулирующий фрагмент сети между пользовательским оборудованием и агрегатором на стороне оператора.

В заключение отметим, что предложенная комбинация механизмов аппаратного обеспечения и виртуальной контейнерно-основанной среды выполнения дает возможность быстро и гибко создавать различные сценарии в контексте использования H-CRAN. Все необходимые компоненты собраны в интегрированную среду, что позволяет без труда конструировать требуемые структуры. Гибкость тестовой среды всесторонне поддерживает современное видение операторской сети.



ЧАСТЬ IV
ТЕРАГЕРЦОВЫЕ СЕТИ ДОСТУПА
СИСТЕМ СВЯЗИ 6G



ГЛАВА 12
ХАРАКТЕРИСТИКИ КАНАЛОВ СВЯЗИ

Терагерцовый диапазон является последним диапазоном частот, который обладает свойствами радиораспространения. В этом диапазоне все еще существуют достаточно сильные компоненты отражения, которые позволяют устанавливать уверенную связь, даже если линия прямой видимости между передатчиком и приемником заблокирована. Сохранение этих свойств позволяет рассматривать нижнюю часть терагерцового спектра в качестве основного диапазона частот для сетей связи 6G.

В терагерцовом диапазоне частот блокировки имеют гораздо более ярко выраженный характер, а длина волны позволяет создавать фазированные антенные решетки с еще более узкой диаграммой направленности. Помимо этого, данный диапазон частот обладает дополнительными свойствами, которые могут как оказать дополнительное отрицательное влияние на системы связи, так и позволить внедрить новые функциональные возможности в такие системы. В этой главе мы рассмотрим свойства терагерцовых каналов связи, определим базовые обменные соотношения и продемонстрируем методы эффективного использования таких свойств для различных приложений.

Изложение главы начнем со свойств терагерцовых каналов связи в параграфе 12.1. Окна прозрачности, в которых возможна эффективная связь в терагерцовом диапазоне частот, вводятся в параграфе 12.2, где также рассмотрены базовые характеристики терагерцовых каналов связи, включая ОСШ, скорость по Шеннону, вероятность ошибочного приема и пропускную способность.

12.1. Особенности канала связи в терагерцовом диапазоне

Общее уравнение для принимаемой мощности в терагерцовом диапазоне можно представить в виде

$$P_{Rx}(f, d) = P_{Tx}(f) - L_P(f, d) - L_A(f, d), \quad (12.1)$$

где P_{Tx} – мощность передаваемого сигнала; P_{Rx} – мощность принятого сигнала; L_P – потери распространения в свободном пространстве; L_A – потери молекулярного поглощения (атмосферная абсорбция/поглощение). В дополнение к двум компонентам потерь распространения на принимаемый сигнал может также оказывать влияние атмосферный шум. Обратим внимание, что в данной главе усиления на передающей и приемной сторонах не учитываются.

Атмосферное поглощение

Наличие атмосферных потерь поглощения L_A , обусловленных внутренними колебаниями молекул, является уникальной особенностью терагерцового диапазона. Молекулярное поглощение электромагнитных волн – это эффект, возникающий, когда частота волны близка к резонансной частоте для внутренних колебательных частот молекул в атмосфере. Вибрирующие молекулы поглощают электромагнитную энергию и преобразуют ее в кинетическую энергию. Способность молекулы поглощать энергию определяется их физическими свойствами (структурой связей, пространственной ориентацией и пр.) и измеряется коэффициентом поглощения. В стандартных условиях среды молекулы водяного пара доминируют в поглощении в терагерцовом диапазоне частот.

Потери атмосферного поглощения характеризуются выражением

$$L_A(f, d) = \frac{1}{\tau(f, d)}, \quad (12.2)$$

где f – частота, d – расстояние между передатчиком и приемником, τ – коэффициент «прозрачности» среды, который может быть определен с помощью закона Бира–Ламберта:

$$\tau(f, d) = e^{-K(f)d} = e^{-\sum_{G,I} K_{G,I}(f)d}, \quad (12.3)$$

где K – коэффициент поглощения, который оценивается путем суммирования индивидуальных коэффициентов поглощения $K_{G,I}$ молекул одного изотополога I газа G . Коэффициент поглощения может быть рассчитан по формуле

$$K_{G,I}(f) = \frac{p}{p_0} \frac{T_0}{T} Q_{G,I} \sigma_{G,I}(f), \quad (12.4)$$

где p и T – текущие давление и температура среды; p_0 и T_0 – стандартные давление и температура; $Q_{G,I}$ – объемная плотность молекул, измеряемая в молекулах/м³; $\sigma_{G,I}$ – сечение поглощения, измеряемое в м²/молекулу. Последние два параметра определяются формулами

$$Q_{G,I} = \frac{pL}{RT} Q_{G,i}, \quad \sigma_{G,I}(f) = S_{G,I} G_{G,I}(f), \quad (12.5)$$

где R – газовая постоянная (8,31 Дж/К·моль); L – постоянная Авогадро (число молекул на моль, $6,02 \cdot 10^{23}$ моль⁻¹); $S_{G,I}$ – линейная интенсивность поглощения; $G_{G,I}$ – форма спектрального компонента. Как $S_{G,I}$, так и $G_{G,I}$ табулированы в базе молекулярной спектроскопии (NITRAN, High-Resolution Transmission). Некоторые дополнительные поправки к $G_{G,I}$ необходимы, когда рассматриваются особые условия (давление/температура). В этом случае сначала необходимо определить резонансную частоту каждого изотополога I . Важно отметить, что увеличение давления приводит к линейному увеличению резонансной частоты. Однако для обычных диапазонов давления – это результирующее изменение L_A практически отсутствует.

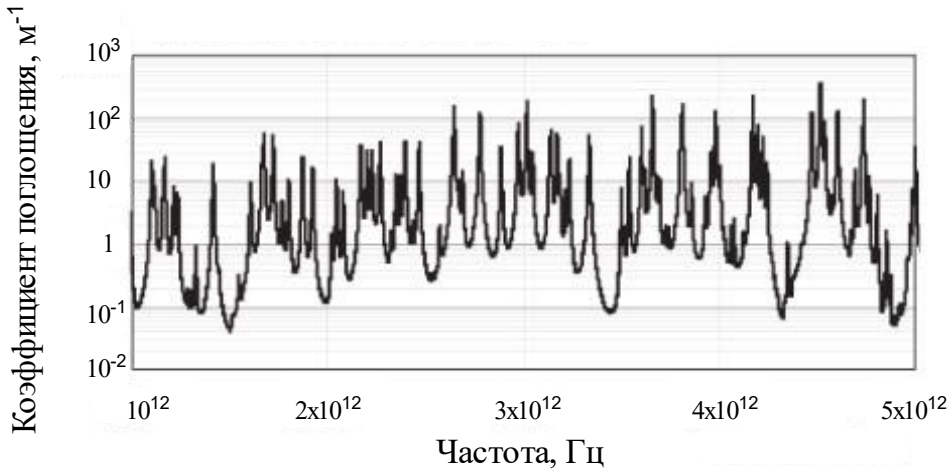


Рис. 12.1. Коэффициент атмосферного поглощения

График коэффициента поглощения K для $T = 296$ К и 1,8 % концентрации водяного пара для 0,1–5 ТГц показан на рис. 12.1. Расчет проведен с помощью приведенных выше формул с использованием

информации, содержащейся в базе данных HITRAN. Одним из важных наблюдений является то, что атмосферное поглощение весьма чувствительно к частоте. На некоторых частотах поглощение очень велико и в итоге может привести к чрезвычайно малой дальности связи. Отметим, что в так называемых окнах прозрачности (см. параграф 12.2) поглощение значительно меньше, чем в остальной части диапазона. Заметим, что коэффициент K не зависит от расстояния d , но потери молекулярного поглощения, L_A , зависят. Таким образом, чтобы обеспечить максимально возможную дальность связи, значение K должно быть минимальным.

Атмосферный шум

Молекулы преобразуют часть поглощенной энергии в кинетическую энергию. В то же время часть энергии может переизлучаться в канал связи, создавая так называемый шум молекулярного поглощения. Спектральная плотность мощности P_N молекулярного шума (также называемого атмосферным) определяется выражением

$$P_N(f, d) = k_B [N_M(f, d) + N_A(f)], \quad (12.6)$$

где k_B – постоянная Больцмана; N_M – эквивалентная шумовая температура молекулярного поглощения; N_A – температура других источников шума на частоте f .

$$N_M(f, d) = T\varepsilon(f, d) = T[1 - \tau(f, d)], \quad (12.7)$$

где $\varepsilon(f, d)$ – излучающая способность среды. Традиционно наиболее важным компонентом N_A считается тепловой шум приемника. На данный момент, насколько известно, существует точная модель шума для устройств на основе графена, но она ожидаемо очень низкая и может не приниматься в расчет при использовании таких систем. Отметим, что представленная модель шума может быть использована только для очень малых расстояний между передатчиком и приемником.

Определенная выше спектральная плотность мощности шума молекулярного поглощения для $d = 0,01, 0,1, 1$ м и $T = 296$ К в некоторых диапазонах частот значительно ниже, чем в других диапазонах (рис. 12.2). При «прозрачности» среды менее 94,5 % шум молекулярного поглощения значительно увеличивается, достигая максимального значения –204 дБ/Гц. Хотя это значение кажется

низким, но при интегрировании на очень большой полосе пропускания канала может привести к высокой мощности шума.

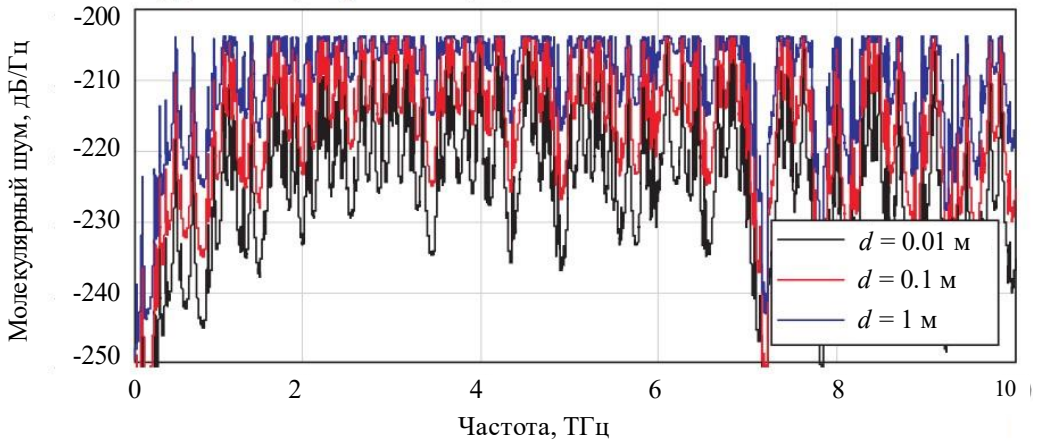


Рис. 12.2. Молекулярный шум в терагерцовом диапазоне частот

Заметим, что когда коэффициент «прозрачности» ниже 5,5%, шум молекулярного поглощения относительно невелик. Это означает, что доминирующим компонентом шума на приемнике является тепловой шум со спектральной плотностью мощности -174 дБм/Гц.

Потери распространения в свободном пространстве

Потери при распространении в свободном пространстве – это еще один параметр, сильно влияющий на мощность принимаемого сигнала. Как будет показано далее, для определенного набора частот в терагерцовом диапазоне потери распространения являются доминирующим фактором, определяющим достижимую дальность связи.

Потери распространения в свободном пространстве могут быть описаны стандартной моделью Фрииса (см. главу 2):

$$L_p(f, d) = \left(\frac{4\pi f d}{c_0} \right)^2, \quad (12.8)$$

где c_0 – скорость света в интересующей среде (для открытого воздуха $c_0 \approx 3 \cdot 10^8$ м/с); d – расстояние между передатчиком и приемником; f – частота. Как можно заметить, увеличение частоты приводит к квадратичному увеличению потерь распространения в свободном пространстве. Когда рассматриваются окна прозрачности, потери

распространения в свободном пространстве становятся доминирующим фактором, ограничивающим зону покрытия передатчика. Из этого также следует, что для увеличения дальности связи необходимо использовать окна прозрачности с более низкими частотами.

На графиках рис. 12.3 показаны общие потери (сумма молекулярного поглощения и потерь при распространении в свободном пространстве) для расстояний $d = 0,1, 0,01, 1$ м, $T = 296$ К и 1,8 % пара в диапазоне 0,1–5 ТГц. Очевидно, что в окнах прозрачности молекулярное поглощение практически отсутствует.

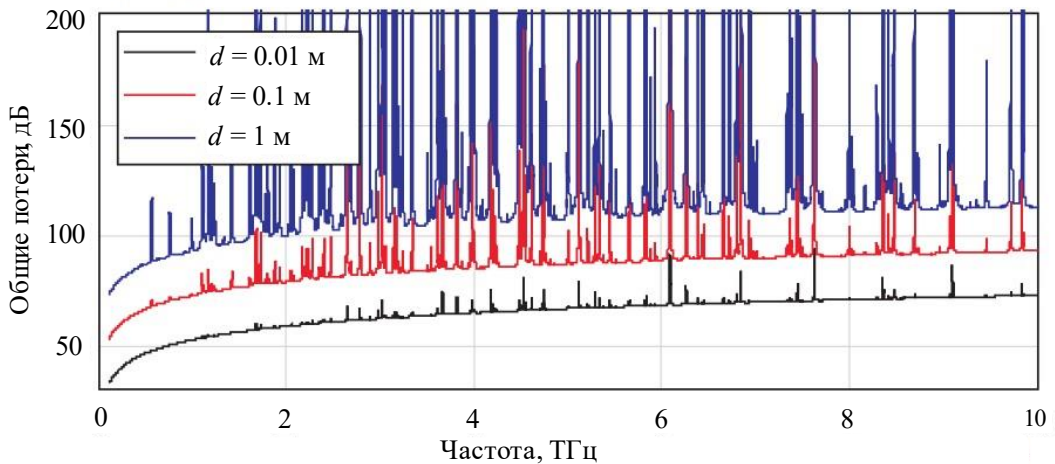


Рис. 12.3. Общие потери распространения в терагерцовом диапазоне частот

12.2. Окна прозрачности и характеристики канала связи

Несмотря на довольно сложные характеристики канала связи терагерцового диапазона частот, существует способ избежать сильного ослабления сигнала с помощью окон прозрачности. В таких окнах «прозрачность» среды $\tau(f, d)$ в (12.3) никогда не бывает меньше 95 % для определенного значения расстояния d . Это означает, что в окнах прозрачности и молекулярные потери на поглощение, и уровень шума настолько малы, насколько это возможно. Ниже в параграфе обсуждаются характеристики окон прозрачности и взаимосвязь между пропускной способностью канала, доступной полосой пропускания, мощностью передачи и дальностью связи.

Характеристики окон прозрачности

Существует множество окон прозрачности в терагерцовом диапазоне. Однако накопительное свойство молекулярных потерь поглощения и шума не позволяет однозначно определить окна прозрачности, так как они зависят от расстояния. Выбор этих окон, как обсуждалось выше, осуществляется на основе коэффициента «прозрачности» среды τ .

В табл. 12.1 сведены свойства окон прозрачности, доступных в диапазоне частот 0,1–10 ТГц, причем число этих окон уменьшается с расстоянием d . Например, на расстоянии 0,01 м в терагерцовом диапазоне имеется более 15 таких окон, первые три окна лежат в диапазонах 0,1–0,54 ТГц ($B = 440$ ГГц), 0,57–0,72 ТГц ($B = 150$ ГГц) и 0,78–0,93 ТГц ($B = 190$ ГГц), а на расстоянии 1 м остается только шесть окон. Эти первые три окна являются единственной возможностью для расширения дальности связи, так как остальные три окна расположены в верхней части терагерцового диапазона, где уровень потерь в свободном пространстве существенно влияет на связь на больших расстояниях. При увеличении дальности связи полоса частот третьего окна резко уменьшается, в то время как первое и второе окна остаются практически неизменными. Столбец, относящийся к эквивалентной длительности импульса, будет рассмотрен ниже, при анализе пропускной способности терагерцового канала связи.

Для иллюстрации выигрыша от использования первого окна прозрачности на рис. 12.4 показана спектральная плотность мощности $1/L_P L_A N_m$ как функция частоты для $d = 0,01, 0,1, 1$ м, $T = 296$ К и 1,8 % концентрации водяного пара в воздухе. Из графиков видно, что, используя все первое окно прозрачности, можно достичь расстояния связи до нескольких сантиметров. Далее эту возможность рассмотрим более подробно. Отметим, что здесь и далее в данной главе мы не учитываем усиление передающей и приемной антенн, что играет основную роль для достижения необходимой дальности связи – до нескольких десятков метров.

Свойства окон прозрачности

Номер	Расстояние, м	Частота, ТГц	Ширина, ГГц	Импульс, пс
1	0,01	0,10–0,56	457	1,36
	0,1	0,10–0,55	445	1,40
	1	0,10–0,54	440	1,48
2	0,01	0,56–0,75	192	3,23
	0,1	0,57–0,74	173	3,59
	1	0,63–0,72	95	6,53
3	0,01	0,75–0,99	237	2,62
	0,1	0,76–0,98	218	2,85
	1	0,76–0,98	126	4,92
4	0,01	7,00–7,33	330	1,88
	0,1	7,01–7,31	300	2,07
	1	7,07–7,23	160	2,59
5	0,01	7,64–7,98	340	1,83
	0,1	7,72–7,96	240	259
	1	7,75–7,88	130	3,88
6	0,01	8,01–8,28	270	2,30
	0,1	8,04–8,15	110	5,64
	1	8,04–8,15	80	7,76

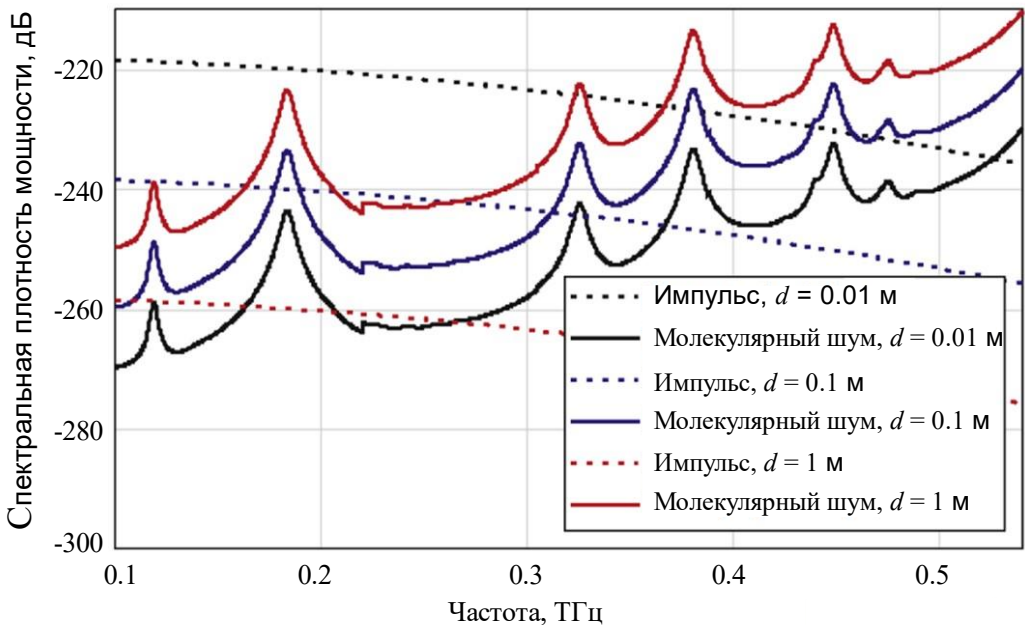


Рис. 12.4. Спектральная плотность мощности

Скорость по Шеннону терагерцового канала связи

Полагая шум гауссовским, теоретически максимальную пропускную способность канала связи C окон прозрачности можно оценить с помощью теоремы Шеннона по формуле

$$C(d) = \int_B \log_2 \left(1 + \frac{P(f)}{L_P(f,d)L_A(f,d)N_M(f,d)} \right) df, \quad (12.9)$$

где B – полоса пропускания канала; P – мощность передаваемого сигнала; L_P – потери распространения в свободном пространстве; N_M – молекулярный шум. Поскольку мы рассматриваем только окна прозрачности, влияние потерь поглощения L_A пренебрежимо мало, а влияние молекулярного шума также практически отсутствует. В расчетах была принята стандартная среда, т.е. $T = 296$ К, $p = 760$ мм рт. ст., 1,8 % молекул пара.

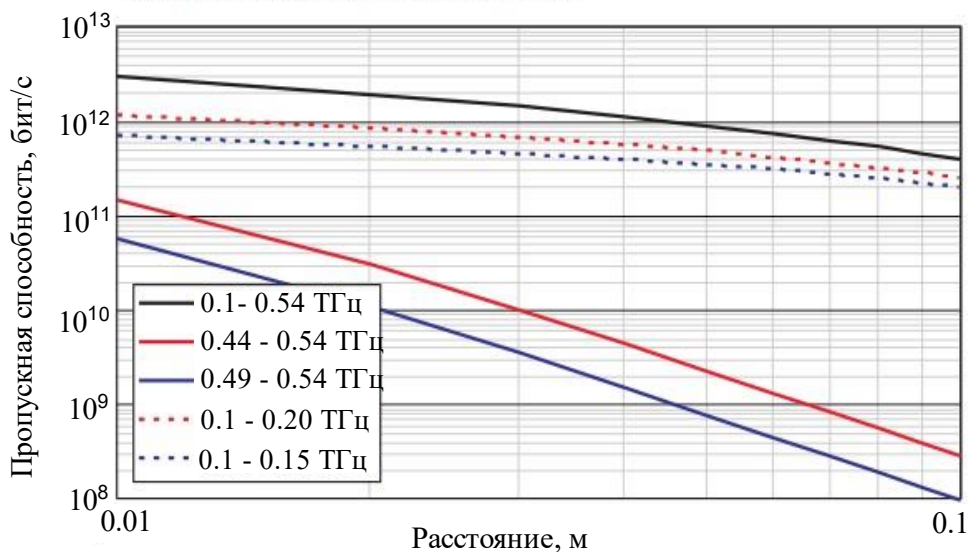


Рис. 12.5. Пропускная способность канала как функция расстояния

На рис. 12.5 показана пропускная способность канала как функция расстояния в логарифмическом масштабе для пяти каналов в первом окне прозрачности, 0,1–0,15 ТГц и 0,1–0,2 ТГц (поддиапазоны 50 ГГц и 100 ГГц в начале окна), 0,44–0,54 и 0,49–0,54 ТГц (поддиапазоны 50 ГГц и 100 ГГц в конце окна) и всей полосы окна 440 ГГц, предполагая 26,5 нВт общей излучаемой мощности. Вспоминая результаты для ОСШ, где значения для «лучшего» канала

50 ГГц в начале окна падают ниже 10 дБ на расстоянии около $d = 0,1$ м, мы ограничиваем максимальное расстояние этой величиной. Обратим внимание, что при использовании всей ширины окна 440 ГГц можно получить чрезвычайно высокую пропускную способность в несколько Тбит/с на относительно небольших расстояниях до нескольких сантиметров, например, около 2 Тбит/с на расстоянии 0,01 м. На расстоянии 0,02 м, соответствующем ОСШ 10 дБ, используя все окна, можно получить скорость примерно 0,7 Тбит/с. Уменьшая пропускную способность канала при использовании прежней мощности, можно достичь больших расстояний за счет меньшей пропускной способности. Наиболее эффективным способом будет использование каналов в начале окна. Действительно, видим, что сохранение пропускной способности каналов при перемещении их в конец окна уменьшает достижимое расстояние более чем в 10 раз и такая же тенденция наблюдается и для пропускной способности канала. Как и ожидалось, каналы 100 ГГц превосходят поддиапазоны 50 ГГц в обоих случаях. Далее, оба поддиапазона в начале окна превосходят поддиапазоны в конце окна, а между ними в скорости канала связи становится больше с расстоянием и может достигать нескольких порядков. Этот эффект также обусловлен потерями при распространении, при этом молекулярный шум оказывает незначительное влияние. В абсолютных числах все рассмотренные каналы обеспечивают более чем достаточную пропускную способность для всех современных услуг связи. Окно шириной 440 ГГц, а также каналы 50 ГГц и 100 ГГц в начале окна также могут быть использованы для приложений, требовательных к скорости передачи данных на расстояниях до нескольких сантиметров, например, для обмена данными между ядрами многоядерных процессоров или других вычислительных компонентов (см. главу 14). Учитывая дополнительно усиление антенн на приеме и передаче, можно увидеть, что достижимые расстояния могут составлять десятки метров для сотовых систем связи 6G.

Зависимость ОСШ от расстояния

Независимо от типа модуляции и канального кодирования, одним из основных факторов, влияющих на достижимую скорость передачи данных, является ОСШ, показанный на рис. 12.6 как функция расстояния для пяти каналов в первом окне прозрачности, 0,1–0,15 ТГц и 0,1–0,2 ТГц (поддиапазоны 50 ГГц и 100 ГГц в начале окна), 0,44–0,54 и 0,49–0,54 ТГц (поддиапазоны 50 ГГц и 100 ГГц в

конце окна) и всей полосы окна 440 ГГц. Прежде всего, видим, что дальность связи в несколько сантиметров может быть достигнута при работе в первом окне прозрачности. При использовании всего окна ОСШ падает ниже 10 дБ на расстоянии около 0,02 м. Далее, видим, что существует значительная разница между каналами в начале окна и в его конце. Причина заключается в доминирующем влиянии потерь распространения в свободном пространстве, поскольку потери на поглощение и молекулярный шум дают довольно ограниченный вклад в ОСШ. Используя канал 50 ГГц в начале диапазона, можно достичь расстояния до 0,1 м, сохраняя ОСШ выше 10 дБ и используя чрезвычайно малую мощность передатчика. Эти результаты показывают, что можно обеспечить связь на достаточно больших расстояниях, что позволит создать сотовые сети связи.

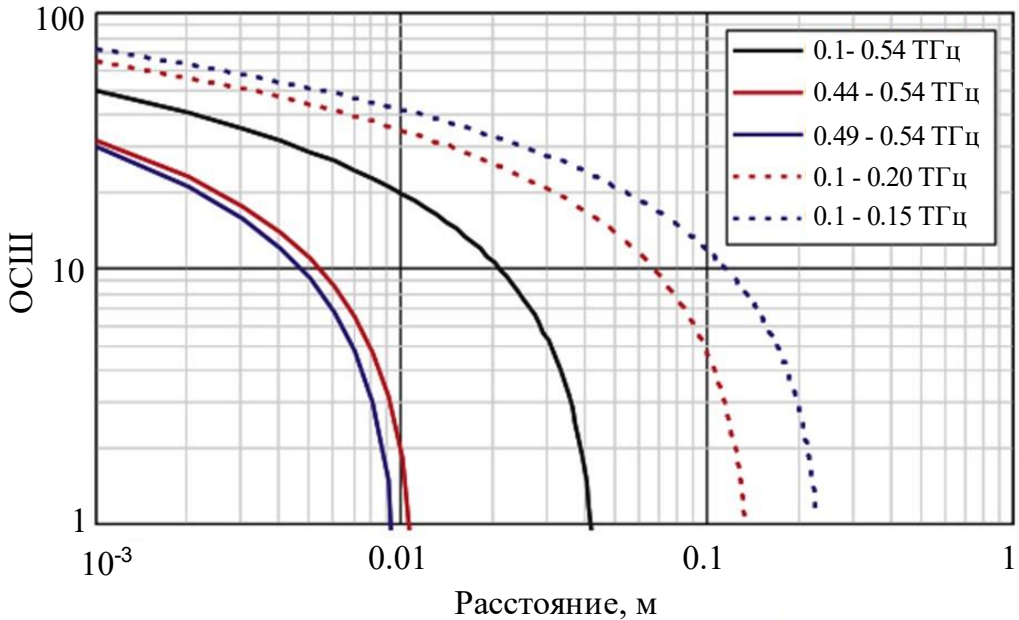


Рис. 12.6. ОСШ как функция расстояния для нескольких каналов

Вероятность ошибочного приема и пропускная способность терагерцового канала связи

Теперь перейдем к численному анализу вероятности битовой ошибки (BER) с помощью имитационного моделирования. Основным моментом исследования является выбор между правильным приемом 1 и 0 (в последовательности 10), вызванный различными

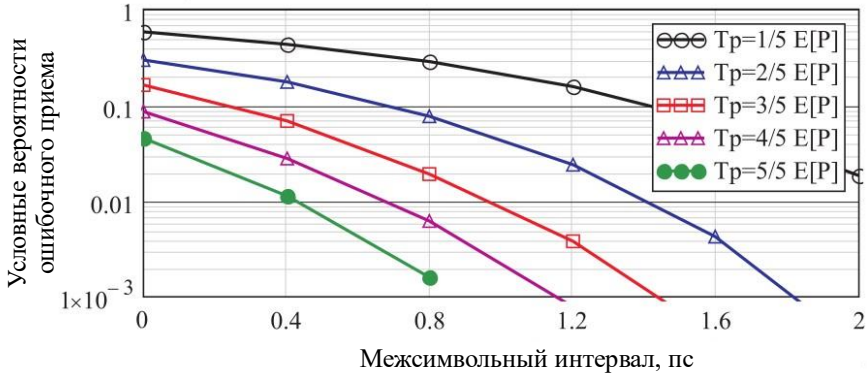
значениями порога обнаружения мощности P_T и межсимвольного времени τ и простейшей модуляционно-кодовой схемы ООК (On-Off Shift Keying). Напомним, что в ООК логическая 1 представлена наличием импульса в канале, 0 – отсутствием импульса. Результаты анализа, в основном качественного, будут использованы в главе 14 для оценки BER для различных приложений терагерцовых сетей связи.

На рис. 12.7 показаны значения уловной вероятности ошибочного приема 0 после единицы в последовательности 10, $p_{10,E}$, для $d = 0,05$ м (ОСШ ≈ 1 дБ) и $d = 0,07$ м (ОСШ < 1 дБ) для времени релаксации $T_R = 5$ пс (время между посылками символов) как функция межсимвольного интервала времени ν и порога обнаружения мощности T_p . Важно отметить, что значения T_p выбраны относительно средней мощности ослабленного сигнала на определенном расстоянии, $E[P] = \int_0^{4t} S^2(t, d) dt$, начиная с $E[P]/5$ до $E[P]$ с шагом $E[P]/5$. Как видно из графиков на рис. 12.7, вероятность ошибочного приема 0 после 1 в последовательности 10 уменьшается как функция межсимвольного интервала τ и порога обнаружения мощности T_p . Как ожидалось, значение $p_{10,E}$ снижается экспоненциально в ответ на увеличение межсимвольного интервала ν , в то время как пропускная способность уменьшается линейно. Эффект влияния порога обнаружения мощности P_T также положительный. Например, для $\nu = 0,8$ пикосекунд изменение значения T_p с $E[P]/5$ до $E[P]$ приводит к уменьшению вероятности $p_{10,E}$ на три порядка. Однако этот эффект нельзя рассматривать изолированно, пренебрегая вероятностью неправильного приема 1 в последовательности 10, поскольку увеличение порога обнаружения энергии должно отрицательно сказаться на вероятности $p_{1,E}$.

Поскольку в ряде приложений приемники и передатчики должны быть чрезвычайно простыми и не использовать сложных алгоритмов коррекции ошибок (см. главу 14), BER оказывается наиболее важной для анализа метрикой. На рис. 12.8 показаны вероятности ошибочного приема 0 в последовательности 10, $p_{10,E}$, и приема 1, $p_{1,E}$, а также скорость битовых ошибок (Bit error rate, BER), p_E , для для расстояния $d = 0,05$ м, $T_R = 5$ пс для нескольких значений ν и P_T . Напомним, что ν не влияет на $p_{1,E}$, т.е. эта вероятность является постоянной для любого выбранного значения ν (зеленая линия на рис. 12.8). Мы видим, что, рассматривая BER как интересующую нас метрику, увеличение межсимвольного интервала ν приводит к

уменьшению BER. Заметим, что значение P_T , оптимизирующее BER для фиксированного значения v , зависит от расстояния между передатчиком и приемником. Для $d = 0,05$ м, $v = 1,2$ пс, $T_R = 5$ пс это значение находится в диапазоне между $0,5 \times E[P]$ и $0,6 \times E[P]$.

а



б

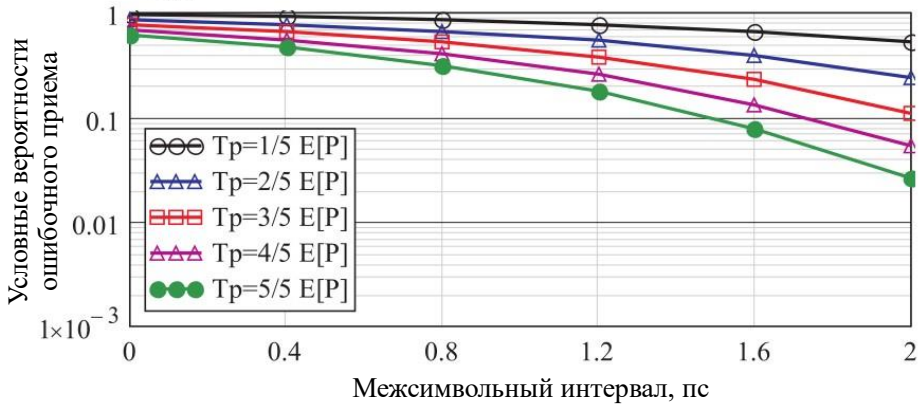


Рис. 12.7. Условные вероятности ошибочного приема:
а – $d = 0,05$ м; б – $d = 0,07$ м

Пропускная способность

Напомним, что значение межсимвольного временного интервала v влияет не только на вероятность битовой ошибки, но и на пропускную способность канала связи, графики которой для нескольких значений v , $d = 0.05$ м и $T_R = 5$ пс, показана на рис. 12.9а. Как и ожидалось, пропускная способность линейно уменьшается в ответ на линейное увеличение v . рис. 12.9б демонстрирует эффективную пропускную способность, определяемую формулой $T(1 - p_E)$, т.е., как

доля правильно принятых бит. Анализируя данные, представленные на рисунках, можно сделать ряд важных наблюдений. Прежде всего, рассматривая пропускную способность как целевую метрику, видим, что межсимвольный интервал ν не зависит от расстояния d для малых значений ОСШ. Таким образом, фиксируя значение $\nu = 0$ пс, получаем наилучшую пропускную способность, если значение порога обнаружения энергии выбрано оптимальным образом. В свою очередь, оптимальное значение P_T зависит от расстояния и для $d = 0,05$ м составляет примерно $0,8 \times E[P]$. Вместе с тем можно заметить, что разница значения пропускной способности для разных значений ν незначительна. Например, максимальная разница для межсимвольных интервалов $\nu = 0$ пс и $\nu = 1,2$ пс достигается при $P_T \approx 0,8$ и составляет примерно 23 % (0,06 Тбит/с).

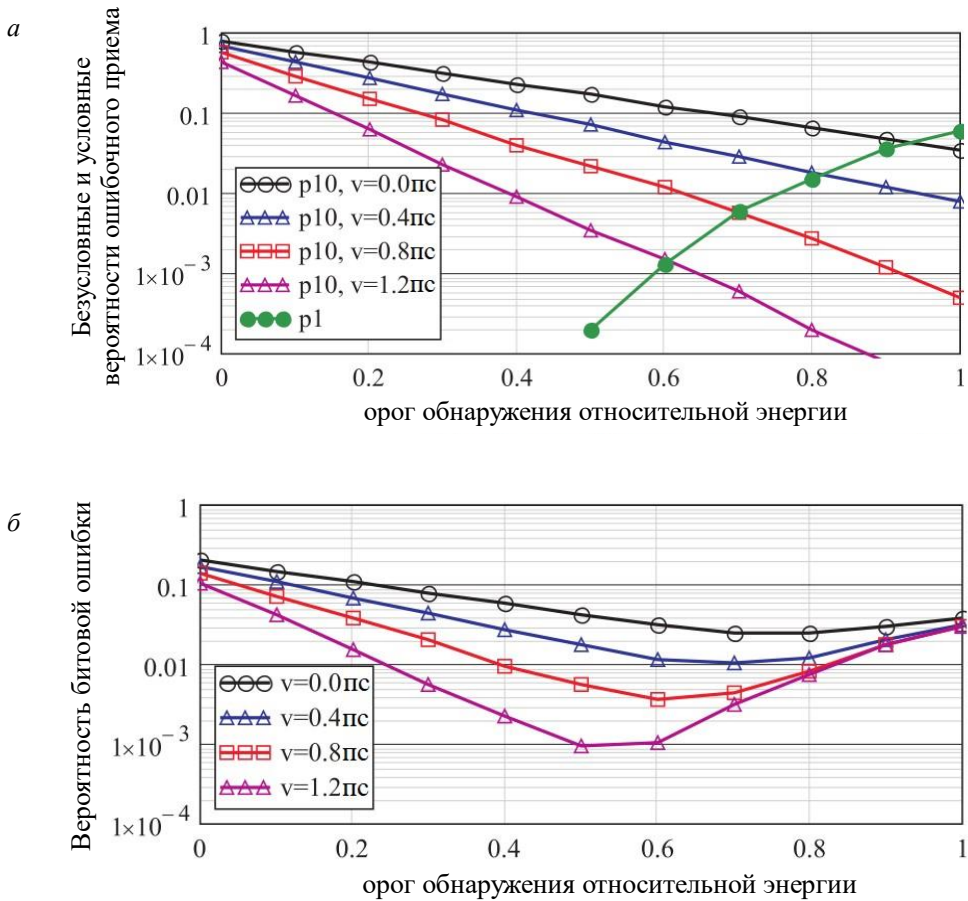


Рис. 12.8. Безусловные и условные вероятности ошибочного приема

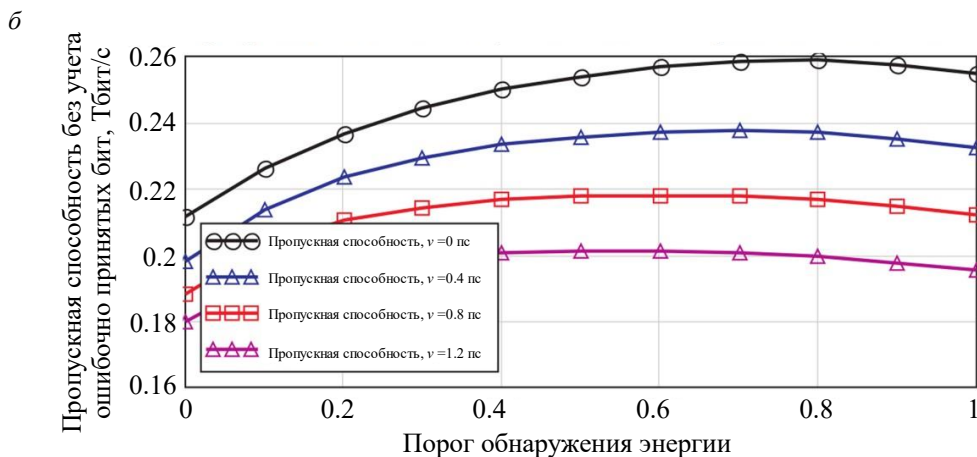
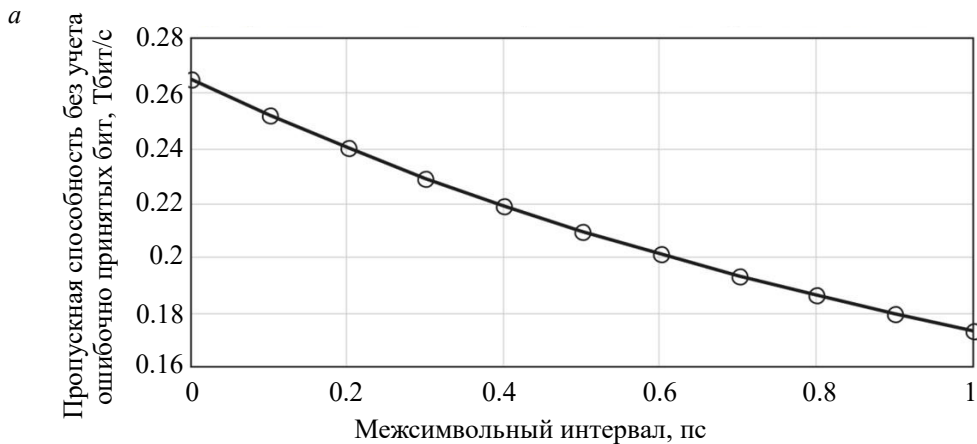


Рис. 12.9. Пропускная способность:

a – без учета ошибочно принятых бит; *б* – с учетом ошибочно принятых бит

Отметим, что обсуждаемые здесь результаты основаны на ряде упрощающих предположений, тем не менее, удалось выявить ряд важных свойств системы, а для нахождения оптимальных значений P_T в зависимости от расстояния d необходимо применять более сложные модели молекулярного шума и распространения в частотной и временной областях. Кроме того, упрощенные аналитические модели могут быть использованы для описания поведения системы для более широкого диапазона входных параметров.

12.3. Особенности устройств связи в терагерцовом диапазоне частот

В данном параграфе кратко изложены некоторые важные ограничения и особенности устройств связи для терагерцового диапазона частот.

А. Температура антенны

Прежде всего, любая радиосвязь — это передача энергии на расстояние. Из-за малых размеров антенны является важным максимальное количество энергии, которое может быть рассеяно с одного антенного элемента для терагерцовых частот, при этом не приводя к перегреву антенны. В любой антенне некоторая часть подводимой радиочастотной энергии поглощается и превращается в тепло, которое необходимо рассеивать.

Определим общее количество тепла, рассеиваемое квадратным элементом патч-антенны с физическими размерами $\lambda/(3 \times 2)$, где λ — длина волны. Параметр $\lambda/2$ в этой формуле относится к резонансному размеру антенного элемента, а дополнительный множитель 3 учитывает корректировку размеров антенны из-за диэлектрической подложки, на которой она размещена, например керамической. В этом случае, с учетом теплового излучения по закону Стефана–Больцмана и конвекции между элементом антенны и окружающей средой, полная рассеиваемая мощность при комнатной температуре, P_{diss} , определяется следующей формулой:

$$P_{diss} = \frac{\lambda^2}{36} (\sigma(T_a^4 - T_r^4) + h_{air}(T_a - T_r)), \quad (12.10)$$

где T_a и T_r — температура антенны и комнатная температура; h_{air} — коэффициент теплопроводности воздуха. Уравневав P_{diss} с количеством полной потребляемой мощности, $P_{cons} = P_{Tx}(1 - \eta)$, где P_{Tx} — мощность передачи и η — коэффициент эффективности антенны, приходим к следующему соотношению:

$$P_{diss} \geq P_{cons} \rightarrow \sigma \cdot (T_a^4 - T_r^4) + h_{air} \times T_a - (T_r \times h_{air} + 36P_{Tx} \times (1 - \eta/\lambda^2)) \geq 0, \quad (12.11)$$

из которого может быть определено значение температуры T_a .

На рис. 12.10 показана равновесная температура одноэлементной патч-антенны для терагерцового диапазона частот при $\eta = 0.95$ и $h_{air} = 100 \text{ Вт/м}^2 \cdot \text{К}$. Как видно из графиков на рис. 12.10, для желаемой мощности излучения 0 дБ температура имеет допустимые

значения (менее 50 °С) только до значения частоты 300 ГГц. Эта же граница сохраняется до значения частоты 1 ТГц для $P_{Tx} = 10$ дБм и до 3 ТГц для $P_{Tx} = 20$ дБм. Таким образом, традиционные металлические антенны поддерживают мощность передачи, сравнимую с более низкими частотами, только в нижнем терагерцовом диапазоне. Если требуется использовать более высокую частоту или мощность, то необходимо использовать большее количество элементов для распределения тепловой нагрузки (так называемые фазированные антенные решетки).

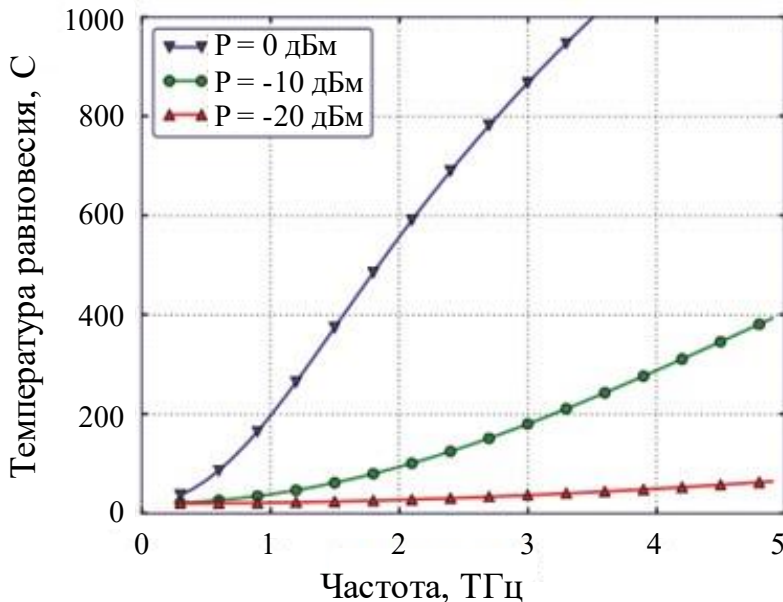


Рис. 12.10. Температура излучающего элемента

Б. Диапазон связи сотового доступа

Рассмотрим достижимую дальность связи для терагерцовых сетей доступа. Сравниваются два случая: 1) антенные решетки как на приеме, так и на передаче представляют собой 2D-фазированные решетки из $M \times M$ элементов и 2) антенная решетка передатчика представляет собой 2D-решетку из $M \times M$ элементов, а антенна приемника является почти всенаправленной, т.е. не имеет усиления. Предположим, что удалось достичь идеального выравнивания лучей. Для проведения анализа максимально допустимые потери связи (дБ) рассчитываем, используя стандартную модель распространения

$$L = P_{Tx} + (173 - 10 \log_{10}(B)) + G_{Tx} + G_{Rx} - S, \quad (12.12)$$

где P_{Tx} – мощность передачи, дБ; B – выделенная полоса частот, G_{Tx} и G_{Rx} – усиление на передаче и приеме, дБ; S – целевое значение ОСШ на пользовательском устройстве. Фиксируя частоту f , находим дальность связи по формуле

$$d = 10^{\frac{L}{20} + 7,38 - \log_{10}(f)}. \quad (12.13)$$

На рис. 12.11 максимальная дальность связи, рассчитанная из соотношения (12.3) для значений $P_{Tx} = 0$ дБм, $S = 5$ дБ, $B = 10$ ГГц. Как видно из графиков на этом рисунке, дальность связи в нижнем терагерцовом диапазоне ограничена десятками метров для случая направленных антенн как на передающей, так и приемной сторонах, и менее двух метров для направленной антенны только на одной стороне. Таким образом, качественно показано, что для сетей связи 6G даже в нижнем диапазоне терагерцовых частот требуются значительные усиления на обеих сторонах беспроводного тракта.

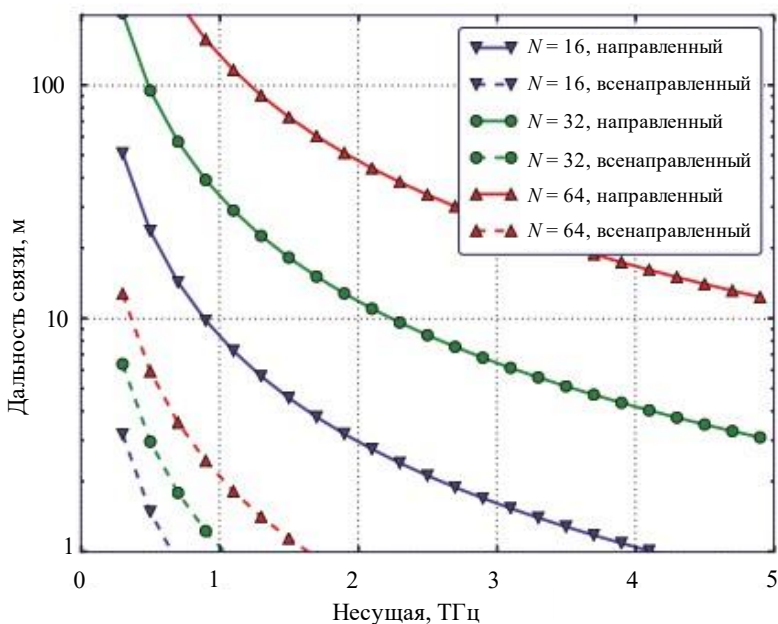


Рис. 12.11. Максимальная дальность связи

В. Дальность связи для множественного доступа

Как сказано выше, для канала связи в терагерцовом диапазоне частот требуется значительное усиление антенн как на передатчике,

так и приемнике. Однако это требование создает проблему для множественного доступа к среде (например, используемого в стандарте IEEE 802.11), так как при использовании направленных антенн в таких условиях возникают проблемы с отсутствием слышимости, несущей в некоторых направлениях. Чтобы обойти эту проблему, можно применить технику разнесения спектра, проще говоря, увеличение времени передачи в N раз может значительно улучшить вероятность корректного приема. Однако в этом случае возникают дополнительные накладные расходы, что теоретически ограничивает применимость этой схемы для трафика с ультрамалой задержкой, определенного для сетей 5G и 6G.

Для простоты анализ проведен, используя свойства терагерцового канала связи, определенного для множественного доступа к среде (DCF, Distributed Coordination Function), используемой в IEEE 802.11. В этом случае граница задержки на прикладном уровне соответствует максимальному значению длительности возможной передачи (TXOP, Transmit Opportunity), причем это значение должно быть приведено в соответствие числу активных пользователей. Максимальная длительность TXOP, в свою очередь, ограничивает длительность кадра данных, что определяет значение верхней границы спектральной эффективности

$$S \leq \frac{t_{Data}}{t_{TXOP}} = \frac{t_{TXOP} - 3t_{RTS} - 3t_{SIFS}}{t_{TXOP}} = 1 - 3 \frac{t_{RTS} + t_{SIFS}}{t_{TXOP}}. \quad (12.14)$$

Следуя стандарту IEEE 802.11ad, задаем длительность короткого межкадрового интервала (SIFS, Short Inter-Frame Space) как 4 OFDM-символа, каждый из которых имеет длительность τ . Каждый символ OFDM требует N_{FFT} поднесущих, и, следовательно, $\tau = \frac{N_{FFT}}{B}$, где B – полоса частот канала связи. Определяем длительность кадра запроса на отправку (RTS, Request-to-Send) теми же 4 символами, но повторяемыми G_{Rx} раз, чтобы компенсировать отсутствие усиления антенны на приемнике. В результате максимальная спектральная эффективность определяется соотношениями

$$S_{max} = 1 - 3 \frac{t_{RTS} + t_{SIFS}}{t_{TXOP}} = 1 - 12\tau \frac{1 + G_{Rx}}{t_{TXOP}} = 1 - 12\tau \frac{N_{FFT}(1 + G_{Rx})}{t_{TXOP}B}. \quad (12.15)$$

На рис. 12.12 показана зависимость между спектральной эффективностью и дальностью связи для трех значений длительности TXOP при ширине полосы частот канала 10 ГГц. Как видно из графиков, рассмотренный метод множественного доступа крайне

неэффективен при работе с требовательными к задержкам узлами. Например, эффективная дальность связи для $t_{ТХОР} = 0,1$ мс (10 пользователей, каждый с рекомендованной в стандарте ИМТ-2020 задержкой в 1 мс) ограничена всего 15 м при мощности передатчика 0 дБм и усилении антенны 15 дБ. В то же время, когда ограничения на задержку не такие строгие и составляют $t_{ТХОР} = 10$ мс, обсуждаемая схема может быть применена с минимальными ограничениями.

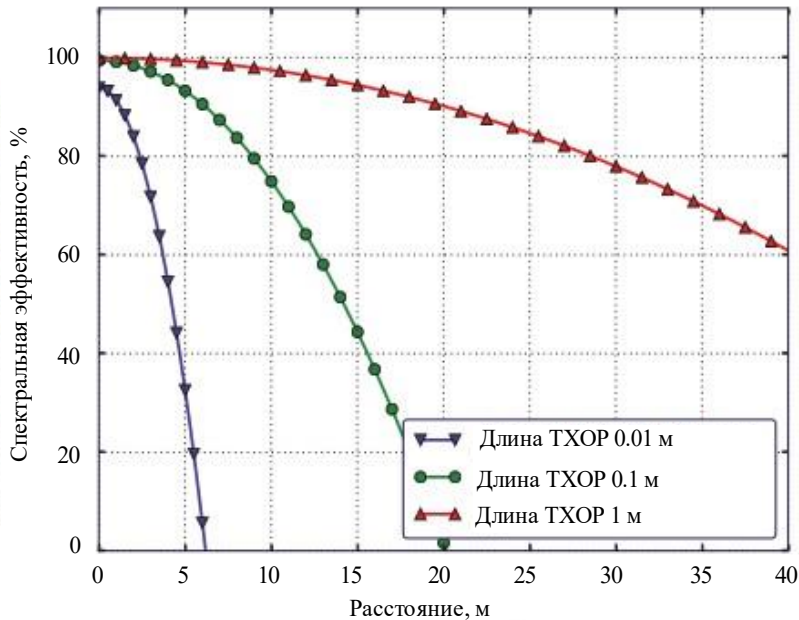


Рис. 12.12. Спектральная эффективность

Подводя итог, отметим, что не критичный к задержке трафик может быть эффективно обслужен в терагерцовых сетях с децентрализованным множественным доступом, а поддержка критичных к задержкам типов приложений (например, Тактильный интернет) может быть обеспечена только в том случае, когда антенны на передатчике и на приемнике являются направленными. Поэтому разработка эффективного и надежного протокола доступа к среде для терагерцовой связи, способного работать в случае двусторонней направленности, является одним из важнейших исследовательских вопросов.

ГЛАВА 13

ПРОБЛЕМЫ И РЕШЕНИЯ ДЛЯ СЕТЕЙ СВЯЗИ 6G В ТЕРАГЕРЦОВОМ ДИАПАЗОНЕ ЧАСТОТ

Помимо проблем, связанных с увеличением потерь распространения, атмосферной абсорбцией и блокировкой, терагерцовый диапазон частот предоставляет и новые возможности для разработчиков технологии связи 6G. В этой главе мы обсудим негативные свойства терагерцовых каналов связи, а также возможности использования их положительных свойств. В параграфе 13.1 обсуждаются эффекты микромобильности, которые могут приводить к обрывам связи и дан анализ производительности терагерцовых каналов при использовании различных методов поиска луча. Далее в 13.2 рассмотрены методы защиты информации на физическом уровне, используя свойства сверхнаправленности фазированных антенных решеток. Наконец, в параграфе 13.3 рассмотрен метод доступа на «последнем метре».

13.1. Эффект микромобильности

Рассмотрим связь типа *точка–точка* между стационарной точкой доступа в терагерцовом диапазоне (THz-AP, THz Access Point) и абонентским мобильным устройством с возможностью подключения в терагерцовом диапазоне (THz-UE, THz User Equipment). Как THz-AP, так и THz-UE оснащены одинаковыми планарными 2D антенными решетками с $N(\alpha)$ элементами каждая. Используется модель распространения для терагерцовой связи и диаграмма излучения в виде конуса, где коэффициент усиления G является функцией угла направленности антенны $\alpha - G = 2/(1 - \cos(\alpha/2))$. Предполагаем постоянную спектральную плотность мощности шума, равную -174 дБм/Гц.

Система работает следующим образом. После завершения процедуры поиска луча в момент времени $t_0 = 0$ предполагается, что лучи идеально выровнены. Процесс микромобильности UE влияет на поиск луча и может привести к падению ОСШ ниже допустимого уровня в некоторый момент времени $t_1 > t_0$. Затем поиск луча начинается снова. Для простоты мы предполагаем иерархический механизм поиска луча, где время поиска T_S зависит от времени переключения решетки δ , количества элементов передающей и приемной антенных решеток N_T и N_R соответственно, $T_S = (N_T + N_R)\delta = 2\delta N(\alpha)$.

Метриками, представляющими интерес, являются ОСШ как функция времени, $SNR(t)$, и средняя пропускная способность канала связи по Шеннону $E[C]$, где $C = B \log_2(1 + SNR(t))$.

Измерения и иллюстрация эффекта

При проведении измерений iPhone 5SE выступал в качестве АУ, а программное обеспечение Sensor Kinetics Pro использовалось для регистрации изменений положения и поворотов устройства во время испытаний. Пользователь оставался неподвижным в течение всего испытания (никаких крупномасштабных перемещений) и постоянно держал устройство двумя руками. Были выбраны три типовые модели мобильности.

1. Просмотр видео. АУ использовалось для просмотра видео на YouTube.

2. Имитатор полета. АУ использовалось для игры в авиасимулятор, где управление осуществляется с помощью вращения устройства.

3. Игра с мячом. АУ использовалось для игры, в которой пользователь стремится удержать мяч на поверхности как можно дольше. Управление осуществляется с помощью движений и вращений.

Продолжительность всех испытаний составила 1 мин. Результаты измерений мощности для каждого из шаблонов были усреднены по 20 различным образцам. Результаты измерений микромобильности АУ были использованы для моделирования ОСШ и пропускной способности терагерцового канала между THz-AP и THz-UE. Для численного исследования была выбрана полоса шириной 50 ГГц в диапазоне 275–325 ГГц. Расстояние до канала связи равнялось – 2 м, а мощность передачи $P_{Tx} = 0$ дБм.

На рис. 13.1 представлен пример поведения ОСШ для диаграммы направленности игры с мячом и трех различных значений угла направленности антенны, $\alpha = 5^\circ, 10^\circ, 20^\circ$. Из графика на рисунке видно, что каждый из углов направленности антенны имеет максимальный уровень ОСШ, достигаемый при идеальном согласовании излучения антенны. Эти значения составляют 45,45 дБ, 33,41 дБ и 21,39 дБ соответственно. Можно сделать вывод, что меньшее значение α приводит к большему значению ОСШ благодаря более высокому коэффициенту усиления антенны. Между тем

меньшее α приводит к менее частым и серьезным ухудшениям ОСШ из-за малой мобильности АУ.

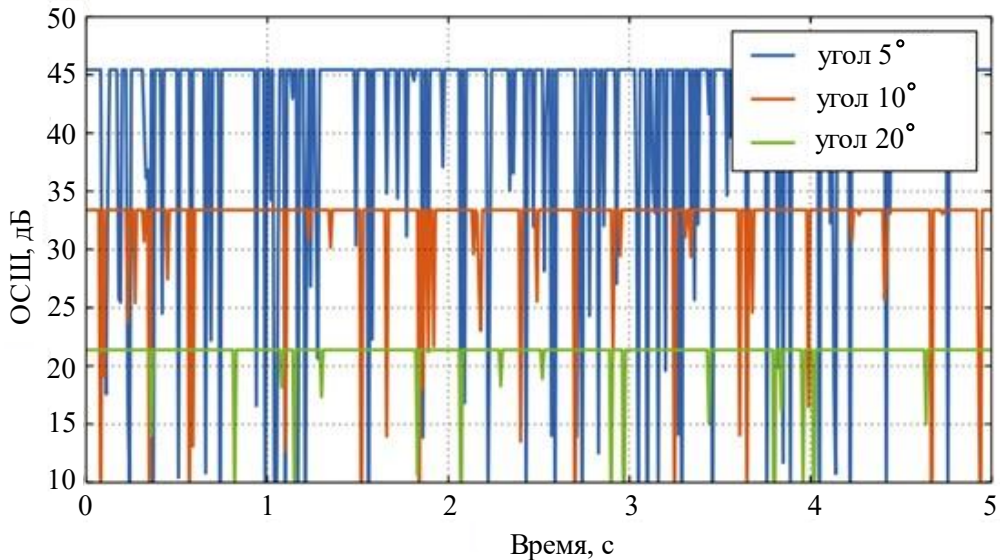


Рис. 13.1. ОСШ для нескольких направленностей антенны

На рис. 13.2 показана усредненная по времени пропускная способность (скорость) терагерцового канала связи в зависимости от α для трех выбранных схем микромобильности и теоретической верхней границы для полностью статичной линии связи. Этот рисунок иллюстрирует, что каждая из выбранных моделей имеет свое «оптимальное» значение угла направленности антенны: при всех меньших значениях α спонтанное ухудшение ОСШ начинает доминировать над дополнительным выигрышем и средняя пропускная способность быстро уменьшается. Для выбранного набора шаблонов эти значения составляют около 8° , 11° и 22° .

Важно отметить, что для значений α больше «оптимальных» пропускная способность начинает быстро убывать и отклонение в несколько градусов может даже привести к снижению пропускной способности на порядки. Таким образом, правильная настройка ширины луча в соответствии с изменениями в шаблоне подвижности чрезвычайно важна для разработки высокоскоростных и надежных терагерцовых каналов связи. Это подразумевает необходимость дальнейшего исследования моделей микромобильности и их интеграцию в методики оценки производительности таких каналов связи.

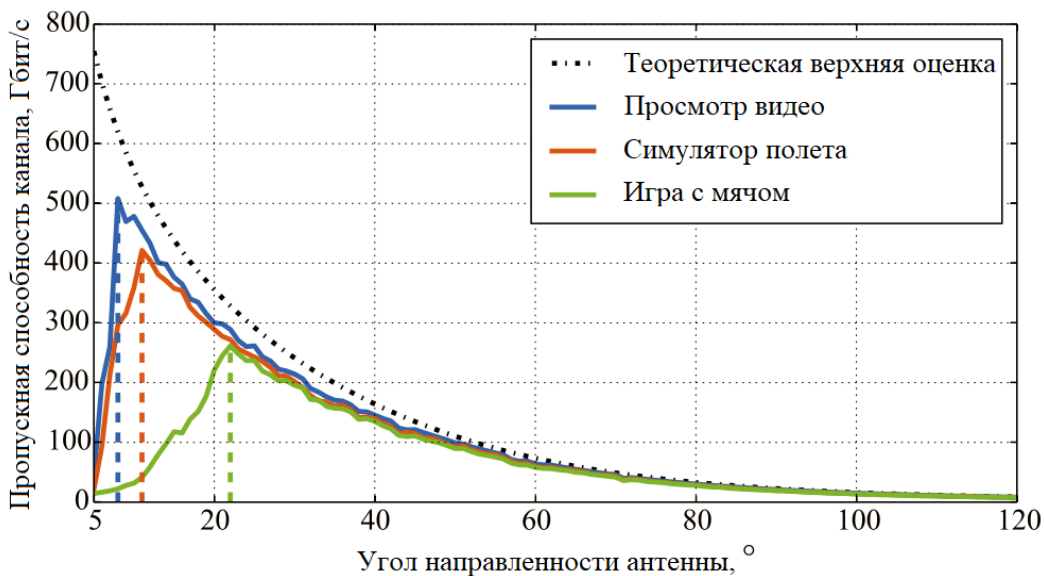


Рис. 13.2. Пропускная способность канала связи по Шеннону

Предположения о системной модели

Использование очень узких диаграмм направленности терагерцового излучения в мобильной связи ставит под сомнение эффективность процедур управления лучом для отслеживания подвижности узлов. Для портативных и носимых терагерцовых устройств необходимо учитывать не только обычную макромобильность (ходьба человека, движение автомобиля и т.д.), но и гораздо менее предсказуемую микромобильность (малые смещения и вращения). Такие быстрые перемещения пользовательского устройства могут вызвать частые смещения высоконаправленных терагерцовых лучей, что, в свою очередь, приведет к возможным перебоям и снижению пропускной способности канала.

Для анализа эффектов микромобильности рассмотрим систему в следующих предположениях.

1. *Модель развертывания и распространения.* Рассматривается один канал связи между стационарной THz-AP и портативным THz-UE, на расстоянии d метров друг от друга (рис. 13.3). Для моделирования узких терагерцовых лучей предполагаем, что и THz-AP, и THz-UE оснащены 2D антенными решетками из $N_A \times N_A$ и $N_U \times N_U$ ($N_A \gg N_U$) элементов, создающих 3D секторные диаграммы направленности излучения антенны с коэффициентами усиления $G_A = N_A^2$,

$G_U = N_U^2$. Соответствующие углы диаграммы направленности излучения антенн для THz-AP и THz-UE задаются величинами $\alpha = \frac{102\pi}{180N_A}$ и $\beta = \frac{102\pi}{180N_U}$.

Применяется терагерцовая модель распространения, для которой ОСШ в THz-UE, S_U , определяется формулой

$$S_U = P_A G_A G_U \frac{c^2}{16\pi^2 f^2 N_0(B)} d^{-2} e^{-Kd}, \quad (13.1)$$

где f – несущая частота; B – полоса пропускания; K – коэффициент поглощения для полосы; $N_0(B)$ – уровень шума.

2. *Модель микромобильности.* Предполагается, что микромобильность THz-UE представляет собой комбинацию процессов случайных перемещений $x(t)$ и $y(t)$ по осям OX и OY вместе со случайными процессами вращения $\phi(t)$ и $\theta(t)$ (см. рис. 13.3). Эти СП моделируются как независимые случайные блуждания и параметризуются средним значением смещения за одну секунду: $\Delta x, \Delta y, \Delta \phi$ и $\Delta \theta$.

Для простоты мы не учитываем ни микроперемещения вдоль оси OZ $z(t)$ на рис. 13.3, ни повороты устройства вокруг этой оси $\xi(t)$. Причина заключается в том, что для рассматриваемого сценария ни перемещения вдоль пути линии связи на несколько сантиметров, ни повороты перпендикулярно линии связи не оказывают заметного влияния на производительность на канальном уровне.

3. *Схемы поиска луча.* Разработка эффективных механизмов поиска луча для терагерцовой связи все еще продолжается, поэтому рассматриваемая здесь математическая модель не зависит от выбора конкретной процедуры поиска луча. Предполагается лишь, что ее продолжительность постоянна и равна T_B .

На практике величина T_B зависит от многих факторов, включая используемые алгоритмы поиска луча, размеры антенных решеток и время, необходимое для оценки одной конфигурации лучей. Чтобы учесть эти факторы в численном анализе, мы моделируем последовательный поиск луча (аналогично IEEE 802.11ad) со сложностью $T_B = (N_A^2 + N_U^2)\delta$, где δ – так называемая задержка направления луча, т.е. время, необходимое терагерцовому узлу для сканирования в одном выбранном направлении, плюс время для изменения выбранного направления. Можно использовать и другие формулы для вычисления величины T_B при моделировании других процедур поиска луча.

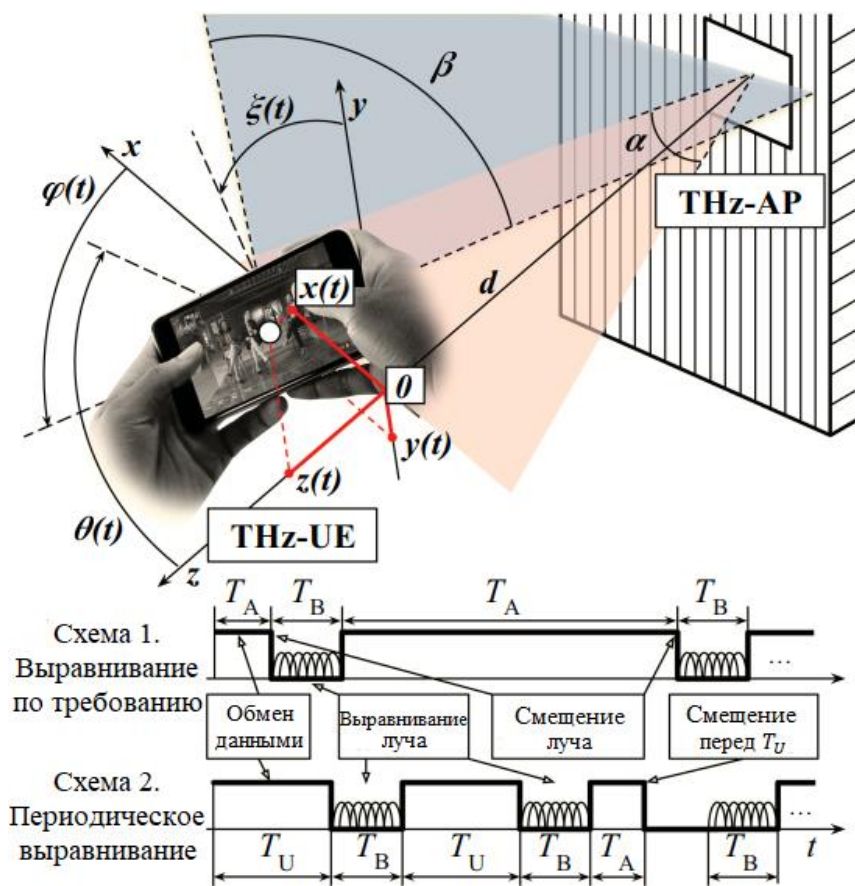


Рис. 13.3. Иллюстрация модели терагерцовой системы

В нашем исследовании анализируются две схемы, показанные на рис. 13.3.

- *Схема 1 – поиск по требованию.* Процедура поиска луча запускается каждый раз, когда микроподвижность THz-UE приводит к потере связи. Эта схема подходит для беспроводных локальных сетей.
- *Схема 2 – периодический поиск.* Процедура перестройки луча выполняется периодически с периодом $T_U + T_B$. Эта схема отражает системы сотового типа с централизованным управлением.

В следующем разделе предложена математическая модель для анализа и сравнения характеристик этих схем в присутствии микроподвижности THz-UE. Для простоты мы предполагаем, что обнаружение состояния потери связи (т.е. потеря луча) происходит мгновенно. Однако схема может быть расширена для учета ненулевого времени обнаружения потери связи путем введения дополнительного компонента в величину T_B .

Метод анализа системы

Ниже рассматривается подход к анализу и сравнению характеристик схемы 1 и схемы 2 в условиях микромобильности ТНz-УЕ. Задача состоит в нахождении функции плотности с.в. времени, в течение которого микромобильность вызывает потерю луча. Полученный результат применяется для оценки производительности терагерцового канала связи.

Несовпадение лучей из-за микромобильности

1. *Компоненты микромобильности.* Предположим, что в момент $t = 0$ лучи ТНz-АР и ТНz-УЕ идеально выровнены. Для нахождения компонентов микромобильности $\phi(t)$, $\theta(t)$, $x(t)$, $y(t)$ используем процесс броуновского движения, описываемый стохастическим уравнением

$$\frac{\partial c(x,t)}{\partial t} - D \frac{\partial^2 c(x,t)}{\partial x^2} = 0, \quad (13.2)$$

где «концентрация» $c(x, t)$ интерпретируется как вероятность наличия диффундирующей точки в координате x в момент t , $D = (\Delta x)^2 / 2\Delta t$ – коэффициент диффузии, Δx – смещение, соответствующее приращению времени Δt . Заметим, что компоненты случайных процессов связаны с коэффициентами диффузии D_ϕ , D_θ , D_x , D_y . За единицу времени выберем секунду и определим $D_x = (\Delta x)^2 / 2$.

2. *Микромобильность $x(t)$ и $y(t)$.* Как и в случае $N_A \gg N_U$, луч ТНz-АР более узкий, чем луч ТНz-УЕ. Тогда смещение произойдет, когда либо $x(t)$, либо $y(t)$ покинут пучок ТНz-АР с геометрическими границами $[-M_{XY}, +M_{XY}]$, где $M_{XY} = d \tan(\alpha/2) = d \tan\left(\frac{102\pi}{360N_A}\right)$.

Рассмотрим сначала $x(t)$. Плотность с.в. времени первого пролета (FPT, First Passage Time) до симметричных границ $\pm M_{XY}$ имеет следующий вид:

$$f_X(t) = \sum_{n=-\infty}^{\infty} \frac{M_{XY}[4n+1] \left[\exp\left(-\frac{M_{XY}^2}{4D_x t}\right) \right]^{(4n+1)^2}}{\sqrt{\pi D_x t^3}}, \quad t > 0. \quad (13.3)$$

Интегрируя (13.3), получаем формулу для ФР:

$$F_X(t) = \sum_{n=-\infty}^{\infty} \frac{\left[\operatorname{erf} \left(\frac{M_{XY} \sqrt{\frac{8n^2+4n+1}{D_X t}}}{\sqrt{2}} \right) - 1 \right]}{\sqrt{4n^2+2n+1/2}}, \quad (13.4)$$

где $\operatorname{erf}(x)$ – функция Лапласа (известная также как функция ошибок).

Аналогично определяем плотность вероятности ФРТ для $y(t)$, $f_Y(t)$, и тогда время, необходимое для потери соединения из-за изменения $x(t)$ или $y(t)$, находится как минимум двух с.в. с плотностью вероятности

$$f_{T_{XY}}(t) = f_X(t)[1 - F_Y(t)] + f_Y(t)[1 - F_X(t)]. \quad (13.5)$$

3. *Микромобильность $\phi(t)$ и $\theta(t)$.* Теперь учтем в анализе угловую подвижность. Потеря связи, вызванная вращением, происходит, когда ТНz-АР оказывается вне поворачивающегося луча ТНz-УЕ. Поэтому угол, на который должен повернуться ТНz-УЕ, чтобы вызвать рассогласование, зависит от текущего местоположения ТНz-УЕ, $x(t)$ и $y(t)$. В то же время для любого смещения Δx и Δy граничный угол лежит в границах $M_{\phi\theta} \in \left[\frac{\beta}{2} - \frac{\alpha}{2}, \frac{\beta}{2} + \frac{\alpha}{2} \right]$, $\alpha = \frac{102\pi}{180N_A}$, $\beta = \frac{102\pi}{180N_U}$ (см. рис. 13.3).

Мы используем верхнее ограничение на

$$M_{\phi\theta}, M_{\theta\phi} = \frac{102\pi}{360} (1/N_U + 1/N_A)$$

так, что, когда $\phi(t)$ или $\theta(t)$ достигают $\pm M_{\phi\theta}$, луч теряется. Наконец, мы можем охарактеризовать ФРТ плотности вероятности для $\phi(t)$ и $\theta(t)$ – $f_\phi(t)$ и $f_\theta(t)$ – определены ранее, а $f_{T_{\phi,\theta}}(t)$ получается с помощью выражения для определения плотности вероятности минимума СВ.

4. *Агрегированная микромобильность.* Плотность вероятности с.в. T_A , определяющей ФРТ для достижения любого из граничных условий $\pm M_{XY}$ или $\pm M_{\phi\theta}$, может быть найдена как плотность минимума двух с.в. по формуле

$$f_{T_A}(t) = f_{T_{XY}}(t)[1 - F_{T_{\phi,\theta}}(t)] + f_{T_{\phi,\theta}}(t)[1 - F_{T_{XY}}(t)]. \quad (13.6)$$

Вычисление плотности непосредственно по формуле (13.6) представляет большую сложность. Учитывая, что среднее и дисперсия с.в. ФРТ конечны и равны $M_{XY}^2/2D_X$ и $M_{XY}^4/6D_X^2$, то для случая с одним

компонентом микромобильности можно использовать приближение логнормальным распределением с теми же первыми моментами, т.е.

$$f_{T_{X,Y}}(t) = \frac{1}{2\sqrt{2\pi}\sigma_y\sigma_x t} \left[\sigma_x e^{-\frac{(\mu_x - \log(t))^2}{2\sigma_y^2}} \left[\operatorname{erf}\left(\frac{\mu_x - \log(t)}{\sqrt{2}\sigma_x}\right) + 1 \right] + \right. \\ \left. + \sigma_y e^{-\frac{(\mu_x - \log(t))^2}{2\sigma_x^2}} \left[\operatorname{erf}\left(\frac{\mu_x - \log(t)}{\sqrt{2}\sigma_y}\right) + 1 \right] \right], t > 0, \quad (13.7)$$

где $\mu_x = \log\left(\frac{2M_{XY}}{\Delta x}\right) - \frac{1}{2}\log\left(\frac{5}{3}\right)$, $\mu_y = \log\left(\frac{2M_{XY}}{\Delta y}\right) - \frac{1}{2}\log\left(\frac{5}{3}\right)$,
 $\sigma_x = \sigma_y = \sqrt{\log(5/3)}$, а также $\mu_\phi = \log\left(\frac{2M_{XY}}{\Delta\phi}\right) - \frac{1}{2}\log\left(\frac{5}{3}\right)$,
 $\mu_\theta = \log\left(\frac{2M_{XY}}{\Delta\theta}\right) - \frac{1}{2}\log\left(\frac{5}{3}\right)$, $\sigma_\phi = \sigma_\theta = \sqrt{\log(5/3)}$.

Теперь, чтобы найти ФРТ для агрегированной мобильности, достаточно использовать выражение для определения плотности минимума случайных величин, где ФР $F_{T_{X,Y}}(t)$ и $F_{T_{\phi,\theta}}(t)$ могут быть получены интегрированием соответствующих плотностей вероятности.

Метрики терагерцовой связи с микромобильностью

Теперь определим а) долю времени, в течение которого ТНz-UE находится в состоянии отсутствия связи, включая время, затраченное на перестройку луча и б) среднюю спектральную эффективность и пропускную способность терагерцового канала.

Схема 1. Рассмотрим схему, в которой процедура поиска луча вызывается только тогда, когда ТНz-UE попадает в состояние отсутствия связи. Доля времени отсутствия связи равна $p_{O,1} = T_B/(T_A + T_B)$, где T_B – длительность процедуры поиска луча и может быть представлена в виде

$$p_{O,1} = \int_0^\infty \frac{T_B}{t+T_B} f_{T_A}(t) dt. \quad (13.8)$$

Чтобы получить среднюю спектральную эффективность $E[L_1]$ и пропускную способность терагерцового канала $E[C_1]$, напомним, что с вероятностью $p_{O,1}$ эти величины равны нулю и определяются по формуле (13.10). В случае, когда лучи выровнены определяются по формуле Шенона

$$L_{max} = \log_2(1 + S_U), \quad C_{max} = B \log_2(1 + S_U); \quad (13.9)$$

$$E[L_1] = (1 - p_{O,1})L_{max}, \quad E[C_1] = (1 - p_{O,1})C_{max}. \quad (13.10)$$

Среднее время до первой потери луча определяется формулой

$$E[T_1] = \int_0^{\infty} t f_{T_A}(t) dt. \quad (13.11)$$

Схема 2. Рассматриваем схему, в которой поиск луча вызывается через фиксированный интервал времени T_U . Величина $T_{O,2}$ продолжительности отсутствия связи за время $T_U + T_B$ имеет вид

$$T_{O,2} = \begin{cases} \frac{T_B}{T_U + T_B}, T_A \geq T_U, \\ \frac{T_U + T_B - T_A}{T_U + T_B}, T_A < T_U, \end{cases} \quad (13.12)$$

и

$$Pr\{T_A < T_U\} = \int_0^{T_U} f_{T_A}(t) dt = F_{T_A}(T_U). \quad (13.13)$$

Условная плотность вероятности с.в. времени до потери луча при условии, что оно меньше T_U , равна $f_{T_A}(t)/F_{T_A}(T_U)$. Тогда доля времени в состоянии отсутствия связи может быть получена путем взвешивания средних в виде

$$p_{O,2} = \frac{T_B [1 - F_{T_A}(T_U)]}{T_U + T_B} + \int_0^{T_U} \frac{T_U + T_B - t}{T_U + T_B} f_{T_A}(t) dt. \quad (13.14)$$

Тогда средние значения спектральной эффективности и пропускной способности терагерцового канала для схемы 2 вычисляются по формулам

$$E[L_2] = (1 - p_{O,2})L_{max}, \quad E[C_2] = (1 - p_{O,2})C_{max}. \quad (13.15)$$

В отличие от схемы 1 среднее время до первой потери луча $E[T_2]$ в случае схемы 2 может быть больше, чем величины T_A . Учитывая, что T_2 состоит из нескольких последовательных интервалов длительностью $T_U + T_B$ плюс соответствующая доля интервала T_A , получаем

$$E[T_2] = \frac{1 - F_{T_A}(T_U)}{F_{T_A}(T_U)} (T_U + T_B) + \frac{1}{F_{T_A}(T_U)} \int_0^{T_U} t f_{T_A}(t) dt. \quad (13.16)$$

Численный анализ

В этом разделе представлены результаты численного анализа. В соответствии со стандартами IEEE 802.15.3d (см. главу 1)

$f = 0,3$ ТГц, $B = 50$ ГГц и $P_A = 20$ дБм. Задержка смены конфигурации антенной решетки δ установлена в 5 мкс. Для оценки влияния декартовой и угловой микромобильности мы полагаем, что $\Delta x = \Delta y$ и $\Delta \phi = \Delta \theta$. Для иллюстрации взяты значения Δx и $\Delta \phi$ из натурального эксперимента. В том числе $\Delta x = 0,1$ м и $\Delta \phi = 4^\circ$ характерны для игры на смартфоне, а $\Delta x = 0,01$ м; $\Delta \phi = 3^\circ$ соответствуют сценарию просмотра видео.

Влияние периода перестройки луча, T_U

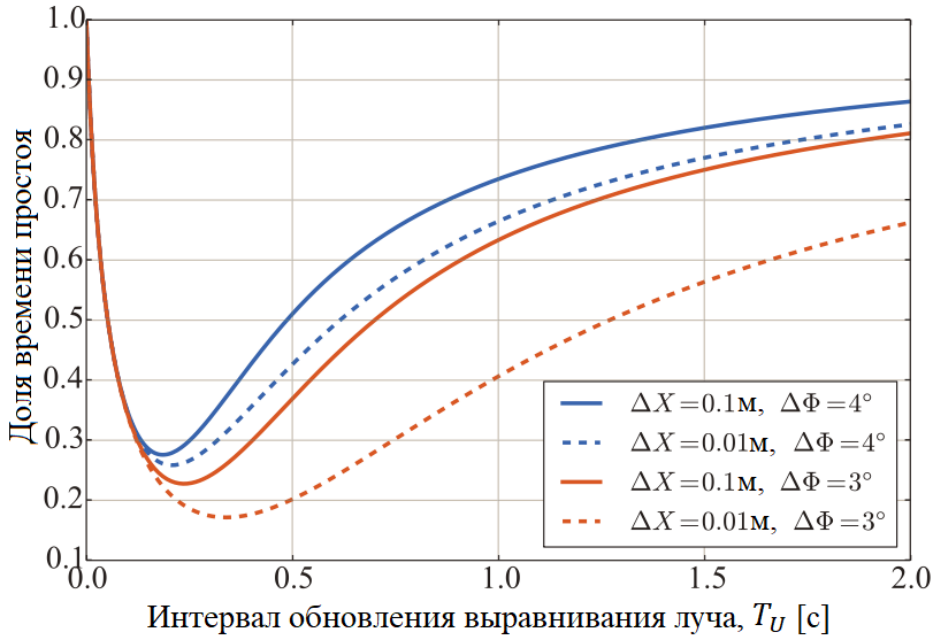
Начнем с рис. 13.4а, на котором представлена доля времени отсутствия связи в зависимости от T_U для схемы 2. Из графиков на рисунке видно, что малое значение T_U приводит к тому, что система тратит большую часть времени на поиск луча, а слишком большие значения $T_U > 1$ с также приводят к снижению пропускной способности системы. Таким образом, существует некоторое оптимальное значение T_U для каждой из конфигураций, а типичное значение лежит в пределах около 200–300 мс и уменьшается с уровнем микромобильности.

Влияние размера антенной решетки при THz-UE, N_U

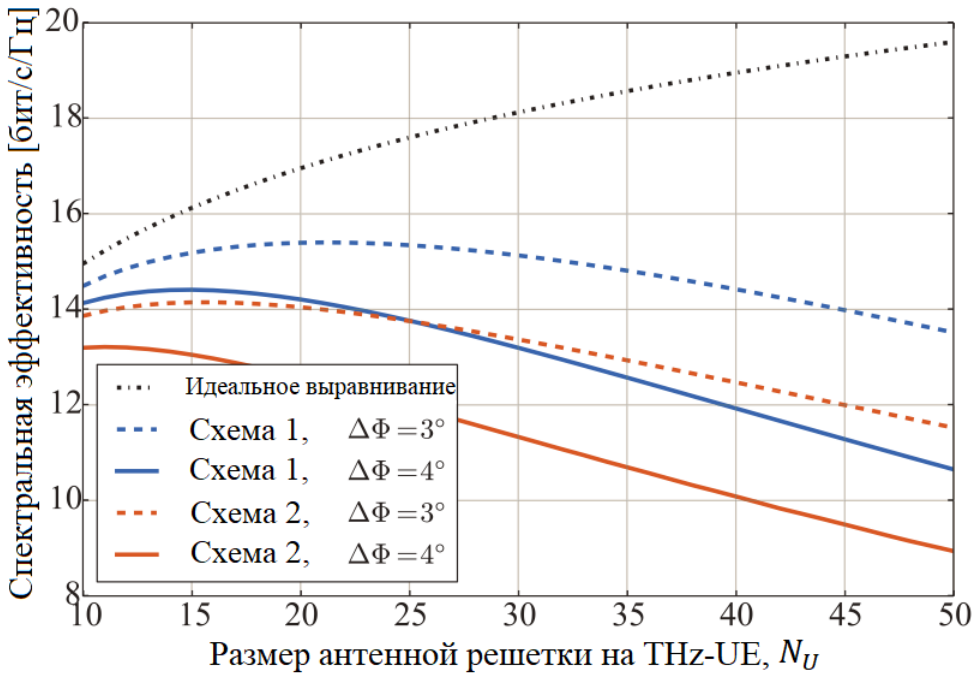
Перейдем к рис. 13.4б, иллюстрирующему влияние N_U на спектральную эффективность SE канала связи при $N_A = 100$ и $\Delta x = 0,1$ м. Для схемы 2 используются «оптимальные» значения T_U , показанные на рис. 13.4а. Анализируя представленные данные, видно, что относительно малые значения угловой микромобильности $\Delta \phi = 3^\circ$ приводят к снижению SE до 70% по сравнению с теоретическим пределом $N_U = 50$.

Можно обсудить проблему оптимизации, поскольку большее значение N_U улучшает SE во время передачи данных, но одновременно увеличивает частоту и длительность поиска луча. Например, схема 1 достигает наивысшей производительности при $N_U = 21$ для $\Delta \phi = 3^\circ$ и при $N_U = 15$ для $\Delta \phi = 4^\circ$. Большой уровень угловой микромобильности делает меньшее значение N_U предпочтительным и снижает максимальную SE – 15,4 бит/с/Гц против 14,4 бит/с/Гц. Аналогичные эффекты наблюдаются и для схемы 2. В заключение отметим, что схема 1 не намного превосходит схему 2, сохраняя максимальную спектральную эффективность примерно на 1 бит/с/Гц выше.

a



b



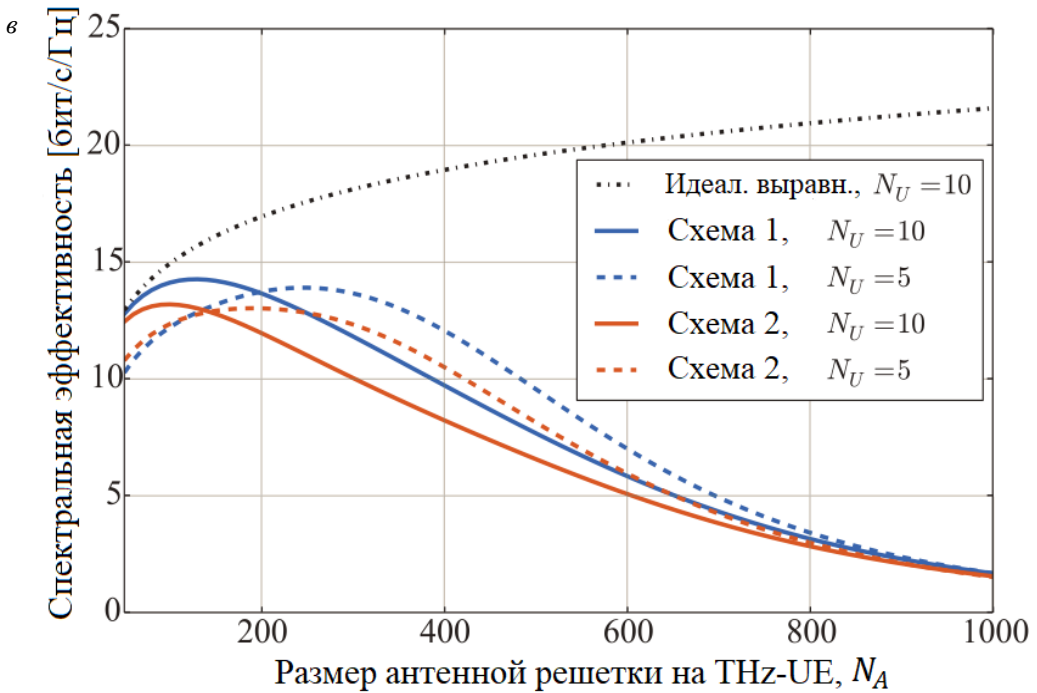


Рис. 13.4. Иллюстрация настройки параметров системы (пояснения в тексте)

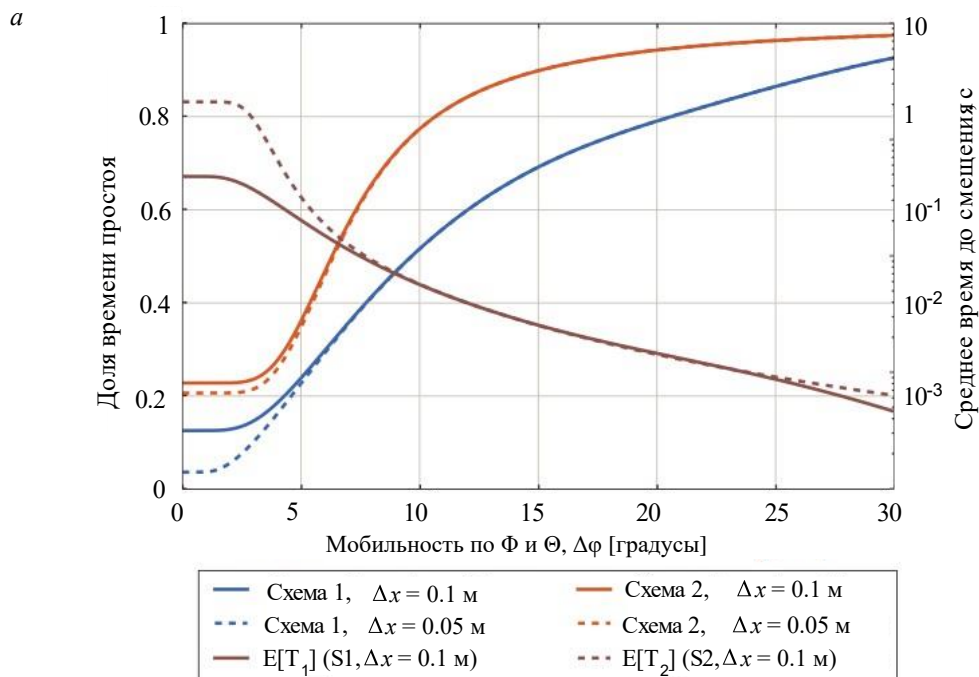
Влияние размера антенной решетки при THz-AP, N_A

На рис. 13.4в показаны средние значения SE в зависимости от N_A . Помимо эффектов аналогичных наблюдаемым на рис. 13.4б, отмечаем, что выбранное значение N_U влияет на «оптимальное» значение N_A . Для схемы 2 спектральная эффективность достигает максимума при $N_A = 190$ и $N_U = 5$, а также при $N_A = 100$ и $N_U = 10$. Таким образом, для получения максимальной спектральной эффективности N_A и N_U должны быть оптимизированы совместно для каждого набора характеристик микромобильности. Графики на рисунке также показывают, что сверхмассивные антенные решетки ($N_A \approx 1024$) в настоящее время неактуальны для мобильных систем, поскольку они характеризуются как малым значением времени в состоянии связи T_A , так и высокими накладными расходами на поиск луча T_B .

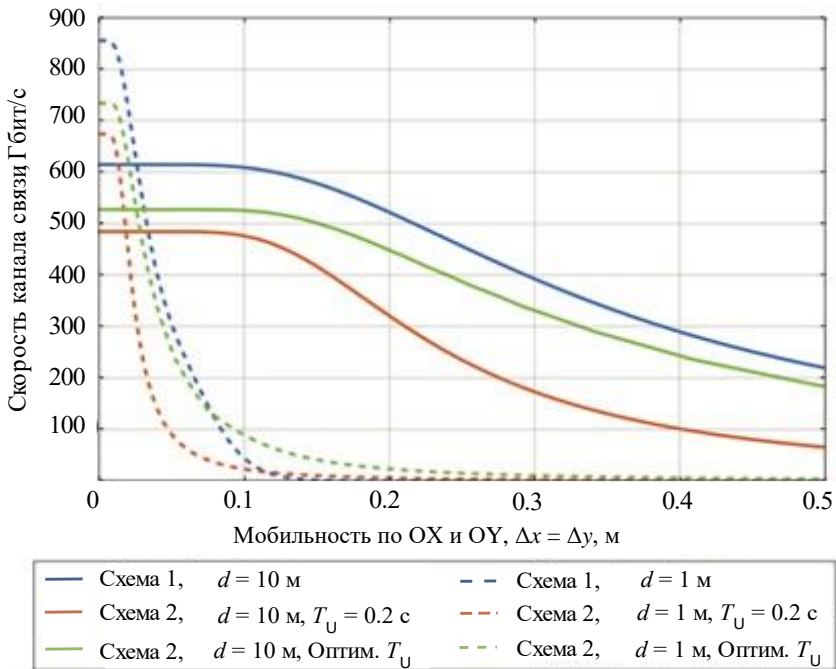
Влияние угловой микромобильности

На рис. 13.5а показаны доля времени простоя (левая ось) и среднее время до первой потери луча (правая ось) в зависимости от величины $\Delta\varphi$. Значение N_A задано равным 100, а значение N_U – равным 20. Начнем с доли времени простоя системы. Для обеих схем существует режим $\Delta\varphi < 3^\circ$, где графики параллельны оси ОХ, так как при малом значении $\Delta\varphi$ производительность системы ограничена в основном декартовой микромобильностью (смещениями по осям координат), поэтому незначительные изменения $\Delta\varphi$ не оказывают существенного влияния.

Напротив, для $\Delta\varphi > 10^\circ$ влияние угловой микромобильности преобладает. Это объясняет, что кривые для $\Delta x = 0,1$ м и $\Delta x = 0,01$ м ведут себя аналогично для обеих схем. Отмечаем, что схема 1 превосходит схему 2 по доле времени в состоянии отсутствия связи. Противоположная тенденция наблюдается для среднего времени до смещения, где $E[T_2] > E[T_1]$ для всех значений $\Delta\varphi$. Выигрыш наиболее заметен при $\Delta\varphi = 6^\circ$.



б



в

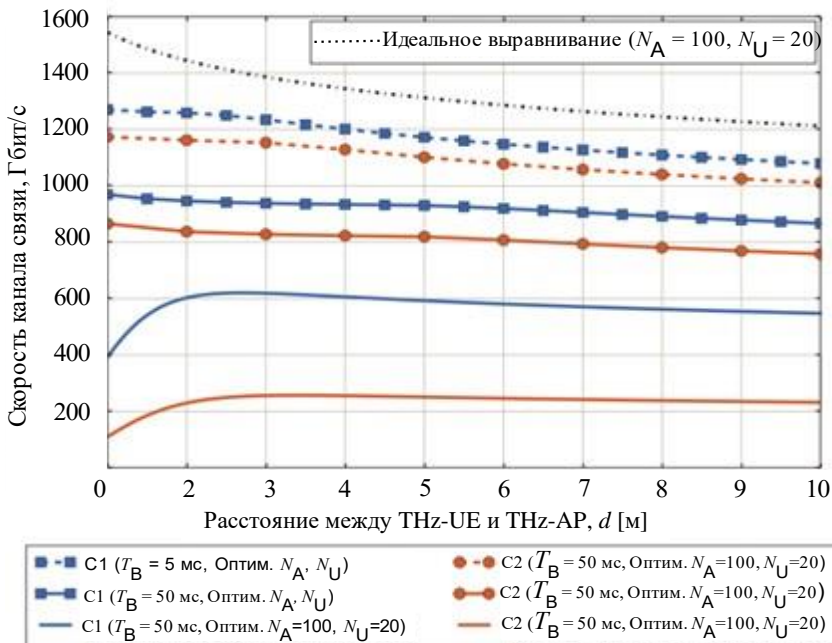


Рис. 13.5. Влияние эффектов микромобильности на характеристики системы:

a – доля времени простоя; b – скорость канала связи в зависимости от мобильности по OX и OY; $в$ – скорость канала связи в зависимости от расстояния между THz-UE и THz-AP

Влияние декартовой микромобильности

На рис. 13.5б представлена средняя пропускная способность как функция Δx при $N_A = 100$ и $N_U = 20$. Здесь для $d = 10$ м наблюдаем эффект, аналогичный эффекту на рис. 13.5а, т.е. малые значения Δx (например, $\Delta x < 0,1$ м) не оказывают существенного влияния на метрику, поскольку система в основном ограничена относительно высокой угловой микромобильностью $\Delta\phi = 6^\circ$. В то же время при $d = 1$ м эффект уже не проявляется, так как луч, генерируемый THz-AP на расстоянии 1 м, имеет ширину менее 2 см. Следовательно, даже небольшие смещения, вызванные декартовой микромобильностью, приводят к немедленному снижению пропускной способности (например, с 850 Гбит/с при $\Delta x = 0,02$ м до 50 Гбит/с при $\Delta x = 0,1$ м для схемы 1). Поэтому особенно важно учитывать декартову микромобильность на малых расстояниях между THz-BS и THz-UE.

Наконец, сравним схему 2 при $T_U = 0,2$ со схемой 1, когда T_U численно оптимизируется для каждого из значений Δx , следуя подходу на рис. 13.4а. Из данных рисунка видно, что схема 2 значительно превосходит первую во всем диапазоне значений Δx и для обоих значений расстояний между THz-BS и THz-UE. Этим сравнением мы подчеркиваем важность эффекта микромобильности в терагерцевых системах связи и оптимизации параметров системы для предполагаемой модели микромобильности, когда это возможно.

Влияние времени поиска луча

Продолжим обсуждение необходимости оптимизации системы с учетом микромобильности, следуя иллюстрации на рис. 13.5в, который демонстрирует максимально достижимую пропускную способность терагерцевого канала, когда все параметры N_A , N_U и T_U оптимизированы в соответствии с эффектами, показанными на рис. 13.4.

Из анализа рисунка следуют два основных наблюдения. Во-первых, кривые для $T_B = 50$ мс на 350 Гбит/с ниже, чем кривые для $T_B = 5$ мс. Таким образом, разработка надежной и эффективной процедуры поиска луча имеет решающее значение для мобильной терагерцевой связи. Во-вторых, видна разница около 0,6–1,1 Тбит/с между теоретической кривой и кривой для практического значения $T_B = 50$ мс (соответствует $N_A = 100$, $N_U = 20$ и $\delta = 5$ мс). Таким образом, важно учитывать микромобильность при оценке производительности мобильных терагерцевых систем, а игнорирование эффекта

приводит к существенному завышению достижимой скорости передачи данных.

Наконец, сравним пропускную способность системы для схем 1 и 2 при значениях $T_B = 50$ мс и оптимизированных N_A и N_U при $T_B = 50$ мс, $N_A = 100$ и $N_U = 20$. Как видно из рис. 13.5в, схема 1 обеспечивает большую пропускную способность примерно на 500 Гбит/с. Выигрыш особенно заметен при $d < 2$ м, где декартова микромобильность приводит к ухудшению спектральной эффективности на меньших расстояниях, несмотря на малые потери распространения. Последнее наблюдение подтверждает необходимость учета микромобильности ТНз-УЕ не только в исследованиях пропускной способности, но и при проектировании приемопередатчиков для сотовых систем связи, работающих в терагерцовом диапазоне частот.

13.2. Методы защиты от прослушивания на физическом уровне

Для снижения влияния крайне малой апертуры антенных элементов системы связи терагерцового диапазона будут использовать антенные решетки с чрезвычайно узкими диаграммами направленности, обеспечивающие заметное усиление как на передающей, так и на приемной стороне. Кроме того, терагерцовые частоты естественным образом подвержены блокировке распространения как стационарными, так и подвижными объектами, включая стены зданий, транспортные средства, мебель и даже человеческие тела. Это приводит к сложной динамической многолучевой среде распространения между точкой доступа и пользовательским устройством с одним компонентом прямой видимости и несколькими отраженными и рассеянными путями вне прямой видимости. Для обеспечения непрерывной передачи данных необходимо использовать механизм управления лучом для выбора пути, который характеризуется наилучшим соотношением ОСШ.

Использование чрезвычайно направленных узких лучей дает преимущества для безопасности физического уровня связи в терагерцовом диапазоне частот. Причина в том, что злоумышленник должен физически находиться в пределах передающего луча, чтобы перехватить и декодировать передаваемые данные. Использование узких лучей вместе с кодированием, специфичным для безопасности физического уровня, позволяет существенно снизить вероятность прослушивания данных уже в миллиметровом диапазоне частот в 5G. Тем

не менее полная безопасность передачи все равно не может быть достигнута, поскольку «зона прослушивания» – это область, где злоумышленник может успешно прослушивать данные, – все еще относительно велика.

В последнее время было предложено несколько методик для решения этой проблемы путем разделения канала связи между сетью и АУ, используя несколько точек доступа или несколько разнесенных антенных решеток на одной точке доступа. В том числе показано, что предполагаемое расстояние между несколькими антенными решетками в терагерцовом диапазоне недостаточно для значительного пространственного разнесения лучей. В свою очередь, поскольку условия канала быстро меняются, обмен данными между несколькими точками доступа значительно увеличивает сложность системы, требуя синхронизации в реальном времени между несколькими терагерцовыми точками доступа.

Для преодоления вышеупомянутых проблем рассмотрим усовершенствованный метод построения безопасной системы связи путем использования нескольких путей распространения между точкой доступа и АУ. Следует обратить внимание на многолучевой подход, при котором узел передает различные доли конфиденциальных данных по различным путям распространения, доступным в данный момент времени приемнику (рис. 13.6). Данные кодируются таким образом, что приемник может декодировать сообщение только в том случае, если он успешно получит все части передаваемых данных. Мы показываем, что, хотя такая схема немного снижает пропускную способность канала по сравнению с базовой однолучевой схемой, вероятность прослушивания сообщений резко снижается, даже если злоумышленники действуют в кооперативном режиме.

Системная модель

1. *Сценарий развертывания.* Рассматривается один канал связи между точкой доступа терагерцового диапазона, расположенной на высоте h_A (например, на фонарном столбе или на стене) и АУ, расположенного на высоте h_U и находящегося на расстоянии d метров от точки доступа. АУ окружают два типа объектов: а) люди, действующие как блокаторы путей распространения терагерцовых волн, б) злоумышленники, прослушивающие все данные, передаваемые по путям распространения, проходящим через их местоположение.

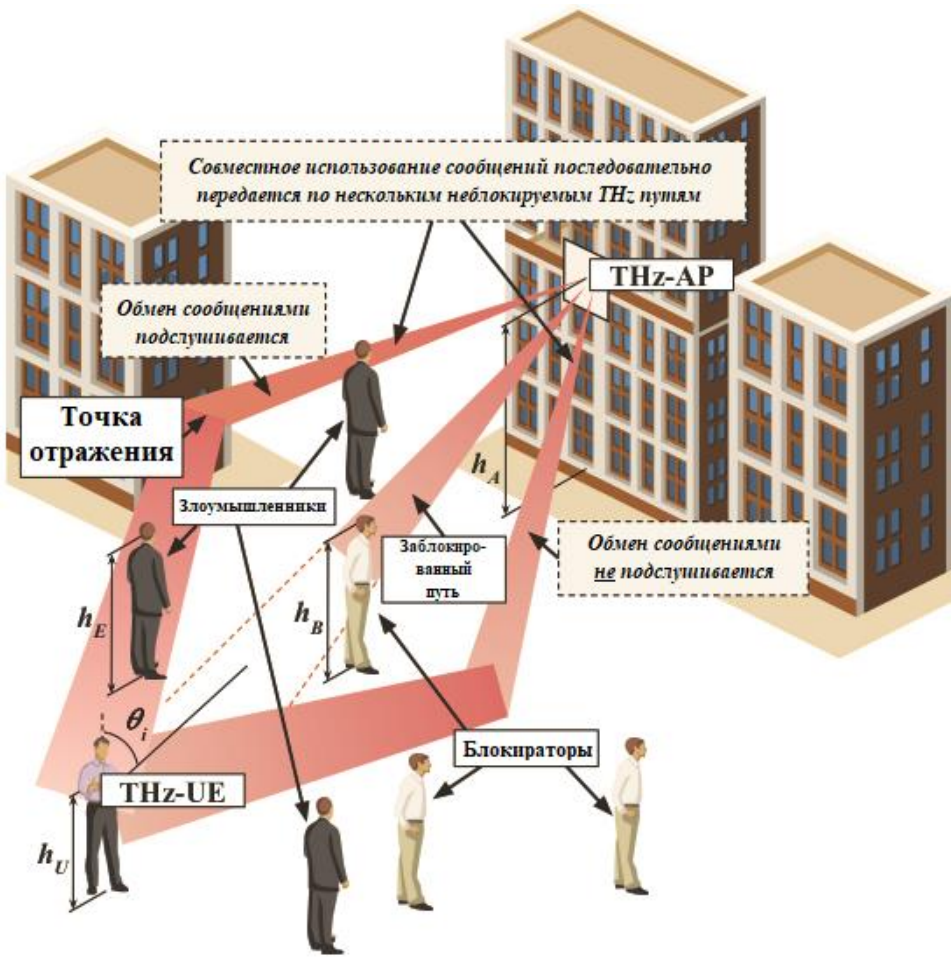


Рис. 13.6. Сценарий взаимодействия в терагерцовом диапазоне частот

Предполагается, что люди распределены в соответствии с процессом PPP (Poisson Point Process) в круге с плотностью μ единиц на квадратный метр. Человеческие тела моделируются цилиндрами с высотой h_B и радиусом основания r_B . Злоумышленники также моделируются в соответствии с PPP с плотностью λ . И блокаторы, и злоумышленники моделируются как неподвижные объекты в течение сессии передачи данных, что достаточно реалистично ввиду малой длительности передачи на таких высоких скоростях.

2. *Модель безопасности и злоумышленников.* Сравняются два решения для безопасной передачи данных в терагерцовом диапазоне. В первом решении, названном однолучевым, конфиденциальные данные передаются в виде одного сообщения по самому «лучшему» пути, т.е. пути, который в данный момент ассоциируется с

наибольшим значением ОСШ. Мы сравниваем базовый однолучевой подход с альтернативным многолучевым решением, где сообщение, содержащее конфиденциальные данные, разделяется на $M \leq N$ путей, имеющих наибольшее значение ОСШ. Используется безопасное кодирование так, чтобы сообщение могло быть декодировано только в том случае, если получены все части данных.

Роль злоумышленников заключается в нарушении безопасности передачи в терагерцовом диапазоне путем прослушивания. Предполагаются злоумышленники, которые прослушивают все сообщения, проходящие через их АУ, но не блокируют сигнал и не изменяют данные. Предполагается, что злоумышленники действуют совместно, то есть считается, что сообщение прослушивается, если оно перехвачено хотя бы одним злоумышленником. Считается, что злоумышленники успешно прослушивают сообщение, если они захватывают все части сообщения.

3. *Метрики эффективности.* Интерес представляют характеристики зависимости вероятности прослушивания p_E от пропускной способности канала связи на беспроводном интерфейсе $C(x)$ между точкой доступа и АУ, расположенными на расстоянии x в 2D-пространстве. Первая определяется как вероятность того, что поле злоумышленников способно прослушать текущую передачу. Для однолучевой схемы пропускная способность канала связи предполагается равной скорости Шеннона по пути с наибольшим ОСШ из всех доступных незаблокированных путей. Для многолучевой схемы пропускная способность канала связи определяется как скорость по Шеннону для M незаблокированных путей, используемых для обмена данными, т.е.

$$C(x) = \frac{B}{M} \sum_{i=1}^M \log_2(1 + S_i(x)), \quad (13.17)$$

где S_i , $i = 1, 2, \dots, M$ – значения ОСШ на расстоянии x на пути i . Наконец, интерес представляет безопасная скорость передачи, определяемая как скорость передачи, на которой невозможно прослушивание злоумышленниками.

Анализ системы

В этом параграфе рассматривается модель для анализа зависимости между вероятностью прослушивания и скоростью передачи данных. Сначала найдем вероятности блокировки и прослушивания

для отдельных путей, а затем проведем анализ интересующих метрик.

А. Зоны прослушивания при многолучевом распространении

Для того чтобы прослушать сообщение, злоумышленник должен физически находиться в пределах передающего луча. Поэтому площадь зоны прослушивания – двумерной зоны на уровне земли вокруг АУ, в которой должен находиться злоумышленник, чтобы оставаться в пределах передающего луча, – может быть рассчитана на основе параметров разворачивания. Рассмотрим зону прослушивания для пути i в виде прямоугольника $l_{E,i} \times w_{E,i}$, где $w_{E,i}$ – ширина (в метрах) луча точки доступа в непосредственной близости от АУ и $l_{E,i}$ определяется высотой луча точки доступа вблизи АУ и высотой расположения прослушивающего устройства (рис. 13.7).

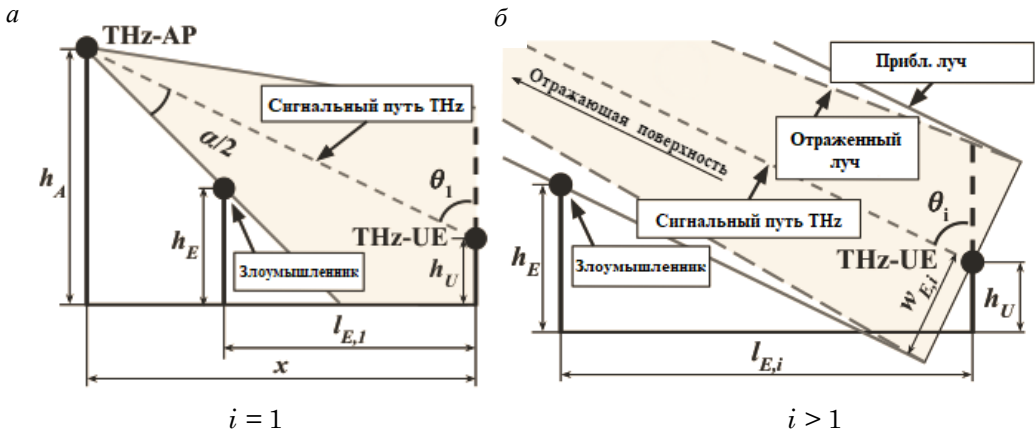


Рис. 13.7. Иллюстрация зон прослушивания (пояснения в тексте)

Для модели антенны ширина зоны прослушивания для пути LoS $w_{E,1}$ может быть рассчитана по формуле (рис. 13.7а)

$$w_{E,1} = d \tan(\alpha/2), \quad (13.18)$$

где d – расстояние между точкой доступа и АУ в 3D-пространстве и α – ширина луча точки доступа.

Длина зоны прослушивания для случая LoS $l_{E,1}$ определяется из геометрических соображений следующим образом:

$$l_{E,1} = x - (h_A - h_E) \tan(\theta_1 - \alpha/2), \quad (13.19)$$

где x – расстояние между точкой доступа и узлами АУ в 2D-пространстве, h_A – высота точки доступа, h_E – высота расположения прослушивающего устройства, и θ_1 – угол зенита поступления луча LoS.

Отметим, что кластерная модель распространения 3GPP (см. главу 2) не дает точного количества отражений. Получено приближение, предполагающее одно отражение и отсутствие рассеяния для пути, при котором отсутствует прямая видимость (nLoS, non-line-of-sight), что, вообще говоря, ограничивает размер зоны прослушивания. На практике зона может быть меньше, так как на пути nLoS может быть более одного отражения.

Следует подчеркнуть, что плотность принимаемой мощности от пути nLoS обратно пропорциональна ширине луча, проходящего по этому пути. Поэтому ширина зоны прослушивания для nLoS-пути может быть рассчитана по формуле

$$w_{E,i} = d \tan(\alpha/2) \sqrt{P_{S,1}/(P_{S,i}K)}, \quad (13.20)$$

где $P_{S,1}$ – доля мощности LoS луча, $P_{S,i}$ – доля мощности i -го пути nLoS, определяемая кластерной моделью распространения 3GPP. Здесь K – дополнительное затухание, вызванное отражением от объекта, которое для обычных материалов составляет не менее 10 дБ.

Следовательно, длина зоны прослушивания $l_{E,i}$ может рассчитываться по формуле (см. рис. 13.7б):

$$l_{E,i} = (h_E - h_U) \tan(\theta_i) + w_{E,i}/\cos(\theta_i), \quad (13.21)$$

где h_U – высота расположения АУ; θ_i – с.в. угла зенита поступления луча (ЗОА) для выбранного пути.

Для поля расположения злоумышленников (PPP с плотностью λ) мы определим вероятность прослушивания для пути i , $p_{E,i}$, как вероятность того, что хотя бы один злоумышленник находится в соответствующей зоне, т.е.

$$p_{E,i} = 1 - e^{-\lambda w_{E,i} l_{E,i}}, \quad i = 1, 2, \dots, N. \quad (13.22)$$

Б. Зоны блокировки для многолучевого распространения

Пусть $p_{B,i}$ – вероятность того, что i -й путь между точкой доступа и АУ заблокирован. Сначала рассмотрим блокировку пути LoS, $p_{B,1}$. Для заданного расстояния x между точкой доступа и АУ всегда существует так называемая зона блокировки прямой видимости, как показано на рис. 13.8. Из геометрических соображений ширина и длина зоны блокировки LoS вычисляются по формулам

$$w_{B,1} = 2r_B, l_{B,1} = \left(x \frac{h_B - h_U}{h_A - h_U} + r_B \right), \quad (13.23)$$

где r_B – радиус блокатора.

Используя результаты предыдущих глав книги, получаем следующее выражение для вероятности блокировки LoS:

$$p_{B,1} = 1 - e^{-2\mu r_B \left[x \frac{h_B - h_U}{h_A - h_U} + r_B \right]}. \quad (13.24)$$

Пусть $\theta_i(x)$ являются с.в. угла зенита поступления луча ZOA. Рассмотрим вероятность блокировки для i -го пути $p_{B,i}$ и заметим, что кластерной 3GPP-модели распространения (см. главу 2) ZOA пути i может определена распределением Лапласа с функцией плотности (см. главу 2):

$$f_{\theta_i}(y; x) = \frac{1}{2b_i(x)} e^{-\frac{|y - a_i(x)|}{b_i(x)}}, \quad y \in [0; \pi), \quad i = 2, 3, \dots, \quad (13.25)$$

где y – соответствующее значение ZOA, $a_i(x)$, $b_i(x)$, $i = 2, 3, \dots, N$, – параметры распределения Лапласа. Эксперименты показывают, что $a_i(x)$ не зависит от номера пути i и зависит только от расстояния x , т.е. $a_i(x) = a_j(x)$ для $i, j = 2, 3, \dots, N$. Более того, среднее значение ZOA для всех путей совпадает с ZOA пути LoS. Напротив, $b_i(x)$ не зависит от расстояния и зависит только от индекса пути i .

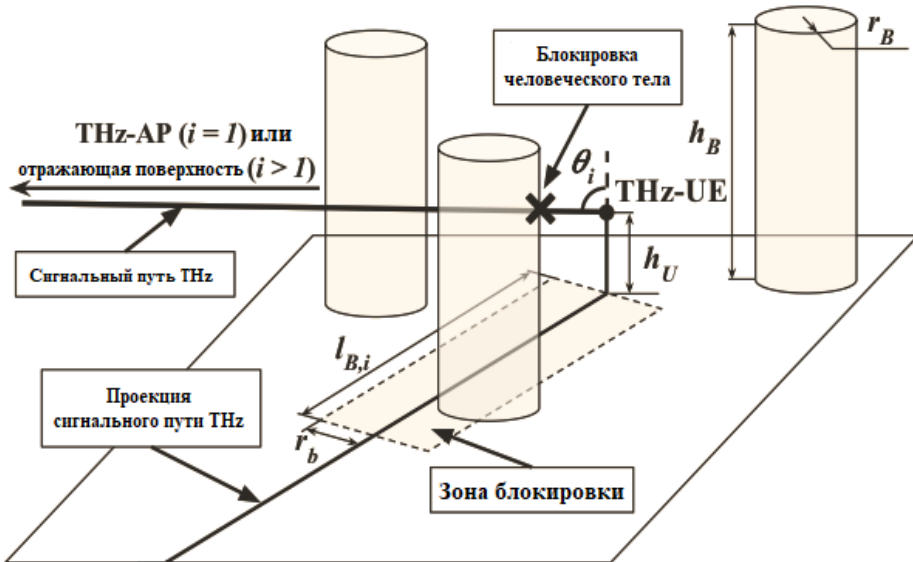


Рис. 13.8. Иллюстрация зоны блокировки распространения сигнала

Аналогично пути LoS, для любого данного значения ZOA y_i пути i , $i = 2, 3, \dots, N$, мы определяем зону блокировки пути i . Используя геометрические соотношения (см. рис. 13.8), для пути i имеем следующую вероятность блокировки:

$$p_{B,i}(y; x) = 1 - e^{-2\lambda_B r_B (\tan y_i (h_B - h_U) + r_B)}. \quad (13.26)$$

Учитывая, что задана плотность вероятности с.в. θ_i , окончательно получаем

$$p_{B,i}(x) = \int_0^\pi f_{\theta_i}(y; x) p_{B,i}(y; x) dy. \quad (13.27)$$

В. Мощность и скорость путей распространения

Определив вероятности блокировки путей распространения между точкой доступа и АУ, перейдем к анализу пропускной способности этих путей. Сначала определим величину принимаемой мощности. С учетом сделанных предположений можно показать, что доля мощности по i -му пути $P_{S,i}$ следует логнормальному закону с плотностью вероятности

$$f_{P_{S,i}}(y; x) = \frac{1}{y d_{i\sqrt{2\pi}}} e^{-\frac{(\ln y - c_i)^2}{2d_i^2}}, \quad i = 1, 2, \dots, N, \quad (13.28)$$

с параметрами c_i , d_i , которые не зависят от расстояния x и зависят только от номера пути.

Получаемая мощность на пути i может быть рассчитана по формуле

$$P_i(x) = P_{S,i} 10^{(P_T + G_A + G_U - T(x))/10}, \quad i = 1, 2, \dots, N, \quad (13.29)$$

где P_T – мощность передачи, T – потери пути, G_A и G_U – коэффициенты усиления передающей антенной решетки точки доступа и приемной антенной решетки АУ. Подставляя потери пути для терагерцевых частот, получаем

$$P_i(x) = P_{S,i} 10^{\frac{P_T + G_A + G_U - 20 \log_{10} f_c - 20 \log_{10} x - 10kx / \ln(10) - 147.55}{10}}, \quad (13.30)$$

где f_c – частота и k – коэффициент поглощения.

Таким образом, получаемая мощность на пути i имеет плотность распределения

$$f_{P_i}(y; x) = f_{P_{S,i}}(y/A(x); x), \quad (13.31)$$

где $A(x) = 10^{\frac{P_T + G_A + G_U - 20 \log_{10} f_c - 20 \log_{10} x - 10kx / \ln(10) - 147.55}{10}}$.

Найдя принятую мощность, скорость Шеннона для пути i можно записать в виде

$$C_i(x) = B \log_2(1 + P_i(x)/N_0), \quad i = 1, 2, \dots, N, \quad (13.32)$$

где N_0 – тепловой шум в полосе пропускания B .

Заметим, что $C_i(x)$ – нелинейная функция совместной плотности P_i , $i = 1, 2, \dots, N$. Напомним, что плотность вероятности св Y , $w(y)$, выраженная как функция $y = \phi(x)$ другой св X с плотностью вероятности $f(x)$ есть

$$w(y) = \sum_{\nu j} f(\psi_j(y)) |\psi_j'(y)|, \quad (13.33)$$

где $x = \psi_j(y) = \phi^{-1}(x)$ – это обратные функции.

Обратная ветвь функции скорости и ее производная определяются в виде:

$$\psi(y) = N_0(2^{y/B} - 1), \quad \psi'(y) = N_0 \log(2) 2^{y/B} / B, \quad (13.34)$$

что приводит к плотности вероятности в следующем виде:

$$f_{C_i}(y; x) = \frac{A(x) \log(2) 2^{\frac{y}{B}}}{B(2^{\frac{y}{B}} - 1) d_{i\sqrt{2\pi}}} e^{-\frac{\left(\ln \frac{N_0(2^{y/B} - 1)}{A(x)} - c_i\right)^2}{2d_i^2}}, \quad i = 1, 2, \dots \quad (13.35)$$

Г. Пропускная способность и вероятность прослушивания

Перейдем к нахождению пропускной способности канала связи и вероятности прослушивания для однолучевой и многолучевой схем передачи.

1. *Пропускная способность по Шеннону.* Напомним, что в однолучевой схеме передачи точка доступа и АУ всегда работают, используя путь с наибольшей принимаемой мощностью. Сортируя пути в порядке убывания мощности, получаем следующую формулу для пропускной способности по Шенону для однолучевой схемы:

$$C(x) = \sum_{i=1}^M [\prod_{j=1}^{i-1} p_{B,j}] (1 - p_{B,i}) C_i(x). \quad (13.36)$$

Обратим внимание, что в (13.36) скорость определяется как взвешенная сумма слагаемых. Теперь можно найти среднее значение пропускной способности

$$E[C(x)] = \sum_{i=1}^M [\prod_{j=1}^{i-1} p_{B,j}] (1 - p_{B,i}) \int_0^\infty f_{C_i}(y; x) y dy. \quad (13.37)$$

Окончательно, пропускную способность многолучевой схемы для случая $M = 2$ можно найти по формуле

$$C(x) = (1 - p_{B,1})(1 - p_{B,2})\frac{1}{2}[C_1(x) + C_2(x)] + \\ + (1 - p_{B,1})p_{B,2}C_1(x) + p_{B,1}(1 - p_{B,2})C_2(x). \quad (13.38)$$

Аналогично можно получить формулы и для случаев $M > 3$. Однако при больших значениях M (например, $M = 20$) вычислительная сложность по точным формулам возрастает. В этом случае, напомним, что при $M \leq N$ среднее значение угла поступления луча ZOA на пути i совпадает с LoS ZOA, можно предположить, что $p_{B,i} \approx p_{B,j} = p_B$, для $i, j = 1, 2, \dots, N$. Тогда вероятность того, что i из M путей не заблокированы можно задать биномиальным распределением

$$q_{M,i} = \binom{M}{i} i (1 - p_B)^i p_B^{M-i}, \quad i = 1, 2, \dots, M. \quad (13.39)$$

Используя это приближение, для $M = 2$ получаем, что

$$C(x) = q_{2,1} \frac{C_1(x)}{2} + q_{2,1} \frac{C_2(x)}{2} + q_{2,2} \left(\frac{C_1(x)}{2} + \frac{C_2(x)}{2} \right), \quad (13.40)$$

и, для $M = 3$,

$$C(x) = \frac{1}{3} q_{3,1} \sum_{i=1}^3 C_i(x) + \frac{1}{3} q_{3,2} \left[\frac{1}{2} (C_1(x) + C_2(x)) + \frac{1}{2} (C_2(x) + \right. \\ \left. + C_3(x)) + \frac{1}{2} (C_1(x) + C_3(x)) \right] + \frac{1}{3} q_{3,3} \sum_{i=1}^3 C_i(x). \quad (13.41)$$

Теперь рассмотрим вклад пути i в общую пропускную способность. Заметим, что, когда j путей не заблокированы и LoS-путь является одним из них, доля времени, используемая для передачи, составляет $1/j$. Общее число комбинаций выбора пути j из M путей равно $\binom{M}{j}$, а число раз, когда путь i появляется в этих комбинациях, равно $\binom{M-1}{j-1}$. Суммируя все возможные значения числа незаблокированных путей, взвешенные с соответствующими биномиальными вероятностями, получаем следующее выражение для пропускной способности многолучевой схемы:

$$C(x) = \sum_{i=1}^M \sum_{j=1}^M q_{M,j} \frac{\binom{M-1}{j-1}}{\binom{M}{j}} \frac{1}{j} C_i(x) = \frac{1-p_B^M}{M} \sum_{i=1}^M C_i(x). \quad (13.42)$$

Теперь среднее значение скорости на расстоянии x можно найти по формуле

$$E[C(x)] = \frac{1-p_B^M}{M} \sum_{i=1}^M \int_0^\infty f_{C_i}(y; x) y dy. \quad (13.43)$$

2. *Вероятность прослушивания.* Для оценки вероятности прослушивания необходимо учесть, что зона прослушивания на пути LoS отличается от остальных путей. Кроме того, напомним, что интегрес представляет вероятность прослушивания при условии, что хотя бы один путь не заблокирован на АУ. Очевидно, эта вероятность вычисляется по формуле $1 - \prod_{i=1}^M p_{B,i}$.

Заметим, что вероятность того, что путь i используется, равна $(1 - p_{B,i}) \prod_{j=1}^{i-1} p_{B,j}$, и тогда вероятность того, что путь i используется и прослушивается, равна $(1 - p_{B,i}) p_{E,i} \prod_{j=1}^{i-1} p_{B,j}$. Суммируя по незаблокированным путям, получаем формулу для вероятности прослушивания

$$p_E = \frac{\sum_{i=1}^M \left((1-p_{B,i}) p_{E,i} \prod_{j=1}^{i-1} p_{B,j} \right)}{1 - \prod_{i=1}^M p_{B,i}}, \quad (13.44)$$

где $p_{E,i}$ – вероятность прослушивания на пути i .

Аналогично вероятность прослушивания для многолучевой схемы можно найти, учитывая вероятности блокировки различных путей $p_{B,i}$, $i = 1, 2, \dots, M$. Пусть $v_{M,i}$ – вероятность того, что i путей не заблокированы, и пусть $v_{M,0}$ – вероятность того, что не существует незаблокированных путей. Тогда

$$\begin{aligned} v_{1,0} &= p_{B,1}, \quad v_{1,1} = (1 - p_{B,1}) p_{E,1}, \\ v_{2,0} &= p_{B,1} p_{B,2}, \quad v_{2,1} = v_{1,1} p_{B,2} + v_{1,0} (1 - p_{B,2}) p_{E,2}, \end{aligned} \quad (13.45)$$

...

и тогда справедлива следующая рекуррентная формула:

$$v_{M,i} = v_{M-1,i-1} (1 - p_{B,M}) p_{E,M} + v_{M-1,i} p_{B,M}. \quad (13.46)$$

Теперь вероятность прослушивания для многолучевой схемы можно найти по формуле

$$p_E = \frac{\sum_{i=1}^M v_{M,i}}{1 - \prod_{i=1}^M p_{B,i}}, \quad (13.47)$$

а «безопасная скорость» (скорость передачи информации при условии отсутствия прослушивания) принимает вид $C_S(x) = (1 - p_E) C(x)$.

Предполагая, что $p_{B,i} \approx p_{B,j} = p_B$ для $i, j = 1, 2, \dots, N$, и различая вероятности прослушивания пути LoS и других путей, т.е. $p_{E,1} = p_{E,L}$, $p_{E,i} = p_{E,nL}$ можно получить упрощенную приближенную формулу

для вероятности прослушивания. Для любого числа i незаблокированных путей вероятность того, что путь LoS заблокирован, равна $\frac{M-1}{M} \frac{M-2}{M-1} \times \dots \times \frac{M-i}{M-i+1} = \frac{M-i}{M}$, а вероятность того, что путь LoS не заблокирован, равна i/M . Окончательно имеем, что

$$p_E \approx \frac{\sum_{i=1}^M q_{M,i} \left(\frac{M-i}{M} p_{E,nL}^i + \frac{i}{M} p_{E,L} p_{E,nL}^{i-1} \right)}{1 - \prod_{i=1}^M p_{B,i}}. \quad (13.48)$$

Численный анализ

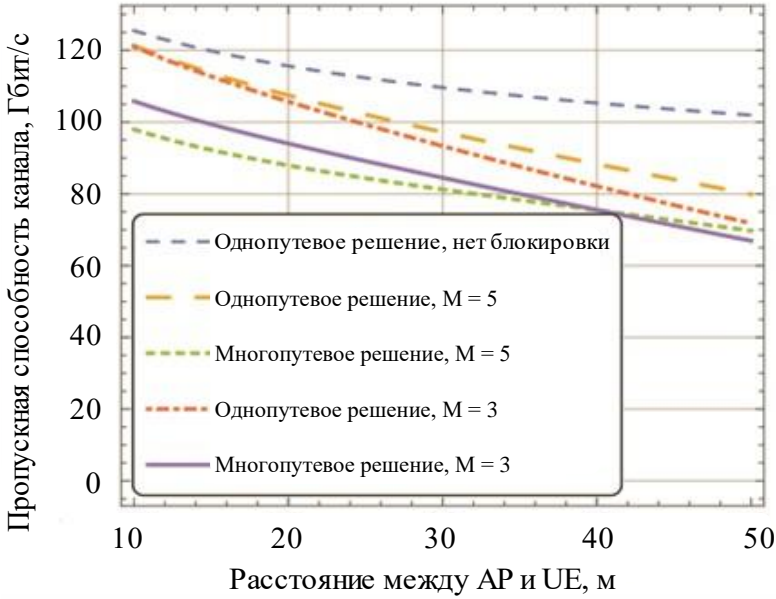
Далее проанализируем зависимость пропускной способности от вероятности прослушивания для однолучевой и многолучевой схем.

1. *Пропускная способность канала связи.* На рис. 13.9а показана пропускная способность канала в зависимости от расстояния между точкой доступа и АУ для обеих схем и $M = 3, 5$. Также показана верхняя граница пропускной способности для базовой схемы, предполагающая использование одного пути распространения и отсутствие блокаторов, т. е. для случая $\mu = 0$. Ожидаемо, что в присутствии блокаторов пропускная способность обеих схем должна снижаться. Более того, как видно из графиков, разница между однолучевой и многолучевой схемами заметна только для расстояний 40–50 м между точкой доступа и АУ. При $d > 50$ м путь LoS характеризуется большой вероятностью блокировки, поэтому влияние этого пути становится менее значимым.

Многолучевое распространение, присущее терагерцовым системам связи, играет важную роль в пропускной способности канала связи. Как видно из графиков рис. 13.9а, возможность использовать больше путей приводит к лучшей пропускной способности по сравнению с однолучевой схемой для всех рассматриваемых значений d . Более того, разница увеличивается с расстоянием, так как вероятность использования пути LoS уменьшается из-за возможной блокировки. Это поведение также справедливо для многолучевой схемы и для больших расстояний между точкой доступа и АУ, начиная с $d \approx 40$ м. Для меньших значений d наблюдается обратный эффект. Основной причиной является зависимость размеров зон блокировки путей от расстояния. Когда d мало, вероятность того, что все пути не заблокированы, высока, как для $M = 3$, так и для $M = 5$. Это означает, что вклад пути LoS в общую пропускную способность выше для

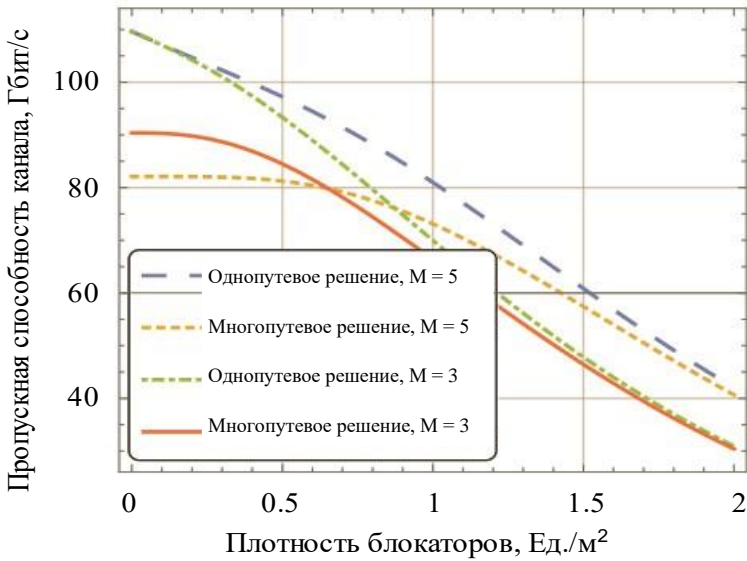
$M = 3$. С увеличением d вероятность блокировки пути LoS возрастает, тем самым снижая роль этого пути для пропускной способности канала связи.

а



$\mu = 0$

б



$d = 30$ м

Рис. 13.9. Пропускная способность канала для однолучевой и многолучевой схем (пояснения в тексте)

Рассмотрим теперь влияние плотности блокаторов μ на пропускную способность (рис. 13.9б). При малых значениях μ пропускная способность совпадает, так как путь LoS используется практически без перерывов. При больших значениях μ пропускная способность начинает резко снижаться, так как вероятность одновременной блокировки трех путей намного меньше соответствующей вероятности для $M = 5$. Кроме того, разница между пропускными способностями каналов для рассматриваемых схем уменьшается с ростом плотности блокаторов и значения практически совпадают для $\mu = 1,5 \text{ 1/м}^2$. Аналогично на рис. 13.9а наблюдаем, что для малых значений μ многолучевая схема с $M = 3$ превосходит схему с $M = 5$.

2. Вероятность прослушивания. Перейдем к анализу вероятности прослушивания для диапазона параметров системы, иллюстрированных на рис. 13.10. На рис. 13.10а показана вероятность прослушивания для обеих схем в зависимости от плотности блокаторов μ . Как можно заметить, для широкого диапазона значений μ , многолучевая схема обеспечивает существенный прирост производительности при небольшом снижении пропускной способности, как показано на рис. 13.10а для значений $M = 3$ и $M = 5$. Для $\mu = 1,5$ и $M = 5$ потеря в пропускной способности составляет лишь 2–3 Гбит/с, в то время как разница в вероятности прослушивания составляет более 600 %, уменьшаясь со значения около 0,06 для однолучевой схемы до менее чем 0,01 для многолучевой схемы. Можно также заметить, что большее количество доступных путей M приводит к меньшей вероятности прослушивания для обеих схем.

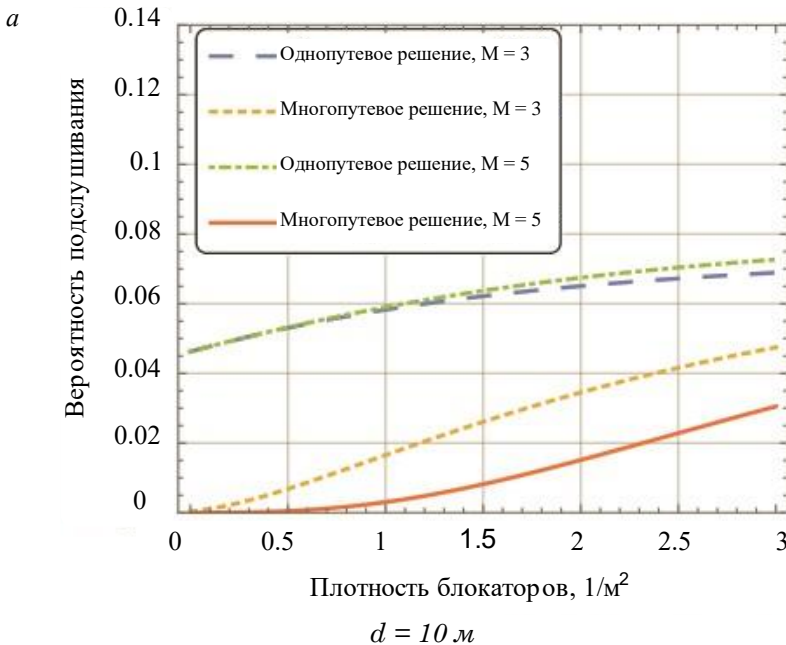
Изучим теперь поведение обеих схем в зависимости от плотности прослушивающих устройств λ , как показано на рис. 13.10б для двух значений числа антенных элементов на точке доступа, $L_A = 512$ и $L_A = 2048$. Одной из важнейших тенденций здесь является то, что увеличение L_A приводит к улучшению вероятности прослушивания как для однолучевой, так и многолучевой схемы. Основная причина в том, что система с большим значением L_A приводит к меньшему размеру зоны прослушивания и, таким образом, к меньшим значениям индивидуальных вероятностей прослушивания по соответствующим путям. Для $L_A = 2048$ выигрыш наблюдается во всем диапазоне плотности прослушивающих устройств, а для $L_A = 512$ обе кривые сходятся уже при $\lambda \approx 0,03 \text{ ед./м}^2$.

Вероятность прослушивания для однолучевой и многолучевой схем показана на рис. 13.10в как функция расстояния между точкой

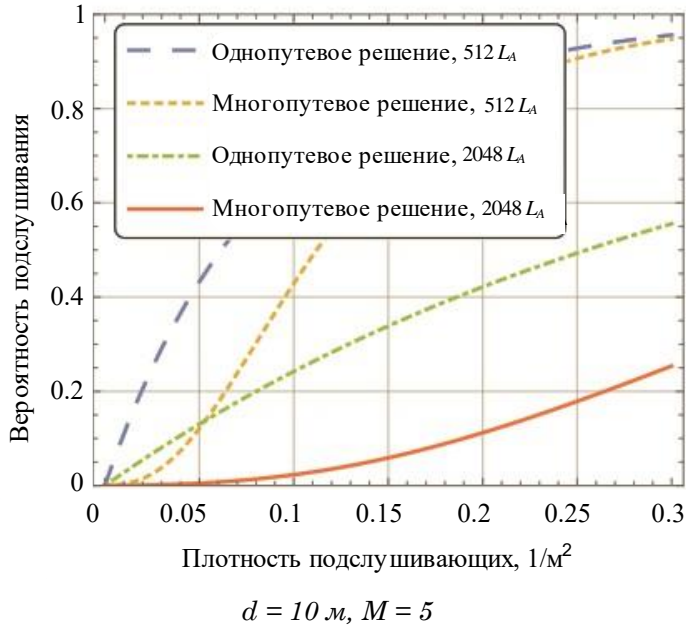
доступа и АУ, для двух значений L_A . Для $L_A = 2048$ заметный выигрыш наблюдается во всем диапазоне d , а для $L_A = 512$ кривые сходятся при $d \approx 50$ м.

3. «Безопасная скорость». Наконец, исследуем безопасную скорость C_S , как функцию плотности блокаторов. Данная метрика объединяет две предыдущие (C и p_E), что позволяет сравнивать две схемы в рамках одного графика. Многолучевая схема достигает значительно большей степени безопасности в широком диапазоне параметров системы. Выигрыш варьируется от примерно 4 Гбит/с для низкой плотности прослушивающих $\lambda = 0,03$ до более чем 10 Гбит/с для $\lambda = 0,1$. Аналогичные выводы наблюдаются и для безопасной скорости как функции других параметров системы, таких как расстояние и количество антенных элементов, L_A .

В итоге можно сделать вывод, что многолучевая схема достигает значительного выигрыша по сравнению с однолучевой схемой как в вероятности прослушивания, так и в безопасной скорости за счет незначительного снижения пропускной способности канала.



б



в

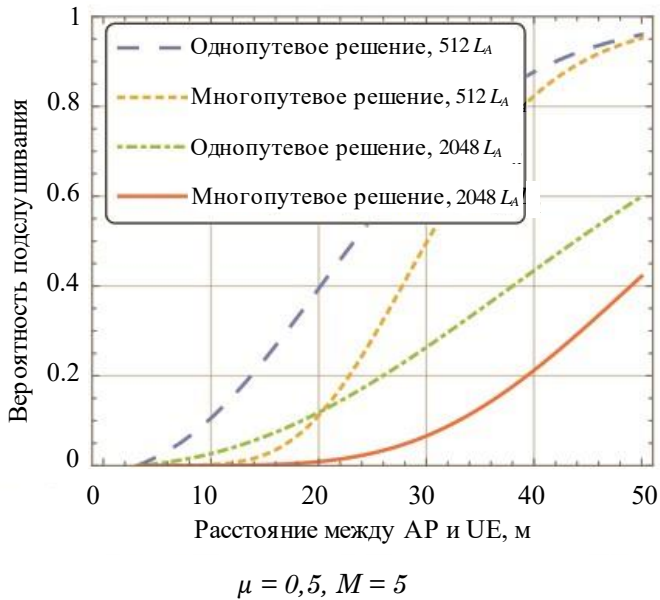


Рис. 13.10. Иллюстрация зон прослушивания (пояснения в тексте)

13.3. Доступ на последнем метре

Основываясь на последних достижениях в разработке терагерцовых приемопередатчиков, рассмотрим терагерцовую систему связи

с высокой скоростью передачи данных на «последнем метре» внутри помещений, использующую существующую инфраструктуру Ethernet для подключения к Интернету и разгрузки беспроводных сотовых и Wi-Fi-сетей связи. Эта концепция хорошо интегрируется в сетевую инфраструктуру и представляет собой эффективное использование текущих систем доступа и их интеграцию в будущие беспроводные системы связи.

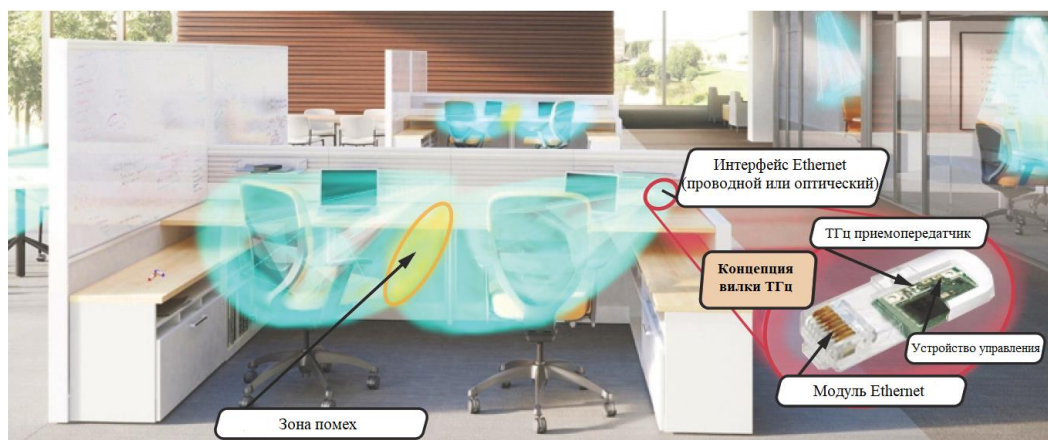


Рис. 13.11. Иллюстрация терагерцового доступа на последнем метре

Рассмотрим рис. 13.11, иллюстрирующий применение терагерцовой точки доступа, которая использует розетку Ethernet, подключенную к стандартной инфраструктуре Ethernet для отвода трафика, и в определенных условиях среды передачи может получать энергию по технологии Power-over-Ethernet (PoE). Если разместить такую точку доступа рядом с офисным столом, пользователь сможет надежно подключить ноутбук или планшет к мультигигабитному беспроводному соединению в любом месте на столе или в нескольких метрах вокруг него.

Инфраструктура Ethernet уже развернута в офисах и других помещениях внутри зданий, что позволяет эффективно и экономично перегружать данные в/из терагерцового интерфейса с высокой скоростью. В настоящее время инфраструктура Ethernet используется крайне недостаточно из-за популярности технологий Wi-Fi. Предложенная система поможет пользователям избежать зависимости объединит преимущества проводных и беспроводных сетей, а также в полной мере воспользоваться высокими скоростями, предлагаемыми терагерцовым частотным диапазоном. В сочетании с терагерцовым

широкополосным доступом Ethernet-технология обеспечивает действительно широкополосный беспроводной доступ внутри помещений.

Внутри помещений наблюдается сложная среда распространения радиоволн, в связи с чем архитектурные решения значительно различаются в разных зданиях и даже разных комнатах. Потенциально большое количество мебели в сочетании со стенами и движущимися объектами, такими как двери и даже люди, вызывают значительное ослабление сигнала связи. Традиционные системы связи внутри помещений, такие как Wi-Fi, справляются с этими проблемами благодаря достаточным проникающим свойствам радиочастот микроволнового диапазона. При переходе в терагерцовый диапазон объекты становятся непрозрачными, и даже небольшие объекты, такие как кружка на столе, могут препятствовать связи в пределах прямой видимости. Чтобы оценить потенциал рассматриваемой архитектуры, необходимо понять основные преимущества и ограничения распространения терагерцового сигнала внутри помещений.

В связи со сложностью среды распространения в помещении для оценки показателей эффективности системы мы применяем двухэтапную методику, включающую натурные эксперименты и имитационное моделирование.

- *Измерения распространения внутри помещений.* На первом этапе измеряется распространение терагерцовых волн (0,3–3 ТГц) в типичном офисном помещении, фокусируясь на отражении и рассеянии волн от таких материалов, как бетон, пластик, твердая древесина и др.

- *Оценка методом трассировки лучей.* На втором этапе используются данные измерений для параметризации имитатора трассировки лучей. В этом инструменте моделируются типичные офисные сценарии, а затем определяется производительность системы беспроводного доступа в терагерцовом диапазоне частот, используя в качестве показателей пропускную способность и ОСШ.

BER и пропускная способность

Для проведения измерений использовалось устройство TeraView Mini Pulse2, способное передавать/принимать сигнал в терагерцовом диапазоне частот от 60 ГГц до 4 ТГц с временным сдвигом 8,3 фемто-секунда и частотным разрешением 5,9 ГГц. Основное внимание

уделялось свойствам рассеяния типичных материалов в офисных помещениях, включая стекло, пластик, древесину, бетон и алюминий. Для получения данных были разработаны два испытательных стенда. Основной – использовался для измерения количества полученной энергии, отраженной/рассеянной от образца с учетом угла падения и угла отражения/рассеяния. Второй испытательный стенд представляет собой передачу в условиях прямой видимости того же сигнала на идентичное расстояние, но без каких-либо препятствий на пути. Измеренные значения энергии, полученные на первом стенде, согласуются путем нормализации с данными, полученными на втором, чтобы устранить эффекты распространения. В результате получены первичные данные об отражающих и рассеивающих свойствах материалов, показанные на рис. 13.12.

Как видно из данных рис. 13.12, чем более гладкий материал, тем выше пиковый отклик в районе пути отражения. Шероховатость поверхности уменьшает энергию отраженного пути, распределяя энергию по полю диффузного рассеяния. Результаты измерений показывают, что отраженный путь имеет наибольшую энергию, но также подтверждают, что шероховатые поверхности, такие как бетон, имеют довольно плоский отклик по всему фронту рассеивания. Видно, что алюминий является лучшим отражателем среди рассматриваемых материалов. Однако стекло, пластик и твердый картон ненамного хуже. Одна из причин заключается в том, что терагерцовые сигналы не проникают через алюминий, в то время как, например, пластик пропускает часть терагерцового сигнала. Вместе с тем бетон имеет существенно иные свойства, чем другие материалы. Из рис. 13.12 видно, что только на относительно низких частотах бетон имеет значительную отраженную составляющую.

Таким образом, показано, что моделирование распространения в терагерцовом диапазоне частот представляет собой сложную задачу, для решения которой необходимо учитывать как сильно отражающие, так и сильно рассеивающие материалы. Отметим, что даже на очень высоких частотах, таких как 3 ТГц, все еще существует сильная отраженная компонента от типичных офисных материалов. Это означает, что связь nLoS через первое отражение может быть осуществима даже на терагерцовых частотах.

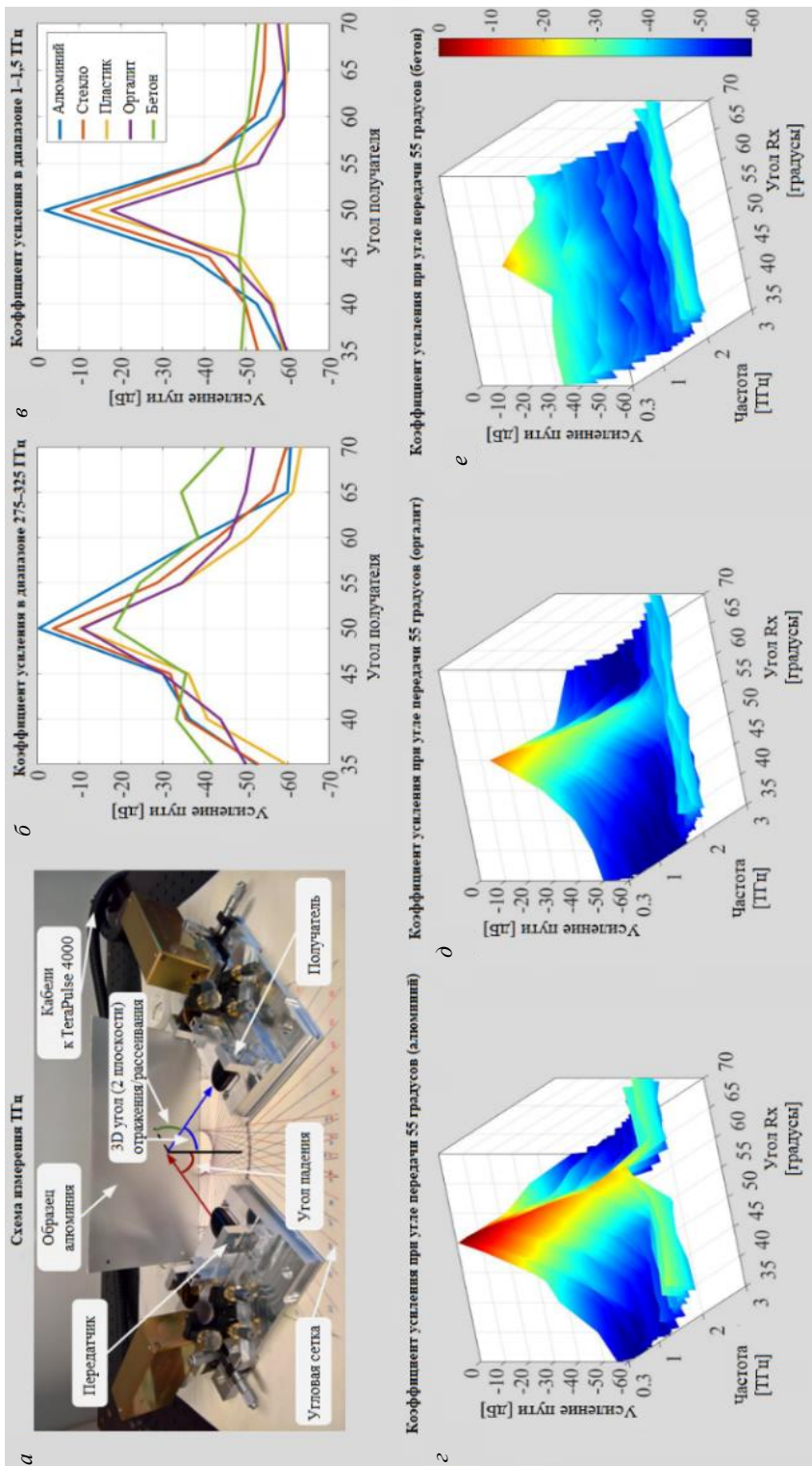


Рис. 13.12. Свойства отражения различных материалов:

а – схема измерения ПЧ; *б* – коэффициент усиления в диапазоне 275–325 ГГц;

в – коэффициент усиления в диапазоне 1–1,5 ГГц;

г-е – коэффициенты усиления при угле 55 градусов при разных материалах (алюминий, оргалит, бетон)

Моделирование канала на основе трассировки лучей

Для оценки работы терагерцовых точек доступа внутри помещений был разработан и реализован метод моделирования трассировки лучей, основанный на разбиении поверхности на миниатюрные сегменты. Размер каждого сегмента сопоставим с длиной волны, поэтому сегменты можно рассматривать как точечные приемопередатчики, принимающие часть энергии от передатчика и отражающие часть ее на приемник или другой точечный приемопередатчик на другой поверхности.

Проанализируем типичную офисную комнату размером $6 \times 4 \times 3$ м со стенами, окном, столами, потолком и полом из различных материалов, как показано на рис. 13.13 для двух сценариев. Первый сценарий «IEEE» учитывает текущее состояние терагерцовой электроники и в целом следует стандарту IEEE 802.15.3d, рассмотренному в предыдущей главе. Во втором сценарии делается прогноз производительности терагерцовой связи внутри помещений с учетом потенциального прогресса в терагерцовых устройствах. Интерес представляет связь как при наличии LoS, так и при отсутствии (nLoS) прямой видимости. Сценарий LoS является идеальным случаем, когда всегда выбирается путь с наименьшими потерями.

Рассматриваются случаи подключения как ноутбука, так и мобильного устройства. В первом случае устройство располагается на деревянном столе на 10 см ниже и 50 см дальше от терагерцовой точки доступа. В случае мобильного подключения устройство может быть расположено в любой точке комнаты на высоте 1 м. Метрики, представляющие интерес, — пропускная способность и ОСШ покрытия комнаты, а также профиль задержки мощности (PDP, Power Delay Profile) поверхностей, вносящих вклад в принимаемый сигнал в фиксированной точке комнаты.

Профиль PDP для задействованных поверхностей в случае подключения ноутбука на частоте 0,3 ТГц показан на рис. 13.13. Как видно, LoS-компонента доминирует, а отражения от ближайших двух поверхностей (деревянного стола и окна) примерно на 20 дБ слабее. Отраженные лучи от бетонных стен сильно рассеиваются и на 75–120 дБ ниже, чем LoS-компонент, что делает их влияние незначительным. Заметим, что уровень рассеяния существенно влияет на структуру PDP от выбранной поверхности. Агрегированный PDP

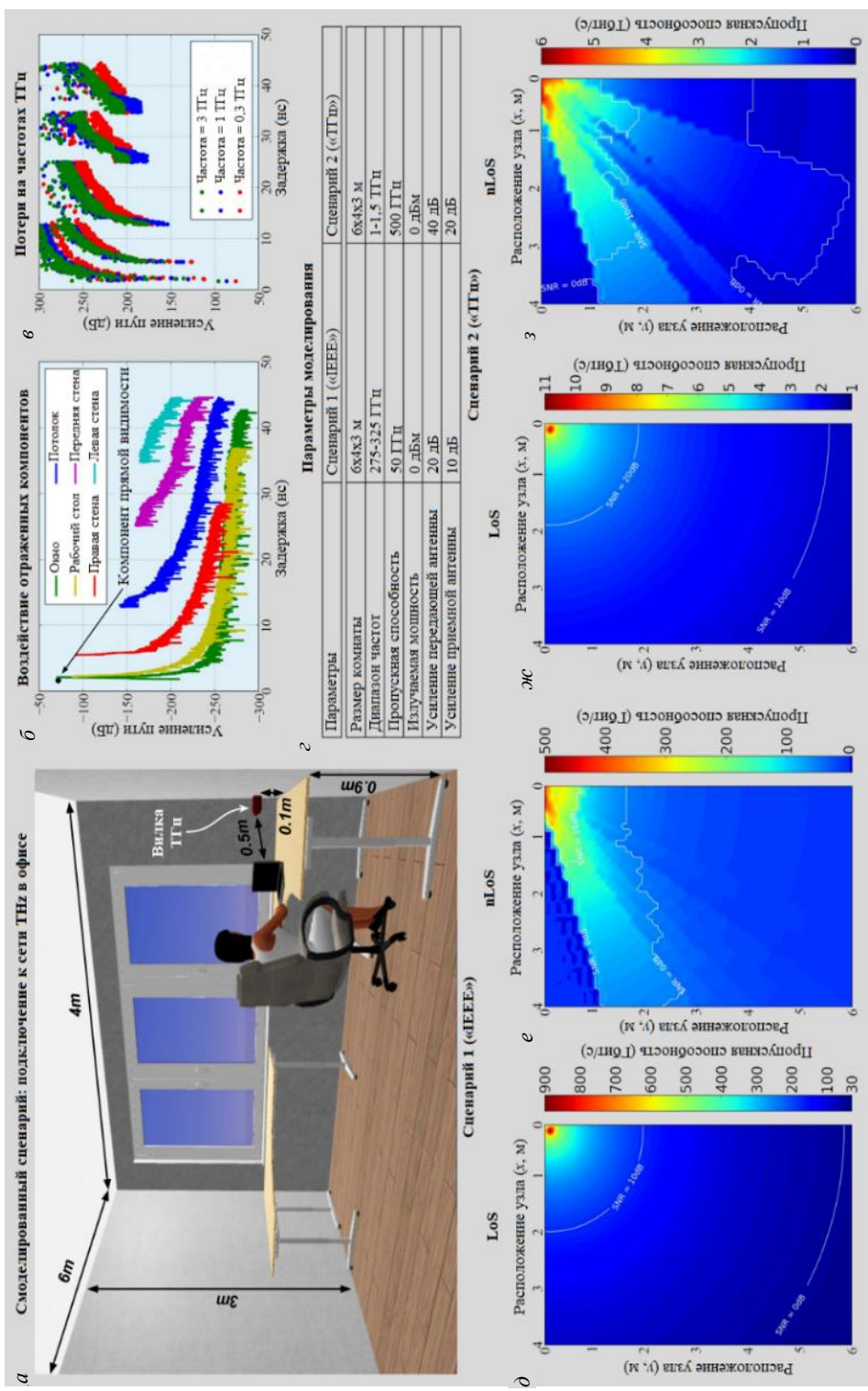


Рис. 13.13. Свойства распространения терагерцовых волн в комнате:

а – сценарий подключения к сети THz в офисе; **б** – воздействие отраженных компонентов; **в** – потери на частотах ТГц;

г – параметры моделирования; **д-з** – пропускная способность сети в зависимости от расположения узла

показан на рис. 13.13. Важно отметить, что качественно схожая картина наблюдается для всех выбранных частот, независимо от того, принадлежат они окну прозрачности (0,3 ТГц) или нет (1 и 3 ТГц). Как и ожидалось, LoS доминирует в общей структуре PDP. Кроме того, в принятом сигнале четко выделяются пять пиков (первые отражения от поверхностей), каждый со своим хвостом рассеяния. Эти компоненты могут позволить обеспечить связь в состоянии nLoS.

На рис. 13.13 показана средняя пропускная способность для сценария «IEEE» между терагерцовой точкой доступа и мобильным узлом, расположенным внутри комнаты, для случаев LoS и nLoS. Для случая nLoS мы предполагаем наличие препятствия рядом с мобильным узлом, блокирующего LoS (например, голова, рука и т.д.). Как можно заметить, даже при отсутствии LoS можно обеспечить надежную связь, используя мощность наилучшего прямого отражения: минимальный ОСШ составляет около 0 дБ в пределах 1–3 м от терагерцовой точки доступа (см. рис. 13.13). Этот результат предполагает, что отраженный компонент от окна не блокируется. Отметим, что данный подход требует точного отслеживания местоположения узла и возможности динамической настройки диаграмм направленности передающей и приемной антенн на терагерцовой точке доступа, что является важной исследовательской задачей, которая будет рассмотрена далее. В то же время поддерживается достаточно высокая скорость передачи данных, поскольку теоретическая пропускная способность канала по Шеннону превышает 100 Гбит/с. Если пессимистически предположить 90% потери в эффективности вследствие использования неоптимальных механизмов физического и канального уровней (схемы модуляции и кодирования, протокол множественного доступа к среде и т.д.), то это приведет к ожидаемой пропускной способности около 10 Гбит/с.

Наконец, мы рассмотрим результаты оценки терагерцового сценария, предполагающего наличие эффективных направленных антенн как на терагерцовой точке доступа, так и на мобильном узле. Как видно из данных рис. 13.13, уровень ОСШ значительно выше как для LoS, так и для nLoS даже при десятикратном расширении полосы пропускания и более высоких потерях распространения. Хотя концепция использования сверхузких диаграмм направленности сегодня может показаться идеалистичной, она иллюстрирует потенциал в этой области для достижения теоретической пропускной способности канала связи.

Дорожная карта внедрения

Предложенная терагерцовая система связи на последнем метре допускает постепенное развитие и может быть реализована в четыре этапа с возрастающими сложностью и удобством использования.

– *Фаза I. Доступ для стационарных устройств.* Разработка системы беспроводного удлинителя для обслуживания соединений точка–многоточие со стационарными узлами.

– *Фаза II. Доступ для мобильных устройств.* В систему включены мобильность внутрикомнатных узлов и несколько терагерцовых точек доступа. Основное внимание уделяется обеспечению управления терагерцовым лучом на беспроводном участке доступа и соответствующим механизмам для поддержки хэндовера.

– *Фаза III. Устранение помех и множественный доступ.* Предусмотрена поддержка нескольких передатчиков и приемников в одном канале связи. В том числе решаются проблемы, создаваемые помехами, возникающими при одновременной передаче.

– *Фаза IV. Интеграция.* Беспроводной доступ в терагерцовом диапазоне интегрируется как одна из технологий радиодоступа в концепцию гетерогенной сети HetNet для систем сотовой связи 6G.

Фаза I. Доступ для стационарных устройств

На этом этапе рассматриваются вопросы направленных антенн, слежения за лучом, модуляции, кодирования и работы в режиме plug-and-play. Направленные антенны (реализованные, например, с помощью фазированных антенных решеток) необходимы для увеличения эффективной апертуры антенны приемника. Передатчик также должен использовать направленные антенны, чтобы компенсировать ограниченную мощность передатчика. На миллиметровых волнах уже реализованы миниатюрные фазированные решетки, но достигнутый прогресс в терагерцовом диапазоне все еще ограничен.

Возможно, уже на данном этапе необходимо использовать несколько антенных решеток на терагерцовой точке доступа и/или мобильные терагерцовые модули, чтобы уменьшить вероятность блокировки и увеличить вероятность нахождения хорошего пути распространения. При использовании направленных антенн поиск направления луча становится критически важным. Для поддержки мобильности пользователей требуется быстрое и точное отслеживание луча, особенно в условиях nLoS. Хотя механически управляемые антенны для 0,3 ТГц уже разработаны, для практического использования требуется более быстрое электронное управление лучом.

Для обеспечения функциональности plug-and-play терагерцовая точка доступа должна восприниматься сетью как обычный коммутатор или мост. Таким образом, предполагается, что устройство не будет иметь IP-адреса, а вся пересылка данных будет осуществляться на канальном уровне. В частных случаях для работы в качестве радиоудлинителя интерес может представлять возможность сохранения структуры кадра Ethernet для беспрепятственного распространения услуг на беспроводной интерфейс.

Фаза II. Доступ для мобильных устройств

На втором этапе необходимо учесть макромобильность пользователей и обеспечить поддержку функций хэндовера. Следует внедрить SDN-контроллер, собирающий данные и управляющий всеми терагерцовыми точками доступа в выделенной области. Роль этого устройства заключается в том, постоянно отслеживать местоположение пользователя и предварительно выбирать новую терагерцовую точку доступа так, чтобы бесшовно передать вызов, когда пользователь будет близок к выходу из текущей зоны обслуживания.

Хэндовер в терагерцовых сетях доступа может быть реализован путем оценки траектории пользователя по конфигурации луча и уровню принимаемой мощности. Поскольку ожидается, что большую часть времени связь будет осуществляться в условиях прямой видимости, а терагерцовая точка доступа стационарна с известным местоположением, вычисления вполне осуществимы. Когда пользователь достигает последней доступной конфигурации антенны, процедура передачи должна быть выполнена путем направления приемного луча от следующей терагерцовой точки доступа.

Фаза III. Устранение помех и множественный доступ

Хотя высокая направленность лучей может привести к ограничению помех при связи вне помещений, интерференция все еще будет играть существенную роль внутри помещений. Например, интерференция может возникнуть, когда два мобильных приемника находятся в непосредственной близости друг от друга.

Накопленный опыт стандартизации систем доступа миллиметрового диапазона частот, IEEE 802.11ad, и 5G NR, показывает, что для систем с направленными антеннами потребуются новые механизмы множественного доступа. Это одна из принципиально сложных областей в исследованиях в области беспроводной связи с направленными антеннами, для которой пока не предложено комплексного решения. Кроме того, отраженные и рассеянные компоненты одного канала связи могут создавать помехи для другого.

Вследствие наличия этих эффектов в сочетании с мобильностью узлов общая структура помех становится совершенно новой, по сравнению с микроволновым диапазоном частот, с быстрым и непредсказуемым снижением качества принимаемого сигнала в случайные моменты времени.

На данном этапе детальное моделирование помех для высоконаправленных терагерцовых антенн необходимо для понимания структуры интерференции и определения методов борьбы с помехами для многопользовательского терагерцового беспроводного доступа. В качестве кандидатов необходимо рассмотреть как традиционные методы борьбы с помехами, такие как управление мощностью и планирование частотного диапазона, так и разработать новые методы, специфичные для плотных терагерцовых сетей доступа.

Фаза IV. Интеграция в гетерогенные сети 6G

Следуя концепции гетерогенных сетей доступа, которая в настоящее время рассматривается как совокупность сетей доступа с централизованным контролем посредством «всегда включенного» интерфейса макробазовой станции (например, LTE), предлагаемая система может быть интегрирована в будущие беспроводные сети. В том числе беспроводной доступ в терагерцовом диапазоне частот может быть изначально реализован аналогично первой фазе внедрения технологии 5G NR посредством контроля со стороны LTE и стать управляемой технологией доступа. Как и для других сетей доступа, для интеграции потребуется система управления сетью для взаимодействия с инфраструктурой мобильных операторов связи. Плоскость передачи данных будет работать, используя высокоскоростное соединение последнего метра, когда пользователь находится вблизи терагерцовой точки доступа, и переключаться на более медленную технологию, когда пользователь покидает зону покрытия.

ГЛАВА 14

ПРИЛОЖЕНИЯ ТЕРАГЕРЦОВЫХ СИСТЕМ СВЯЗИ

На сегодняшний день предполагается, что основное применение терагерцовые системы связи найдут на участке доступа и для отвода трафика в системах сотовой связи 5G/6G. Малая длина волны терагерцового диапазона частот является не только недостатком, вызывая диффузионное рассеяние от шероховатых поверхностей, но и делает такие системы привлекательными для широкого круга других приложений. Мы начнем эту главу с изучения приложений терагерцовых сетей связи в параграфе 14.1. Далее мы рассмотрим несколько нестандартных применений включая связь между компонентами вычислительных систем и связь внутри вычислительных элементов в параграфах 14.2 и 14.3.

14.1. Обзор приложений терагерцовых систем

Постоянный рост требований пользователей к скорости передачи данных в беспроводных сетях приводит к непрерывному увеличению объема ресурсов частотного диапазона, выделяемых на одного пользователя. В связи с тем что уплотнение беспроводных сетей точками доступа имеет очевидные ограничения, этот спрос естественным образом преобразуется в дальнейшее увеличение общей полосы частот, используемой технологиями радиодоступа. Наиболее наглядным примером этой тенденции является эволюция на протяжении последних десятилетий семейства технологий IEEE 802.11 – стандарта беспроводных локальных сетей, получивших название «Wi-Fi». Начиная с каналов 22 МГц в уже устаревшем стандарте IEEE 802.11-1997, выделенная полоса пропускания увеличилась до 2,16 ГГц в IEEE 802.11ad менее чем за двадцать лет. Предложения для следующего поколения этой технологии, IEEE 802.11ay, еще более «агрессивны» и намерены занять полосу до 8 ГГц. Последнее значение даже выше, чем несущая частота устаревших решений Wi-Fi, включая выпущенный в конце 2013 года стандарт IEEE 802.11ac. Однако стремление к дальнейшему увеличению пропускной способности, помимо чисто технической задачи широкополосного радио, сталкивается с рядом проблем. Из-за частотного регламента в диапазонах ниже 100 ГГц не осталось блоков частот шире 10 ГГц. Первые доступные блоки такого размера находятся в диапазоне выше 275 ГГц, что уже находится в нижнем терагерцовом диапазоне.

Терагерцовый диапазон — это в некоторой степени уникальная частотная область. В то время как диапазоны выше и ниже терагерцового диапазона (а именно микроволновые и миллиметровые волны, а также инфракрасный диапазон) уже хорошо исследованы, терагерцовый диапазон остается одной из наименее изученных зон в спектре электромагнитных волн. Данный диапазон находится между частотными областями, характеризующимися подходами к генерации электромагнитного сигнала на основе осцилляторов и фотонных эмиттеров, что приводит к сложности генерации сигнала на терагерцовых частотах. Вторым недостатком диапазона, ограничивающим его использование, является высокий уровень потерь на молекулярное (атмосферное) поглощение, вызванный резонансом определенных молекул (кислород, вода) в среде распространения. Кроме того, потери в свободном пространстве на терагерцовых частотах изначально высоки, что требует использования чрезвычайно направленных антенн для обеспечения приемлемой дальности связи.

В то же время огромный объем доступной полосы пропускания (теоретически до нескольких терагерц, т.е. на три порядка выше, чем в нерегулируемых областях на миллиметровых частотах) приводит к потенциальной пропускной способности в несколько Тбит/с, что позволяет передавать трафик практически любого мыслимого приложения. В отличие от видимого света терагерцовые волны все еще проникают через некоторые тонкие объекты и, таким образом, способны передавать данные даже в отсутствие прямой видимости. Например, осуществлять передачу и прием данных смартфоном в кармане одежды пользователя. Наконец, благодаря длине волны порядка сотен микрометров, терагерцовая связь может быть использована для обеспечения взаимодействия между миниатюрными объектами и тем самым снабдить появляющиеся микро- и наноустройства коммуникационными возможностями.

Появление терагерцовой связи дает возможность реализовать несколько групп потенциальных приложений. Некоторые из них представляют собой развитие существующих, в то время как другие практически неосуществимы без терагерцовой связи. Далее перечислены основные приложения и сценарии использования терагерцовой связи, а также кратко описана их мотивация и предполагаемые преимущества.

1. *Информационный душ*. Присущий терагерцовым сотам небольшой диапазон связи вдохновил исследовательское сообщество на поиск сценариев, в которых небольшие (радиус не более нескольких метров) и чрезвычайно высокоскоростные (до Тбит/с) соты могут быть использованы наиболее эффективным способом. Эту группу идей обычно называют «информационным душем» или реже – «душем данных» (information shower). Концепция предполагает размещение терагерцовых точек доступа в зонах с высокой проходимостью (например, у входа на станцию метро, у входов в общественные здания, в залах торговых центров и т.д.). При такой стратегии развертывания каждый из пользователей может получить большой объем данных (до нескольких гигабайт), просто проходя мимо данной точки доступа. Такие информационные души могут быть использованы для бесперебойной доставки обновлений программного обеспечения, а также других типов тяжелого трафика, таких как высококачественное видео (например, фильм для просмотра в поезде). Поскольку время контакта между пользовательскими устройствами очень мало (несколько секунд), внедрение систем информационного душа в существующую сетевую архитектуру требует частичной переработки нескольких уровней для обеспечения быстрого объединения и аутентификации узлов, а также своевременной доставки контента в соответствующий информационный душ и его кэширования там. Было продемонстрировано, что в определенных сценариях внедрение всего нескольких терагерцовых точек доступа и перенаправление на них всего тяжелого и устойчивого к задержкам трафика, когда это возможно, позволяет существенно разгрузить макромасштабную сеть (например, WLAN или сотовую).

2. *Мобильный доступ*. Применимость терагерцовых коммуникаций в типичных сценариях использования (например, доступ к WLAN внутри помещений) ограничена из-за значительных потерь при распространении. Эта проблема может быть решена путем обмена пропускной способности терагерцовых точек доступа на покрытие, в основном за счет уменьшения используемой полосы пропускания и переноса всей связи с частот выше 1 ТГц на так называемые нижние терагерцовые несущие, расположенные около 300 ГГц. В результате появляется возможность создавать надежные беспроводные каналы связи на расстоянии десятков метров при сохранении пропускной способности в десятки гигабит в секунду, что делает возможным применение точек доступа в терагерцовом диапазоне,

подобных Wi-Fi-системам, или даже фемтосотам для сотового доступа. Это приложение является одним из наиболее желательных, но в то же время очень сложным из-за требований надежного отслеживания луча и эффективного управления доступом к среде.

3. *Связь, чувствительная к безопасности.* Возможность создания высоконаправленных лучей с помощью антенных решеток миниатюрных размеров в сочетании с высокой пропускной способностью терагерцовых каналов связи дает ряд преимуществ для использования в целях обеспечения безопасности. Типичный сценарий специального назначения представляет собой связь с многочисленными объектами (солдаты, бронетранспортеры, танки и т.д.), образующими терагерцовую самоорганизующуюся (ad hoc) сеть. Основным преимуществом по сравнению с более низкими частотами является физическая невозможность подслушать или даже заметить передачу для любого узла, расположенного вне луча передатчика. Поэтому безопасность передаваемых данных может быть обеспечена не только правильной схемой шифрования, но и геометрическими особенностями сети. Идея использования направленных терагерцовых антенн для повышения безопасности линий связи специального назначения недавно была расширена и для гражданского использования. Конкретные примеры варьируются от банкоматов с беспроводной аутентификацией до загрузки данных в киосках. Природа терагерцовых каналов с ограниченным диапазоном связи и узкими лучами также обеспечивает основу для методов защиты на физическом уровне, которые вносят некоторый «шум» в сигнал перед передачей. Конфиденциальность сообщения в этом случае обеспечивается тем, что только адресат, на который указывает передатчик, имеет достаточно хороший канал для декодирования данных.

4. *Эквивалентные оптоволокну беспроводные линии связи.* Стратегия беспроводных сетей следующих поколений (5G, 6G и далее) предусматривает появление многочисленных высокоскоростных малых сот, работающих в спектре миллиметровых волн. Такой подход считается практически единственным возможным решением для быстрого увеличения спектральной эффективности и, таким образом, поддержки новых приложений, требующих высокой пропускной способности для виртуальной и дополненной реальности, голографической связи и др. Помимо ряда проблем, связанных с минимизацией помех, поддержкой хэндовера, возникает задача обеспечения надежных фронтальных и транзитных каналов, где проводное или

оптическое решения не всегда являются наилучшим вариантом. Например, известны решения, предполагающие развертывание высокоскоростных малых сот на уличных фонарях, где постоянный источник питания предусмотрен конструкцией, а канал передачи данных – нет. В таком случае беспроводная передача данных (fronthaul/backhaul) рассматривается как перспективная альтернатива оптоволоконным линиям связи. Более того, предложения о применении БПЛА для расширения зоны покрытия и емкости сотовой сети вызывают необходимость обеспечения беспроводной связи и для них.

Предполагаемая скорость передачи данных в миллиметровых сотах (канал доступа) составляет порядка нескольких гигабит в секунду, поэтому пропускная способность канала транзитной связи должна быть в несколько раз выше, чтобы гарантировать надежную и своевременную доставку данных от множества пользователей, обслуживаемых в этой соте. Учитывая это, полоса пропускания, доступная в терагерцовых частотах, является привлекательным ресурсом, а высокие потери распространения в терагерцовом диапазоне (транзитные линии могут быть значительно длиннее фронтальных) могут быть компенсированы чрезвычайно высокой направленностью антенн. Возможность создания мультигигабитных скоростных беспроводных линий связи в нижнем терагерцовом диапазоне на расстояние до 1 км была недавно экспериментально подтверждена.

5. *Связь с миниатюрными устройствами.* Возможность создания микромасштабных приемопередатчиков, работающих в терагерцовом диапазоне, естественно, приводит к желанию оснастить такой функцией микро- и нанороботов. Эти так называемые наноустройства обычно включают весьма примитивную логику из-за ограничений по размеру и энергопотреблению и, таким образом, технически не способны выполнить какую-либо сложную задачу в одиночку. Напротив, соединившись с множеством других наноустройств, они могут образовать сеть, способную помочь человеку в различных областях, начиная от зондирования окружающей среды и медицины и заканчивая услугами по защите населения и оказанию помощи во время стихийных бедствий (PPDR, Public Protection and Disaster Relief). В отличие от крупномасштабных систем терагерцовой связи, которые предполагается строить на основе массивных антенных решеток из сотен или даже тысяч элементов, микро- и наноустройства

должны быть оснащены очень простыми радиомодулями с несколькими или даже одной антеннами. Следовательно, дальность связи в этих наносетях будет ограничена сантиметрами. В то же время пропускная способность линий связи на несколько порядков выше, чем предполагаемая скорость передачи данных, что открывает возможность для тривиальных схем модуляции и кодирования, таких как модуляция ООК (On-Off Keying).

6. *Внутрикристалльные и межкристалльные сети.* В связи с некоторыми ограничениями дальнейшего увеличения частоты процессора горизонтальное масштабирование вычислительной мощности путем включения все большего количества вычислительных ядер на кристалл сегодня рассматривается как основной подход. К сожалению, в большинстве типичных вычислительных задач отсутствует так называемый параллелизм данных, поэтому ядра должны постоянно взаимодействовать друг с другом, обмениваясь общими данными и синхронизируя свои действия. Более того, эти взаимодействия должны быть чрезвычайно быстрыми и надежными. Поэтому обеспечение базовой связи между чипами/ядрами/регистрами/кэшами/пр. имеет решающее значение для достижения желаемого уровня вычислительной производительности. Естественно, что с ростом числа коммуникационных узлов производители начинают все чаще использовать решения из мира телекоммуникаций, где многие проблемы, с которыми они сталкиваются сегодня (например, синхронизация времени, доступ к каналам, маршрутизация и т.д.), уже решены. Эта тенденция в основном охватывается идеей «сеть на чипе» и предполагает использование сетевых технологий для разработки эффективных решений для обмена данными между различными объектами в рамках одного компьютера. В то время как научное сообщество рассматривает варианты использования микроволнового и миллиметрового диапазонов частот для связи между компонентами компьютерных систем, связь между элементами на кристалле и между кристаллами может быть обеспечена только с помощью значительно меньших по размеру приемопередатчиков субмиллиметрового масштаба, что вызывает необходимость использования терагерцового диапазона частот. Терагерцовая электроника на основе графена является одним из основных факторов, обеспечивающих создание массивных многоядерных беспроводных сетей на чипе.

14.2. Связь между компонентами вычислительных систем

Обычная архитектура для связи между различными устройствами, расположенными на материнской плате или подключенными к ней, схематично показана на рис. 14.1а. С точки зрения коммуникаций все интерфейсы представляют собой многополосные шины. Например, интерфейсы QuickPath Interconnect (QPI) и HyperTransport, используемые Intel и AMD соответственно для соединения аппаратных компонентов, в частности процессоров друг с другом, могут обеспечивать скорость до нескольких сотен Гбит/с (например, 400 Гбит/с на 20+20 полосах, работающих на частоте 3,2 ГГц для QPI). Стандартизированный в 2015 году интерфейс PCIe v4.0 предлагает 15,755 Гбит/с на полосу, что составляет 252,064 Гбит/с на 16 полос при использовании «агрессивного» кодирования 128b/130b. Помимо коммуникационных интерфейсов на рис. 14.1а показана система охлаждения, используемая в современных компьютерах. Она включает один или несколько вентиляторов и радиаторов, а также металлические тепловые трубки, проходящие через ряд компонентов, требующих охлаждения. Тепловые трубки передают тепловую энергию от этих компонентов к радиаторам и вентиляторам.

Открытая беспроводная архитектура, показанная на рис. 14.1б, основана на беспроводных соединениях вида *точка–точка* вместо проводов или многополосных шин. Таким образом, каждая из плат оснащается высокоскоростными беспроводными передатчиками. Такая архитектура уменьшает число и сложность физических соединений между платами, тем самым упрощая общую архитектуру системы. В то же время она влечет за собой ряд недостатков и проблем, в основном связанных с распространением электромагнитных волн внутри корпуса компьютера и недостаточной пропускной способностью воздушного интерфейса. Для того чтобы конкурировать с проводными технологиями, пропускная способность должна быть очень высокой, что вряд ли возможно на частотах ниже нескольких десятков гигагерц. Это фактически перевешивает преимущества, присущие беспроводной связи в открытом пространстве, такие как мультимедиа, реализующее соединения вида *точка–многоточие*. Вместе с тем использование более высоких частот приведет к появлению направленных антенн и потребует прямой видимости

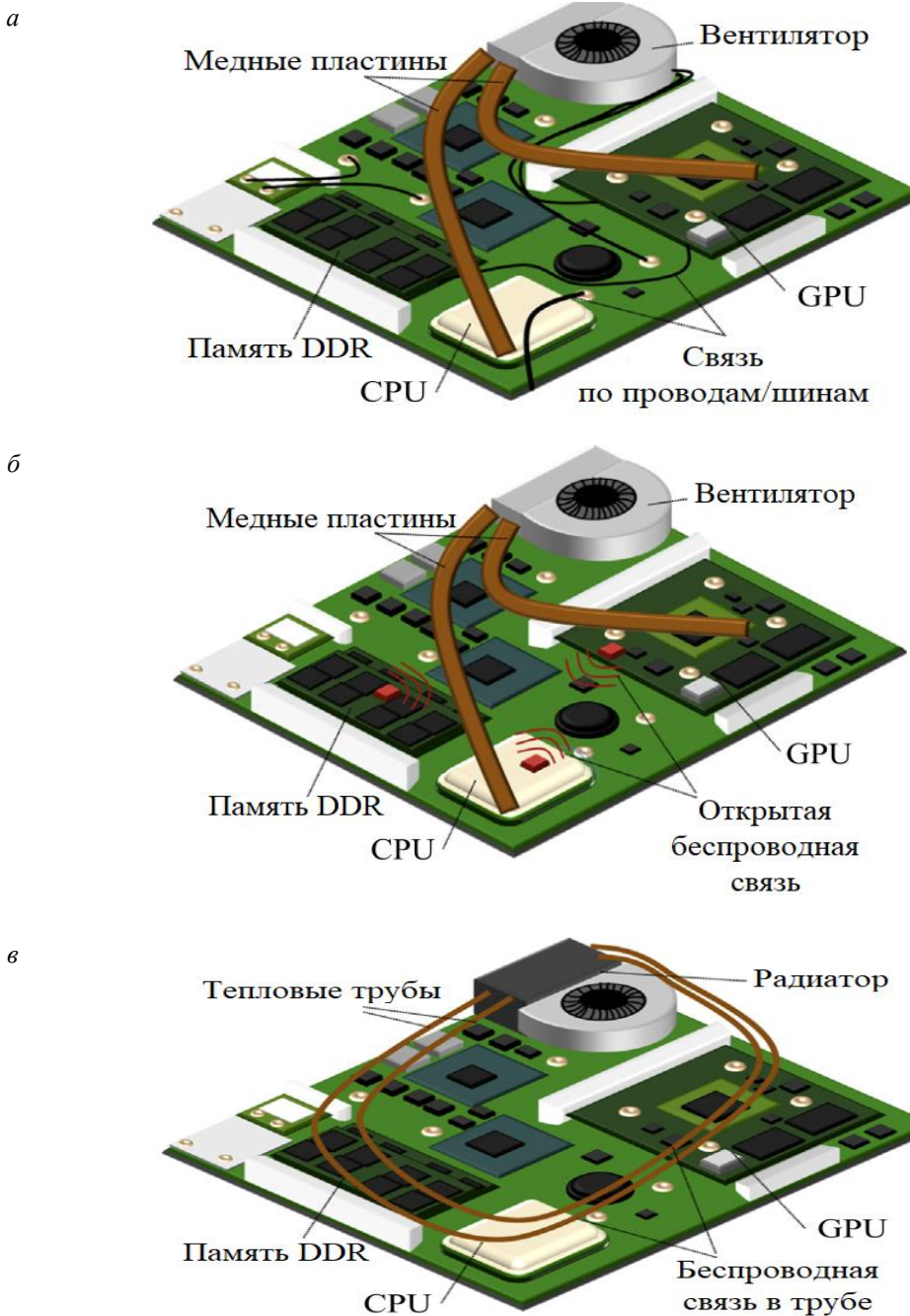


Рис. 14.1. Варианты реализации передачи данных между устройствами, расположенными на материнской плате:

а – обычная архитектура; *б* – открытая беспроводная архитектура; *в* – предлагаемая архитектура

между передатчиком и приемником, что вряд ли возможно в современных системах с высокой степенью интеграции. И последнее, но не менее важное, – такие открытые беспроводные линии связи могут создавать значительный объем помех для других проводных каналов связи, а в некоторых случаях даже для вычислительных чипов и других электронных компонентов.

Для решения выявленных проблем предлагается использовать архитектуру, показанную на рис. 14.1в, где функция охлаждения дополняется обменом информацией через ту же среду (полые трубки). Хотя предлагаемое решение классифицируется как беспроводное из-за использования диапазона 0,1–10 ТГц, оно концептуально схоже с оптическим, требующим использования «волноводов» для обеспечения связи между узлами. Использование волноводов обеспечивает отсутствие помех между линиями связи. Поскольку влияние терагерцового излучения на кремниевую электронику в настоящий момент не до конца изучено, изоляция путей распространения гарантирует, что терагерцовое излучение не повлияет на функциональность компонентов компьютера.

В предлагаемой архитектуре в точках, где полая труба соприкасается с компонентом или охлаждающим блоком, прикрепленным к ней (вход), и выходит из компонента (выход), установлены терагерцовые приемопередатчики для обеспечения связи между интересующими компонентами. Предполагается, что сигналы не распространяются через полые охлаждающие трубки. Проблема создания терагерцовых излучателей активно изучается в литературе и уже имеется ряд прототипов. Наиболее распространенным принципом является использование лазера для облучения определенных полупроводников. Размер этих антенн может быть сделан небольшим, что позволяет интегрировать их в будущие интегральные схемы. Более того, уже предложен ряд конструкций интегральных приемопередатчиков для нижнего терагерцового диапазона (до 1 ТГц). Производительность предлагаемого дизайна в основном зависит от следующих параметров: а) типа охлаждения; б) материала оболочки трубы; в) расстояния между приемником и передатчиком; г) кривизны трубы.

Благодаря использованию простой схемы модуляции ООК алгоритмическая сложность предлагаемой системы очень низкая, так как сигнал может быть передан с помощью импульсного приемопередатчика и принят с помощью простого детектора энергии

(см. главу 12). Учитывая технические проблемы реализации, такая система может быть реализована путем использования аналогового детектора энергии, состоящего из квадратичного интегратора, порогового детектора и, возможно, полосового фильтра, что может быть предпочтительнее цифрового приемника для реализации предлагаемой системы.

Предложенная на рис. 14.1б архитектура характеризуется схожими недостатками, включая необходимость использования сложных мультиплексоров для обеспечения сетевого взаимодействия между более чем двумя устройствами, подключенными напрямую. Оба варианта лучше всего подходят для каналов связи *точка-точка*. Когда число устройств увеличивается, вместо создания Mesh-сети между устройствами можно использовать архитектуру на основе звезды с мультиплексором в центре. В дополнение к обеспечению связи между участвующими объектами такая система позволяет охлаждать вычислительные компоненты.

Экспериментальные данные

В этом параграфе оценивается эффективность передачи данных в предложенной системе. Рассматриваются как прямые, так и изогнутые полые трубы длиной l , кривизной α и диаметром d . Сначала описывается испытательный стенд для измерения сигналов при распространении внутри полых труб, а затем представлена методика анализа коэффициента битовых ошибок BER и пропускной способности модуляции ООК (On/Off Keying). Далее обсуждаются полученные результаты. Характеристики системы охлаждения можно найти в литературе, а здесь основное внимание уделяется беспроводной связи.

Для того чтобы оценить характеристики передачи информации в предлагаемой системе связи терагерцового диапазона, были проведены измерения на стенде, которые осуществлялись для медных труб различной длины, диаметра и кривизны. Измерительным прибором был TeraView Mini Pulse, способный проводить измерения в терагерцовом диапазоне примерно от 60 ГГц до 4 ТГц (рис. 14.2). Характеристики измерительного прибора представлены в табл. 14.1. Передача и прием терагерцовых импульсов основаны на лазерном возбуждении полупроводников. Лазерные импульсы длиной около

100 фс направляются на антенны приемника и передатчика. Лазерный импульс заставляет терагерцовые антенны либо принимать, либо передавать терагерцовые импульсы. В целях исследования поведения сигнала во временной области терагерцовые импульсы регистрируются через интервалы определенной длины.

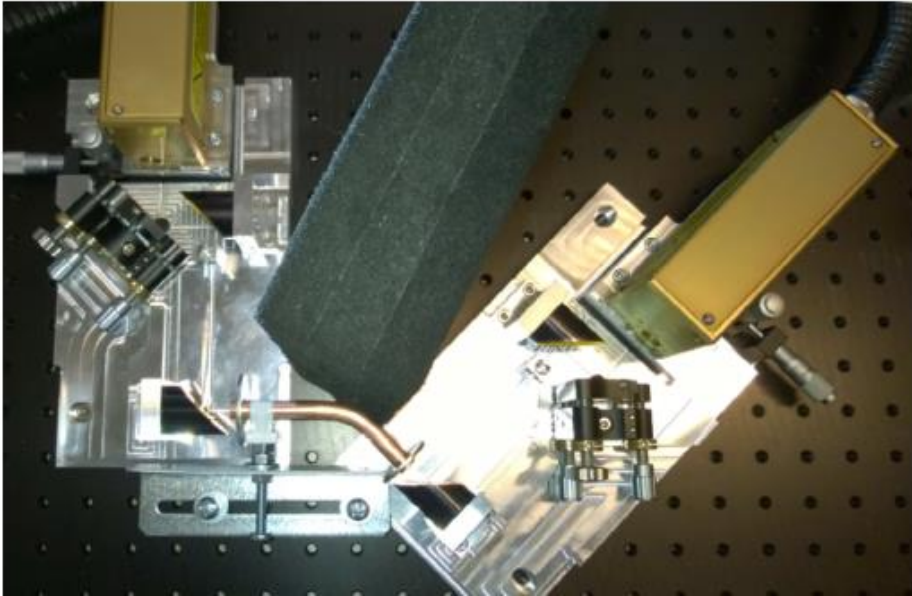


Рис. 14.2. Экспериментальная установка

Таблица 14.1

Характеристики измерительного устройства

Устройство	TeraView MiniPulse
Полоса пропускания	~ 4 ТГц
Разрешение по времени	8.3 фс
Частотное разрешение	5.9 ГГц
Длина импульса (лазер)	~ 100 фс
Длительность импульса, ТГц	< 500 фс

Для того чтобы захватить весь импульс, в методе спектроскопии во временной области ТГц (THz-TDS) используется линия задержки лазерных импульсов, т.е. прием (или передача) сигнала задерживается на один временной интервал в течение желаемого окна обнаружения. Это требует около 4000 импульсов на каждый обнаруженный импульс. В реальных измерениях все результаты усредняются по 1000 импульсам, т.е. каждый результат получается примерно при 4106 ТГц импульсах. После получения сигнала во временной

области можно получить его поведение в частотной области, выполнив преобразование Фурье.

Идея эксперимента заключается в том, чтобы проверить, можно ли использовать обычные медные трубки в качестве терагерцовых волноводов. Предпочтение отдадим медным трубкам из-за их доступности, низкой стоимости и широкого применения в системах охлаждения. Внутренняя поверхность не подвергалась специальной обработке. Были выбраны два внешних диаметра трубок 10 мм и 12 мм с внутренним диаметром 8 мм и 10 мм. Для обоих вариантов были измерения для трубок различной длины. Кроме того, измерены трубки различной кривизны от 0° до 90° с интервалом в 30° для трубок длиной 14 см. Результаты измерений приведены в табл. 14.2. Измерения проводились в сухом воздухе (1,5 % водяного пара). Подставленные данные обеспечивают нижнюю границу производительности системы как для пассивных методов охлаждения, так и активных методов, использующих теплоносители. Отметим, что в реальных условиях необходимо обеспечить идеальную изоляцию охлаждающего компонента на входе и выходе трубок. Тем не менее ожидается, что полученные результаты позволят дать приблизительную оценку производительности тепловых трубок в данном сценарии. Для систем с жидкостным охлаждением беспроводная связь практически невозможна, поэтому данный метод охлаждения не рассматривался.

Таблица 14.2

Измерения: «N» – узкий, 10 мм, «W» – широкий, 12 мм

Медная труба	8 см	10 см	12 см	14 см
Прямой	N/W	N/W	N/W	N/W
Изогнутый	--	--	--	W
Изогнутый	--	--	--	W
Изогнутый	--	--	--	W

Далее представлены результаты, показывающие, что медные трубки являются потенциальным кандидатом для терагерцовых волноводов, поскольку они а) эффективно отражают терагерцовое излучение, б) широко используются в системах охлаждения. Также показано, что отражения внутри трубок значительно увеличивают BER. Это можно компенсировать увеличением межсимвольного времени, уменьшая при этом теоретическую скорость передачи данных.

Методология оценки BER и пропускной способности

Импульсная модуляция ООК была предложена как простейшее решение для терагерцовой связи. В ООК передатчик генерирует импульс для передачи логической «1» и молчит для передачи логического «0». В то же время приемник применяет детектор энергии, чтобы определить тип принимаемого символа. Следует подчеркнуть, что представленные ниже результаты обеспечивают нижнюю границу производительности системы, оставляя возможность для дальнейших усовершенствований.

При оценке производительности на уровне канала исследованы две метрики – BER и пропускная способность. Заметим, что при условии отсутствия разделительного расстояния между импульсами теоретическая пропускная способность для модуляции ООК равна $1/\Delta t$, где Δt – длительность импульса. Поскольку устройство имеет полосу пропускания 4 ТГц, мы не можем достичь значения пропускной способности, превышающего 4 Тбит/с. Более того, если требуется минимизировать коэффициент ошибок на декодере, необходимо обеспечить достаточное разделение между символами τ , чтобы позволить беспроводному каналу прийти в нормальное состояние после предыдущей передачи логической «1». Таким образом, пропускная способность канала связи может быть оценена как $1/(\tau + \Delta t)$.

Даже когда межсимвольный интервал τ не равен нулю, существует вероятность того, что текущий символ будет декодирован с ошибкой. Более того, чем меньше τ , тем выше межсимвольная интерференция (ISI, Inter-Symbol Interference). Канал также несимметричен, т.е. вероятность ошибки при получении либо «0», либо «1» после «0» пренебрежимо мала по сравнению с вероятностью ошибки при получении «0» после «1». Таким образом, возникает задача максимизации пропускной способности системы, $S = [(1 - BER)/(\tau + \Delta t)]$, если известен отклик канала на приемнике для последовательности «10». Для решения этой задачи оценивается BER для заданных τ и Δt .

Традиционный подход к оценке BER основан на использовании ОСШ, однако он требует знания потерь на пути распространения и явно предполагает аддитивный гауссовский белый шум (AWGN, Additive Gaussian White Noise) в приемнике. Оценка потерь на пути внутри трубки круговой конфигурации является сложной задачей, так как существует множество многопутевых компонентов, поступающих в приемник со сдвигом во времени. Более того, по сравнению с

обычным каналом AWGN, доминирующим эффектом, ограничивающим пропускную способность канала связи, в рассматриваемом случае является ISI.

Далее для получения BER использован приближенный метод. Во-первых, заметим, что временной отклик канала усреднен по большому числу измерений и их амплитуда на несколько порядков превышает уровень шума. Во-вторых, ISI намного больше, чем длительность импульса Δt . Наконец, напомним, что для порогового детектора значение BER зависит от порога и достигает оптимального значения, если порог находится на пересечении плотности вероятностей энергетических уровней, соответствующих символам «1» и «0».

Для получения достоверных оценок плотности распределения, соответствующих «1» и «0», число измерений должно быть порядка нескольких тысяч, что невозможно получить на практике. Однако может быть сгенерирован мешающий сигнал для интересующего нас символа для каждого из возможных $2N$ символов комбинации путем их наложения. Таким образом, для каждой комбинации известен отклик канала, определяющий ISI. Затем добавляется соответствующий полезный сигнал (принятый импульс для «1» и молчание для «0») и оценивается общее количество энергии на приемнике. В итоге получаем два массива уровней энергии, рассчитанных для всех рассмотренных значений R , для которых строятся эмпирические плотности распределения энергии в приемнике для обоих символов. Несмотря на то, что сложность предложенного метода составляет $\mathcal{O}(2^N R)$, сложность одной реализации достаточно мала: достаточно получить сумму массива, состоящего из $2N$ элементов, что можно сделать не медленнее, чем $\mathcal{O}(N)$. Таким образом, предложенное приближение является вычислительно выполнимой задачей по сравнению с получением плотностей вероятностей на основе анализа тысяч реализаций.

В эксперименте значение N выбрано равным 15, что обеспечивает компромисс между точностью аппроксимации плотностей вероятностей и вычислительной сложностью. Поскольку расположение передатчика и приемника неидеально даже для прямых трубок, в экспериментах может отсутствовать линия прямой видимости и результирующий символ представляет собой суперпозицию отраженных компонентов. Тогда можно принять, что распределение энергии для символа «1» приближается распределением Рэля, соответствующим состоянию отсутствия прямой видимости. Было установлено, что

ISI для символа «0» соответствует гауссовскому распределению при любом значении межсимвольного времени. На последнем этапе эксперимента численно находится точка пересечения плотностей, обеспечивающая оптимальное значение порога обнаружения символа.

Вероятность ошибки и пропускная способность

В данном параграфе изложены результаты оценки BER (Bit Error Rate). Начнем с анализа рис. 14.3, где BER представлен как функция межсимвольного времени $\tau + \Delta t$. Отметим, что применяется усреднение по скользящему окну для частичной компенсации колебаний BER, вызванных корреляцией между текущим сигналом и помехами от ранее принятых сигналов. Для широких и узких трубок сначала иллюстрируем BER для различных длин прямой трубки (рис. 14.3а и рис. 14.3б), а затем для различных углов изгиба трубки (рис. 14.3в и рис. 14.3г).

Обсуждая качественные эффекты, наблюдаемые на рис. 14.3а и рис. 14.3б, можно заметить, что, помимо колебаний, общая тенденция для BER определяется значением межсимвольного интервала. Другими словами, ожидается, что когда импульсы посылаются реже, соответствующий BER будет ниже. В случае прямой трубки отмечаем, что более короткие трубки имеют более низкие значения BER для тех же межсимвольных интервалов. Это логично, поскольку меньшая длина приводит к меньшему количеству отражений и, следовательно, меньшему ослаблению сигналов, поступающих в приемник. Напомним, что значения BER ниже 10^{-18} обеспечивают идеальные условия канала, гарантирующие надежную связь *точка–точка*. Такое значение BER достигается для диапазонов межсимвольного интервала приблизительно 2,5–3,0 нс и 2,0–2,5 нс для узких и широких труб соответственно. Таким образом, предлагаемая система обладает достаточной надежностью передачи данных при использовании прямых труб.

Анализируя данные, представленные на рис. 14.3в и рис. 14.3г, видим, что дисперсия полученных результатов значительно больше, чем на предыдущих рисунках, чему есть два объяснения. Во-первых, кривизна трубки вызывает сложные отражения в месте изгиба, что должно привести к значительно большей дисперсии измеренного отклика канала и, как следствие, потребовать более обширного измерения. Другим объяснением служит сложность самих экспериментов, в частности их повторяемость. Действительно, для получения

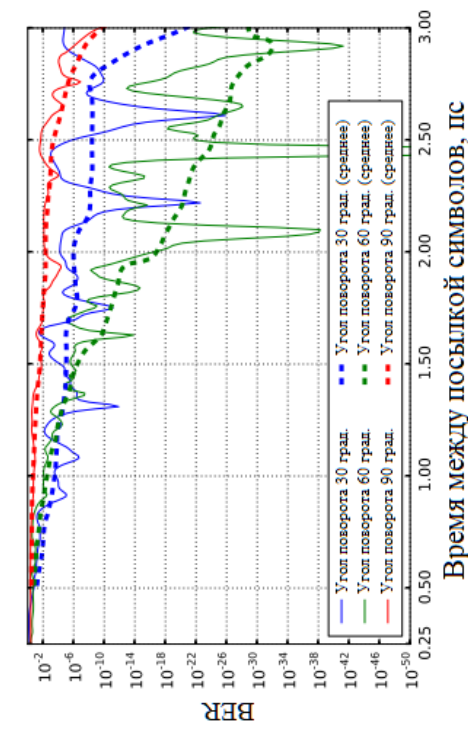
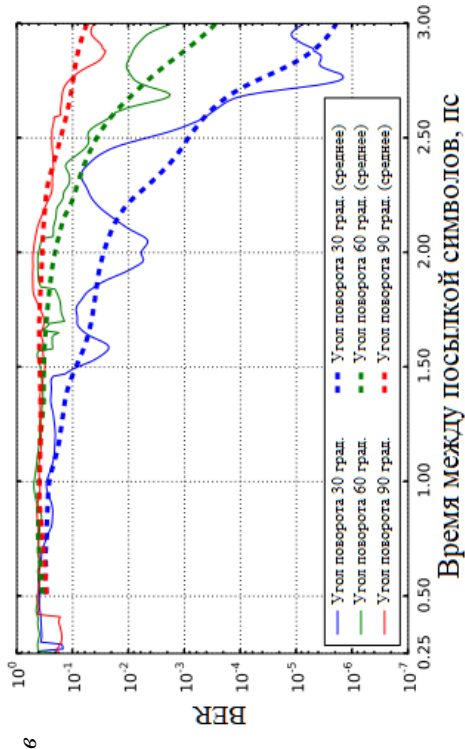
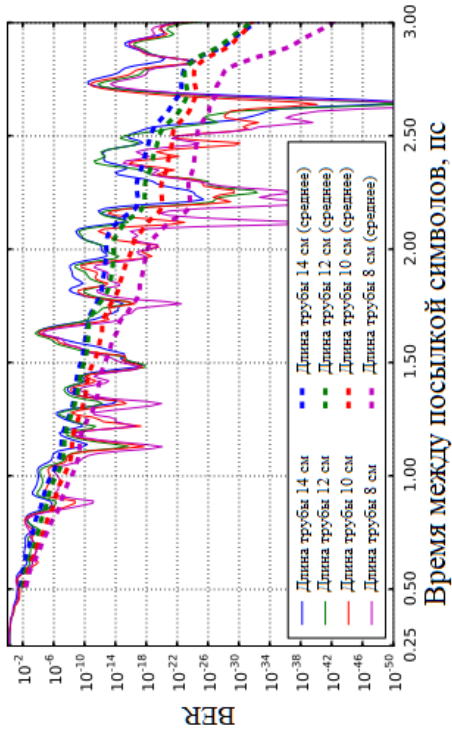
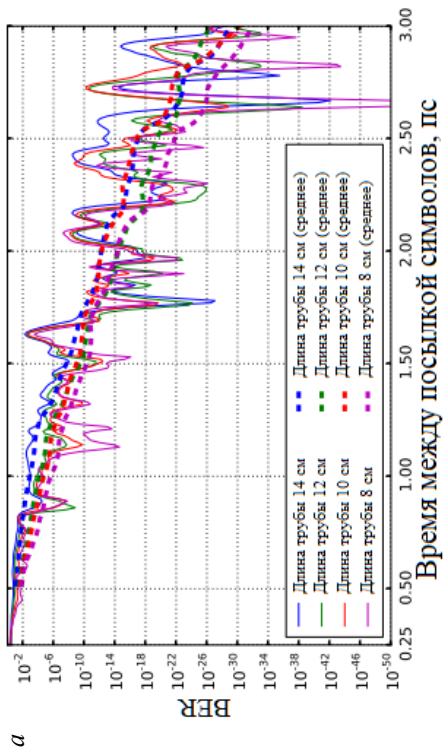


Рис. 14.3. BER как функция системных параметров:

a – прямая узкая трубка; *б* – прямая широкая трубка;
в – искривленная узкая трубка; *г* – искривленная широкая трубка

согласованных результатов требуется идеальный изгиб и позиционирование передатчиков и приемников. Несовершенства в изгибе трубки также могут привести к особым конфигурациям, когда лучи отражаются обратно к излучателю. Вклад этих эффектов крайне сложно оценить без аналитической модели распространения терагерцовых волн внутри трубок. Тем не менее представленные ниже практические результаты дают качественную оценку эффекта изгиба в целом и позволяют понять влияние кривизны.

Анализируя рис. 14.3в и рис. 14.3г, видим, что для широкой трубки лучший BER получается при кривизне 60° , чем при 30° , что противоречит ожиданиям. Кроме того, количественная разница в BER между изогнутой широкой и изогнутой узкой трубками чрезвычайно велика, т.е. изогнутая узкая трубка работает намного хуже, чем изогнутая широкая трубка. Однако даже с учетом этих двух особенностей, можно извлечь пользу из проведенных экспериментов. Во-первых, обратим внимание, что количественные значения BER для изогнутой широкой трубки ненамного ниже, чем для прямой широкой трубки. Также заметим зависимость между углом изгиба и соответствующим BER – чем менее прямой является трубка, тем хуже будут значения BER. Наконец, углы изгиба более 30° приводят к высоким значениям BER (более 10^{-10} для разумных межсимвольных временных интервалов). Поэтому предложенную систему следует использовать для прямых трубок и трубок с малым углом изгиба. Если система применяется для изогнутых трубок, она должна включать в себя канальное кодирование для компенсации высоких значений BER.

Помимо надежности предлагаемой системы большое значение имеет ее производительность с точки зрения достижимой скорости передачи данных. Для того чтобы исследовать этот вопрос, рассчитано теоретическое значение пропускной способности для данного канала по формуле $S = [(1 - BER)/(\tau + \Delta t)]$. Результаты показаны на рис. 14.4, где видим, что для больших значений межсимвольного интервала (начиная с 1 пс) влияние неправильного приема символов канала на пропускную способность пренебрежимо мало, так как значения BER составляют порядка 10^{-6} . Таким образом, все рассчитанные кривые следуют тенденции, заданной исходными данными. В то же время, при приближении к более низким значениям межсимвольных интервалов, на результаты пропускной способности влияют две

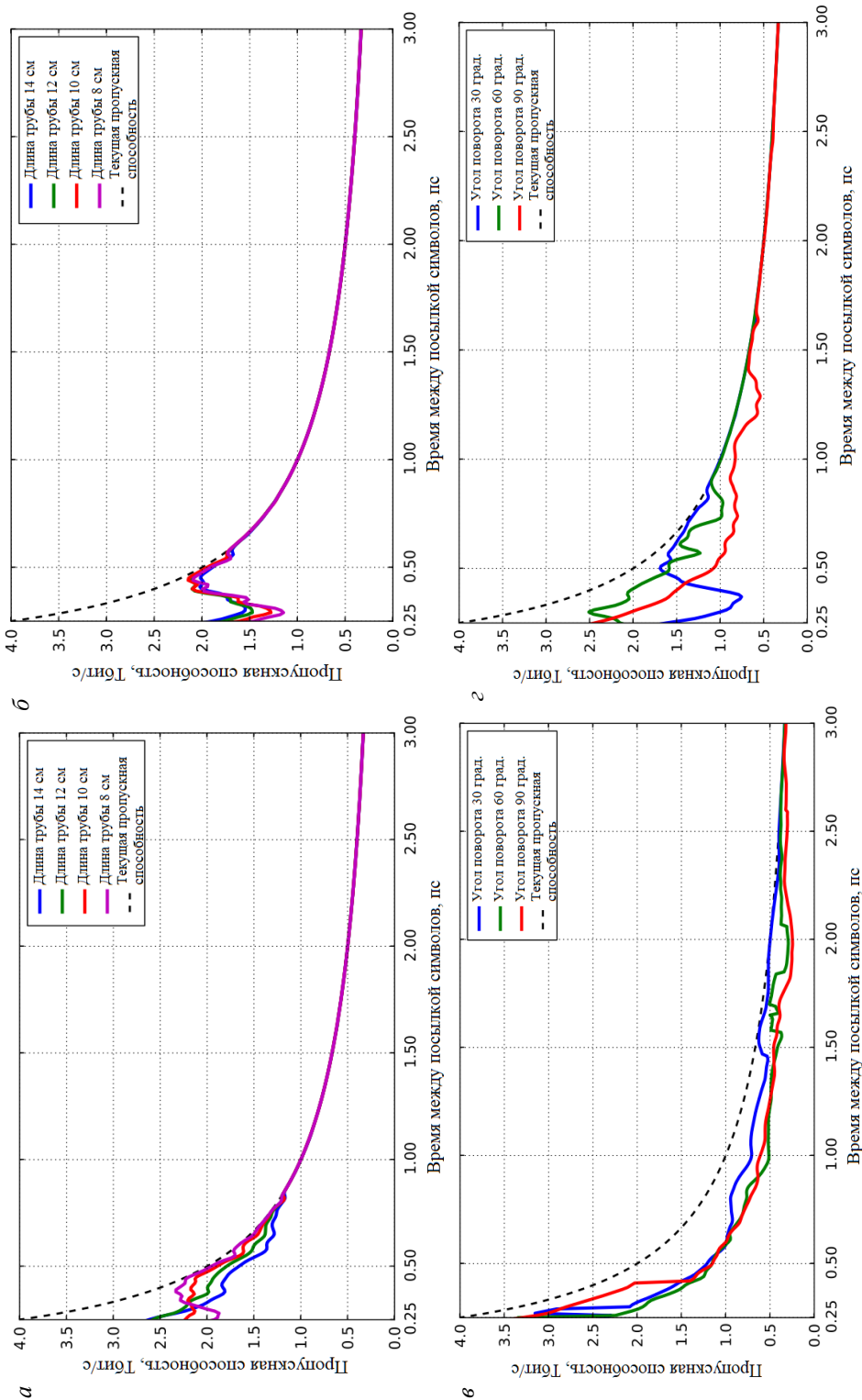


Рис. 14.4. Пропускная способность как функция системных параметров:

а – прямая, узкая трубка; *б* – прямая широкая трубка;

в – искривленная узкая трубка; *г* – искривленная широкая трубка

тенденции: увеличение пропускной способности $1/(\tau + \Delta t)$ и уменьшение значения $(1 - BER)$ за счет уменьшения межсимвольного интервала. Данная комбинация приводит к тому, что теоретическая пропускная способность имеет максимум в интервале 0,25–0,5 пс. Соответствующее значение пропускной способности составляет около 2 Тбит/с, что превышает скорость передачи данных современных проводных интерфейсов для связи между компонентами компьютера и абсолютно достаточно для любого реалистичного типа трафика. Оптимальная пропускная способность для выбранных расстояний и кривизны показана в табл. 14.3.

Таблица 14.3

Оптимальная пропускная способность канала связи, Тбит/с

Тип	Длина			
	8 см	10 см	12 см	14 см
Прямой, узкий	2,51	2,52	2,53	2,56
Прямой, широкий	2,17	2,15	2,12	2,03
Тип/кривизна	0°	30°	60°	90°
Изогнутый, узкий, 14 см	3,56	3,17	2,73	3,16
Изогнутый, широкий, 14 см	2,03	1,7	2,49	2,38

14.3. Связь внутри вычислительных элементов

Несмотря на усовершенствование технологического процесса, позволяющего в ближайшем будущем интегрировать в чип (например, центральный процессор) все больше ядер, с ростом числа ядер обеспечение эффективной синхронизации памяти между компонентами чипа становится все более сложной задачей. Многоуровневые иерархии памяти, наряду с многоядерными архитектурами, образуют распределенную подсистему памяти, которая должна обеспечить согласованность и когерентность между блоками памяти, чтобы сделать ее пригодной для использования разработчиками программного обеспечения. Хотя свойства архитектуры памяти зависят от ее реализации, наблюдаются две основные тенденции: а) размеры кэшей растут, заставляя их конкурировать с другими устройствами за площадь кристалла, б) коммуникационная инфраструктура переходит от простых соединений на основе шин, которые страдают от проблем масштабируемости, к более сложным. Недавно предложенная парадигма проектирования 3D-чипов может решить первую проблему, распределив ядра и кэш-память между разными слоями.

Между тем проблема эффективных коммуникаций все еще остается нерешенной.

Предполагаемым решением проблемы является концепция «сеть на кристалле» (NoC, Network-on-Chip), предлагающая заменить коммуникационную шину сетью с более сложной топологией. Хотя использование эффективных топологий, таких как полностью связанные графы или «малые миры» (small worlds), может поддерживать задержку на приемлемом уровне и уменьшить площадь, используемую для коммуникационной инфраструктуры, они также создают дополнительные проблемы, такие как маршрутизация и энергоэффективность.

В качестве альтернативы, позволяющей избавиться от проводов, недавно предложенная парадигма беспроводных NoC (WNoC, Wireless Network-on-Chip) обещает заменить проводные соединения беспроводным аналогом. В результате все компоненты микросхемы будут иметь только одно беспроводное приемопередающее устройство вместо нескольких проводных соединений. Концепция WNoC упрощает размещение компонентов, может поддерживать существующие протоколы когерентности памяти и позволяет передавать сообщения широкого спектра. Последние достижения в области беспроводной связи малого радиуса действия, особенно терагерцовой технологии, предоставляют значительное количество ресурсов для обеспечения высокой пропускной способности и низкой задержки трафика, передаваемого внутри кристаллов.

Современное состояние исследований

Проследив эволюцию центрального процессора (ЦП) за последние десятилетия, выявлен ряд тенденций (рис. 14.5). Начиная с 1970-х годов метрики росли линейно, обеспечивая постепенное увеличение производительности процессора. Несмотря на то, что число транзисторов на чип продолжает расти благодаря прогрессу в производственном процессе, другие характеристики сегодня сталкиваются с целым рядом проблем.

Первая проблема связана с максимальной потребляемой мощностью процессора, ограниченной а) количеством тепла, которое необходимо отвести от процессора; б) эффективностью систем охлаждения. Одним из аспектов проектирования, влияющих на

количество рассеиваемого тепла, является рабочей частота. Максимальная производительность одноядерного процессора является функцией тактовой частоты процессора, которая не может вырасти намного выше 4 ГГц при комнатной температуре без принципиального изменения материала, из которого изготовлены транзисторы. Численно однопоточные процессоры достигли максимальной производительности примерно в 2008–2012 годах, дальнейшее увеличение вычислительной мощности происходило по линейной шкале. Значения энергопотребления, тактовой частоты и однопоточной производительности, достигнутые к 2010 году, являются максимально возможными при нынешнем состоянии кремниевой электроники, обеспечивающей компромисс между сложностью конструкции систем охлаждения и производительностью процессора. Любое дальнейшее увеличение энергопотребления процессора потребует значительного прогресса в этих областях.

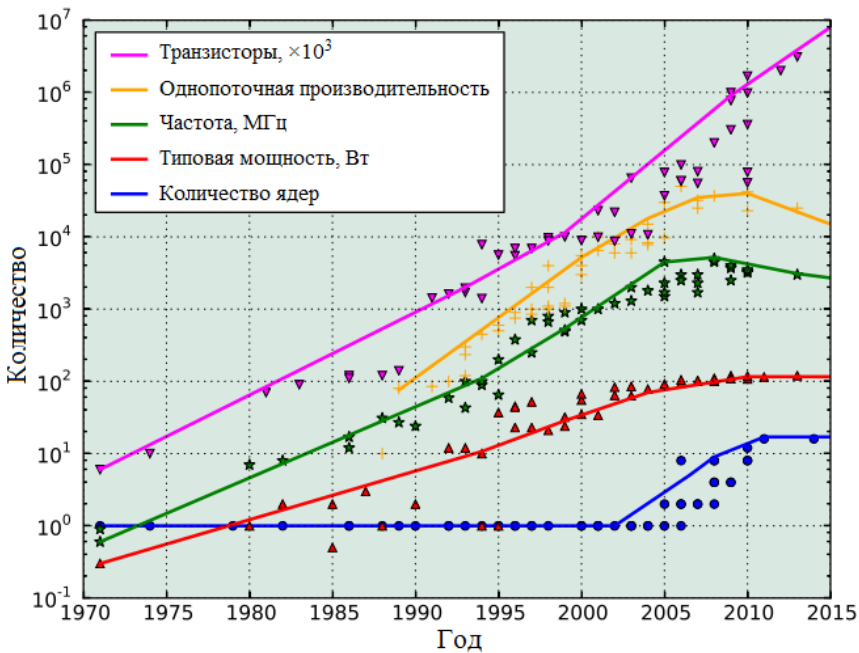


Рис. 14.5. Эволюция характеристик ЦП

Еще одна проблема, с которой сталкивается индустрия ЦП, – это недостаточная производительность памяти с произвольным доступом (RAM, Random Access Memory). Основная проблема заключается в том, что производительность обычного процессора в несколько раз опережает производительность памяти RAM и шины CPU-RAM даже

при использовании нескольких банков памяти. Эта проблема была решена путем внедрения механизмов многоуровневого кэширования, когда промежуточные результаты вычислений процессора временно хранятся в кэш-памяти – небольшом хранилище, обладающем значительно более высокой скоростью и меньшей латентностью (задержкой) по сравнению с оперативной памятью. Интеграция высокоскоростной кэш-памяти на чипе в процессор Intel 80486 еще в 1989 году радикально повысила скорость выполнения обычных операций.

Для решения упомянутых проблем основные производители процессоров, Intel и AMD перешли на многоядерный дизайн, представив в 2005 году двухъядерные Intel Pentium D и Athlon 64 X2 соответственно. Многоядерные процессоры позволили идти в ногу с растущими вычислительными потребностями последнего десятилетия, сохраняя при этом компромисс между мощностью/частотой/производительностью, достигнутый для одного ядра. Проектирование многоядерных процессоров привнесло свои проблемы, которые не были очевидны до тех пор, пока число ядер было ограничено несколькими единицами. Одной из основных проблем, с которой сегодня сталкивается промышленность, является эффективная реализация коммуникационной инфраструктуры, связывающей растущее число ядер и позволяющей эффективно их использовать.

Задачи проектирования NoC/WNoC

За последнее десятилетие производители пытались решить проблему реализации процессоров путем соединения вычислительных элементов с помощью сложных топологий сетей, таких как Mesh или «малый мир». Этот подход известен как парадигма «сеть на кристалле» (NoC). Основными целями исследования являются быстрая маршрутизация данных через NoC, энергоэффективность и интеграция в существующие архитектуры ЦП (центральный процессор). В частности, традиционные алгоритмы маршрутизации на основе таблиц поиска должны быть заменены более быстрыми решениями, реализованными в аппаратной логике. Сложность проводных коммутаторов и буферов данных должна быть как можно ниже, чтобы снизить энергопотребление, нагрев и использование пространства.

Существует множество примеров успешного применения концепции NoC, например, в графических процессорах (GPU, Graphics

Processing Unit). Причина необычайного роста производительности таких систем в основном связана с природой задач, позволяющих осуществлять их эффективное распараллеливание, что приводит к простым элементам обработки и четко определенным и заранее известным схемам передачи трафика. Применение концепции NoC к CPU является более сложным процессом, поскольку выполняемые задачи сильно различаются по своей специфике, уровню распараллеливания и, следовательно, могут потребовать интенсивного обмена трафиком между вычислительными элементами в произвольные моменты времени, что предъявляет дополнительные требования к проектированию внутрипроцессорной коммуникационной инфраструктуры.

В последние годы было создано несколько прототипов решений, включая процессор Sony/Toshiba/IBM Cell (12 ядер), чип Tiler TILE64 (64 ядра) и прототип Intel TeraFLOPS (80 ядер), а также процессоры Intel Xeon и AMD EPYC для серверных решений. Однако предлагаемое проводное подключение с помощью нескольких коммутаторов применимо к принципиально новым системам, разработанным с учетом низкопроизводительных ядер. При применении к существующим ЦП общего назначения парадигма NoC требует радикальных изменений в архитектуре, операционных системах и низкоуровневом программном обеспечении. В том числе структура кэша и связанные с ней протоколы когерентности данных должны быть переработаны. Это накладывает большие риски на производителей процессоров, поскольку все принципиально новые конструкции имеют свои «проблемы зрелости». Дополнительные шины, необходимые для топологии, занимают ценную площадь на чипе, уменьшая количество транзисторов, доступных для других компонентов. Наконец, требования к задержкам при обмене информацией внутри процессора являются жесткими и ограничивают использование сетей с несколькими точками приемопередачи.

Концепция WNoC, в которой проводные интерфейсы заменяются миниатюрными беспроводными приемопередатчиками, решает некоторые проблемы проводных NoC. Беспроводная связь между внутренними элементами снимает проблему площади, занимаемой шинами. Дополнительными преимуществами этих систем являются миниатюрные приемопередатчики, широкополосный характер радиointерфейса, позволяющий осуществлять связь *одна точка–многоточие* и низкое энергопотребление. В общей форме WNoC все

еще опираются на парадигму многошаговых сетей с несколькими точками приемопередачи, требующей модификации внутренней логики процессора и ставящую новые задачи, связанные с протоколами доступа к среде, повторным использованием частот и подавлением помех. Однако при применении только к отдельным интерфейсам, таким как интерфейс между кэшами ядра и общим кэшем последнего уровня, этот подход обеспечивает обратную совместимость с текущей конструкцией процессора.

Современные беспроводные предложения WNoC в основном сосредоточены на системах миллиметровых волн (mmWave) в диапазоне 30–300 ГГц, предлагающих пропускную способность каналов связи до нескольких Гбит/с. Однако заявленные скорости достигаются только при использовании сложных схем модуляции и кодирования, которые могут быть невыполнимы для приемопередатчиков с ограниченным уровнем сложности. Для решения проблемы логично использовать еще более высокие частоты, предлагающие такую же или более высокую пропускную способность с более простыми схемами модуляции и кодирования, выигрывая за счет беспроводной природы связи и дальнейшей миниатюризации антенн.

Таким образом, естественным следующим шагом является терагерцовый диапазон частот: 0,1–10 ТГц. Используя этот сверхширокий диапазон, антенны размером в сотни микрон могут быть использованы для передачи и приема данных со скоростью до нескольких Тбит/с. В результате последних исследований имеются детальное представление о специфике терагерцового распространения, прототипы микроантенн и интегрированные в чип приемопередатчики. Показано, что, используя поддиапазон 0,1–0,54 ТГц и всего 0,1 Дж на символ, можно достичь пропускной способности в 2 Тбит/с на расстоянии до 3 см. Однако до сих пор нет понимания того, какие скорости достаточны для связи внутри ЦП.

Электронные и фотонные терагерцовые технологии

Для обеспечения связи в терагерцовом диапазоне необходимо понять возможности и проблемы реализации, связанные с основными блоками системы связи в терагерцовом диапазоне – усилители, множители частот, цифроаналоговые преобразователи. С точки зрения устройств требуемые характеристики приемопередатчиков включают относительно высокую выходную мощность, высокую

чувствительность приемников и низкий уровень шума. Такие системы также должны преодолевать высокие потери в тракте в терагерцовом диапазоне. Ниже показаны некоторые возможности, ограничения и проблемы электронных и нанофотонных технологий для терагерцовой связи (см. рис. 14.6).

Продолжающийся рост производительности цифровой технологии CMOS (Complementary-symmetry/Metal-Oxide Semiconductor) за счет увеличения размеров привел к тому, что частота среза f_T и максимальная частота колебаний f_{max} устройства достигли нескольких сотен ГГц. Традиционно при разработке устройств для высоких частот можно было использовать улучшение технологии генераторов для достижения более высокой частоты и уменьшения площади схемы путем простого масштабирования размеров и значений компонентов микросхемы (транзисторы, конденсаторы и индукторы). Хотя такой подход и эффективен, когда рабочая частота намного ниже частоты отсечки, на терагерцовых частотах его эффективность падает вследствие дополнительных ограничений, накладываемых потерями, возникающими в металлических структурах на кристалле. Таким образом, существует два возможных пути преодоления проблем, связанных с созданием высокоэффективных, недорогих и энергосберегающих терагерцовых компонентов на кристалле. Первый подход основан на поиске новой технологии, которая предлагает лучшие перспективы масштабирования и более высокую собственную частоту. Альтернативный подход заключается в использовании нелинейности устройства для эффективной генерации энергии на гармониках высшего порядка.

Подходы, которые не только полагаются на характеристики отдельных устройств, но и используют новые концепции схем для генерации, излучения и управления терагерцовыми частотами, широко применяются в науке и промышленности. Например, уже продемонстрирован генератор на основе CMOS с частотой 220 ГГц и усилением сигнала четвертой гармоники до 870 ГГц. Кроме того, терагерцовые компоненты должны соответствовать ограничениям по площади и мощности для применения в системах на кристалле. Интеграция терагерцовой технологии с существующей цифровой CMOS-технологией на одной матрице может использовать эффект масштаба для создания полностью интегрированной и экономически эффективной системы на кристалле. Вкратце отметим несколько

инноваций в технологиях устройств и схемотехнике, которые произошли за последнее десятилетие в области терагерцовой связи (рис. 14.6).

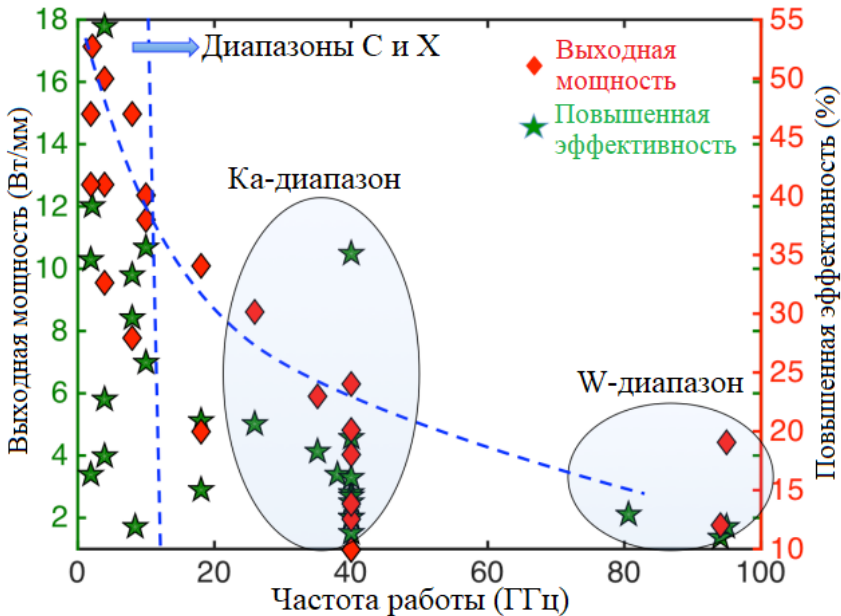


Рис. 14.6. Эволюция терагерцовой технологии в 2010–2020 годы

65-нанометровая цифровая CMOS-технология была использована для успешной демонстрации полностью интегрированного приемопередатчика на 260 ГГц с эквивалентной изотропно излучаемой мощностью (EIRP, Equivalent Isotropically Radiated Power) 5 дБм для беспроводной связи между кристаллами. Общая потребляемая мощность цепи передатчика составляет 688 мВт, в то время как цепь приемника потребляет 465 мВт. Также известны варианты использования передовых CMOS-технологий для внутри- и межкристальной связи в терагерцовом диапазоне, где продемонстрирован полностью интегрированный ООК-приемопередатчик, работающий на частоте 210 ГГц с потребляемой мощностью 421 мВт.

Также реализована улучшенная техника компоновки для минимизации внутренних паразитных наводок. Были продемонстрированы различные критически важные функции схем для микросистем следующего терагерцового поколения, такие как фазовая блокировка, управление лучом и генерация коротких импульсов в кремнии. Было экспериментально продемонстрировано управление

лучом на частоте выше 300 ГГц в 65-нм CMOS технологии с излучаемой мощностью 0,8 мВт и максимальным углом управления 50°. Технологии кремний-германия (SiGe) обладают превосходными высокочастотными характеристиками и возможностью интеграции с CMOS-технологией и поэтому являются привлекательными для реализации терагерцовых компонентов на кристалле. На частоте 240 ГГц продемонстрирован модулятор ООК с энергопотреблением 30 мВт, выполненный по 130-нм технологии SiGe BiCMOS. Были измерены выходная мощность 6 дБм и скорость передачи данных 13,3 Гбит/с. Также продемонстрирован полностью интегрированный приемопередатчик с использованием 130-нм технологии SiGe BiCMOS. Приемопередатчик работает на расстоянии 10 см с измеренной EIRP –11 дБм и имеет общую потребляемую мощность 380 мВт.

Недавно был продемонстрирован новый метод генерации терагерцового излучения с использованием технологий Si и SiGe. Показано, что он позволяет излучать высокие уровни мощности при частоте сигнала, которая на текущий момент времени не может быть достигнута при использовании обычных источников и фазированных решеток. В методике исследователи реализовали 2D фазированную решетку, работающую на частоте 338 ГГц с рассеиваемой мощностью 1,54 Вт с использованием 65 нм CMOS-процесса. Эта уникальная технология использует коллективную производительность большого числа синхронизированных или когерентных источников для достижения недостижимых ранее характеристик фазового шума (–93 дБс/Гц). Поскольку основная идея основана на генераторах с задержкой, она может быть распространена на другие полупроводниковые технологии для достижения еще более высокого уровня мощности на более высоких рабочих частотах. Например, технологии III-V, такие как арсенид галлия (GaAs), арсенид галлия-индия (InGaAs), нитрид галлия (GaN) и фосфид индия (InP), имеют более высокую подвижность электронов и, следовательно, демонстрируют более высокие частоты. Таким образом, технологии составных полупроводников могут обеспечить более высокую рабочую частоту и достаточную выходную мощность, что необходимо для терагерцовых компонентов.

Также были созданы микроволновые монолитные интегральные схемы (ММИС, Monolithic Microwave Integrated Circuit) с использованием 50-нм составных InGaAs приборов для достижения беспроводной

передачи данных со скоростью 25 Гбит/с на частоте 220 ГГц, и было сообщено об излучаемой мощности от $-3,4$ до $-1,4$ дБм. Технология GaN стала предпочтительной технологией для реализации высокочастотных усилителей мощности для различных коммерческих и других продуктов на высоких частотах. Технология имеет большое напряжение пробоя материала (3 мВ/см) и высокую подвижность электронов, что позволяет GaN достичь высокого коэффициента полезного действия. Прогресс, достигнутый за последнее десятилетие в выходной мощности и эффективности добавленной мощности GaN-транзисторов с высокой подвижностью электронов (HEMT, High Electron Mobility Transistors), показан на рис. 14.6. Монолитные интегральные усилители мощности на основе GaN, работающие в диапазоне 75–100 ГГц с выходной мощностью от 350 до 840 мВт и с эффективностью добавления мощности более 15 % продемонстрированы с использованием GaN HEMT с длиной канала 120 нм.

На частотах выше 200 ГГц зарегистрированы усилители мощности на основе InP с плотностью мощности в диапазоне 25–100 мВт/мм для HEMT. В частности, усилитель на основе монолитной интегральной схемы (MMIC, Monolithic Microwave Integrated Circuit) для терагерцового диапазона с использованием технологии InP способен обеспечить пиковую выходную мощность 3 мВт на частоте 650 ТГц. Эффективность GaN HEMTs по добавленной мощности сопоставима с эффективностью с InP HEMT в W-диапазоне в пределах 15–20 %.

Наноплазмонные технологии основаны на управлении потоком электромагнитного излучения для масштабов меньше длины волны света, преодолевая тем самым дифракционный предел. Благодаря очень высокой скорости распространения электромагнитного излучения в веществе можно создать терагерцовые устройства, работающие на частоте выше 1 ТГц. Например, двумерный углеродный материал графен поддерживает распространение поверхностных плазмонных поляритонов (SPP, Surface Plasmon Polaritons) в широком спектре частот от микроволнового до инфракрасного диапазона. SPP – это связанные электронно-световые колебания на границе диэлектрика и металла, которые могут распространяться со скоростью света.

В отличие от металлов SPP в графене электрически и химически управляем. В графене длина распространения плазмонов может составлять несколько микрометров, а скорость их распространения имеет нижнюю границу $vf/2$, где $vf = 8105$ м/с – скорость Ферми

для фермионов Дирака в графене. Длительное время жизни плазмонов и их очень высокая скорость распространения делают графен идеальной платформой для реализации плазмонных волноводов для систем на кристалле и сверхширокополосных антенн для беспроводной связи. Более того, такие наноантенны и приемопередатчики с графеновыми SPP обеспечивают уменьшение размеров в сотни раз по сравнению с обычными микрополосковыми антеннами, сохраняя или даже превосходя их по полосе пропускания и коэффициенту усиления. Уже получены экспериментальные результаты по возбуждению и распространению SPP в графеновых наноструктурах на различных диэлектрических материалах, таких как SiO_2 , нитрид бора и алмазоподобный углерод. Однако отметим, что проблемы, связанные с электрическим возбуждением SPP в графене при комнатной температуре, должны быть решены до того, как наноплазмонные технологии могут быть использованы на практике.

Проблемы интеграции

За последнее десятилетие интеграция полупроводниковых технологий активно исследовалась в основном для разработки оптических устройств в сочетании с цифровыми схемами CMOS. В настоящее время разработано несколько методов интеграции соединений III-V с кремнием. Например, техника захвата соотношения сторон (ART, Aspect Ratio) использует переходы с высоким соотношением сторон для захвата нитевидных дислокаций из латентно несовпадающего материала, что позволяет получить высококачественные слои. Другой подход предполагает перенос слоя III-V на кремний, покрытый тонким диэлектриком. Этот подход похож на технологию «кремний на изоляторе», которая хорошо известна в промышленности. С помощью композитного буферного слоя InP/InAlAs толщиной 1 нм были изготовлены транзисторы InGaAs на кремнии. Ожидается, что многообещающие результаты совместной интеграции разнородных технологий будут полезны для предлагаемой беспроводной внутричиповой связи следующего поколения.

Все графеновые терагерцовые компоненты могут быть объединены с массивами терагерцовых антенн, также реализованных в графене, для достижения динамического формирования луча. Кроме того, четная симметрия ветвей переноса электронов и дырок в графене позволяет эффективно и компактно генерировать четные

гармоники. Влияние паразитных свойств на уровне устройства на эффективность умножения частоты в графеновых полевых транзисторах было теоретически исследовано. Используя графеновые полевые транзисторы с антенной, исследователи продемонстрировали терагерцовые детекторы при комнатной температуре на частоте 0,3 ТГц. Графен был также успешно использован для создания, перестраиваемого магнитным полем терагерцового детектора – ИК излучения с рабочей частотой от 0,7 ТГц до 33 ТГц. Несмотря на то, что графен является относительно новой технологией, его совместимость с кремниевой CMOS-технологией и уникальная двумерная физика делают его жизнеспособной технологией для устройств беспроводной связи следующего поколения.

Беспроводная 3D-архитектура x86 CPU

Анализируя тенденции развития архитектуры процессоров, можно заметить существенный разрыв между академическими и промышленными усилиями. Пытаясь минимизировать время выхода на рынок, основные производители консервативны и, скорее всего, продолжают дальнейшее развитие архитектуры x86 с иерархическими кэшами, используемыми для совместного обмена данными внутри процессора. Вместе с тем академические и долгосрочные промышленные проекты активно продвигают дизайн WNoC с прямой связью между ядрами как реальное решение для создания процессоров с сотнями вычислительных элементов. Мы рассмотрим гибридную архитектуру процессора с беспроводным интерфейсом *ядро–кэш*, которая, с одной стороны, полностью совместима с существующей архитектурой, включая механизмы когерентности кэша, и не требует кардинальных изменений в процессе изготовления процессора и разработки программного обеспечения нижнего уровня, а с другой стороны, масштабируется до нескольких десятков или даже сотен вычислительных элементов на одном кристалле. Такая эволюционная архитектура может побудить академические и промышленные круги инвестировать больше ресурсов в эту область и создать действующий прототип в течение ближайших нескольких лет.

Текущая и предлагаемая архитектура процессора показаны на рис. 14.7. В этой архитектуре можно разделить ядра и общий кэш третьего уровня (LLC, Last Level Cache) в рассматриваемом случае L3 на два слоя, расположенных один над другим, и заменить

проводную шину связи, соединяющую основные компоненты существующего процессора, беспроводным сверхширокополосным каналом связи. Такой подход позволяет избавиться от топологии шины между ядрами и общим кэшем, которая вскоре может стать узким местом, препятствующим увеличению числа вычислительных элементов. В то же время пространство на нижнем уровне освобождается для хранения большего объема кэш-памяти L3, которая сегодня занимает около 40% площади кристалла и должна масштабироваться соответствующим образом с увеличением числа ядер. Таким образом, даже с одним дополнительным слоем объем кэш-памяти L3 может быть удвоен. Обратим внимание, что рассматриваемая архитектура не меняет способ организации кэш-памяти, позволяя повторно использовать существующие протоколы когерентности кэша.

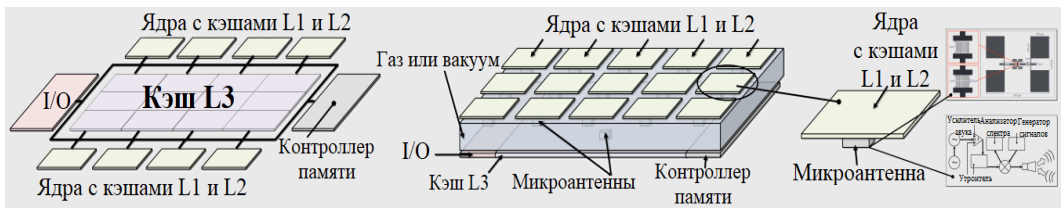


Рис. 14.7. Текущая и перспективная архитектуры x86

В качестве технологии для организации передачи трафика между ядрами и кэшем предлагается беспроводная связь в терагерцовом диапазоне как наиболее перспективная технология, позволяющая создавать антенны размером в сотни микрон и интегрированную приемопередающую электронику. Для уменьшения потерь на распространение во внутреннем пространстве между двумя чипами предлагается сделать поверхность чипов из мягкого металла с высоким коэффициентом отражения, например меди, а внутреннее пространство заполнить инертным газом, не имеющим линий поглощения в терагерцовом диапазоне частот. Использование хороших теплопроводников, таких как инертный газ, позволяет использовать обычные вентиляторы или охлаждающие трубки, расположенные в верхней части процессора. Предполагается, что терагерцовая технология будет достаточной для внутрочипового трафика между ядрами и общим кэшем L3. В дальнейшем утверждение подтверждается путем оценки применимости терагерцового диапазона для предлагаемой архитектуры.

Производительность терагерцового канала

Ожидается, что условия распространения в рассматриваемой системе связи будут резко отличаться от распространения в свободном пространстве. В частности, многочисленные отражения и рассеяния со всех сторон не только влияют на количество принимаемой энергии, но и вызывают искажения принимаемого сигнала в результате межсимвольной интерференции. Известные модели многолучевых каналов в основном являются крупномасштабными и характеризуют работу канала на больших расстояниях разделения по сравнению с внутричиповыми коммуникациями. Вместе с тем существующие популярные коммерческие системы многолучевого моделирования не учитывают специфику терагерцового диапазона, в том числе молекулярное поглощение. Поэтому была применена оригинальная система многолучевого моделирования, специально разработанная для терагерцовых частот.

Инструмент разработан на языках программирования C++ и Python, где первый используется в большинстве модулей для обеспечения эффективности вычислений, а второй – для сбора и обработки полученных результатов. Инструмент использует методологию трассировки лучей и основан на тесселяции поверхности на миниатюрные сегменты. Размер каждого сегмента сопоставим с длиной волны, поэтому сегменты можно рассматривать как точечные приемопередатчики, принимающие часть энергии от передатчика и отражающие/рассеивающие часть ее на приемник или другой точечный приемопередатчик на другой поверхности. Такой подход позволяет оценить все возможные пути отражения и рассеяния без необходимости дополнительной фильтрации на этапе постобработки. Кроме того, уникальной особенностью инструмента является возможность учитывать рассеяние порядка более двух за приемлемое время, что обычно нехарактерно для современных коммерческих решений.

Интересующая среда распространения представлена параллелепипедом размером $20 \times 20 \times 2$ мм с целевым передатчиком, представляющим приемопередатчик интерфейса кэш-памяти L3 и расположенным в центре верхней поверхности, как показано на рис. 14.8. Рассмотрены три варианта целевого приемника: «центр», «сторона» и «угол», представляющие потенциальные места расположения приемопередатчиков интерфейса кэш-памяти L2. Рассмотрен канал с

полосой пропускания в 1 ТГц, 0,5–1,5 ТГц. В связи с высокой сложностью перспективного оборудования для связи в терагерцовом диапазоне была использована простая модуляция ООК (см. главу 12).

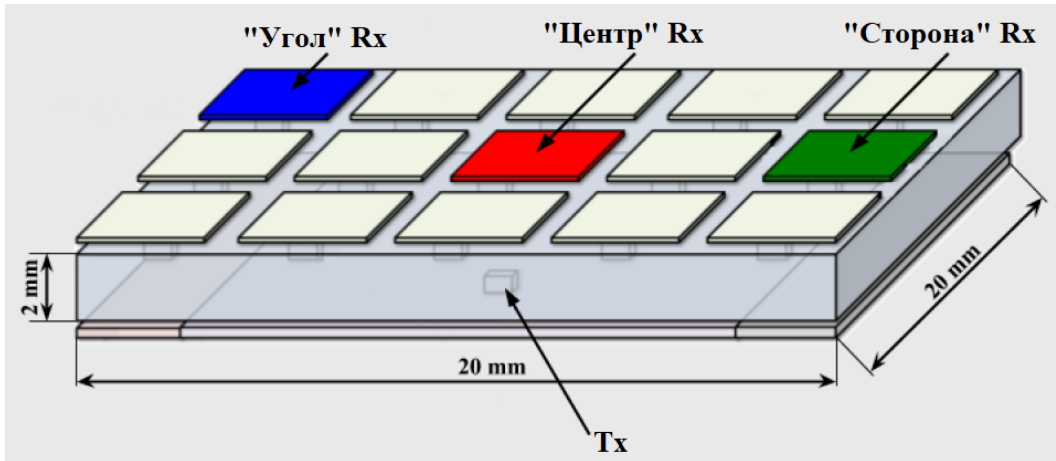


Рис. 14.8. Среда распространения

Начнем анализ с представления импульсного отклика канала на рис. 14.9, где точки представляют затухание компонентов, поступающих на приемники. Ожидаемо, что вариант «центр» характеризуется наилучшими условиями канала, так как он имеет четкий компонент прямой видимости с затуханием около 30 дБ. Два других варианта, «сторона» и «угол», имеют схожее ослабление компонента прямой видимости, хотя и немного хуже для «угла» из-за большего расстояния. В то же время «угловое» расположение в целом является наихудшим из-за наличия множества вторичных компонентов, отраженных и рассеянных от стенок.

Применим полученные характеристики канала для анализа формы терагерцового импульса на приемниках. На рис. 14.10 представлена амплитуда принятого сигнала как функция времени для «центрального», «бокового» и «углового» вариантов. Сигнал, соответствующий «угловому» приемнику, является наиболее ослабленным, в то время как ослабление двух других положений приемника практически одинаково. Как можно заметить, «угловой» приемник испытывает худшие условия распространения, так как отраженные и рассеянные компоненты остаются заметными даже спустя 100 пс после приема основного компонента.

Теперь используем полученные сигналы для оценки BER для схемы модуляции ООК. Поскольку многолучевые отражения сильно коррелированы с передаваемым сигналом, вторичные компоненты могут представлять собой шум или использоваться как полезный сигнал. Поэтому традиционный подход к оценке BER как функции выбранной схемы модуляции и среднего отношения ОСШ не может быть применен для рассматриваемого канала. Ниже оценен BER как функция передачи символов, используя подход, рассмотренный в параграфе 14.2, т.е. путем анализа всех возможных комбинаций передаваемых символов канала. Рассматриваются 10 предыдущих символов, что является разумным компромиссом между точностью и вычислительной сложностью. Для заданной комбинации бит, например «0111010101», и частоты повторения t строится результирующий сигнал на приемнике и оценивается вероятность его ложного детектирования.

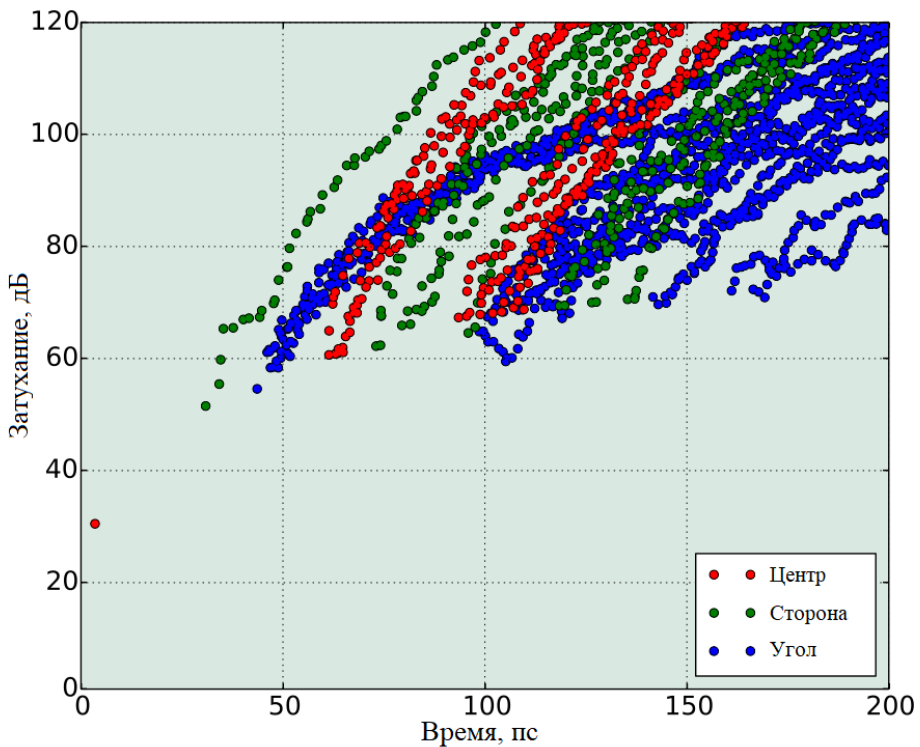


Рис. 14.9. Импульсный отклик канала

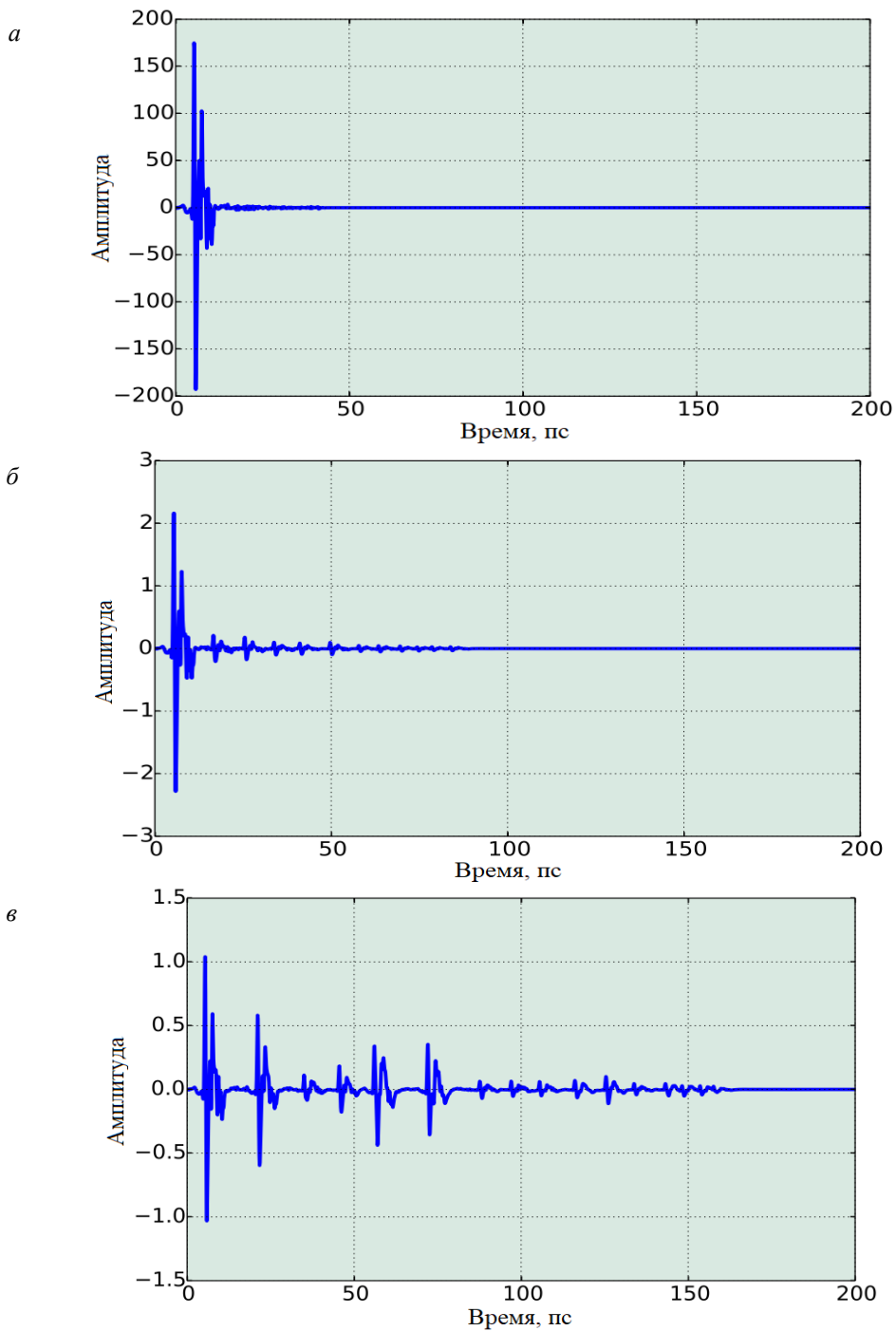


Рис. 14.10. Отклик канала во временной области:
a – ядро, расположенное по центру; *б* – ядро, расположенное на стороне;
в – угловое расположение ядра

На рис. 14.11а представлены результаты, где BER показан как функция межсимвольного расстояния, а мгновенное значение BER сопровождается приближением. Ожидается, что наилучший BER по всем рассматриваемым межсимвольным интервалам наблюдается для позиции «центр». Однако относительная разница между «боковым» и «угловым» положениями гораздо меньше, чем разница между «центральной» и «боковой» вариантами. Этот эффект объясняется тем, что на BER для позиции «угол» сильно влияют отраженные и рассеянные компоненты. Тем не менее BER для позиции «угол» становится меньше 10^{-9} после 36 пс.

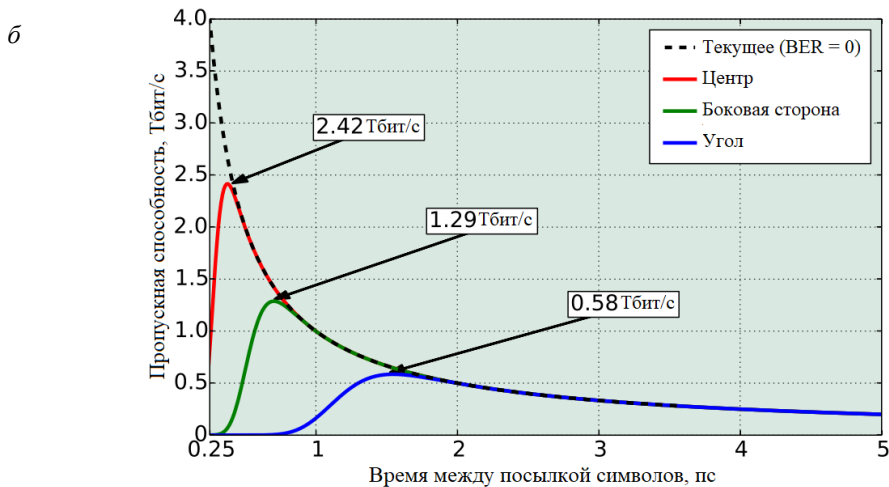
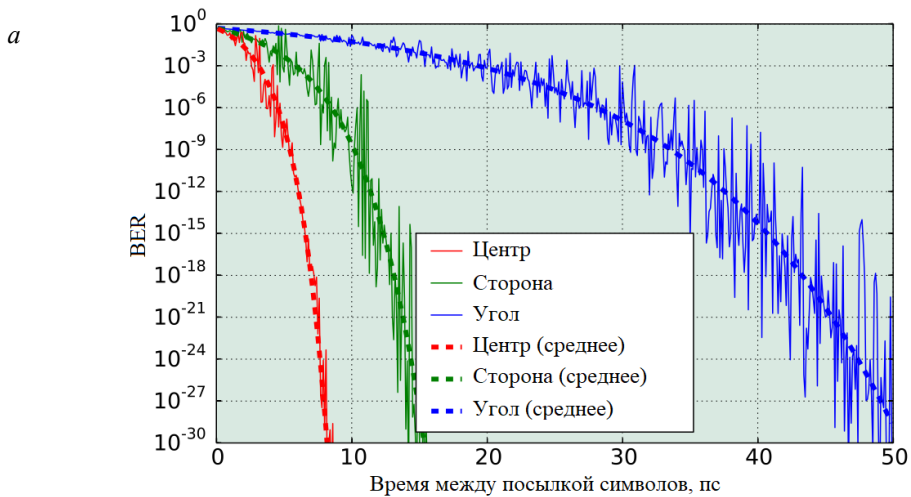


Рис. 14.11. Характеристики канала связи:
 а – вероятность ошибки; б – достижимая пропускная способность

Завершаем анализ производительности физического уровня, представляя среднюю пропускную способность канального уровня как функцию внутрисимвольной длительности: $T = [1 - BER(\tau)]^l /$, где τ – межсимвольная длительность, а l соответствует длине блока, установленной в данном случае на 64 бита. Результаты показаны на рис. 14.11б. Как можно заметить, пропускная способность начинается с нуля для всех рассматриваемых позиций, так как значение $BER \approx 10^{-3}$ близко к 1. С увеличением межсимвольной длительности пропускная способность растет, так как снижение пропускной способности канала компенсируется более быстрым снижением BER. Однако после определенного момента, примерно соответствующего BER, пропускная способность достигает максимального значения и начинает снижаться до нуля, поскольку BER уже пренебрежимо мал и внутрисимвольная длительность теперь доминирует в компромиссе.

Анализ показывает, что пропускная способность канала связи между кэш-памятью L2 и L3 варьируется от 580 Гбит/с до 2422 Гбит/с, в зависимости от расположения приемопередатчика кэш-памяти L2. В дальнейшем будем использовать первое значение в качестве консервативной оценки пропускной способности уровня канала внутричиповой терагерцовой связи.

Характеристики трафика на интерфейсе L2-L3

Для того чтобы сделать выводы о максимальном числе подключаемых ядер, оценив производительность MAC-архитектуры, необходимо определить характеристики входного трафика и интерфейсов. В этом разделе сначала охарактеризуем объем трафика с помощью прямых измерений, а затем характер основного трафика. Наконец, представим результаты бенчмаркинга интерфейсов, предоставляя характеристики задержки.

Закрытый характер разработки процессоров x86 не позволяет полностью положиться на один подход для выявления свойств внутрипроцессного трафика и его точных моделей. Необходимо рассмотреть все потенциальные подходы, чтобы сформулировать общее понимание, на которое можно опираться, чтобы прийти к детальному описанию свойств трафика.

Анализ на уровне микроархитектуры с использованием общедоступной документации обеспечивает первый шаг к модели трафика. Фиксация набора алгоритмов и архитектурных решений помогает

определить инструменты, необходимые на последующих этапах. Во-первых, анализ функциональности процессора, и в частности протокола когерентности кэша, позволяет понять влияние различных подсистем и принять решение об уровне детализации имитационных моделей. Моделирование на уровне системы с точностью до цикла, выполненное с учетом всех основных механизмов, реализованных в современных ЦП, позволяет понять мелкомасштабный характер трафика на различных внутрипроцессорных интерфейсах. Абсолютные значения трафика, полученные с помощью этого подхода, могут, однако, отличаться от реальности из-за абстракций моделирования и нераскрытых «ноу-хау» в реализации. Реальные измерения добавляют понимание точных значений объема трафика на внутренних интерфейсах. Измерения также необходимы для определения задержек между различными уровнями кэша и параметризации имитаций. Используя знания, предоставляемые всеми тремя подходами, можно составить подробную модель трафика для современных процессоров.

Объем трафика

Измерения трафика L2-L3 проведены с помощью программного обеспечения Intel PCM. В качестве тестовой системы использованы восьмиядерный процессор Intel Core i7-5960X с архитектурой Haswell и 20 МБ кэша L3. Использование восьмиядерного процессора облегчает процесс калибровки для экстраполяции трафика. PCM позволяет считывать значения встроенных счетчиков в процессорах Intel, включая количество промахов кэша L2. Поскольку за каждым промахом L2 следует запрос на чтение кэша L3, преобразуем измеренные значения в расчетный трафик на интерфейсе L2-L3 с помощью функции $C = L2M * LD / t$, где L2M – количество промахов L2 за раунд измерения, LD = 64 байта – размер кэш-линии, а t – время измерения, равное 10 минутам.

Для тестирования выбраны типичные приложения: многопользовательская игра, шифрование данных с помощью AES-128 и декодирование видео H.264. Для имитации максимальной и минимальной возможных нагрузок разработаны два синтетических теста, выполняющих чтение данных из бесконечного массива, расположенного в оперативной памяти (рис. 14.12).

Тест «1В» эмулирует «хороший стиль программирования», когда данные считываются последовательно байт за байтом. В этом случае процессор использует иерархию кэширования, где данные хранятся в кэш-линиях по 64 байта. Тест «64В» считывает каждый 65-й байт, делая кэширование данных неэффективным и генерируя гораздо больший трафик на интерфейсе L2-L3. Этот тест считается «плохим стилем программирования».

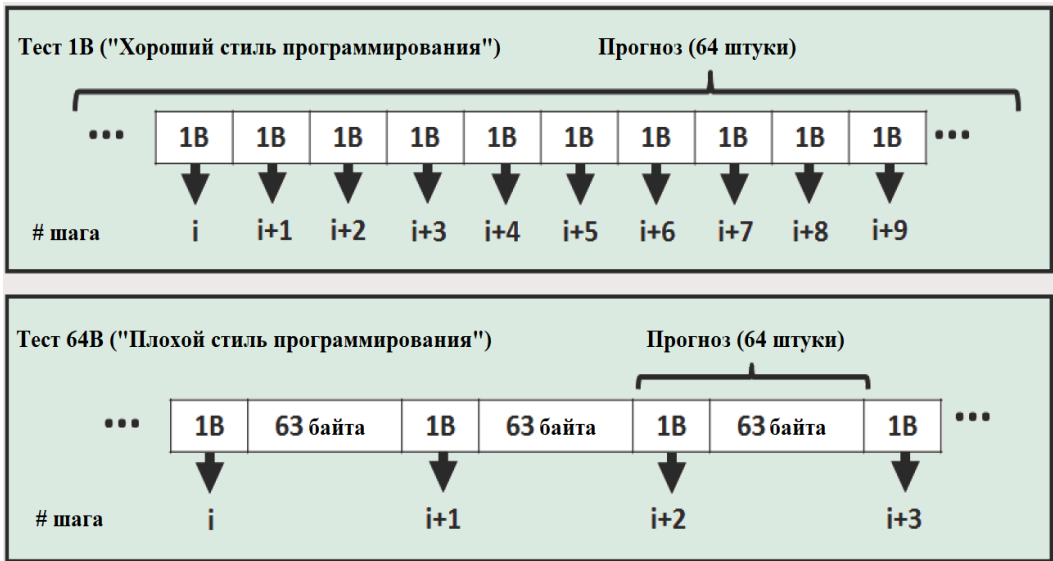


Рис. 14.12. «Хороший» и «плохой» стили программирования

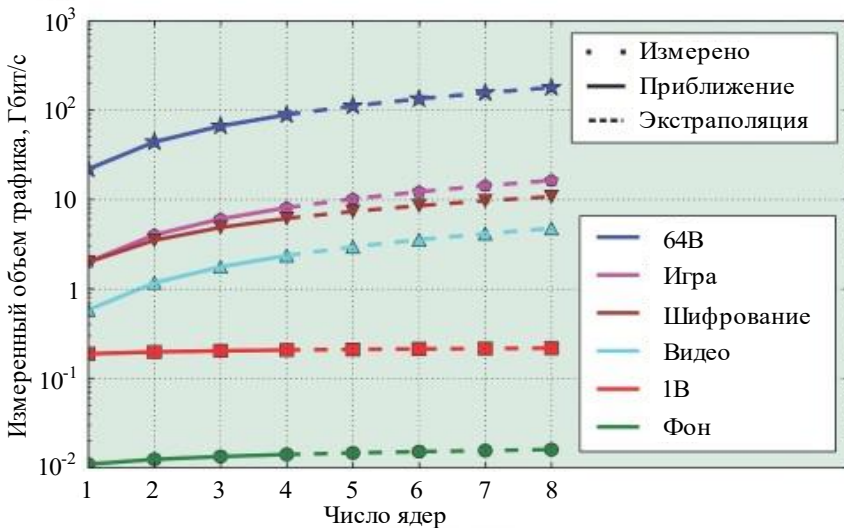


Рис. 14.13. Измеренные значения трафика на интерфейсе L2-L3

На рис. 14.13 показаны оценки трафика на интерфейсе L2-L3. Как и ожидалось, тест «64В» приводит к наибольшей нагрузке. Используются измерения для 1, 2, 4 ядер для построения экстраполяции в виде AnB для 8 ядер, где A и B – константы и n – число ядер. Эти кривые показаны пунктирными линиями. Как можно заметить, рассчитанный трафик незначительно отклоняется от эмпирических измерений для 8 ядер, что означает достаточную уверенность в расчетных результатах. Далее приведены экстраполяция данных на массивный многоядерный сценарий и сравнение с достижимой пропускной способностью беспроводного интерфейса.

Стохастические свойства трафика

Производительность системы, особенно характеристики MAC-уровня, чувствительны не только к средней нагрузке трафика, но и стохастическим характеристикам потока. К сожалению, ни анализ на уровне микроархитектуры, ни прямые измерения не способны предоставить детальную структуру трафика на транзакционном уровне, поскольку оба вносят погрешности, сравнимые с измеренным значением. Одним из подходов, позволяющим получить достаточно высокое разрешение, является имитация процессора с точностью до цикла. Сегодня существует ряд имитаторов, поддерживающих архитектуру x86, MARSS, Gem5, zSim, SST. Все они являются гибкими инструментами, позволяющими детально фиксировать время событий, что делает их подходящими для нашей задачи. Имитаторы zSim и SST предназначены для моделирования очень больших систем, включающих сотни ядер, и, по сравнению с MARSS и Gem5, не имеют детальной функциональности управления.

В рассматриваемом исследовании типичная архитектура Intel x86 реализована в Gem5, включая все основные возможности и компоненты архитектуры Intel. Выбранный размер кэша и другие параметры типичны для современных процессоров общего назначения, однако незначительные отклонения от реальных значений конкретных систем не должны качественно изменить полученный временной ряд. Подсистема кэша предполагалась инклюзивной с размером/латентностью 64 КБ/2нс, 2 МБ/12нс и 20 МБ/30нс в кэшах L1, L2 и L3. Модель учитывает задержки, связанные с поиском информации, и эмулирует возможность конвейеризации. Смоделированы системы с 1, 2, 3, 4, 8 и 16 ядрами, где тактовая частота была

установлена на 3,0 ГГц. Заметим, что тактовая частота дает только количественный эффект и полученные результаты могут быть масштабированы практически на любую рабочую частоту. В многоядерных конфигурациях количество одновременно выполняемых тестов устанавливалось равным количеству работающих ядер. В целом было выполнено более 70 тестов.

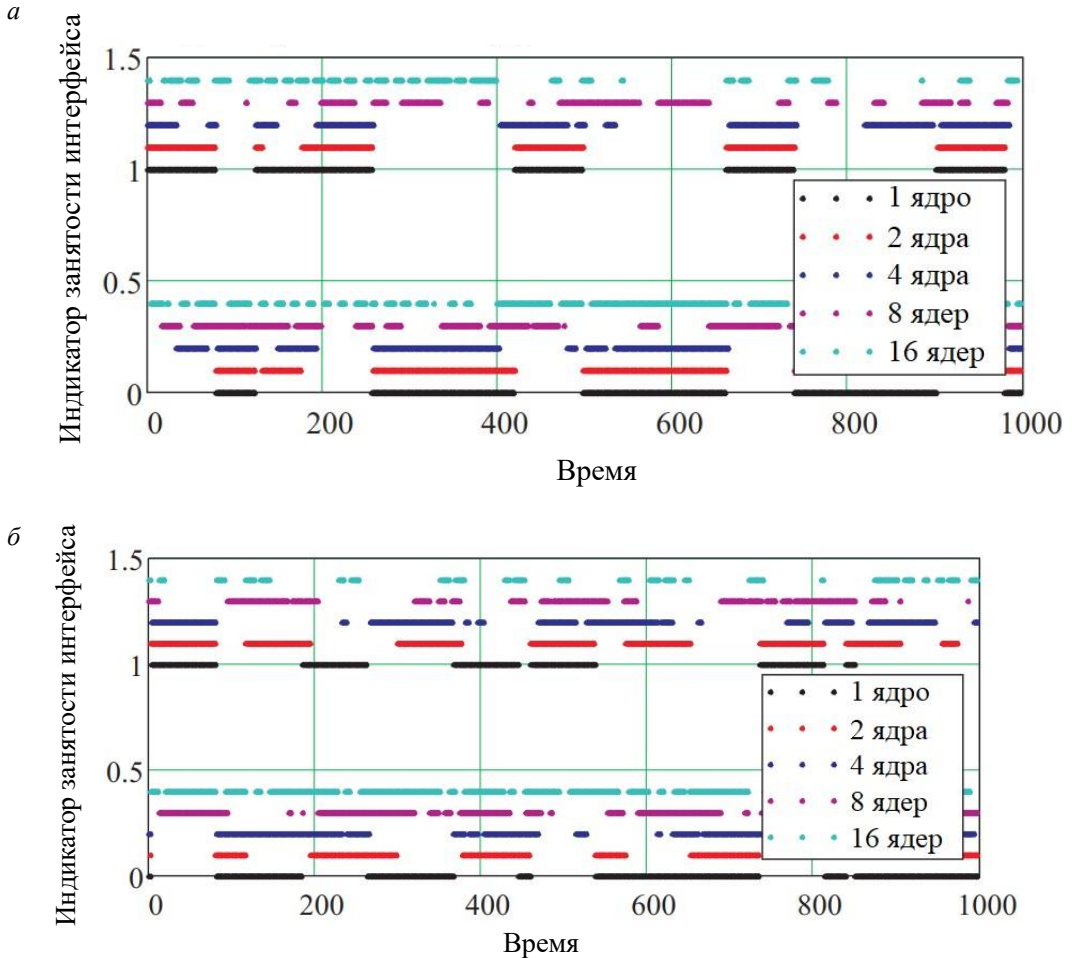
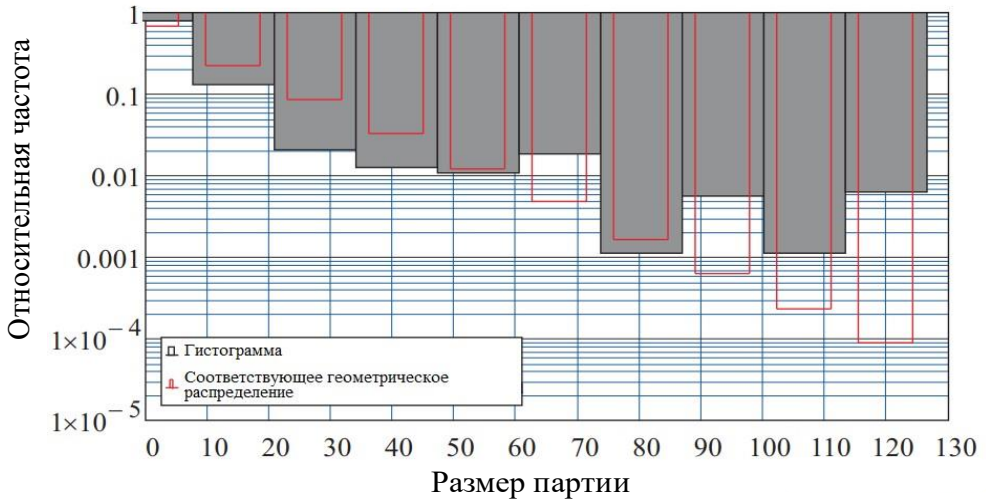


Рис. 14.14. Стохастическая структура трафика на интерфейсе L2-L3:
a – тест «1В»; *б* – тест «шифрование»

Для моделирования типичной нагрузки выбраны два теста, включая простую процедуру чтения (тест «1В») и более комплексный тест с шифрованием AES, включающий деление и умножение. На рис. 14.14 показаны временные ряды трафика на интерфейсе L2-L3 для двух тестов, «запись» и «шифрование», для различных вариантов

числа ядер с индикаторами занятости интерфейса $I_A + b$, где A – событие занятости интерфейса, b – константа, добавленная для того, чтобы различать данные для разного числа ядер. Анализируя представленные результаты, можем сделать два качественных вывода: а) трафик на обоих интерфейсах имеет стохастическую структуру; б) трафик имеет четко идентифицируемые партии и промежутки между ними.

а



б

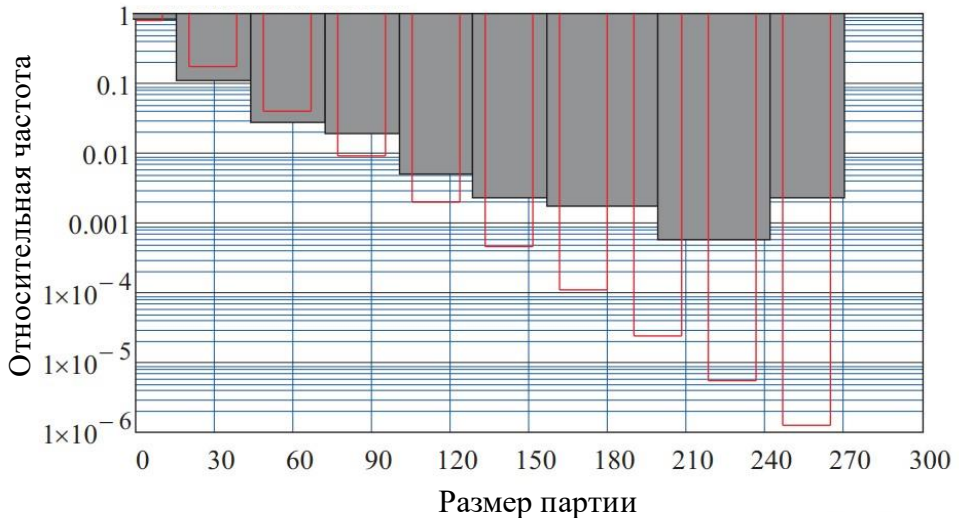


Рис. 14.15. Гистограммы частот размера партии и интервала между партиями:

а – тест «шифрование», размеры партий;

б – тест «шифрование», длительность интервала между партиями

На рис. 14.15 представлены гистограммы относительных частот для пакетных и межпакетных временных интервалов, в качестве примера взят тест «Шифрование» и 16 ядер. Как можно заметить, для 16 ядер гистограммы имеют четко выраженное геометрически убывающее поведение. Соответствующие приближения показывают, что геометрическое распределение может обеспечить высокую точность как для размеров партии, так и межпакетных интервалов. В результате можно исследовать производительность интерфейса разделяемого кэша в перспективных системах, используя процесс поступления вида Geo, т.е. геометрически распределенные интервалы между партиями. Эти результаты могут быть использованы для изучения характеристик MAC-уровня системы для реалистичной схемы трафика.

Параметризация задержки

Для завершения параметризации модели для оценки производительности MAC необходимо определить допустимую задержку на интерфейсе L2-L3. Поскольку инструмент Intel PCM не может дать точных оценок задержки на определенном интерфейсе, разработана отдельная методика для анализа задержки, использующая счетчики производительности, встроенные в процессоры Intel, но основанная на последовательных «зацикленных» обращениях к памяти, сопровождаемых мелкозернистыми измерениями времени. В том числе измерения задержек выполняются путем прохода по структуре связанного списка, а пропуски кэша достигаются с помощью внесения дополнительных смещений.

Концептуально схожие подходы были описаны ранее в литературе. Однако в предложенном инструменте было реализовано несколько важных усовершенствований, направленных на повышение точности данных. В частности, время измеряется с помощью встроенного высокоточного таймера. В каждом эксперименте применялись взаимодействия только с определенным уровнем кэша. Четыре программируемых счетчика, имеющихся в рассматриваемом процессоре, позволяют подсчитывать каждое попадание на уровни кэша L1, L2, L3. Несмотря на все меры, принятые для обеспечения корректности измерений, все же возможны случаи, когда количество попыток доступа не совпадает с количеством попаданий на нужный уровень

кэша. Это происходит из-за продвинутых алгоритмов предсказания кэша, реализованных в процессорах Intel: например, нужные данные были загружены из кэша нижнего уровня в кэш высокого уровня или часть рабочего массива была аннулирована, чтобы сохранить свободное место для новых данных, которые потенциально могут быть запрошены. Таким образом, при оценке статистики учитывались только те циклы измерений, в которых количество попыток доступа было равно количеству обращений.

Измерения задержки, как правило, очень чувствительны к любым фоновым процессам, запущенным в операционной системе. Поэтому для обеспечения точности результатов были предприняты определенные действия. Во-первых, влияние фоновых процессов в основном было устранено путем реализации программы тестирования как модуля загрузчика GNU GRUB 2.0. Поэтому тестирование начинается еще до загрузки операционной системы. Кроме того, тест разработан для работы в одном потоке, что позволило избежать неоднозначностей, вызванных переключением контекста. Наконец, влияние компилятора было смягчено за счет использования встроенного ассемблера, а язык C использовался только для ввода и вывода.

Результаты тестов задержки для операций чтения и записи представлены на рис. 14.16. Заметим, что измеренная задержка в основном вызвана механизмом поиска кэш-линий. Требования к обмену информацией между ядром и кэшем L3 чрезвычайно жесткие – 13 нс для чтения и 25 нс для записи. Время передачи одной кэш-линии размером 64 байта через канал 0,5–1,5 ТГц с длительностью импульса 800 фс составляет лишь 0,4 нс, поэтому дополнительная задержка, вызванная беспроводным интерфейсом (кодирование, распространение, декодирование), не должна оказывать существенного влияния на общий бюджет задержки, пока количество активных ядер невелико. Напротив, когда число активных ядер становится большим, протокол управления доступом к среде может резко повлиять на производительность задержки, поскольку время ожидания отправки фрагмента данных может быть значительным. В параграфе 14.3 проведен анализ масштабирования рассматриваемой системы вплоть до сценария массивной многоядерной системы и его влияния на величину задержки.

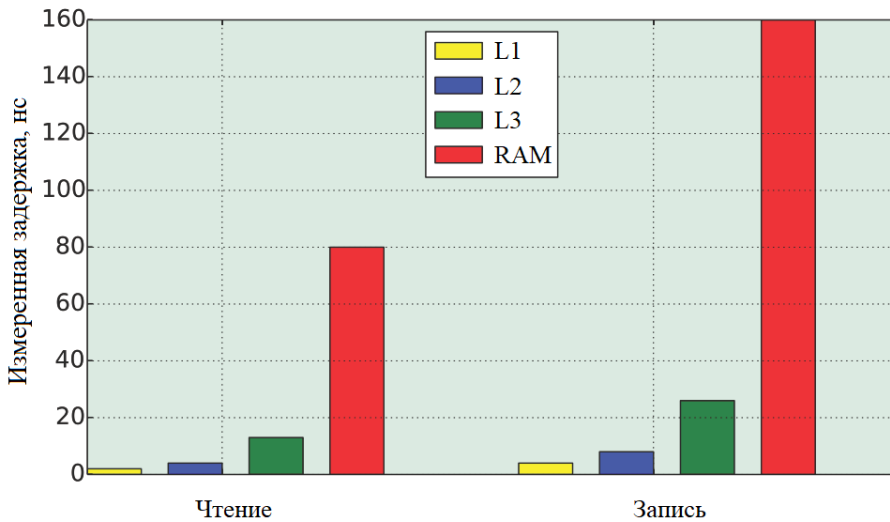


Рис. 14.16. Задержка на интерфейсе L2-L3

Масштабирование процессора

В данном параграфе используем полученные по каналам, трафику и задержкам результаты для исследования масштабируемости рассматриваемой архитектуры ЦП с внутричиповыми коммуникациями на терагерцовой частоте. Сделана попытка оценить максимальное число ядер, которое может быть поддержано с точки зрения границ пропускной способности и задержки.

Для анализа производительности на уровне системы необходимо определить конкретный протокол управления доступом к среде. Протоколы MAC со случайным доступом затруднительно использовать для рассматриваемого приложения из-за их хорошо известных ограничений производительности в условиях интенсивного трафика и необходимости чрезвычайно быстрой генерации случайных чисел для разрешения коллизий. Кроме того, специфика среды многоядерных процессоров с четко определенным числом взаимодействующих объектов делает ее в целом более подходящей для централизованного или гибридного протокола MAC.

Не ставя целью сравнение протоколов MAC для внутрикристальной проводной связи, предложены два упрощенных решения множественного доступа с временным разделением (TDMA, Time-Division Multiple Access) и схожей сигнализацией (рис. 14.17). Оба

решения реалистично предполагают достаточно точную синхронизацию времени между узлами. В предлагаемых решениях все время делится на кадры. Предполагается две категории трафика «uplink» (UL, восходящая линия связи от ядра к кэшу L3) и «downlink» (DL, нисходящая линия связи от кэша L3 к ядру). Основное различие между решениями заключается в подходе к мультиплексированию данных в канале. Первое решение предлагает так называемое потоково-ориентированное мультиплексирование, при котором каждый кадр сначала разделяется на два равных подкадра (один для восходящего канала, другой для нисходящего), а подкадр восходящего канала далее делится на N слотов, где N – число ядер. При втором так называемом ориентированном на ядро мультиплексировании кадр сначала делится на N подкадров, где N – число ядер. В свою очередь, каждый подкадр далее делится на два равных слота – один для восходящей линии связи, другой – для нисходящей.

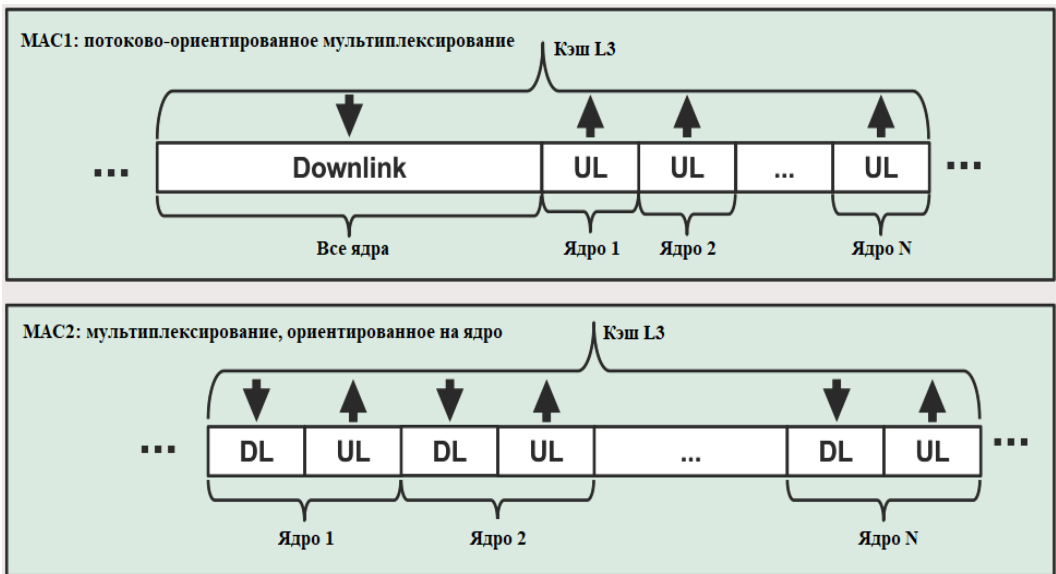


Рис. 14.17. Протоколы доступа к среде

Исследуется наихудший сценарий с максимальной нагрузкой, соответствующей тесту «64В». Длительность слота устанавливается равной времени, необходимому для передачи одной кэш-линии размером 64 байта в/из ядра. Длительность слота определяется отдельно для «центральной», «боковой» и «угловой» позиций.

Сначала проведем анализ масштабируемости процессора с точки зрения общей пропускной способности, обеспечиваемой на уровне МАС. На рис. 14.18 показан расчетный объем трафика в зависимости от количества поддерживаемых ядер для различных вариантов их расположения внутри процессора (582, 1287 и 2422 Гбит/с для «углового», «бокового» и «центрального» расположения соответственно). На основании полученных результатов можно прийти к выводу, что с точки зрения общей пропускной способности предложенная архитектура может быть масштабирована до 250 активных ядер для углового расположения. Поскольку в реальных массивных многоядерных процессорах только небольшая часть ядер будет иметь условия канала, аналогичные угловым, этот результат является пессимистической оценкой, подразумевающей, что реальное значение может быть намного больше.

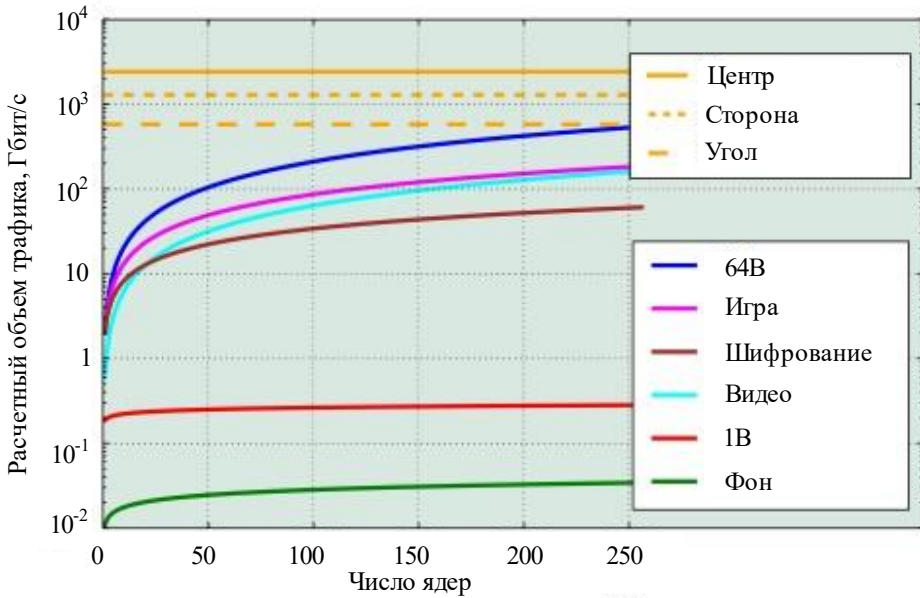


Рис. 14.18. Расчетный объем трафика как функция числа ядер

Оценка на основе пропускной способности может обеспечить оптимальную верхнюю границу числа поддерживаемых ядер, даже если все ядра находятся в пессимистичных «угловых» позициях. Причина в том, что с ростом числа активных ядер увеличивается общая длительность МАС-кадра, что приводит к дополнительной задержке доступа. Кроме того, стохастическая природа трафика может еще

больше повлиять на производительность обслуживания ядер. Рассмотрим результаты моделирования масштабирования процессора с учетом особенностей распространения в канале и процесса поступления трафика, а также деталей протоколов MAC.

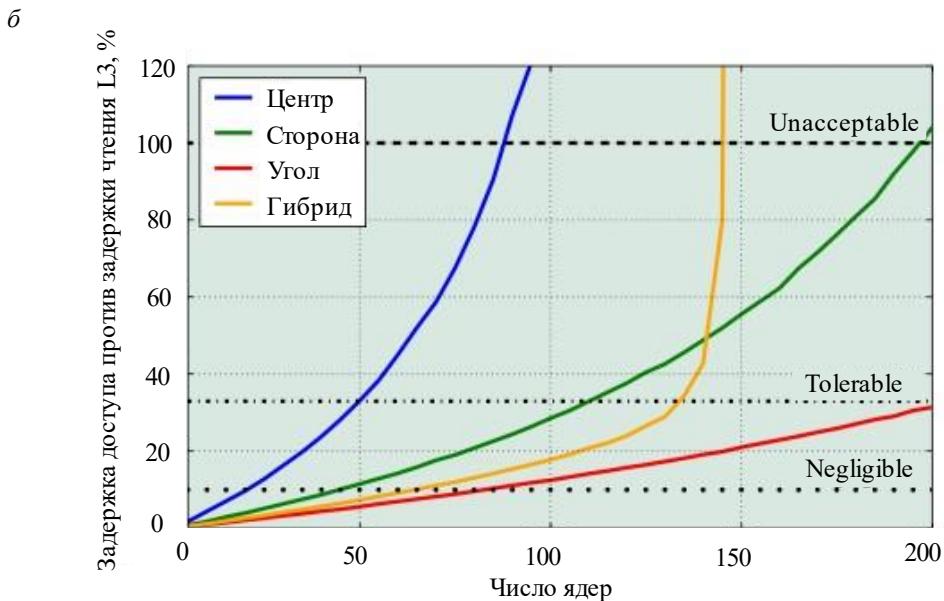
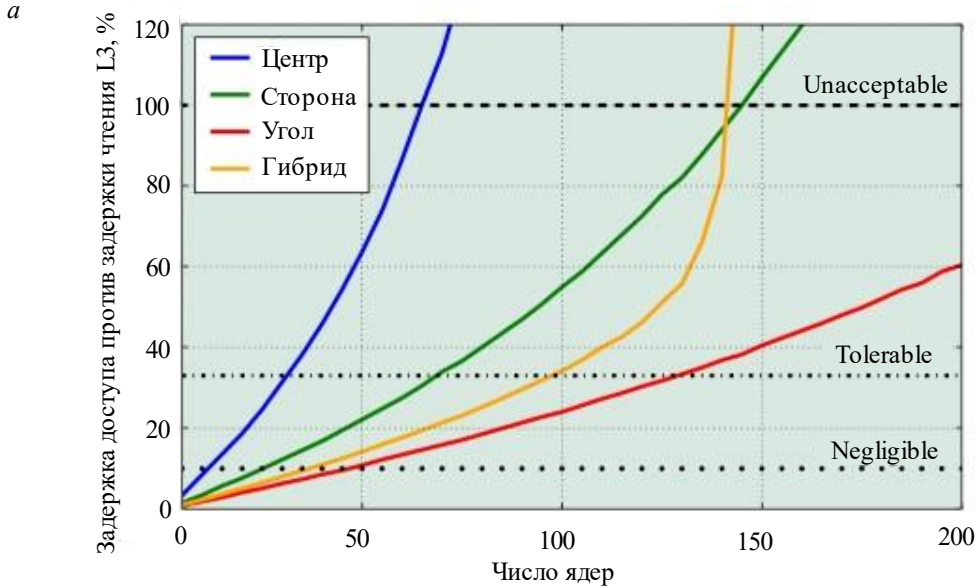


Рис. 14.19. Задержки доступа как функция числа ядер:
а – операции чтения; б – операции записи

На рис. 14.19 показана доля МАС в задержке беспроводного доступа по отношению к задержке доступа между кэшами второго и третьего уровней, оцененной ранее для операций чтения (нисходящий канал) и записи (восходящий канал). Сначала вводятся границы для задержки беспроводного доступа между кэшами второго и третьего уровней, установленные на 10, 33 и 100 % от задержки измеренной задержки для так называемых незначительных, терпимых и неприемлемых режимов соответственно. Причина такой классификации заключается в том, что дополнительный интерфейс не должен оказывать существенное влияние на общую величину задержки. Рассмотрены четыре случая: «центральный», «боковой» и «угловой», т.е. варианты размещения ядер, представленные ранее, а также сбалансированный «гибридный». В последнем случае система состоит из 4 «угловых» ядер, $4\lceil\sqrt{N}\rceil - 4$ «боковых» ядер, а остальные ядра находятся в «центре».

Анализируя полученные результаты, видим, что «угловое» развертывание не может масштабироваться для массивных многоядерных сценариев, так как задержка беспроводного канала становится больше, чем вся измеренная задержка для доступа L3 уже после 65 и 87 ядер для операций чтения/записи соответственно. Напротив, условия «центрального» канала позволяют системе масштабироваться до 130–200 ядер для операций чтения/записи без нарушения «допустимого» уровня. Важно отметить, что «гибридная» система хорошо масштабируется, достигая 101–134 ядер для операций чтения/записи при сохранении задержки ниже $1/3$ от задержки доступа L2-L3. Заметим, что пределы масштабируемости по задержке намного жестче, чем по пропускной способности (т.е. 101 ядро против 250 ядер), что означает, что требования к задержке должны учитываться в исследованиях масштабируемости процессора.

ЗАКЛЮЧЕНИЕ

Завершать работу над книгой всегда тяжелее, чем начинать. Несколько месяцев уходит на работу с издательством. Мы очень признательны нашему редактору за проявленное терпение к нашим капризам и правке текста, который Панкратова Ирина Леонидовна знает лучше нас, авторов. Без нее книга могла бы и не увидеть свет.

В заключение мы обращаемся к молодым ученым и специалистам, которых ждет исключительно интересное будущее. В телекоммуникациях будущее планируется, и его название уже известно – «Сети 2030». Так определено в рекомендациях Международного союза электросвязи, и пройдет всего около десяти лет до того, как человечество будет жить в новом цифровом мире, ядром которого будут эти сети. Этот мир пока не построен, сейчас он проходит начальный этап разработки и моделирования основных особенностей. Участвовать в этом процессе надо, будучи вооруженными знаниями и опытом моделирования сетей предыдущих поколений.

Надеемся, что наша книга поможет движению вперед – к сетям будущих поколений. Технологии меняются, но фундаментальные основы и методы решения проблем сохраняются на долгое время. Мы показали подходы к созданию системных и соответствующих им математических моделей и уверены, что на ближайшее десятилетие они сохранят свою актуальность. На сегодняшний день нам ясно, как изменились и были дополнены базовые теоретические дисциплины, которые теперь надо осваивать молодым специалистам. Эти изменения прежде всего коснулись теории телетрафика, теории массового обслуживания и методов стохастической геометрии. Совокупность новых методов и подходов мы называем теорией стохастического анализа мобильных сетей связи. Книга по теории ресурсных систем массового обслуживания уже написана нашими коллегами (URL: <https://link.springer.com/content/pdf/10.1007/978-3-030-83132-5.pdf>), а мы надеемся, что хватит сил написать еще одну книгу по итогам наших исследований сетей будущих поколений.

Жизнь продолжается!

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

2021 год

1. Begishev V., Sopin E., Moltchanov D., Kovalchukov R., Samuylov A., Andreev S., Koucheryavy Y., Samouylov K. Joint Use of Guard Capacity and Multiconnectivity for Improved Session Continuity in Millimeter-Wave 5G NR Systems // *IEEE Transactions on Vehicular Technology*. 2021. Vol. 70, no. 3. P. 2657-2672. <https://doi.org/10.1109/tvt.2021.3061906>.
2. Begishev V., Sopin E., Moltchanov D., Pirmagomedov R., Samuylov A., Andreev S., Koucheryavy Y., Samouylov K. Performance Analysis of Multi-Band Microwave and Millimeter-Wave Operation in 5G NR Systems // *IEEE Transactions on Wireless Communications*. 2021. Vol. 20, no. 6. P. 3475-3490. <https://doi.org/10.1109/twc.2021.3051027>.
3. 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. <https://doi.org/10.1109/TVT.2021.3113617>.
4. Dudin A., Dudina O., Dudin S., Samouylov K. Analysis of multi-server queue with self-sustained servers // *Mathematics*. 2021. Vol. 9, no. 17, article-no. 2134. P. 1–18. <https://doi.org/10.3390/math9172134>.
5. Dudin A., Dudina O., Dudin S., Samouylov K. Analysis of single-server multi-class queue with unreliable service, batch correlated arrivals, customers impatience, and dynamical change of priorities // *Mathematics*. 2021. Vol. 9 no. 11, article-no. 1257. <https://doi.org/10.3390/math9111257>.
6. Eckhard J., Petro V., Moltchano D., Koucheryav Y., Kurne T. Channel Measurements and Modeling for Low-Terahertz Band Vehicular Communications // *IEEE Journal on Selected Areas in Communications*. 2021. Vol. 39, no. 6. P. 1590–603. <https://doi.org/10.1109/jsac.2021.3071843>.
7. Gapeyenko M., Moltchanov D., Andreev S., Heath R.W. Line-of-Sight Probability for mmWave-based UAV Communications in 3D Urban Grid Deployments // *IEEE Transactions on Wireless Communications*. 2021. Vol. 20, no. 10. P. 6566–6579. <https://doi.org/10.1109/TWC.2021.3075099>.

8. Khakimov A., Mokrov E., Poluektov D., Samouylov K., Koucheryavy A. Evaluating the quality of experience performance metric for UAV-based networks // *Sensors*. 2021. Vol. 21, no. 17, article-no. 5689. <https://doi.org/10.3390/s21175689>.
9. Kokkonieniemi J., Jornet J., Petrov V., Koucheryavy Y., Juntti M. Channel Modeling and Performance Analysis of Airplane-Satellite Terahertz Band Communications // *IEEE Transactions on Vehicular Technology*. 2021. Vol. 70, no. 3. P. 2047-2061. <https://doi.org/10.1109/tvt.2021.3058581>.
10. Meshcheryakov Y., Melman A., Evsutin O., Morozov V., Koucheryavy Y. On Performance of PBFT Blockchain Consensus Algorithm for IoT-Applications With Constrained Devices // *IEEE Access*. 2021. Vol. 9. P. 80559–80570. <https://doi.org/10.1109/access.2021.3085405>.
11. Moltchanov D., Gaidamaka Y., Ostriкова D., Beschastnyi V., Koucheryavy Y., Samouylov K. Ergodic Outage and Capacity of Terahertz Systems Under Micromobility and Blockage Impairments // *IEEE Transactions on Wireless Communications*. 2021. <https://doi.org/10.1109/TWC.2021.3117583>, принята к публикации 15.10.2021.
12. Naumov V., Samouylov K. Resource system with losses in a random environment // *Mathematics*. 2021. Vol. 9, no. 21, article-no. 2685. P. 1–10. <https://doi.org/10.3390/math9212685>.
13. Pokorny J., Ma K., Saafi S., Frolka J., Villa J., Gerasimenko M., Koucheryavy Y., Hosek J. Prototype Design and Experimental Evaluation of Autonomous Collaborative Communication System for Emerging Maritime Use Cases // *Sensors*. 2021. Vol. 21, no. 11. P. 3871. <https://doi.org/10.3390/s21113871>.
14. Rinaldi F., Maattanen H., Torsner J., Pizzi S., Andreev S., Iera A., Koucheryavy Y., Araniti G. Broadcasting Services Over 5G NR Enabled Multi-Beam Non-Terrestrial Networks // *IEEE Transactions on Broadcasting*. 2021. Vol. 67, no. 1. P. 33–45. <https://doi.org/10.1109/tbc.2020.2991312>.
15. Rykova T., Göktepe B., Schierl T., Samouylov K., Hellge C. Analytical model and feedback predictor optimization for combined early-HARQ and HARQ // *Mathematics*. 2021. Vol. 9, no. 17, article-no. 2104. P. 1–19. <https://doi.org/10.3390/math9172104>.
16. Skibinska J., Burget R., Channa A., Popescu N., Koucheryavy Y. COVID-19 Diagnosis at Early Stage Based on

Smartwatches and Machine Learning Techniques // IEEE Access. 2021. Vol. 9. P. 119476–119491. <https://doi.org/10.1109/access.2021.3106255>.

17. Stepanov N., Moltchanov D., Begishev V., Turlikov A., Koucheryavy Y. Statistical Analysis and Modeling of User Micromobility for THz Cellular Communications // IEEE Transactions on Vehicular Technology. 2021. Vol. 71, no. 1. P. 725–738. <https://doi.org/10.1109/TVT.2021.3124870>.

18. Stusek M., Moltchanov D., Masek P., Mikhaylov K., Hosek J., Andreev S., Koucheryavy Y., Kustarev P., Zeman O., Roubicek M. LPWAN Coverage Assessment Planning Without Explicit Knowledge of Base Station Locations // IEEE Internet of Things Journal. 2022. Vol. 9, no. 6. P. 4031–4050. <https://doi.org/10.1109/jiot.2021.3102694>.

19. Yang M., Zhu H., Qian H., Koucheryavy Y., Samouylov K., Wang H. Peer Offloading with Delayed Feedback in Fog Networks // IEEE Internet of Things Journal. 2020. Vol. 8, no. 17. P. 13690–13702. <https://doi.org/10.1109/JIOT.2021.3067919>.

20. Yang M., Zhu H., Wang H., Koucheryavy Y., Samouylov K., Qian H. An Online Learning Approach to Computation Offloading in Dynamic Fog Networks // IEEE Internet of Things Journal. 2021. Vol. 8, no. 3. P. 1572–1584. <https://doi.org/10.1109/jiot.2020.3015522>.

2020 год

21. Gapeyenko M., Samuylov A., Gerasimenko M., Moltchanov D., Singh S., Akdeniz M., Aryafar E., Andreev S., Himayat N., Koucheryavy Y. Spatially-Consistent Human Body Blockage Modeling: A State Generation Procedure // IEEE Transactions on Mobile Computing. 2020. Vol. 19, no. 9. P. 2221–2233. <https://doi.org/10.1109/tmc.2019.2918318>.

22. Kavuri S., Moltchanov D., Ometov A., Andreev S., Koucheryavy Y. Performance Analysis of Onshore NB-IoT for Container Tracking During Near-the-Shore Vessel Navigation // IEEE Internet of Things Journal. 2020. Vol. 7, no. 4. P. 2928–2943. <https://doi.org/10.1109/jiot.2020.2964245>.

23. Naumov V., Gaidamaka Y., Samouylov K. Computing the Stationary Distribution of Queueing Systems with Random Resource Requirements via Fast Fourier Transform // Mathematics. 2020. Vol. 8, no. 5. P. 772. <https://doi.org/10.3390/math8050772>.

24. Petrov V., Gapeyenko M., Moltchanov D., Andreev S., Heath R. Hover or Perch: Comparing Capacity of Airborne and

Landed Millimeter-Wave UAV Cells // IEEE Wireless Communications Letters. 2020. Vol. 9, no. 12. P. 2059–2063. <https://doi.org/10.1109/lwc.2020.3012552>.

25. Petrov V., Moltchanov D., Koucheryavy Y., Jornet J. Capacity and Outage of Terahertz Communications With User Micro-Mobility and Beam Misalignment // IEEE Transactions on Vehicular Technology. 2020. Vol. 69, no. 6. P. 6822–6827. <https://doi.org/10.1109/tvt.2020.2988600>.

26. Petrov V., Yarkina N., Moltchanov D., Andreev S., Samouylov K. Session-Level Reliability Analysis for Multi-Service Communication in Autonomous Vehicular Fleets // IEEE Access. 2020. Vol. 8. P. 174629–174642. <https://doi.org/10.1109/access.2020.3024790>.

27. Pirmagomedov R., Ometov A., Moltchanov D., Lu X., Kovalchukov R., Olshannikova E., Andreev S., Koucheryavy Y., Dohler M. Applying Blockchain Technology for User Incentivization in mmWave-Based Mesh Networks // IEEE Access. 2020. Vol. 8. P. 50983–50994. <https://doi.org/10.1109/access.2020.2979135>.

28. Sadovaya Y., Solomitckii D., Mao W., Orhan O., Nikopour H., Talwar S., Andreev S., Koucheryavy Y. Ray-Based Modeling of Directional Millimeter-Wave V2V Transmissions in Highway Scenarios // IEEE Access. 2020. Vol. 8. P. 54482–54493. <https://doi.org/10.1109/access.2020.2980987>.

29. 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. <https://doi.org/10.1109/twc.2020.2973375>.

30. Solomitckii D., Koucheryavy Y., Semkin V., Karttunen A., Petrov V., Nguyen S., Nikopour H., Haneda K., Andreev S., Talwar S. Characterizing Radio Wave Propagation in Urban Street Canyon With Vehicular Blockage at 28 GHz // IEEE Transactions on Vehicular Technology. 2020. Vol. 69, no. 2. P. 1227–1236. <https://doi.org/10.1109/tvt.2019.2959127>.

31. Stusek M., Moltchanov D., Masek P., Mikhaylov K., Zeman O., Roubicek M., Koucheryavy Y., Hosek J. Accuracy Assessment and Cross-Validation of LPWAN Propagation Models in Urban Scenarios // IEEE Access. 2020. Vol. 8. P. 154625–154636. <https://doi.org/10.1109/access.2020.3016042>.

32. Tafintsev N., Moltchanov D., Andreev S., Yeh S., Himayat N., Koucheryavy Y., Valkama M. Handling Spontaneous Traffic Variations in 5G+ via Offloading Onto mmWave-Capable UAV “Bridges” // *IEEE Transactions on Vehicular Technology*. 2020. Vol. 69, no. 9. P. 10070–10084. <https://doi.org/10.1109/tvt.2020.3005253>

33. Tafintsev N., Valkama M., Moltchanov D., Gerasimenko M., Gapeyenko M., Zhu J., Yeh S., Himayat N., Andreev S., Koucheryavy Y. Aerial Access and Backhaul in mmWave B5G Systems: Performance Dynamics and Optimization // *IEEE Communications Magazine*. 2020. Vol. 58, no. 2. P. 93–99. <https://doi.org/10.1109/mcom.001.1900318>.

34. Wirdatmadja S., Jornet J.M., Koucheryavy Y., Balasubramaniam S. Channel Impulse Analysis of Light Propagation for Point-to-Point Nano Communications Through Cortical Neurons // *IEEE Transactions on Communications*. 2020. Vol. 68, no. 11. P. 7111–7122. [10.1109/TCOMM.2020.3012477](https://doi.org/10.1109/TCOMM.2020.3012477)

2019 зод

35. Ateya A., Muthanna A., Vybornova A., Algarni A., Abuarqoub A., Koucheryavy Y., Koucheryavy A. Chaotic salp swarm algorithm for SDN multi-controller networks // *Engineering Science and Technology, an International Journal*. 2019. Vol. 22, no. 4. P. 1001–1012. <https://doi.org/10.1016/j.jestch.2018.12.015>.

36. Begishev V., Moltchanov D., Sopin E., Samuylov A., Andreev S., Koucheryavy Y., Samouylov K. Quantifying the Impact of Guard Capacity on Session Continuity in 3GPP New Radio Systems // *IEEE Transactions on Vehicular Technology*. 2019. Vol. 68, no. 12. P. 12345–12359. <https://doi.org/10.1109/tvt.2019.2948702>.

37. Galinina O., Militano L., Andreev S., Pyattaev A., Johnsson K., Orsino A., Araniti G., Iera A., Dohler M., Koucheryavy Y. Analyzing Competition and Cooperation Dynamics of the Aerial mmWave Access Market // *IEEE Access*. 2019. Vol. 7. P. 87192–87211. <https://doi.org/10.1109/access.2019.2923787>.

38. Gapeyenko M., Petrov V., Moltchanov D., Akdeniz M., Andreev S., Himayat N., Koucheryavy Y. On the Degree of Multi-Connectivity in 5G Millimeter-Wave Cellular Urban Deployments // *IEEE Transactions on Vehicular Technology*. 2019. Vol. 68, no. 2. P. 1973–1978. <https://doi.org/10.1109/tvt.2018.2887343>.

39. Gerasimenko M., Moltchanov D., Gapeyenko M., Andreev S., Koucheryavy Y. Capacity of Multiconnectivity mmWave Systems With

Dynamic Blockage and Directional Antennas // IEEE Transactions on Vehicular Technology. 2019. Vol. 68, no. 4. P. 3534–3549. <https://doi.org/10.1109/tvt.2019.2896565>.

40. Kovalchukov R., Moltchanov D., Samuylov A., Ometov A., Andreev S., Koucheryavy Y., Samouylov K. Evaluating SIR in 3D Millimeter-Wave Deployments: Direct Modeling and Feasible Approximations // IEEE Transactions on Wireless Communications. 2019. Vol. 18, no. 2. P. 879–896. <https://doi.org/10.1109/twc.2018.2886188>.

41. 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. <https://doi.org/10.1109/mcom.2019.1800663>.

42. 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. <https://doi.org/10.1109/access.2019.2900195>.

43. Mäkitalo N., Aaltonen T., Raatikainen M., Ometov A., Andreev S., Koucheryavy Y., Mikkonen T. Action-Oriented Programming Model: Collective Executions and Interactions in the Fog // Journal of Systems and Software. 2019. Vol. 157. P. 110391. <https://doi.org/10.1016/j.jss.2019.110391>.

44. Mikhaylov K., Petrov V., Gupta R., Lema M., Galinina O., Andreev S., Koucheryavy Y., Valkama M., Pouttu A., Dohler M. Energy Efficiency of Multi-Radio Massive Machine-Type Communication (MR-MMTC): Applications, Challenges, and Solutions // IEEE Communications Magazine. 2019. Vol. 57, no. 6. P. 100–106. <https://doi.org/10.1109/mcom.2019.1800394>.

45. Moltchanov D., Kovalchukov R., Gerasimenko M., Andreev S., Koucheryavy Y., Gerla M. Socially Inspired Relaying and Proactive Mode Selection in mmWave Vehicular Communications // IEEE Internet of Things Journal. 2019. Vol. 6, no. 3. P. 5172–5183. <https://doi.org/10.1109/jiot.2019.2898420>.

46. Moltchanov D., Ometov A., Koucheryavy Y. Analytical characterization of the blockage process in 3GPP New Radio systems with trilateral mobility and multi-connectivity // Computer Communications. 2019. Vol. 146. P. 110–120. <https://doi.org/10.1016/j.comcom.2019.07.010>.

47. Moltchanov D., Samuylov A., Petrov V., Gapeyenko M., Himayat N., Andreev S., Koucheryavy Y. Improving Session Continuity With Bandwidth Reservation in mmWave Communications // *IEEE Wireless Communications Letters*. 2019. Vol. 8, no. 1. P. 105–108. <https://doi.org/10.1109/lwc.2018.2859988>.
48. Naumov V., Gaidamaka Y., Samouylov K. On Truncation of the Matrix-Geometric Stationary Distributions // *Mathematics*. 2019. Vol. 7, no. 9. P. 798. <https://doi.org/10.3390/math7090798>.
49. Naumov V., Gaidamaka Y., Samouylov K. On Two Interacting Markovian Queueing Systems // *Mathematics*. 2019. Vol. 7, no. 9. P. 799. <https://doi.org/10.3390/math7090799>.
50. Ometov A., Bezzateev S., Davydov V., Shchesniak A., Masek P., Lohan E., Koucheryavy Y. Positioning Information Privacy in Intelligent Transportation Systems: An Overview and Future Perspective // *Sensors*. 2019. Vol. 19, no. 7. P. 1603. <https://doi.org/10.3390/s19071603>.
51. Ometov A., Bezzateev S., Davydov V., Shchesniak A., Masek P., Lohan E., Koucheryavy Y. Positioning Information Privacy in Intelligent Transportation Systems: An Overview and Future Perspective // *Sensors*. 2019. Vol. 19, no. 7. P. 1603. <https://doi.org/10.3390/s19071603>.
52. Ometov A., Moltchanov D., Komarov M., Volvenko S., Koucheryavy Y. Packet Level Performance Assessment of mmWave Backhauling Technology for 3GPP NR Systems // *IEEE Access*. 2019. Vol. 7. P. 9860–9871. <https://doi.org/10.1109/access.2018.2890558>.
53. Ometov A., Petrov V., Bezzateev S., Andreev S., Koucheryavy Y., Gerla M. Challenges of Multi-Factor Authentication for Securing Advanced IoT Applications // *IEEE Network*. 2019. Vol. 33, no. 2. P. 82–88. <https://doi.org/10.1109/mnet.2019.1800240>.
54. Petrov V., Fodor G., Kokkonen J., Moltchanov D., Lehtomaki J., Andreev S., Koucheryavy Y., Juntti M., Valkama M. On Unified Vehicular Communications and Radar Sensing in Millimeter-Wave and Low Terahertz Bands // *IEEE Wireless Communications*. 2019. Vol. 26, no. 3. P. 146–153. <https://doi.org/10.1109/mwc.2019.1800328>.
55. Petrov V., Mikhaylov K., Moltchanov D., Andreev S., Fodor G., Torsner J., Yanikomeroğlu H., Juntti M., Koucheryavy Y. When IoT Keeps People in the Loop: A Path Towards a New Global Utility // *IEEE Communications Magazine*. 2019. Vol. 57, no. 1. P. 114–121. <https://doi.org/10.1109/mcom.2018.1700018>.
56. Pirmagomedov R., Moltchanov D., Ometov A., Muhammad K., Andreev S., Koucheryavy Y. Facilitating mmWave Mesh Reliability in

57. Balasubramaniam S., Wirdatmadja S., Barros M., Koucheryavy Y., Stachowiak M., Jornet J. Wireless Communications for Optogenetics-Based Brain Stimulation: Present Technology and Future Challenges // IEEE Communications Magazine. 2018. Vol. 56, no. 7. P. 218–224. <https://doi.org/10.1109/mcom.2018.1700917>.

58. Begishev V., Petrov V., Samuylov A., Moltchanov D., Andreev S., Koucheryavy Y., Samouylov K. Resource allocation and sharing for heterogeneous data collection over conventional 3GPP LTE and emerging NB-IoT technologies // Computer Communications. 2018. Vol. 120. P. 93–101. <https://doi.org/10.1016/j.comcom.2018.01.009>.

59. Galinina O., Mikhaylov K., Huang K., Andreev S., Koucheryavy Y. Wirelessly Powered Urban Crowd Sensing over Wearables: Trading Energy for Data // IEEE Wireless Communications. 2018. Vol. 25, no. 2. P. 140–149. <https://doi.org/10.1109/mwc.2018.1600468>.

60. Gapeyenko M., Petrov V., Moltchanov D., Andreev S., Himayat N., Koucheryavy Y. Flexible and Reliable UAV-Assisted Backhaul Operation in 5G mmWave Cellular Networks // IEEE Journal on Selected Areas in Communications. 2018. Vol. 36, no. 11. P. 2486–2496. <https://doi.org/10.1109/jsac.2018.2874145>.

61. 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. <https://doi.org/10.1109/tvt.2018.2857215>.

62. Lohan E., Koivisto M., Galinina O., Andreev S., Tolli A., Destino G., Costa M., Leppanen K., Koucheryavy Y., Valkama M. Benefits of Positioning-Aided Communication Technology in High-Frequency Industrial IoT // IEEE Communications Magazine. 2018. Vol. 56, no. 12. P. 142–148. <https://doi.org/10.1109/mcom.2018.1701057>.

63. 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. <https://doi.org/10.1109/access.2018.2793941>.

64. 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. <https://doi.org/10.1109/tvt.2018.2799141>.
65. Orsino A., Kovalchukov R., Samuylov A., Moltchanov D., Andreev S., Koucheryavy Y., Valkama M. Caching-Aided Collaborative D2D Operation for Predictive Data Dissemination in Industrial IoT // *IEEE Wireless Communications*. 2018. Vol. 25, no. 3. P. 50–57. <https://doi.org/10.1109/mwc.2018.1700320>.
66. Petrov V., Andreev S., Gerla M., Koucheryavy Y. Breaking the Limits in Urban Video Monitoring: Massive Crowd Sourced Surveillance over Vehicles // *IEEE Wireless Communications*. 2018. Vol. 25, no. 5. P. 104–112. <https://doi.org/10.1109/mwc.2018.1700415>.
67. Petrov V., Kokkonen J., Moltchanov D., Lehtomaki J., Juntti, M., Koucheryavy Y. The Impact of Interference From the Side Lanes on mmWave/THz Band V2V Communication Systems With Directional Antennas // *IEEE Transactions on Vehicular Technology*. 2018. Vol. 67, no. 6. P. 5028–5041. <https://doi.org/10.1109/tvt.2018.2799564>.
68. Petrov V., Kokkonen J., Moltchanov D., Lehtomaki J., Koucheryavy Y., Juntti M. Last Meter Indoor Terahertz Wireless Access: Performance Insights and Implementation Roadmap // *IEEE Communications Magazine*. 2018. Vol. 56, no. 6. P. 158–165. <https://doi.org/10.1109/mcom.2018.1600300>.
69. Petrov V., Lema M., 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. <https://doi.org/10.1109/jsac.2018.2815419>.
70. Petrov V., Samuylov A., Begishev V., Moltchanov D., Andreev S., Samuylov K., Koucheryavy Y. Vehicle-Based Relay Assistance for Opportunistic Crowdsensing Over Narrowband IoT (NB-IoT) // *IEEE Internet of Things Journal*. 2018. Vol. 5, no. 5. P. 3710–3723. <https://doi.org/10.1109/jiot.2017.2670363>.
71. Slezak C., Semkin V., Andreev S., Koucheryavy Y., Rangan S. Empirical Effects of Dynamic Human-Body Blockage in 60 GHz Communications // *IEEE Communications Magazine*. 2018. Vol. 56, no. 12. P. 60–66. <https://doi.org/10.1109/mcom.2018.1800232>.

72. Solomitchkii D., Gapeyenko M., Semkin V., Andreev S., Koucheryavy Y. Technologies for Efficient Amateur Drone Detection in 5G Millimeter-Wave Cellular Infrastructure // *IEEE Communications Magazine*. 2018. Vol. 56, no. 1. P. 43–50. <https://doi.org/10.1109/mcom.2017.1700450>.

2017 зод

73. Gapeyenko M., Samuylov A., Gerasimenko M., Moltchanov D., Singh S., Akdeniz M., Aryafar E., Himayat N., Andreev S., Koucheryavy Y. On the Temporal Effects of Mobile Blockers in Urban Millimeter-Wave Cellular Scenarios // *IEEE Transactions on Vehicular Technology*. 2017. Vol. 66, no. 11. P. 10124–10138. <https://doi.org/10.1109/tvt.2017.2754543>.

74. Gerasimenko M., Moltchanov D., Andreev S., Koucheryavy Y., Himayat N., Yeh S., Talwar S. Adaptive Resource Management Strategy in Practical Multi-Radio Heterogeneous Networks // *IEEE Access*. 2017. Vol. 5. P. 219–235. <https://doi.org/10.1109/access.2016.2638022>.

75. Hosek J., Masek P., Andreev S., Galinina O., Ometov A., Kropfl F., Wiedermann W., Koucheryavy Y. A SyMPHOnY of Integrated IoT Businesses: Closing the Gap between Availability and Adoption // *IEEE Communications Magazine*. 2017. Vol. 55, no. 12. P. 156–164. <https://doi.org/10.1109/mcom.2017.1700028>.

76. Makitalo N., Ometov A., Kannisto J., Andreev S., Koucheryavy Y., Mikkonen T. Safe, Secure Executions at the Network Edge: Coordinating Cloud, Edge, and Fog Computing // *IEEE Software*. 2018. Vol. 35, no. 1. P. 30–37. <https://doi.org/10.1109/ms.2017.4541037>.

77. 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. <https://doi.org/10.1109/access.2017.2758840>.

78. Ometov A., Sopin E., Gudkova I., Andreev S., Gaidamaka Y., Koucheryavy Y. Modeling Unreliable Operation of mmWave-Based Data Sessions in Mission-Critical PPDR Services // *IEEE Access*. 2017. Vol. 5. P. 20536–20544. <https://doi.org/10.1109/access.2017.2756690>.

79. Orsino A., Ometov A., Fodor G., Moltchanov D., Militano L., Andreev S., Yilmaz O., 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. <https://doi.org/10.1109/mcom.2017.1600443cm>.

80. Orsino A., Samuylov A., Moltchanov D., Andreev S., Militano L., Araniti G., Koucheryavy Y. Time-Dependent Energy and Resource Management in Mobility-Aware D2D-Empowered 5G Systems // IEEE Wireless Communications. 2017. Vol. 24, no. 4. P. 14–22. <https://doi.org/10.1109/mwc.2017.1600393>.

81. Paramonov A., Hussain O., Samouylov K., Koucheryavy A., Kirichek R., Koucheryavy Y. Clustering Optimization for Out-of-Band D2D Communications // Wireless Communications and Mobile Computing. 2017. Vol. 2017. P. 1–11. <https://doi.org/10.1155/2017/6747052>.

82. Petrov V., Moltchanov D., Komar M., Antonov A., Kustarev P., Rakheja S., Koucheryavy Y. Terahertz Band Intra-Chip Communications: Can Wireless Links Scale Modern x86 CPUs? // IEEE Access. 2017. Vol. 5. P. 6095–6109. <https://doi.org/10.1109/access.2017.2689077>.

83. Petrov V., Solomitckii D., Samuylov A., Lema M., Gapeyenko M., Moltchanov D., Andreev S., Naumov V., Samouylov K., Dohler M., Koucheryavy Y. Dynamic Multi-Connectivity Performance in Ultra-Dense Urban mmWave Deployments // IEEE Journal on Selected Areas in Communications. 2017. Vol. 35, no. 9. P. 2038–2055. <https://doi.org/10.1109/jsac.2017.2720482>.

84. Semkin V., Solomitckii D., Naderpour R., Andreev S., Koucheryavy Y., Raisanen A. Characterization of Radio Links at 60 GHz Using Simple Geometrical and Highly Accurate 3-D Models // IEEE Transactions on Vehicular Technology. 2017. Vol. 66, no. 6. P. 4647–4656. <https://doi.org/10.1109/tvt.2016.2617919>.

85. Solomitckii D., Gapeyenko M., Szyszkowicz S., Andreev S., Yanikomeroğlu H., Koucheryavy Y. Toward Massive Ray-Based Simulations of mmWave Small Cells on Open Urban Maps // IEEE Antennas and Wireless Propagation Letters. 2017. Vol. 16. P. 1435–1438. <https://doi.org/10.1109/lawp.2016.2641339>.

86. Wirdatmadja S., Moltchanov D., Balasubramaniam S., Koucheryavy Y. Microfluidic System Protocols for Integrated On-Chip Communications and Cooling // IEEE Access. 2017. Vol. 5. P. 2417–2429. <https://doi.org/10.1109/access.2017.2662798>.

87. Andreev S., Galinina O., Pyattaev A., Hosek J., Masek P., Yanikomeroğlu H., Koucheryavy Y. Exploring synergy between communications, caching, and computing in 5G-grade deployments // *IEEE Communications Magazine*. 2016. Vol. 54, no. 8. P. 60–69. <https://doi.org/10.1109/mcom.2016.7537178>.

88. Andreev S., Hosek J., Olsson T., Johnsson K., Pyattaev A., Ometov A., Olshannikova E., Gerasimenko M., Masek P., Koucheryavy Y., Mikkonen T. A unifying perspective on proximity-based cellular-assisted mobile social networking // *IEEE Communications Magazine*. 2016. Vol. 54, no. 4. P. 108–116. <https://doi.org/10.1109/mcom.2016.7452274>.

89. Balasubramaniam S., Jornet J., Pierobon M., Koucheryavy Y. Guest Editorial Special Issue on the Internet of Nano Things // *IEEE Internet of Things Journal*. 2016. Vol. 3, no. 1. P. 1–3. <https://doi.org/10.1109/jiot.2016.2516838>.

90. Curado M., Ganchev I., Kassler A., Koucheryavy Y. Guest editorial: Special issue on mobile wireless networks // *Computer Networks*. 2016. Vol. 107. P. 145–147. <https://doi.org/10.1016/j.comnet.2016.07.003>.

91. Galinina O., Pyattaev A., Johnsson K., Turlikov A., Andreev S., Koucheryavy Y. Assessing System-Level Energy Efficiency of mmWave-Based Wearable Networks // *IEEE Journal on Selected Areas in Communications*. 2016. Vol. 34, no. 4. P. 923–937. <https://doi.org/10.1109/jsac.2016.2544539>.

92. Galinina O., Tabassum H., Mikhaylov K., Andreev S., Hossain E., Koucheryavy Y. On feasibility of 5G-grade dedicated RF charging technology for wireless-powered wearables // *IEEE Wireless Communications*. 2016. Vol. 23, no. 2. P. 28–37. <https://doi.org/10.1109/mwc.2016.7462482>.

93. Moltchanov D., Antonov A., Kluchev A., Borunova K., Kustarev P., Petrov V., Koucheryavy Y., Platunov A. Statistical Traffic Properties and Model Inference for Shared Cache Interface in Multi-Core CPUs // *IEEE Access*. 2016. Vol. 4. P. 4829–4839. <https://doi.org/10.1109/access.2016.2603169>.

94. Ometov A., Olshannikova E., Masek P., Olsson T., Hosek J., Andreev S., Koucheryavy Y. Dynamic Trust Associations Over Socially-Aware D2D Technology: A Practical Implementation

Perspective // IEEE Access. 2016. Vol. 4. P. 7692–7702.
<https://doi.org/10.1109/access.2016.2617207>.

95. Ometov A., Orsino A., Militano L., Araniti G., Moltchanov D., Andreev S. A novel security-centric framework for D2D connectivity based on spatial and social proximity // Computer Networks. 2016. Vol. 107. P. 327–338. <https://doi.org/10.1016/j.comnet.2016.03.013>.

96. Ometov A., Orsino A., Militano L., Moltchanov D., Araniti G., Olshannikova E., Fodor G., Andreev S., Olsson T., Iera A., Torsner J., Koucheryavy Y., Mikkonen T. Toward trusted, social-aware D2D connectivity: bridging across the technology and sociality realms // IEEE Wireless Communications. 2016. Vol. 23, no. 4. P. 103–111. <https://doi.org/10.1109/mwc.2016.7553033>.

97. Orsino A., Moltchanov D., Gapeyenko M., Samuylov A., Andreev S., Militano L., Araniti G., Koucheryavy Y. Direct Connection on the Move: Characterization of User Mobility in Cellular-Assisted D2D Systems // IEEE Vehicular Technology Magazine. 2016. Vol. 11, no. 3. P. 38–48. <https://doi.org/10.1109/mvt.2016.2550002>.

98. Petrov V., Moltchanov D., Kustarev P., Jornet J., Koucheryavy Y. On the Use of Integral Geometry for Interference Modeling and Analysis in Wireless Networks // IEEE Communications Letters. 2016. Vol. 20, no. 12. P. 2530–2533. <https://doi.org/10.1109/lcomm.2016.2610435>.

99. 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. <https://doi.org/10.1109/mcom.2016.7432155>.

100. 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. <https://doi.org/10.1109/tvt.2015.2481795>.

СПИСОК СОКРАЩЕНИЙ

Русскоязычные:

АУ	– абонентское устройство
БПЛА	– беспилотный летательный аппарат
БС	– базовая станция
ВВХ	– вероятностно-временные характеристики
ВЭ	– вертикальная экспозиция
МСЭ	– международный союз электросвязи
ОЗУ	– оперативное запоминающее устройство
ОСП	– отношение сигнала к помехе
ОСШ	– отношение сигнала к шуму
ОТ	– операционные технологии
ПВ	– прямая видимость
ПРВ	– плотность распределения вероятностей
ПУ	– пост управления
ПФ	– производящая функция
РБ	– ресурсный блок
РСМО	– ресурсные системы массового обслуживания
с.в.	– случайная величина
СКО	– среднеквадратичное отклонение
СМО	– система массового обслуживания
СУР	– система уравнений равновесия
СеМО	– сеть массового обслуживания
ТМО	– теории массового обслуживания
ТС	– транспортное средство
ТфОП	– телефонные сети общего пользования
ФАР	– фазированные антенные решетки
ФВ	– фиксированные высоты
ФР	– функция распределения
ФРВ	– функция распределения вероятностей
ЦП	– центральный процессор

Англоязычные:

3GPP	– 3rd Generation Partnership Project
6LoWPAN	– IPv6 over Low power Wireless Personal Area Networks
8-APSK	– 8-Phase Asymmetric Phase Shift Keying
8-PSK	– 8-Phase Shift Keying
ABS	– Aerial Base Stations
AES	– Advanced Encryption Standard

AF	– Array Factor
AFNOR	– Association Francaise de Normalisation
ALOHA	– Additive Links On-line Hawaii Area
AMD	– Advanced Micro Devices
AMF	– Access Mobility Function
ANDSF	– Access Network Discovery and Selection Function
ANSI	– American National Standards Institute
AP	– Access point
API	– Application Programming Interface
APNIC	– Asia Pacific Network Information Centre
AR	– Augmented reality
AWGN	– Additive White Gaussian Noise
AoA	– Angle of Arrival
AoD	– Angle of Departure
BBU	– Baseband unit
BEB	– Binary Exponential Backoff
BER	– Bit Error Rate
BLER	– Block Error Rate
BP	– Bandwidth parts
BPSK	– Binary Phase Shift Keying
BS	– Base Station
BSI	– British Standards Institution
BSS	– Business Support Systems
BUT	– Brno University of Technology
BWP	– Bandwidth Part
BiCMOS	– Bipolar Complementary Metal-Oxide Semiconductor
CAC	– Connection Admission Control
CBG	– Code Block Group
CBR	– Constant Bit Rate
CDF	– Cumulative Distribution Function
CIoT	– Cell IoT
CMOS	– Complementary Metal-Oxide Semiconductor
COBALT	– Contention-Based LTE Transmission
CPU	– Central Processing Unit
CQI	– Channel Quality Indicator
CRC	– Cyclic Redundancy Check
CRRM	– Cooperative Radio Resource Manager
CSI	– Channel State Information
CSMA/CA	– Carrier Sense Multiple Access With Collision Avoidance
CTMC	– Continuous-Time Markov Chain

CTS	– Clear To Send
CoW	– Cell-on-Wheels
D2D	– Device to Device
DCF	– Distributed Coordination Function
DCN	– Data Centre Networks
DES	– Discrete-Event Simulation
DFT	– Discrete Fourier Transform
DL	– Downlink
DNS	– Domain Name System
DYN	– Dynamic Strategy
DYNRES	– Dynamic Strategy with Reservation
EC-GSM	– Extended Coverage GSM
EDCA	– Enhanced distributed channel access
EDGE	– Enhanced Data Rates for GSM Evolution
EIRP	– Equivalent Isotropically Radiated Power
eMBB	– Enhanced Mobile Broadband
eNB	– Evolved Node B
eNodeB	– Evolved Node B
EPC	– Evolved Packet Core
ETSI	– European Telecommunications Standards Institute
EWMA	– Exponentially-Weighted Moving Average
e-SMLC	– Enhanced Serving Mobile Location Center
FDD	– Frequency Division Duplexing
FEC	– Forward Error Correction
FPT	– First Passage Time
GPCPU	– General Purpose Graphics Processing Unit
GPU	– Graphics Processing Unit
GRE	– Generic Routing Encapsulation
GSM	– Global System for Mobile
GSMA	– Global System for Mobile Association
GaAs	– Gallium Arsenide
GAN	– Generative Adversarial Network
H-CRAN	– Heterogeneous Cloud RAN
H2H	– Host-to-Host
HARQ	– Hybrid Automatic Repeat reQuests
HCS	– Header Check Sequence
HEMT	– High Electron Mobility Transistors
HITRAN	– High-Resolution Transmission Molecular Absorption Database
HPBW	– Half-Power Beamwidth

HSPA	– High Speed Packet Access
HTTPS	– Hypertext Transfer Protocol Secure
HetNet	– Heterogeneous Network
IAB	– Integrated Access and Backhaul
IEEE	– Institute of Electrical and Electronic Engineers
IETF	– Internet Engineering Task Force
IGthz	– Terahertz Interest Group
IIoT	– Industrial IoT
IMS	– IP Multimedia Subsystem
IMT-2020	– International Mobile Telecommunications–2020
IPv4	– Internet Protocol version 4
IPv6	– Internet Protocol version 6
IRP	– Isotropically Radiated Power
ISI	– Inter-Symbol Interference
ISM	– Industrial Scientific and Medical
ISO	– International Organization for Standardization
ISP	– Internet Service Provider
ITU	– International Telecommunication Union
InGaAs	– Indium Gallium Arsenide
InP	– Indium Phosphide
InAlAs	– Indium Aluminum Arsenide
IoS	– Internet of Skills
IoT	– Internet of Things
IoWT	– Internet of Wearable Things
LAA	– LTE Licensed Assisted Access
LBT	– Listen Before Talk
LDPC	– Low Density Parity Check
LLC	– Low-Latency Communications
LPN	– Low-Power Node
LPWAN	– Low-Power Wide-Area Networks
LTE	– Long-Term Evolution
LTE-A	– LTE Advanced
LTE-M	– LTE-Machine Type Communication
LTE-U	– LTE-Unlicensed
LoRaWAN	– Long Range Wide Area Network
LoS	– Line of Sight
M2M	– Machine-to-Machine
MAC	– Medium Access Control
MANO	– Management and Orchestration
MCDATA	– Mission Critical Data

MCPTT	– Mission Critical Push-to-Talk
MCS	– Modulation and Coding Scheme
MCVideo	– Mission Critical Video
MIMO	– Multiple-Input and Multiple-Output
MIP	– Mixed-Integer Programming
MIoT	– Mobile Internet of Things
MME	– Mobile Management Entity
MMIC	– Monolithic Microwave Integrated Circuit
mMTC	– massive Machine-Type Communications
MTC	– Machine-Type Communications
MVNO	– Mobile Virtual Network Operators
NB-IoT	– Narrow Band Internet of Things
NFC	– Near Field Communication
NFV	– Network Function Virtualization
nLoS	– Non-Line-of-Sight
NOMA	– Non-Orthogonal Multiple Access
NP	– Non-deterministic Polynomial
NR	– New Radio
NTP	– Network Time Protocol
NoC	– Network-on-Chip
OFDM	– Orthogonal Frequency-Division Multiplexing
OMNeT++	– Objective Modular Network Testbed in C++
ONOS	– Open Network Operating System
ORAN	– Open RAN
OSS	– Operations Support Systems
P2P	– Peer-to-Peer
PAP	– Pairnet Associated Period
PC	– Pairnet Coordinator
PCF	– Policy Control Function
PCIe	– Peripheral Component Interconnect Express
PCM	– Pulse-Code Modulation
PDCCH	– Physical Downlink Control Channel
PDP	– Power Delay Profile
PGW	– Packet Gateway
PMNO	– Physical Mobile Network Operator
PPDR	– Public Protection and Disaster Relief
PPP	– Poisson Point Process
PRACH	– Physical Random Access Channel
PRB	– Primary Resource Block
PRDEV	– Pairnet Device

PSO	– Particle Swarm Optimization
PSP	– Pairnet Setup Period
PUCCH	– Physical Uplink Control Channel
PUSCH	– Physical Uplink Shared Channel
PoE	– Power over Ethernet
PoS	– Proof-of-Stake
PoW	– Proof-of-Work
ProSe	– Proximity Services
QPI	– Quick Path Interconnect
QPSK	– Quadrature Phase-Shift Keying
QoE	– Quality of Experience
QoS	– Quality of Service
R-GW	– Radio Gateway
RACH	– Random Access Channel
RAN	– Radio Access Network
RAR	– Random Access Response
RAT	– Radio Access Technologies
RAW	– Random Access Window
RB	– Resource Block
RDM	– Random Direction Model
RFID	– Radio Frequency Identification
RL	– Ray Tracing
RPGM	– Reference Point Group Mobility
RRH	– Remote Radio Head
RRM	– Radio Resource Management
RTS	– Request to Send
RTT	– Round-Trip Time
RmNet	– Research Methods Network
S-NSSAI	– Single Network Slice Selection Assistance Information
SDN	– Software-Defined Networking
SE	– Spectral Efficiency
SF	– Shadow Fading
SIFS	– Short Inter-Frame Space
SINR	– Signal-to-Interference plus Noise Ratio
SIR	– Signal-to-Interference Ratio
SLA	– Service Level Agreement
SMF	– Session Management Function
SMLC	– Serving Mobile Location Center
SNR	– Signal-to-Noise Ratio
SPP	– Surface Plasmon Polaritons

SR	– Scheduling Request
SSC mode	– Session and Service Continuity mode
SSL	– Secure Sockets Layer
SST	– Slice/Service Type
SiGe	– Silicon–Germanium
TCP	– Transmission Control Protocol
TDD	– Time Division Duplex
TDMA	– Time Division Multiple Access
THz-AP	– THz band Access Point
THz-BS	– THz band Base Station
THz-TDS	– THz band Time-Domain Spectroscopy
THz-UE	– THz band User Equipment
TMIC	– Terahertz Monolithic Integrated Circuit
TSC	– Time-Sensitive Communication
TXOP	– Transmit Opportunity
UE	– User Equipment
UL	– Uplink
UMi	– Urban Micro
UPF	– User Plane Function
URLLC	– Ultra-Reliable Low-Latency Communications
V2X	– Vehicle-to-Everything
VLC	– Visible Light Communications
VNF	– Virtual Network Functions
VR	– Virtual Reality
VoIP	– Voice over Internet Protocol
WCDMA	– Wideband Code Division Multiple Access
WFD	– Wi-Fi Direct
WLAN	– Wireless Local Area Networks
WMAN	– Wireless Metropolitan Area Network
WNoC	– Wireless Network-on-Chip
WPAN	– Wireless Personal Area Network
WSN	– Wireless Sensor Network
WWAN	– Wireless Wide Area Network
ZoA	– Zenith of Arrival
ZoD	– Zenith of Departure

Научное издание

**Молчанов Дмитрий Александрович
Бегишев Вячеслав Олегович
Самуйлов Константин Евгеньевич
Кучерявый Евгений Андреевич**

**СЕТИ 5G/6G:
АРХИТЕКТУРА, ТЕХНОЛОГИИ,
МЕТОДЫ АНАЛИЗА И РАСЧЕТА**

*Редактор И.Л. Панкратова
Технический редактор Н.А. Ясько
Компьютерная верстка Н.А. Ясько
Дизайн обложки И.А. Черновой*

Подписано в печать 20.06.2022 г. Формат 70×100/16. Печать офсетная.
Усл. печ. л. 41,93. Тираж 300 экз. Заказ 1433.

Российский университет дружбы народов
115419, ГСП-1, г. Москва, ул. Орджоникидзе, д. 3

Типография РУДН
115419, ГСП-1, г. Москва, ул. Орджоникидзе, д. 3.
Тел.: 8 (495) 955-08-74. E-mail: publishing@rudn.ru

Для заметок

Для заметок
