

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ»

На правах рукописи



Андреев Сергей Дмитриевич

**РАЗРАБОТКА И ИССЛЕДОВАНИЕ МОДЕЛЕЙ
МНОЖЕСТВЕННОГО ДОСТУПА И АЛГОРИТМОВ
УПРАВЛЕНИЯ ПОТОКАМИ ТРАФИКА ДЛЯ
ГЕТЕРОГЕННЫХ БЕСПРОВОДНЫХ СЕТЕЙ**

Специальность 05.12.13 —

«Системы, сети и устройства телекоммуникаций»

Диссертация на соискание ученой степени
доктора технических наук

Научные консультанты:

доктор технических наук, профессор

Самуйлов Константин Евгеньевич

доктор технических наук, профессор

Тюрликов Андрей Михайлович

Москва — 2018

Оглавление

	Стр.
Список сокращений	8
Введение	12
Глава 1. Основные понятия, определения и структура гетерогенных сетей связи 5G	29
1.1 Предпосылки появления сетей связи 5G и направления их развития	29
1.1.1 Истоки возникновения сетей связи пятого поколения	29
1.1.2 Основные векторы развития радиотехнологий 5G	31
1.2 Возможности использования беспроводного спектра	36
1.2.1 Различные способы организации управления спектром	36
1.2.2 Изменение беспроводного ландшафта в эпоху 5G	39
1.3 Интеграция сетей связи 5G с Интернетом вещей	41
1.3.1 Становление конвергентной экосистемы 5G-IoT	41
1.3.2 Гетерогенные сети 5G для поддержки Индустриального Интернета	46
1.4 Стандартизация сетей связи 5G и анализ ее требований	49
1.4.1 Возникновение Интернета мобильных и надежных вещей	49
1.4.2 Цели и задачи стандартизации систем 5G	52
1.4.3 Особенности различных сценариев применения 5G	55
1.5 Структура последующего изложения	56
Глава 2. Интеграция технологий радиодоступа и уплотнение сети	58
2.1 Предварительные замечания	58
2.2 Варианты интеграции технологий радиодоступа	59
2.2.1 Необходимость интеграции радиотехнологий	59
2.2.2 Типы архитектуры гетерогенной сети	62
2.3 Анализ интегрированной сотовой и локальной сети доступа	67
2.3.1 Совместный учет динамики трафика и размещения узлов	67
2.3.2 Подход к анализу гетерогенных сетей доступа	70

	Стр.
2.3.3	Результаты анализа интегрированной сети доступа 79
2.4	Анализ уплотнения сетей доступа и учет динамики их загрузки . 83
2.4.1	Совмещенные модули доступа LTE и WiFi 83
2.4.2	Моделирование сверхплотных совмещенных покрытий . . 85
2.4.3	Результаты сравнения алгоритмов 94
2.5	Управление доступом и радиоресурсами в интегрированных сетях 101
2.5.1	Различные классы алгоритмов управления трафиком . . . 101
2.5.2	Сравнительный анализ алгоритмов управления трафиком 107
2.6	Выводы 114

Глава 3. Использование прямых соединений между

	устройствами и крайне высоких частот 115
3.1	Предварительные замечания 115
3.2	Целесообразность и варианты использования прямых соединений 116
3.2.1	Обоснование и применение связи «устройство – устройство» 116
3.2.2	Альтернативные радиотехнологии для соединений D2D . . 118
3.2.3	Обнаружение соседей и установление соединения D2D . . 122
3.3	Использование прямых соединений на нелицензированных частотах 125
3.3.1	Изучение связи D2D в нелицензированном спектре 125
3.3.2	Подход к анализу систем D2D 127
3.3.3	Имитационное моделирование и его результаты 136
3.4	Анализ использования крайне высоких частот в сетях нового поколения 143
3.4.1	Перспективы связи на крайне высоких частотах 143
3.4.2	Подход к моделированию связи mmWave 146
3.4.3	Динамика процесса блокирования передачи 148
3.5	Перспективные сценарии применения связи на крайне высоких частотах 154
3.5.1	Моделирование распространения миллиметровых волн . . 154
3.5.2	Изучение реалистичного сценария использования mmWave 158
3.5.3	Полученные результаты и их обсуждение 162
3.6	Выводы 171

Глава 4. Поддержка приложений Интернета вещей и распространение носимых устройств	173
4.1 Предварительные замечания	173
4.2 Способы организации беспроводного доступа вещей в современных сетях	174
4.2.1 Особенности возникновения и развития Интернета вещей .	174
4.2.2 Современные радиотехнологии для связи M2M	180
4.3 Анализ начального доступа в сотовых системах M2M	186
4.3.1 Характеристики работы канала случайного доступа в LTE	186
4.3.2 Анализ производительности сотовых систем M2M	194
4.4 Предпосылки и последствия распространения носимых устройств	203
4.4.1 Возможности современного носимого оборудования	203
4.4.2 Особенности применения носимых устройств в условиях их высокой плотности	205
4.4.3 Альтернативные радиотехнологии для массового размещения носимого оборудования	207
4.5 Повышение энергетической эффективности доступа носимых устройств	212
4.5.1 Сети носимых устройств на основе mmWave-соединений .	212
4.5.2 Моделирование mmWave-системы носимых устройств . . .	214
4.5.3 Основные численные результаты и их интерпретация . . .	223
4.6 Выводы	228
Глава 5. Применение гетерогенных сетей 5G и их перспективные приложения	229
5.1 Предварительные замечания	229
5.2 Управление прямым соединением устройств в сотовых сетях . . .	230
5.2.1 Контроль связи D2D через систему 3GPP LTE	230
5.2.2 Общесистемное исследование работы прямых соединений .	239
5.2.3 Прототипирование системы выгрузки на соединения D2D	246
5.3 Анализ практического использования прямых соединений для обмена данными	253
5.3.1 Прямое взаимодействие с сетевым кодированием	253
5.3.2 Моделирование прямого обмена контентом	256

	Стр.
5.3.3 Моделирование системы распределения контента	263
5.4 Интеграция систем удаленного хранения и обработки данных . .	274
5.4.1 Дальнейшее развитие сетей следующего поколения	274
5.4.2 Перспективные пользовательские приложения и услуги в 5G	277
5.4.3 Эволюция архитектуры систем связи 5G	281
5.5 Выводы	285
Заключение	286
Список литературы	288
Список рисунков	329
Список таблиц	334
Список приложений	335
Приложение А. Анализ интегрированной гетерогенной сети доступа	336
A.1 Особенности слоев гетерогенной сети доступа	336
A.1.1 Модель слоя WLAN	336
A.1.2 Модель пикослоя	337
A.1.3 Модель макрослоя	339
A.2 Подход к аналитическому моделированию	341
A.2.1 Анализ процессов в макрослое	341
A.2.2 Анализ процессов в слое WLAN и пикослое	342
A.3 Стационарное состояние системы	344
A.4 Основные переходы в системе	346
A.4.1 Переходы в макрослое	346
A.4.2 Переходы в слое WLAN	347
A.4.3 Переходы в пикослое	350
A.5 Вспомогательные вычисления	351
A.5.1 Некоторые распределения для макрослоя	352
A.5.2 Некоторые распределения для слоя WLAN	353

А.5.3	Некоторые распределения для пикослоя	356
Приложение Б. Анализ совмещенной гетерогенной сети доступа		360
Б.1	Подход к анализу и общие замечания	360
Б.2	Моделирование динамики системы связи	361
Б.3	Основные вычисления для сети WiFi	365
Б.4	Основные вычисления для сети LTE	368
Б.5	Базовый алгоритм: предпочтительное использование технологии WiFi	372
Б.6	Продвинутый алгоритм: одновременное использование технологий LTE и WiFi	374
Б.7	Важное практическое обобщение	375
Б.8	Вспомогательные вычисления	377
Б.8.1	Интенсивность переходов b_i для процесса $N(t)$ после укрупнения состояний	377
Б.8.2	Средняя площадь пересечения накладываются зон покрытия	379
Приложение В. Анализ интегрированных сотовой системы и сети D2D		381
В.1	Моделирование сети D2D	381
В.1.1	Вероятностная модель	381
В.1.2	Стационарное распределение вероятностей	382
В.1.3	Энергетическое потребление	384
В.2	Моделирование сотовой системы связи	385
В.2.1	Вероятностная модель	385
В.2.2	Стационарное распределение вероятностей	385
В.2.3	Энергетическое потребление	387
В.3	Вспомогательные вычисления	388
В.3.1	Слой D2D	389
В.3.2	Сотовый слой: алгоритм MR	390
В.3.3	Сотовый слой: алгоритм FU	392

Приложение Г. Анализ энергетической эффективности mmWave-системы носимых устройств	394
Г.1 Моделирование сети доступа в целом	394
Г.1.1 Протокол доступа к каналу	394
Г.1.2 Геометрия сети связи	400
Г.1.3 Направленность передачи	401
Г.1.4 Расширения предложенной модели	406
Г.2 Анализ энергетической эффективности сети	409
Приложение Д. Анализ применения сетевого кодирования в системе D2D с сетевым содействием	413
Д.1 Моделирование системы распределения контента	413
Д.1.1 Жизненный цикл контента	414
Д.1.2 Доставка контента получателям	414
Д.1.3 Базовый режим распределения контента	417
Д.1.4 Кодированный режим распределения контента	418
Д.2 Вспомогательные вычисления	421
Д.2.1 Описание фазы роста	421
Д.2.2 Описание фазы сокращения	424
Приложение Е. Документы, подтверждающие внедрение основных результатов диссертационной работы:	427

Список сокращений

3G, 4G, 5G	—	Third, Fourth, Fifth Generation
3GPP	—	Third Generation Partnership Project
ANDSF	—	Access Network Discovery and Selection Function
ATI	—	Announcement Transmission Interval
BI	—	Beacon Interval
BPSK	—	Binary Phase Shift Keying
BRP	—	Beam Refinement Protocol
BTI	—	Beacon Transmission Interval
CBP	—	Contention-Based Period
CCA	—	Clear Channel Assessment
CDF	—	Cumulative Distribution Function
CDN	—	Content Delivery Network
CSI	—	Channel State Information
D2D	—	Device-to-Device
DCF	—	Distributed Coordination Function
DRX	—	Discontinuous Reception
DTI	—	Data Transmission Interval
EAB	—	Enhanced Access Barring
EC-GSM	—	Extended Coverage GSM
ECMA	—	European Computer Manufacturers Association
EDGE	—	Enhanced Data rates for GSM Evolution
eNodeB	—	evolved Node B
EPC	—	Evolved Packet Core
e-SMLC	—	evolved Serving Mobile Location Center
ETSI	—	European Telecommunications Standards Institute
FU	—	Full Utilization
GERAN	—	GSM/EDGE Radio Access Network
GFSK	—	Gaussian Frequency Shift Keying
GSM	—	Global System for Mobile communications
H2H	—	Human-to-Human
HARQ	—	Hybrid Automatic Repeat Request

HetNet	—	Heterogeneous Network
HEW	—	High Efficiency WLAN
IEEE	—	Institute of Electrical and Electronics Engineers
IMS	—	IP Multimedia Subsystem
IMT	—	International Mobile Telecommunication
IoT	—	Internet of Things
IP	—	Internet Protocol
ISM	—	Industrial, Scientific, and Medical
ITU	—	International Telecommunication Union
LAA	—	Licensed Assisted Access
LoS	—	Line-of-Sight
LPWAN	—	Low-Power Wide-Area Network
LTE	—	Long-Term Evolution
LTN	—	Low Throughput Network
LWA	—	LTE-WLAN Aggregation
M2M	—	Machine-to-Machine
MCS	—	Modulation and Coding Scheme
MIMO	—	Multiple Input Multiple Output
MME	—	Mobility Management Entity
mmWave	—	Millimeter-Wave
MR	—	Maximum Rate
NAS	—	Non-Access Stratum
NB-IoT	—	Narrowband IoT
NFV	—	Network Function Virtualization
nLoS	—	non-Line-of-Sight
NR	—	New Radio
OFDM	—	Orthogonal Frequency-Division Multiplexing
OSM	—	Open Street Map
P2P	—	Peer-to-Peer
PBSS	—	Personal Basic Service Set
PDSCH	—	Physical Downlink Shared Channel
PHP	—	Personal Home Page
PLMR	—	Private Land Mobile Radio

PPP	—	Poisson Point Process
PRACH	—	Physical Random Access Channel
ProSe	—	Proximity Services
PSK	—	Pre-Shared Key
PUCCH	—	Physical Uplink Control Channel
PUSCH	—	Physical Uplink Shared Channel
RACH	—	Random Access Channel
RAN	—	Radio Access Network
RAR	—	Random Access Response
RAT	—	Radio Access Technology
RDM	—	Random Direction Model
RFID	—	Radio Frequency Identification
RFTDMA	—	Random Frequency and Time Division Multiple Access
RL	—	Ray Launching
RLNC	—	Random Linear Network Coding
RPMA	—	Random Phase Multiple Access
RRC	—	Radio Resource Control
RRM	—	Radio Resource Management
RSSI	—	Received Signal Strength Indication
RT	—	Ray Tracing
RX	—	Receiver
SCADA	—	Supervisory Control and Data Acquisition
SDN	—	Software-Defined Networking
SINR	—	Signal-to-Interference-plus-Noise Ratio
SIP	—	Session Initiation Protocol
SLS	—	Sector Level Sweep
SNR	—	Signal-to-Noise Ratio
SP	—	Service Period
SQL	—	Structured Query Language
SRD	—	Short Range Devices
SS	—	Spread Spectrum
TX	—	Transmitter
UE	—	User Equipment

UMi	—	Urban Microcellular
UNB	—	Ultra-Narrowband
VoLTE	—	Voice over LTE
WFD	—	WiFi-Direct
WiFi	—	Wireless Fidelity
WiGig	—	Wireless Gigabit
WirelessHD	—	Wireless High Definition
WLAN	—	Wireless Local Area Network
WPAN	—	Wireless Personal Area Network
WSN	—	Wireless Sensor Network

Введение

В настоящее время развитие телекоммуникационных технологий происходит стремительными темпами, что обусловлено все возрастающим числом пользователей систем мобильной связи и их повышенными потребностями в ресурсах беспроводной сети. Во избежание истощения ограниченных частотно-временных радиоресурсов, имеющихся у современных технологий мобильного доступа, международными комитетами по стандартизации ведется напряженная работа с целью принятия до 2020 года новых спецификаций, определяющих работу сетей связи *пятого поколения* (5G). Для повышения качества обслуживания своих пользователей такие *гетерогенные* беспроводные сети все более тесно интегрируют между собой существующие, а также перспективные технологии радиодоступа, что приводит к возникновению принципиально новых научно-технических задач.

Помимо услуг мобильного доступа, которые предполагают передачу разнородного пользовательского трафика, гетерогенные сети 5G обеспечат поддержку множества приложений Интернета вещей, основанных на принципе межмашинного взаимодействия, а также реализуют технологии программно-определяемых сетей и облачных вычислений. В рамках концепции Интернета вещей возникает все более существенное различие между потребительскими и промышленными сценариями на пути к Индустриальному Интернету. Отдельного внимания заслуживает при этом подключение носимых устройств, которые занимают промежуточное положение между пользовательскими системами и сетями межмашинного взаимодействия. Такое многообразие приложений обуславливает важность обеспечения эффективного функционирования гетерогенных сетей пятого поколения.

Для повышения качества обслуживания в системах 5G потребуется согласованное взаимодействие сотовых технологий радиодоступа, функционирующих в лицензированном диапазоне частот, таких как система LTE комитета 3GPP, с семейством технологий локального доступа IEEE 802.11, использующих нелицензированный спектр и широко известных как WiFi. Эти традиционные системы радиодоступа будут дополнены более современными решениями IEEE 802.11ad/ay (WiGig) и новейшими технологиями 3GPP, такими как «Новое Радио» (New Radio), которые реализуют связь в том числе на крайне высоких

частотах. Кроме того, представляет интерес использование прямых соединений между пользовательскими устройствами, что снижает нагрузку на инфраструктурные каналы связи, позволяя перенаправлять трафик из сотовой сети в сети других технологий радиодоступа.

В силу своей исторически сложившейся эволюции различные технологии радиодоступа, входящие в состав гетерогенных беспроводных сетей, подвержены значительной фрагментации и недостаточно интегрированы между собой. Это обстоятельство сдерживает их дальнейшее развитие, а также обуславливает разрыв между требованиями перспективных пользовательских приложений и эффективностью функционирования современных систем мобильного доступа. Для его преодоления требуется такое управление информационными потоками, которое позволяет балансировать загрузку различных технологий радиодоступа в рамках гетерогенной беспроводной сети. При этом возможно как распределение потоков трафика между сотовой сетью и другими (например, локальными) системами радиодоступа, так и их перенаправление на прямые соединения между устройствами.

Таким образом, согласование работы различных технологий радиодоступа и использование прямых соединений «устройство – устройство» требуют создания новой методологии управления потоками трафика и организации множественного доступа, предоставляющей набор методов для контроля за установлением соединения, назначения мощности передачи и разделения ресурса между пользователями. Соответствующие методы должны обеспечить высокую скорость передачи данных в числе прочих требований к сетям 5G, таких как высокая емкость системы (выраженная, например, в числе одновременно обслуживаемых пользователей), низкая задержка доступа в сеть при установлении соединения или передаче данных, а также приемлемая энергетическая эффективность при доставке трафика (определяемая как объем данных, переданный в расчете на затраченную при этом энергию).

Как следствие необходим комплексный подход, учитывающий основные особенности современного этапа развития гетерогенных беспроводных систем мобильного доступа, такие как высокая плотность размещения пользовательских устройств и узлов сетевой инфраструктуры, тесная интеграция разнородных технологий радиодоступа и использование прямых соединений между устройствами, а также применение крайне высоких частот и обеспечение поддержки приложений Интернета вещей с учетом особенностей носимых

устройств. В результате возникает **важная научно-техническая проблема** по созданию моделей и алгоритмов для обеспечения эффективного функционирования гетерогенных беспроводных сетей. Ее всестороннему решению и посвящена тема данной диссертационной работы, что определяет ее актуальность.

Таким образом, **актуальными являются** разработка и исследование *моделей* организации коллективного использования ресурса для гетерогенных беспроводных сетей в условиях высокой плотности размещения устройств с целью оценки показателей качества обслуживания, а также создание эффективных *алгоритмов* управления информационными потоками с целью снижения нагрузки на сотовую сеть методом перенаправления трафика в сети других технологий радиодоступа.

С учетом вышеизложенного целесообразно кратко охарактеризовать **работанность темы исследования**. В частности, значительный вклад в развитие данной тематики внесли следующие российские и зарубежные ученые и исследователи.

- Теория массового обслуживания и теория телетрафика – Г.П. Башарин, В.М. Вишневецкий, Ю.В. Гайдамака, А.Н. Дудин, А.Е. Кучерявый, Е.В. Морозов, В.А. Наумов, К.Е. Самуйлов, С.Н. Степанов, Г.Г. Яновский, V.V. Iversen, F. Kelly и др.
- Методы анализа на основе стохастической геометрии – Р.В. Амбарцумян, И.И. Цитович, А.Н. Ширяев, J.G. Andrews, F. Baccelli, M. Haenggi, R.W. Heath, E. Hossain и др.
- Модели множественного доступа и методы передачи данных – Н.Д. Введенская, Е.А. Крук, А.И. Ляхов, И.Е. Никульский, А.И. Парамонов, А.М. Тюрликов, С.Г. Фосс, Б.С. Цыбаков, M. Dohler, G.V. Giannakis, L. Kleinrock и др.

Кроме того, имеется множество технических публикаций, включая спецификации и руководящие документы по разработке мобильных систем связи и гетерогенных сетей различных поколений вплоть до пятого.

Под *множественным доступом* в работе понимается организация коллективного использования ресурса многими пользователями (см. одноименную статью, подготовленную Б.С. Цыбаковым [Вероятность и математическая статистика : Энциклопедия / Гл. ред. Ю.В. Прохоров. — М. : Большая Рос. Энцикл., 1999. — 910 с.]).

Под *управлением потоками трафика* понимается их перенаправление между различными технологиями радиодоступа в рамках гетерогенной беспроводной сети с целью снижения нагрузки на сотовую сеть, которое далее называется «выгрузкой» трафика.

Объектом исследования является *гетерогенная беспроводная сеть* с набором технологий радиодоступа при высокой плотности размещения устройств, а **предмет** исследования составляют модели множественного доступа для гетерогенной беспроводной сети и алгоритмы управления потоками трафика с целью снижения нагрузки на сотовую сеть радиодоступа.

Целью диссертации является разработка и исследование моделей множественного доступа для оценки показателей качества обслуживания в гетерогенных беспроводных сетях с набором технологий радиодоступа в условиях высокой плотности размещения устройств, а также разработка и исследование алгоритмов перенаправления информационных потоков для снижения нагрузки на сотовую сеть методом выгрузки трафика в сети других технологий радиодоступа.

Для достижения цели в диссертационной работе решаются следующие крупные **задачи**.

1. Разработка и исследование моделей множественного доступа пользователей для гетерогенных беспроводных сетей с набором технологий радиодоступа в условиях высокой плотности размещения устройств с целью оценки параметров качества обслуживания, что включает в себя:
 - моделирование гетерогенной сети с набором интегрированных между собой технологий радиодоступа;
 - моделирование гетерогенной сети с возможностью установления прямых соединений между устройствами;
 - моделирование сети радиодоступа для обслуживания большого числа устройств;
 - моделирование сети радиодоступа для обслуживания множества носимых устройств.
2. Разработка и исследование алгоритмов перенаправления потоков трафика с целью снижения нагрузки на сотовую сеть радиодоступа, а также оценка их характеристик в гетерогенных беспроводных сетях с учетом обслуживания потоков трафика от взаимодействующих устройств, что включает в себя:

- создание алгоритма с разделением пользовательской сессии для совмещенной сети радиодоступа;
- создание алгоритма с одновременным подключением пользователя к нескольким узлам сетевой инфраструктуры;
- создание и реализацию протокола управления системой прямых соединений для гетерогенной сети;
- создание алгоритма распределения трафика пользователей в гетерогенной системе прямых соединений.

Научная новизна диссертации состоит в следующем.

1. Модель гетерогенной сети с набором интегрированных между собой технологий радиодоступа, которая, в отличие от известных, учитывает геометрию размещения узлов сети совместно с особенностями обслуживания потоков трафика (сессий) пользователей во времени.

2. Алгоритм с разделением пользовательской сессии для совмещенной сети радиодоступа, который отличается от известных тем, что позволяет вести передачу трафика по нескольким технологиям радиодоступа одновременно с целью снижения нагрузки на сотовую сеть.

3. Модель гетерогенной сети с возможностью установления прямых соединений между устройствами, отличающаяся от известных совместным учетом особенностей обслуживания потоков трафика во времени и геометрии размещения пользовательских устройств.

4. Алгоритм с одновременным подключением пользователя к нескольким узлам сетевой инфраструктуры, который впервые учитывает возможность блокирования канала прямой видимости подвижными препятствиями при осуществлении пользователями передачи на крайне высоких частотах.

5. Модель сети радиодоступа для обслуживания большого числа устройств, которая отличается от известных тем, что принимает во внимание особенности функционирования протокола случайного множественного доступа, стандартизованного для сотовой системы связи.

6. Модель сети радиодоступа для обслуживания множества носимых устройств, отличающаяся от известных учетом особенностей работы протокола случайного множественного доступа, стандартизованного для локальной системы связи на крайне высоких частотах.

7. Протокол управления системой прямых соединений для гетерогенной сети, а также способ его реализации в рамках модельной сети, которые ранее

не применялись для снижения нагрузки на сотовую сеть методом выгрузки трафика в сети других технологий радиодоступа.

8. Алгоритм распределения трафика пользователей в гетерогенной системе прямых соединений, который, в отличие от известных, основан на применении методов сетевого кодирования с целью повышения доступности контента на соседних пользовательских устройствах.

Теоретическая значимость диссертационной работы обусловлена созданием нового научного направления в области разработки и исследования моделей и алгоритмов для обеспечения эффективного функционирования гетерогенных беспроводных сетей. Методология диссертации развивает и объединяет методы теории массового обслуживания и стохастической геометрии, применяя их к современным гетерогенным сетям, и отличается от известных как своим комплексным подходом, так и учетом основных особенностей текущего этапа развития беспроводных систем. В результате предлагаются новые модели организации множественного доступа для пользователей гетерогенных сетей мобильной связи в условиях высокой плотности размещения устройств, а также создаются эффективные алгоритмы управления потоками трафика с целью снижения нагрузки на сотовую сеть передачи данных. Все это вносит существенный вклад, прежде всего, в исследование путей совершенствования управления информационными потоками.

Полученные в работе научные результаты внедрены в Российском университете дружбы народов (РУДН), Москва; Санкт-Петербургском государственном университете телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ), Санкт-Петербург; публичном акционерном обществе «ГИПРОСВЯЗЬ», Москва; научно-исследовательском институте радио (ФГУП НИИР), Москва; обществе с ограниченной ответственностью «ЭВС», Санкт-Петербург. Часть результатов диссертации получена при выполнении ряда крупных научно-исследовательских проектов, в том числе, при исследованиях по грантам РФФИ (16-11-10227), РФФИ (17-07-00845, 17-07-00142, 15-07-03051, 10-08-01071), а также при выполнении НИР в рамках госзадания (2.882.2017/4.6) и при создании в РУДН модельной сети лаборатории опережающих исследований сетей 5G (начиная с 2017 г.).

Практическая значимость диссертации заключается в разработке для научно-исследовательских, проектных организаций и телекоммуникационных компаний обоснованных рекомендаций, таких как «Методика оценки показате-

телей эффективности для проектирования гетерогенных беспроводных сетей» (внедрена в ПАО «ГИПРОСВЯЗЬ»). Результаты работы также легли в основу учебно-методических комплексов для лекционных курсов, постановок задач для выпускных квалификационных работ бакалавров по направлениям подготовки 02.03.01 «Математика и компьютерные науки» и 02.03.02 «Фундаментальная информатика и информационные технологии», магистров по направлению подготовки 02.04.02 «Фундаментальная информатика и информационные технологии», программа «Управление инфокоммуникациями и интеллектуальные системы» на кафедре прикладной информатики и теории вероятностей факультета физико-математических и естественных наук РУДН.

Методология исследования, применявшаяся в диссертационной работе для решения поставленных задач, включает в себя методы теории вероятностей, теории марковских случайных процессов, теории массового обслуживания и математической теории телетрафика. Для проверки основных положений теоретических исследований использовалось имитационное моделирование, а также проводились натурные испытания в рамках модельной сети.

Положения, выносимые на защиту:

1. Разработанная модель гетерогенной сети с набором интегрированных между собой технологий радиодоступа, учитывающая обслуживание потоков трафика (сессий) от множества пользователей и высокую плотность размещения узлов сетевой инфраструктуры, позволяет оценивать параметры качества обслуживания, такие как среднее число одновременно передающих пользователей и вероятность блокировки их сессий.

2. Предложенный алгоритм с разделением пользовательской сессии применим для управления потоками трафика в гетерогенной сети с двумя совмещенными технологиями радиодоступа, которая характеризуется высокой плотностью размещения узлов сетевой инфраструктуры, и позволяет повысить скорость отправки данных пользователями, а также снизить среднее время их передачи.

3. Разработанная модель гетерогенной сети с двумя одновременно функционирующими технологиями радиодоступа и возможностью установления прямых соединений между соседними устройствами при высокой плотности их размещения применима для оценки параметров качества обслуживания, включающих в себя среднее число одновременно передающих пользователей и вероятность блокировки их сессий.

4. Сформулированный алгоритм с одновременным подключением пользователя к нескольким узлам сетевой инфраструктуры в условиях их плотного размещения, учитывающий блокирование канала прямой видимости препятствиями при передаче на крайне высоких частотах, применим для улучшения параметров качества обслуживания, таких как среднее число одновременно передающих пользователей и вероятность блокировки их сессий.

5. Предложенная модель сотовой сети с конкурентным обслуживанием большого числа пользователей, учитывающая особенности протокола множественного доступа, позволяет изучать показатели эффективности функционирования системы связи, включающие в себя среднее число попыток передачи, а также вероятность успешного установления соединения и задержку доступа в сеть.

6. Разработанная модель сети радиодоступа для обслуживания множества плотно размещенных носимых устройств, учитывающая характер потребления ими энергии и особенности протокола конкурентного доступа, применима для оценки показателей эффективности функционирования системы связи и в частности энергетической эффективности передачи потокового трафика пользователей.

7. Предложенный протокол управления системой прямых соединений для гетерогенной сети с интегрированными между собой технологиями радиодоступа, а также способ его реализации в рамках модельной сети с высокой плотностью позволяют снижать нагрузку на сотовую сеть методом выгрузки трафика в сети других технологий радиодоступа, повышая при этом скорость передачи пользователей.

8. Сформулированный алгоритм распределения трафика пользователей в гетерогенной системе прямых соединений с большим количеством устройств, основанный на применении методов сетевого кодирования и учитывающий структуру передаваемых сессий, применим для управления нагрузкой с целью снижения времени передачи пользовательских данных и повышения вероятности успешного доступа.

Степень достоверности и апробация результатов подтверждается корректным применением математического аппарата, результатами имитационного моделирования, а также широким спектром публикаций и выступлений как на российских, так и на международных конференциях. Основные положения диссертационной работы были представлены и обсуждались на следу-

ющих конгрессах, конференциях и семинарах: IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, 7–10 April 2013; IEEE International Symposium on Information Theory (ISIT), Istanbul, Turkey, 7–12 July 2013; Aachen, Germany, 25–30 June 2017; IEEE Vehicular Technology Conference (VTC-Spring), Dresden, Germany, 2–5 June 2013; Glasgow, UK, 11–14 May 2015; Porto, Portugal, 3–6 June 2018; IEEE World Forum on Internet of Things (WF-IoT), Seoul, Korea, 6–8 March 2014; European Wireless, Budapest, Hungary, 20–22 May 2015; Oulu, Finland, 18–20 May 2016; Dresden, Germany, 17–19 May 2017; IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 23–27 May 2016; Kansas City, USA, 20–24 May 2018; IEEE Global Communications Conference (Globecom), Washington DC, USA, 4–8 December 2016; Singapore, Singapore, 4–8 December 2017; Abu Dhabi, UAE, 9–13 December 2018; Международная конференция «Распределенные Компьютерные и Телекоммуникационные Сети: Управление, Вычисление, Связь», Москва, Россия, 17–21 сентября 2018; семинары РУДН, ГУАП, СПбГУТ.

Основные результаты диссертации изложены в 68 опубликованных работах, в том числе в 6 главах в монографиях, в 12 работах, опубликованных в журналах из перечня ВАК Министерства образования и науки Российской Федерации; в 32 работах, опубликованных в трудах, индексируемых Scopus и Web of Science.

Диссертация состоит из введения, пяти глав, заключения и шести приложений. Полный объем диссертации составляет 438 страниц, включая 116 рисунков и 9 таблиц. Список литературы содержит 360 наименований.

Все основные результаты диссертационной работы **получены автором самостоятельно**. Экспериментальные исследования проведены при его участии и под непосредственным научным руководством.

Диссертационная работа выполнена по специальности 05.12.13 – Системы, сети и устройства телекоммуникаций и соответствует следующим пунктам паспорта специальности: 4, 5, 11, 12, 14.

Краткая характеристика содержания работы

В **первой** главе диссертационной работы введены в рассмотрение перспективные сети связи 5G, для которых сформулированы важные задачи по организации непрерывного беспроводного покрытия и повышению качества восприятия услуг как при обслуживании людей в качестве пользователей, так и при поддержке межмашинного взаимодействия. Существующие на сегодняш-

ний день технологии широкополосного доступа 4G не способны предоставить радиосвязь с требуемыми характеристиками ввиду своей ограниченной емкости и отсутствия необходимых средств обеспечения безотказной работы. Таким образом, разработка новых систем 5G требует значительных изменений сетевой инфраструктуры, а также архитектуры абонентских терминалов, вследствие которых они начинают тесно взаимодействовать при осуществлении всевозможных сетевых функций. В данной главе вначале приведена краткая справка о предпосылках появления систем связи следующего поколения и определены основные векторы их развития.

В частности, описаны такие направления исследований 5G, как интеграция разнородных технологий радиодоступа, использование прямого подключения устройств и повышение их энергетической эффективности. С другой стороны, также рассмотрены различные способы организации управления беспроводным спектром через призму их исторически сложившейся фрагментации, которая замедляет желаемые темпы развития разнообразных технологий 5G. Для преодоления данного препятствия представляет интерес более эффективное использование имеющихся радиочастот, а также выделение для нужд операторов мобильной связи новых спектральных диапазонов, таких как крайне высокие частоты. Совокупность этих и ряда других мер позволяет увеличить как емкость будущих беспроводных сетей доступа, так и объем предоставляемых ими спектральных радиоресурсов.

Далее, в этой главе приведено описание некоторых перспективных технологий 5G и рассмотрен важный процесс интеграции сетей доступа нового поколения с Интернетом вещей. В частности, проведено различие между потребительскими и промышленными услугами связи в рамках конвергентной экосистемы 5G-IoT. Особый упор при этом сделан на более сложные в обслуживании приложения так называемого Индустриального Интернета, которые предъявляют повышенные требования к задержке, доступности и надежности беспроводной связи. С учетом этих и других потребностей, все многообразие предполагаемых сценариев использования технологии 5G-IoT можно условно разделить на три класса с крайне различными характеристиками: (1) расширенный мобильный широкополосный доступ; (2) массовое межмашинное взаимодействие и (3) надежное взаимодействие с низкой задержкой.

Для понимания основных свойств этих трех классов целесообразно различать массовые сценарии Интернета вещей, в которых применяется сравни-

тельно недорогое оборудование (сенсоры, силовые приводы, интеллектуальные измерители, носимые устройства и т.д.), и критические его сценарии, использующие существенно более сложные решения (например, подключенные автомобили, мобильные роботы и дроны). Обеспечение требований таких классов, рассмотренных в данной главе относительно процесса стандартизации сетей связи 5G, является трудоемкой задачей, требующей тесного взаимодействия и интеграции разнородных технологий радиодоступа. Для ее решения необходимо согласованное использование гетерогенных радиотехнологий, которое подробно рассмотрено в последующих главах для ряда типовых сценариев и особенностей перспективных систем 5G.

Во **второй** главе диссертации рассмотрены методы анализа и управления, которые позволяют повысить качество соединения при обслуживании пользователей в гетерогенных сетях 5G. В частности, обсуждены различные варианты организации архитектуры сети связи, а также соответствующие алгоритмы выбора технологии радиодоступа, функционирующие как на стороне абонентских терминалов, так и на стороне сетевой инфраструктуры. Также проведено сравнение эффективности новых способов управления с сетевым содействием и существующих более традиционных подходов. Основной акцент сделан на изучение характеристик передачи данных по линии вверх, которое является более сложным ввиду воздействия интерференции.

Основное различие методов управления с сетевым содействием в сравнении с подходами, связанными с выбором предпочтительной радиотехнологии (на стороне абонентских терминалов) или же с увеличением эффективной зоны покрытия соты (на стороне сети доступа), состоит в том, что последние не в полной мере учитывают текущую интерференционную ситуацию и загрузку сети доступа. Приведенные результаты моделирования показали, что алгоритм управления с сетевым содействием, отслеживающий загрузку системы связи, существенным образом улучшает ее производительность по сравнению с известными подходами. При этом увеличиваются скорость передачи и энергетическая эффективность абонентских терминалов, в особенности тех, которые расположены на краю соты.

Проведенное сравнение схем с сетевым содействием и более сложных механизмов, функционирующих на стороне сетевой инфраструктуры, выявило дополнительные преимущества распределенных методов управления в различных сценариях размещения гетерогенных систем 5G. Полезное направление даль-

нейших исследований состоит в изучении алгоритмов управления на стороне абонентских терминалов, которые в явном виде отслеживают колебания загрузки в сети связи, а также учитывают тип передаваемого трафика. Кроме того, могут быть более точно оценены параметры системы доступа при наличии интерференции от сторонних пользователей, что может потребовать усовершенствованных схем гистерезисного управления для борьбы с неопределенностью, возникающей при оценке скорости передачи в этом случае.

На основании результатов имитационного моделирования в данной главе также предложен подход к математическому моделированию интегрированной гетерогенной системы 5G, который учитывает размещение узлов сети в пространстве и динамику их трафика во времени. Проведенный анализ основан на введении в рассмотрение соответствующего марковского процесса и последующем укрупнении его состояний для получения стационарного распределения вероятностей, а также нахождения характеристик работы слоев макро-, пико- и WLAN как по отдельности, так и совместно. Далее изучены практически важные способы размещения узлов сетевой инфраструктуры и абонентских терминалов, основанные на применении пуассоновских точечных и кластерных процессов. Предложенный подход позволяет учитывать дополнительные технологии радиодоступа и алгоритмы управления.

Непосредственная интеграция и совмещение технологий радиодоступа LTE и WiFi, рассмотренные в этой главе, имеют ряд неоспоримых преимуществ. Например, сотовые системы связи предлагают продвинутые механизмы организации мобильности терминалов, тарификации услуг и обеспечения информационной безопасности, тогда как локальные сети доступа имеют высокую скорость передачи, низкую стоимость соединения и широко распространены. Ожидается, что взаимодействие систем LTE и WiFi позволит эффективно использовать эти специфические особенности, а их интеграция станет еще более тесной в рамках сетей связи 5G вплоть до повсеместного применения совмещенных LTE-WiFi модулей.

В **третьей** главе отмечено, что системы связи 5G будут объединять большое количество различных механизмов и технологий для увеличения своей емкости, а также повышения качества пользовательского соединения. В рамках данного направления основной является тенденция к сокращению фактической зоны обслуживания соты, что может быть достигнуто несколькими способами. Важной альтернативой непосредственному уплотнению сетевого покрытия ста-

новится выгрузка сотового трафика на прямые соединения D2D, позволяющая избежать удорожания инфраструктуры мобильного оператора. В этой главе проведено всестороннее исследование возможности использования прямых соединений «устройство – устройство», а также получены соответствующие характеристики работы сети D2D с сетевым содействием со стороны сотовой инфраструктуры.

При этом особое внимание уделено тому факту, что сотовые системы связи могут использоваться недостаточно в течение длительных интервалов времени, несмотря на свою высокую среднюю загрузку. Как следствие, применен пространственно-временной подход к математическому моделированию сети доступа на уровне сессий с поддержкой соединений D2D. В результате показано, что местоположения абонентских терминалов по отношению друг к другу в значительной степени влияют на производительность системы в целом. В этой связи разработанная модель множественного доступа включает в себя пространственную составляющую для учета размещения узлов сети, а также временную составляющую для отслеживания динамики их трафика и позволяет изучать параметры системы D2D, включая распределение принятой мощности и возникающей при этом интерференции.

Для выяснения практических преимуществ от использования соединений D2D с сетевым содействием в данной главе сконструировано несколько типовых сценариев их применения в современных городских условиях. Например, рассмотрен стандартный сценарий ITU/3GPP, известный как «городские малые соты», с достаточно большим числом абонентских терминалов. Такое плотное размещение устройств выбрано для воссоздания условий, в которых соединения D2D могли бы принести существенную пользу. Полученные таким образом имитационные результаты согласуются с данными аналитического моделирования, повторяя динамику взаимодействия сотовой сети 3GPP LTE для организации инфраструктурных соединений и радиотехнологии IEEE 802.11 (WiFi-Direct) для реализации системы D2D. Они также подтверждают исходное предположение о значительном приросте производительности при выгрузке сотового трафика на прямые соединения.

Далее в этой главе рассмотрено сравнительно недавнее направление исследований, связанное с применением миллиметровых волн для организации соединений в сетях 5G, когда возможность одновременного подключения абонентского терминала к двум и более сотам становится ключевой для поддержа-

ния непрерывности пользовательского соединения. Несмотря на наличие в литературе большого количества технических сведений, посвященных отдельным аспектам обеспечения связи mmWave, изучение ее особенностей в реалистичных условиях применения, а также с точки зрения динамики работы по передаче сессий в настоящее время не завершено. В данной главе отмеченный пробел восполнен результатами имитации распространения радиосигнала в типовом городском сценарии, а также данными детального моделирования работы сети mmWave в целом.

Соответственно, изучена система связи mmWave на уровне сессий, что позволило получить характеристики работы как абонентских терминалов, так и сетевой инфраструктуры. Приведенный анализ рассмотрел функционирование соединений mmWave в реалистичных условиях городской застройки, используя при этом сведения, полученные при помощи имитатора распространения миллиметровых волн. В частности, были учтены особенности передачи радиосигнала на крайне высоких частотах, динамическое блокирование соединений mmWave препятствиями, соизмеримыми с телом человека, а также соответствующие алгоритмы назначения радиоресурсов и управления одновременным подключением. Полученные тем самым результаты позволили сделать выводы о целесообразности применения перспективных сетей mmWave в составе системы 5G, а также оценить ожидаемые параметры качества их работы.

В **четвертой** главе подчеркнута важная тенденция к интеграции сетей 5G с Интернетом вещей, тесно связанная с бурным ростом разнообразия устройств, повсеместно вступающих в межмашинное взаимодействие. Соответственно, технологии радиодоступа эволюционируют с целью обеспечения передачи трафика M2M наряду с доставкой данных от людей в качестве пользователей. Для обслуживания M2M могут использоваться либо специализированные радиотехнологии (например, сотовые системы M2M), либо различные типы данных могут объединяться между собой (например, в услугах мобильного широкополосного доступа). Так или иначе, M2M-приложения занимают центральное место в быстро развивающейся экосистеме 5G-IoT, а их поддержка требует решения принципиально новых задач, связанных, прежде всего, с огромным многообразием устройств и их требований.

Для понимания особенностей межмашинного взаимодействия в данной главе проведено моделирование сценария, предполагающего доступ большого числа M2M-устройств в сеть посредством сотовой системы 3GPP LTE. В резуль-

тате многочисленных коллизий между передачами от различных устройств, вызванных спецификой работы соответствующей процедуры множественного доступа, сеть может испытывать существенные перегрузки, приводящие к значительному повышению энергопотребления устройств. Предложенный математический подход позволил всесторонне проанализировать такую систему с учетом ее основных характеристик, а также осуществить эффективное управление ее параметрами, в итоге ослабляя пагубные последствия внезапных перегрузок. Данные результаты были использованы комитетом 3GPP при разработке теперь уже стандартных механизмов защиты сети доступа от перегрузки.

В продолжение изучения особенностей межмашинного взаимодействия подробно рассмотрены характеристики носимых устройств для сценариев с высокой плотностью размещения. Выяснилось, что традиционные радиотехнологии, такие как WiFi или Bluetooth, оказываются неспособны обеспечить обслуживание в подобных массовых сценариях ввиду своей ограниченной емкости. При этом перспективные решения, функционирующие на крайне высоких частотах (например, в диапазоне 60 ГГц), обладают достаточной емкостью, но испытывают недостаток механизмов управления для контроля интерференции, возникающей между персональными сетями соседних абонентов, и это обстоятельство приводит к повышенному энергопотреблению носимых устройств. Устранение данного недостатка требует разработки эффективных средств координирования интерференции и совместного планирования радиоресурсов.

Чтобы расширить возможности существующих протоколов mmWave-доступа по обслуживанию сценариев с высокой плотностью размещения носимого оборудования, в этой главе проведено соответствующее моделирование mmWave-системы носимых устройств. Его основной задачей стало исследование характеристик производительности сети носимых устройств на системном уровне. Для этих целей разработан новый метод анализа высокоскоростных систем носимых устройств, учитывающий варианты формирования диаграммы направленности, способы управления интерференцией и алгоритмы планирования радиоресурсов. В отличие от других аналогичных моделей, предложенный подход учел особенности организации доступа к среде с упором на пространственно-временную динамику потоков данных и наличие соответствующих процедур управления системой связи.

Предлагаемый таким образом метод анализа можно применить, прежде всего, к технологии IEEE 802.11ad/ay и подобным решениям, позволяя улуч-

шить их работу при обслуживании системы носимых устройств, размещенных с высокой плотностью (например, в переполненном общественном транспорте). В частности, можно преодолеть имеющиеся ограничения, связанные с повышенным энергопотреблением носимого оборудования, при помощи более эффективных алгоритмов организации доступа большого числа устройств, а также более продвинутых способов формирования направленной передачи. В результате использование предложенного здесь подхода к моделированию системы в целом позволило детально исследовать плотные сети носимых устройств на основе mmWave-соединений, выяснить достижимые границы их практической производительности, а также улучшить характеристики их работы.

В **пятой** главе продемонстрировано, что связь D2D, управляемая через систему 3GPP LTE, позволяет существенно повысить производительность как сети в целом, так и отдельных абонентских терминалов. Однако фактическая величина выигрыша зависит от (1) используемого критерия для осуществления выгрузки трафика (соединение D2D должно обладать достаточным качеством и не наносить вреда другим, уже активным соединениям) и (2) наличия достаточного числа взаимодействующих пар устройств, удовлетворяющих заданному критерию выгрузки. При этом применение соединений D2D без эффективного критерия выгрузки может привести к снижению качества обслуживания и даже росту загрузки сети.

Полученные в данной главе результаты подчеркнули высокую вариативность качества D2D-соединений (даже в случае одинаковой их длины) по сравнению с каналами связи 3GPP LTE. В отличие от сотовой сети доступа, в которой скорость передачи данных, как правило, зависит от длины соединения монотонно (при наличии адекватного планирования), в системе D2D «пригодные» соединения должны быть тщательно отобраны для обеспечения надлежащего качества обслуживания. При наличии механизмов сетевого содействия оператор мобильной связи располагает средствами отбора подходящих D2D-соединений, позволяющих повысить скорость передачи данных и сохранить справедливость распределения радиоресурсов. Это достигается, например, использованием соединений D2D, подверженных наименьшей интерференции, а также переводом сессий обратно на сотовые каналы доступа при недостаточном качестве прямой связи.

В таких условиях показано, что использование методов сетевого кодирования обеспечивает значительный рост производительности системы доступа,

особенно когда ее текущая загрузка близка к максимальной емкости. Повышая доступность целевого контента на соседних устройствах, сетевое кодирование позволило поддержать большее число D2D-соединений, что увеличило долю выгружаемого на них трафика. На основании проведенного в этой главе моделирования на системном уровне было установлено, что связь D2D с сетевым кодированием заметно улучшает работу сети доступа при повышении ее загрузки, а также плотности размещения устройств. Несмотря на определенные накладные расходы, связанные с ростом сложности декодирования сообщений, выяснилось, что в практических сценариях наиболее целесообразно применение простейших форм сетевого кодирования.

В дальнейшем результаты проведенного исследования могут успешно применяться при решении двух важных задач. Во-первых, они упрощают разработку эффективных схем распределения контента, позволяющих организовать хранение и распространение востребованной информации так, чтобы применять наибольшее возможное число D2D-соединений. Во-вторых, полученные сведения позволяют установить предельные характеристики системы, использующей прямые соединения для обмена данными, включая ее максимальную емкость и допустимую плотность размещения узлов. Такой анализ полезен для подготовки и предоставления услуг 5G, которые опираются на кэширование и распределение данных по D2D с сетевым содействием. Вообще говоря, кэширование информации на абонентских терминалах и на стороне сети доступа предоставляет мобильным операторам широкую свободу действий.

С учетом данной возможности в этой главе разобран функционал интегрированного узла 5G для вычисления и хранения данных, а также приведены доводы в пользу такого подхода. На основании предложенного технического описания представляет интерес рассмотрение следующих ключевых задач: выбор контента и его расположения при кэшировании, а также выделение достаточного объема вычислительных ресурсов как на пользовательских устройствах, так и на узлах сетевой инфраструктуры. Их решение позволяет обеспечить работу перспективных приложений 5G, таких как сервисы дополненной реальности и услуги удаленных вычислений. Для наглядной классификации подобных сценариев было предложено различать «вытягивающие» и «выталкивающие» сервисы, т.к. они предъявляют принципиально различные требования к качеству обслуживания и характеристикам связи 5G.

Глава 1. Основные понятия, определения и структура гетерогенных сетей связи 5G

1.1 Предпосылки появления сетей связи 5G и направления их развития

1.1.1 Истоки возникновения сетей связи пятого поколения

Стремительное распространение систем беспроводной связи в последние десятилетия в значительной степени изменило возможности доступа к сети Интернет, сделав его практически повсеместным, что повлекло за собой необходимость решения целого ряда исследовательских задач. Разработка мобильных технологий широкополосного доступа четвертого поколения 4G была завершена уже в 2011 году и привела к существенным улучшениям во всех аспектах работы беспроводной системы связи, таких как ее емкость, энергетическая эффективность и показатели качества обслуживания [1]. По мере того как соответствующие сотовые технологии получали все более широкое распространение на практике, внимание исследовательского сообщества своевременно сместилось в сторону проектирования и разработки мобильных сетей пятого поколения 5G.

С учетом исторически сложившегося 10-летнего цикла развития для каждого существующего поколения мобильной связи не удивительно, что создание систем 5G планируется к завершению в 2020 году. Однако первоначально возникли трудности с четким определением того, что должно было последовать за технологиями 4G, и сообщество разработчиков интерпретировало технические требования к новому поколению сетей, в основном, с точки зрения конечных пользователей. Независимо от их текущего местоположения, абоненты нуждаются в надежном беспроводном подключении, а также в доступе к широкому кругу современных мультимедиа-приложений. В этих целях потребовались перспективные наработки, которые предоставили бы возможность иметь целевые скорости передачи данных для достижения адекватного качества получения услуг пользователем.

Повсеместно доступное беспроводное покрытие, предоставляющее требуемое качество обслуживания, стало, таким образом, начальной задачей при разработке систем 5G. Ее решение позволило бы существенно обогатить повседневный опыт использования сетей мобильной связи в свете все возрастающей загрузки данными, большого количества подключенных абонентских терминалов, а также значительно варьирующихся требований беспроводных приложений и услуг. В частности, ожидалось, что технологии связи 5G позволят как отдельным людям, так и их сообществам органично взаимодействовать друг с другом, а также с окружающим их информационным полем и всевозможными сервисами. Такое начальное видение предполагало, что «умное» устройство пользователя сможет взаимодействовать с «умным» облаком для предоставления индивидуального беспроводного сервиса.

К сожалению, современные на тот момент системы 4G не имели возможности для предоставления требуемого повсеместного подключения к сети. Им недоставало единообразия в скоростях передачи данных, задержки зачастую превосходили допустимые, и нередко случались отказы в обслуживании из-за низкого качества беспроводного соединения и повышенного уровня интерференции. В итоге, имеющиеся радиотехнологии были признаны неспособными или же недостаточными для удовлетворения целевых требований, и начался активный поиск новых решений. Сложившаяся ситуация усугублялась по мере того, как миллиарды различных устройств Интернета вещей подключались к сотовым системам доступа, изменяя привычный облик сети Интернет [2]. Все эти технологические проблемы послужили предпосылками к созданию новых беспроводных решений при разработке технологий 5G.

Достижение нового уровня качества обслуживания пользователей требовало основательных изменений как в построении сетевой архитектуры, так и в архитектуре самого абонентского оборудования. Работа по соответствующим направлениям быстро набрала скорость в исследовательском сообществе [3–5]. Отражая широкое разнообразие всевозможных устройств и услуг, от мобильного телефона до подключенного автомобиля, надо было принять во внимание различные альтернативные решения от новых механизмов доступа к беспроводному спектру до продвинутых схем взаимодействия между устройствами. При этом лицензированный спектральный радиоресурс оставался весьма ограниченным и дорогим в эксплуатации, тогда как традиционные способы повышения эффективности его использования приближались к своему теоретическому

пределу. В этой связи существенного прироста производительности следовало ожидать от использования нелицензированного и общего спектра, что стало возможно при совместном использовании гетерогенных технологий радиодоступа.

Таким образом, в начале своего бурного развития технологии связи 5G стремились, в основном, к (1) существенному увеличению доступной емкости беспроводной сети; (2) предоставлению единообразного качества беспроводного соединения, а также (3) повышению уровня качества обслуживания и восприятия услуги. Для достижения этих перспективных целей в числе прочих были изучены подходы к интеграции различных технологий радиодоступа и уплотнению сети, использованию прямых соединений между пользовательскими устройствами и крайне высоких частот, повышению энергетической эффективности при подключении устройств Интернета вещей и носимых устройств. Эти основные направления исследований, рассматриваемые ниже вкратце и далее более подробно в рамках отдельных глав данной работы, привели к быстрому становлению гетерогенных сетей связи 5G.

Сети 5G представляют собой перспективную архитектуру систем беспроводного доступа (см. рисунок 1.1), основанную на использовании различных малых сот и предназначенную для удовлетворения новых требований по емкости и качеству покрытия сети. Такие гетерогенные многослойные покрытия предполагают наличие базового слоя макросот с дополнительными слоями более плотно размещенных сот меньшего размера, таких как пикосоты, фемтосоты, точки доступа WiFi и т.д. Соответственно, базовая станция макросети служит для обеспечения сетевого подключения и поддержки мобильности терминалов на обширных площадях, тогда как недорогие малые соты применяются для повышения качества и емкости сотового покрытия. Мобильные устройства, снабженные возможностями для установления прямого соединения [6], могут также рассматриваться в качестве одного из слоев в рамках такой иерархической гетерогенной структуры.

1.1.2 Основные векторы развития радиотехнологий 5G

Интеграция различных технологий радиодоступа. Наряду с емкостью сети и скоростью сотового соединения, равномерность сетевого покрытия

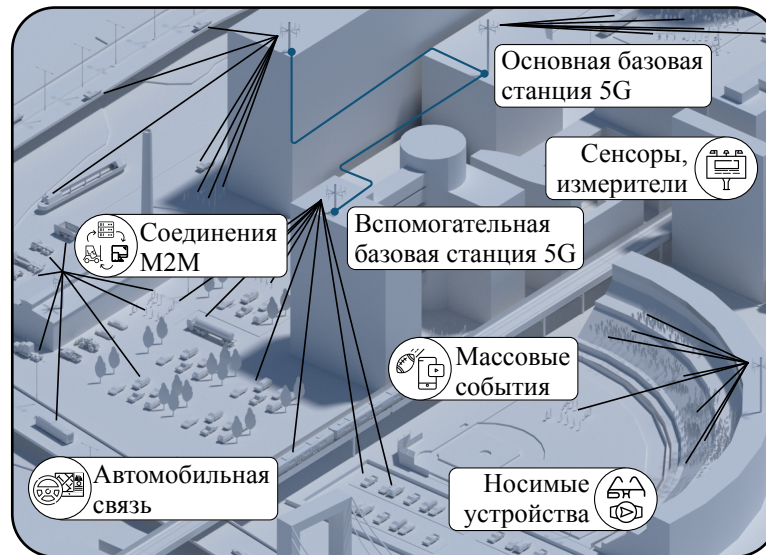


Рисунок 1.1 — Особенности перспективных гетерогенных сетей 5G

является важным фактором, влияющим на качество обслуживания и восприятия услуг пользователем. Абонентские терминалы, размещенные на краю соты, обычно имеют беспроводное соединение более низкого качества ввиду значительной удаленности от обслуживающей их базовой станции, а также из-за повышенного уровня интерференции [7]. Надежность сотового покрытия внутри помещений также остается неудовлетворительной. Нацеленные на обеспечение более единообразного и бесшовного соединения, перспективные системы 5G вынуждены применять агрессивные механизмы переиспользования спектрального ресурса и продвинутые средства управления интерференцией [8]. Поскольку данные технологии были предложены сравнительно недавно, границы их применимости в практических гетерогенных сетях еще не установлены окончательно.

Другим немаловажным аспектом является тот факт, что сотовые системы в лицензированном спектре часто совмещаются географически с сетями, функционирующими на нелицензированных частотах [9] (например, WiFi [10]). С другой стороны, современные абонентские терминалы приобретают возможность использования нескольких технологий радиодоступа одновременно [11; 12]. Все возрастающая популярность таких устройств с поддержкой различных радиотехнологий позволяет, в частности, разгружать сотовые сети при помощи прямых соединений между устройствами в нелицензированном спектре [13]. С учетом ограниченной полосы лицензированных частот возрастает необходимость в эффективной координации всевозможных гетерогенных тех-

нологий радиодоступа. В этой связи распределенные сети в нелицензированном спектре могут воспользоваться поддержкой со стороны централизованных сотовых систем, использующих лицензированные частоты [14].

По мере того как сокращаются зоны покрытия малых сот, эффективные радиусы действия сотовых, локальных и персональных сетей доступа начинают в значительной степени перекрываться. Данное обстоятельство создает возможность одновременного использования нескольких радиотехнологий для повышения качества беспроводной связи. Однако такому совместному использованию сетей доступа было уделено значительно меньше внимания, чем оптимизации работы отдельных радиотехнологий. Таким образом, интеграция различных технологий радиодоступа становится важным направлением исследования в сетях 5G, особенно в свете тенденции к уплотнению сотового покрытия. В результате беспроводные технологии ближнего и дальнего радиуса действия должны взаимодействовать более тесно для достижения желаемых показателей качества обслуживания и восприятия услуги.

Недавние исследования автора, приведенные в главе 2 данной работы, направлены на изучение совместного использования нескольких технологий радиодоступа для повышения производительности гетерогенных сетей 5G. В частности, представляет интерес рассмотрение эффективных методов управления трафиком, связанных с выбором радиотехнологии как на стороне абонентского терминала, так и непосредственно инфраструктурной сотовой сетью. Как следствие, повышается эффективная скорость передачи данных и снижается амплитуда ее колебаний, а средняя спектральная эффективность сохраняется при этом на приемлемом уровне. Интеграция различных технологий радиодоступа в сверхплотных гетерогенных сетях становится крайне востребованным направлением исследований не только для поддержки множества традиционных сценариев мобильной связи, но также для обеспечения перспективных приложений Интернета вещей [15; 16].

Использование прямого подключения устройств. Согласно ряду прогнозов объемы мобильного трафика, передаваемого по беспроводным сетям связи, продолжают существенно возрастать на протяжении последующих лет [17]. Соответственно, развернутые в настоящее время системы широкополосного доступа будут подвержены значительным перегрузкам, вызванным недостаточной емкостью радиосетей и приводящим к резкому снижению качества обслуживания их пользователей. Однако ожидается, что существенного приоро-

ста в спектральной эффективности удастся добиться при использовании прямого подключения абонентских устройств. Вообще говоря, применение связи «устройство – устройство» позволяет кардинально повысить степень переиспользования радиоресурсов в пространстве [5]. В частности, соседние абонентские терминалы могут взаимодействовать напрямую, без вовлечения сетевой инфраструктуры в процесс передачи своих данных.

Вообще говоря, использование прямого соединения пользовательских устройств в рамках централизованной архитектуры сотовой сети становится новым направлением для повышения ее спектральной эффективности [18]. При этом считается, что при помощи интенсивного взаимодействия между абонентскими терминалами можно достичь линейного роста емкости системы связи с увеличением числа задействованных устройств [19]. Это обстоятельство позволяет значительным образом повысить доступную сетевую емкость в перспективных системах 5G. Однако данное направление исследований получило недостаточное внимание в имеющейся технической литературе. Отдельные известные в этой области подходы [20; 21] не учитывают в должной мере множество важных факторов, таких как реальная архитектура сети, практические механизмы обеспечения ее работы, а также характеристики беспроводного канала связи.

Для преодоления ограничений, свойственных имеющимся результатам, в главе 3 проводится всестороннее изучение прямых соединений в перспективных сетях связи 5G. Детально рассматриваются фундаментальные подходы к анализу таких систем относительно их пространственно-временной динамики, а также практические методы управления выгрузкой трафика из инфраструктурной сотовой сети на соединения «устройство – устройство». Данное направление работы приводит к более полному пониманию потенциала гетерогенных сетей 5G, снабженных функционалом прямого подключения устройств, в плане их доступной емкости, качества беспроводного покрытия и надежности установленного соединения. При проведении соответствующего исследования применяются математический анализ и имитационное моделирование.

Полученные сведения позволяют разрабатывать практические механизмы разгрузки сотовых систем 5G при помощи прямого взаимодействия терминалов. Таким образом, результатом проведенного исследования становится значительное увеличение фактической скорости передачи данных, а также повышение пиковой скорости обмена информацией по сравнению с аналогичными показателями в системах, лишенных поддержки прямых соединений. Само по себе

беспроводное взаимодействие «устройство – устройство» является важным режимом функционирования системы не только для передачи мобильного трафика в перспективных сетях, ориентированных на обслуживание традиционных пользователей, но и для обеспечения широкого круга сценариев, возникающих при осуществлении межмашинного взаимодействия в рамках концепции Интернета вещей [22; 23].

Повышение энергетической эффективности устройств. Возрастающая распространенность мультимедиа-приложений и сервисов, порождающих объемные потоки данных, а также более высокая сложность конструкции абонентских терминалов, поддерживающих несколько технологий радиодоступа, в совокупности приводят к значительному энергопотреблению пользовательских устройств. Более того, существует целый ряд сценариев, включающих в себя широкое множество приложений Интернета вещей, в которых тысячи устройств подключаются к одной и той же соте, где высокая энергетическая эффективность работы оконечного оборудования является определяющей. Необходимость поддержки таких массовых сервисов создает дополнительные трудности при оптимизации межмашинного взаимодействия [24]. С учетом ограниченного роста емкости аккумуляторной батареи высокая энергетическая эффективность устройств является важным требованием в сетях 5G.

В то время как подавляющее большинство исследований в развитии технологий 5G было нацелено на повышение спектральной эффективности, рассмотрение энергетической эффективности становится все более актуально в тех сценариях применения сетей 5G, которые ориентированы на обслуживание терминалов с крайне ограниченным ресурсом аккумуляторной батареи [25]. Это, например, устройства Интернета вещей и носимое оборудование. Необходимость обеспечения непрерывного беспроводного подключения не должна при этом приводить к быстрому истощению батареи малогабаритных устройств, а увеличение времени их работы без подзарядки является важной практической задачей. С учетом данного ограничения возникают нетривиальные обменные соотношения между характеристиками производительности отдельных устройств и сети в целом [26; 27].

В главе 4 данной работы исследуются свойства систем 5G, связанные с энергетической эффективностью оконечного оборудования в типовых сценариях Интернета вещей и с учетом специфики современных носимых устройств. Целью исследования здесь является не только продление времени жизни бата-

реи для обеспечения более высокого качества обслуживания, но и повышение эффективности расходования энергетического ресурса в более широком контексте [27; 28]. В частности, полученные результаты могут быть использованы для разработки схем совместного управления мощностью передачи и планированием радиоресурсов с учетом уровня энергетической эффективности при динамической загрузке. При этом важно соблюдать баланс между показателями энергетической и спектральной эффективности системы 5G в различных режимах ее работы.

Характеристики функционирования сетей связи 5G, отражающие снижение энергопотребления подключенных к ним устройств, как правило, определяются соотношением между достигнутой скоростью передачи данных и потребленной при этом мощностью (в битах на Джоуль). В этой связи интерес представляет разработка новых методов повышения энергетической эффективности в системах 5G и проведение соответствующей оптимизации. Это становится особенно важно при осуществлении межмашинного взаимодействия, которое связано с работой малогабаритного оборудования (сенсоры, силовые приводы, интеллектуальные измерители, носимые устройства и т.д.). Использование устройств Интернета вещей становится все более распространенным во многих областях жизнедеятельности, что выдвигает задачу повышения энергетической эффективности такого оборудования на передний план.

1.2 Возможности использования беспроводного спектра

1.2.1 Различные способы организации управления спектром

С течением времени ограниченный беспроводной спектр стал ключевым ресурсом, применяющимся для самых различных целей: от экономических и социальных до научных и культурных. Однако подходы к управлению спектральными радиоресурсами не претерпели значительных изменений на протяжении последних трех десятилетий, в основном, в силу сложности соответствующего процесса и недостаточной зрелости технологий радиодоступа. Это также связано с необходимостью контролировать множество параметров радиоспектра,

включая место и время его использования, радиочастоты, излучаемую мощность и возникающую при этом интерференцию. Координация последней при работе различных пользователей спектра требует особого внимания и всегда оставалась предметом строгого регулирования. Соответственно, известен целый ряд традиционных способов организации управления спектром [29].

«Командный» способ управления спектром. Этот уже устаревший подход, предполагающий статическое выделение спектральных радиоресурсов, возник в связи с ограничениями для исторически первых технологий связи. Здесь регулирующий орган производит назначение радиочастот некоторому пользователю, накладывая при этом жесткие ограничения на их использование. Примеры такого способа управления спектром встречаются в правительственной и военной сфере, а также в системах национальной безопасности, астрономии и авиации. Естественно, использование «командного» подхода зачастую приводило к неэффективному использованию спектра, вызывая трудности в организации работы соответствующих услуг беспроводной связи. Кроме того, назначение радиочастот не было основано на рыночных механизмах ввиду отсутствия конкуренции, что приводило в итоге к недостаточной экономической эффективности. Традиционно в рамках данного подхода права на владение спектром выдавались на конкурсной основе и требовали значительного лоббирования в органах власти. В последнее время вовлеченные в этот процесс правительственные службы лучше понимают значительный рыночный потенциал беспроводного спектра и все чаще используют более гибкие способы его назначения.

Эксклюзивное пользование спектром. Этот способ управления спектром основан на его долгосрочной (порядка 15–30 лет) аренде, закрепляющей определенную полосу частот для использования в четко оговоренных целях. В частности, лицензия на пользование спектральными радиоресурсами задает допустимый диапазон частот, область их использования, а также мощность передачи. В результате работа с радиоспектром подчинилась набору правил, регламентирующих максимальный уровень излучаемой мощности и географический охват территории вещания. Эксклюзивные права на пользование спектром позволяют их владельцам (например, операторам мобильной связи) применять собственные механизмы управления интерференцией. Это приводит к возможности гарантировать определенный уровень качества обслуживания их абонентов, но при этом создает высокий барьер при входе на данный рынок

(порядка миллиардов евро). В отличие от устаревших конкурсов назначение радиочастот регулируется здесь непосредственной продажей спектра: большинство вовлеченных регламентирующих органов на сегодняшний день использует гибкие рыночные подходы (например, аукционы) для распределения подконтрольных им частотных диапазонов.

Общее пользование исходно лицензированным спектром. Согласно данной концепции частотные диапазоны некоторого лицензированного владельца (называемого первичным пользователем) могут быть предоставлены в совместное пользование одному или нескольким нелицензированным (вторичным) пользователям. Важно отметить, что доступ вторичных систем к спектру общего пользования может в ряде случаев осуществляться без оповещения первичного пользователя и требует определенных мер защиты со стороны последнего во избежание нарушения работы первичной системы связи. Этот способ организации управления спектром породил всплеск работ, связанный с исследованием программно-определяемых технологий радиодоступа, когнитивных и адаптивных систем, а также перестраиваемых (реконфигурируемых) сетей, в целях сохранения желаемых свойств динамического доступа к спектру (например, использование полосы защитных интервалов диапазона аналогового телевидения – TV white space). Однако принципиальным ограничением такого подхода к управлению спектром является невозможность предоставления вторичными системами связи каких-либо гарантий относительно качества обслуживания их абонентов в таком спектре общего пользования.

Общее пользование нелицензированным спектром. Когда некоторая полоса радиочастот предоставляется в открытый доступ, ни один из пользователей не может получить ее в свое исключительное распоряжение, так как соответствующие спектральные радиоресурсы должны оставаться доступны для справедливого общего пользования. Примерами подобных частотных диапазонов являются полосы для промышленного, научного и медицинского применения ISM, в которых множество потенциальных пользователей (например, медицинские и сенсорные устройства, микроволновые печи, беспроводные телефоны, сети WiFi и т.д.) получают совместный доступ к беспроводному спектру без дополнительных препятствий. Несмотря на то, что такой нерегулируемый доступ к радиочастотам значительно упрощает вход на соответствующий рынок, он также создает риск возникновения неконтролируемой интерференции, которая, в свою очередь, усложняет обеспечение требуемого качества обслужи-

вания абонентов. Кроме того, множество пользователей общего спектра может попасть в критическую ситуацию, в которой ни один из них не может достичь удовлетворительного результата. Такой результат крайне вероятен в настоящее время, когда бурный рост популярности сетей WiFi и использование их операторами мобильной связи для разгрузки своих сотовых сетей угрожают качеству работы систем, функционирующих на нелицензированных частотах.

1.2.2 Изменение беспроводного ландшафта в эпоху 5G

Для преодоления негативных последствий длительной фрагментации в способах организации управления спектром, которые на сегодняшний день приводят к его крайне неравномерной доступности, возникает острая необходимость в разработке новых механизмов организации доступа к радиочастотам общего пользования [30], а также в предоставлении дополнительных спектральных радиоресурсов для нужд систем мобильной связи. Данное обстоятельство становится особенно важным в свете неуклонного увеличения нагрузки на радиочастоты, связанного со все возрастающими объемами мобильного трафика. Основными стимулами роста количества передаваемых через сети связи пользовательских данных являются высокая популярность мультимедиа-приложений, широкое разнообразие моделей абонентских терминалов, а также большое многообразие ресурсоемких приложений и сервисов [31]. Ожидается, что данные факторы только усугубят текущую ситуацию с течением времени, т.е. в эпоху развития технологий 5G необходимыми становятся срочные меры по повышению емкости сетей доступа и объема спектральных радиоресурсов.

Эти и другие обстоятельства, связанные с увеличением объемов мобильных данных, изложены в отчете ITU-R M.2370, который содержит глобальные прогнозы по росту трафика, собранные из нескольких источников. Соответственно, следует ожидать, что нагрузка на системы международной мобильной связи ИМТ возрастет от 10 до 100 раз с 2020 по 2030 год, и с ней могут не справиться имеющиеся сети 4G. Как следствие, повышение эффективности использования уже доступного системам ИМТ беспроводного спектра может быть недостаточно для достижения требуемых показателей производительности. В этой связи представляется важным как назначение новых радиочастот для экс-

клюдивного использования беспроводными сетями доступа, так и более эффективное применение частот общего пользования для нужд операторов мобильной связи.

В результате обусловленные повсеместным распространением всевозможных пользовательских устройств и готовящиеся обслуживать произведенный ими мобильный трафик гетерогенные сети связи 5G сочетают в себе целый ряд усовершенствований и новых функций, предназначенных для повышения производительности современных беспроводных решений [32]. Они включают в себя более тесную интеграцию различных технологий радиодоступа и уплотнения сети, использование прямых соединений между устройствами для разгрузки сотовой сети и крайне высоких частот для повышения скорости передачи данных, механизмы конвергенции с Интернетом вещей и поддержки носимого оборудования, а также методы передачи с применением массивных многоантенных систем ММО, технологии одновременной передачи и приема данных и многие другие.

Из всех приведенных выше решений радикальное уплотнение сети, массивные системы ММО и связь на крайне высоких частотах, т.е. миллиметровых волнах (mmWave) были названы «большой тройкой» технологий 5G [33]. Решение применять mmWave становится при этом основным нововведением в системах 5G и готовится обеспечить 20-кратное увеличение производительности в рамках требуемого 1000-кратного роста емкости сетей к 2020 году (см. рисунок 1.2). Это связано с тем, что широкие диапазоны достаточно малоиспользуемого спектра доступны в полосе частот от 30 до 300 ГГц (что соответствует длине волны 1–10 мм), а также в меньшей степени от 20 до 30 ГГц. Однако расширение доступного системам мобильной связи спектрального диапазона в сторону крайне высоких частот долгое время считалось затруднительным из-за сложностей, связанных с распространением миллиметровых волн, таких как значительные потери на трассе, низкие коэффициенты дифракции и проникновения, сильный фазовый шум, а также существенное поглощение атмосферой и ослабление в случае дождя, тумана или снега [34].

Сравнительно недавно было показано, что, несмотря на особенности распространения, крайне высокие частоты возможно использовать для систем связи следующего поколения в различных диапазонах [35]. Последовавший за этим ажиотаж в поисках нового пригодного спектра выявил, что наиболее привлекательными частотами являются 28–30 ГГц; 71–76 ГГц; 81–86 ГГц и

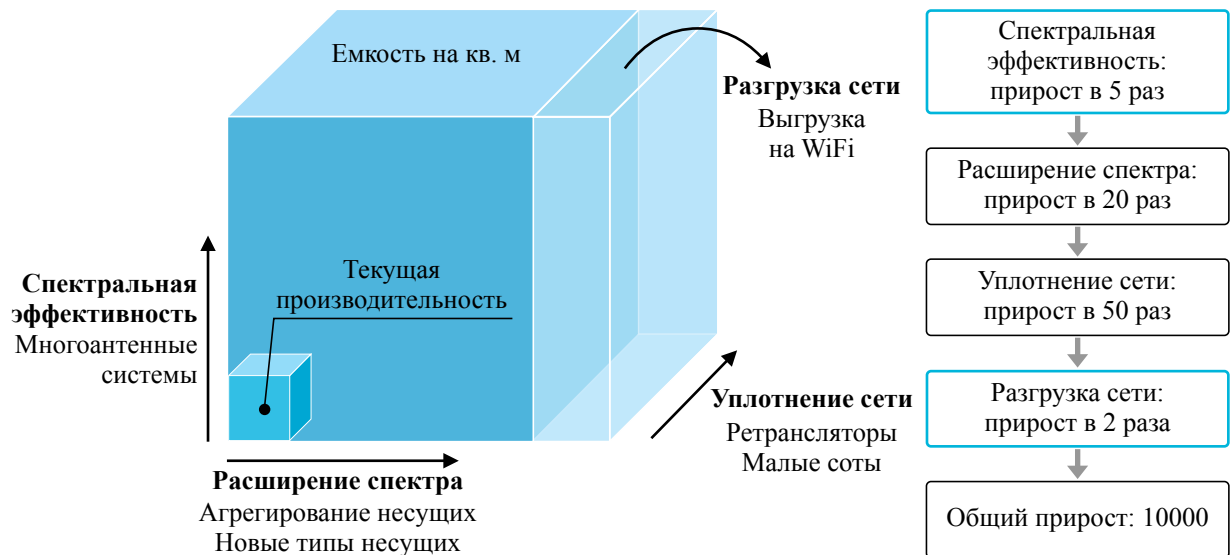


Рисунок 1.2 — Основные способы увеличения емкости систем 5G

92–95 ГГц [36], а также нелицензированные частоты 60 ГГц, уже давно применяющиеся системами связи IEEE. Для новой радиотехнологии 3GPP, получившей название New Radio, представляют интерес частоты 28, 71–76 и 81–86 ГГц (для развертывания емких малых сот и транзитных каналов [37]), тогда как технологии IEEE 802.11ad (WiGig) и 802.11ay используют 60 ГГц (для связи ближнего радиуса действия) [38]. К другим системам связи mmWave для применения, в основном, внутри помещений можно отнести IEEE 802.15.3c, WirelessHD и ECMA-387.

1.3 Интеграция сетей связи 5G с Интернетом вещей

1.3.1 Становление конвергентной экосистемы 5G-IoT

Технологии широкополосной связи плотно укоренились в качестве важной составной части повседневной жизнедеятельности современного общества и приобрели практически ту же степень важности, что и водоснабжение или электричество. На сегодняшний день беспроводные системы радиодоступа обеспечивают широкий диапазон всевозможных услуг связи, от профессиональных

и развлекательных сетей до различных промышленных приложений. Дальнейшее развитие имеющихся радиотехнологий все больше стимулирует прогресс и внедряется в рыночные сегменты за пределами телекоммуникационной сферы, такие как промышленная автоматизация, управление производственными процессами, автомобилестроение, интеллектуальные транспортные системы, энергетика, теле- и радиовещание, а также здравоохранение в числе многих других [39].

Для ускорения темпов своего экономического и социального развития [40] системы ИМТ следующего поколения должны решить новые для них крупные задачи: (1) поддержка очень высоких скоростей передачи данных; (2) подключение большого числа устройств различных типов; (3) обеспечение более высокого качества разнообразных услуг связи, а также (4) повышение экономической доступности и снижение эксплуатационных расходов [41]. Имея в виду эти цели, а также ожидая рост объемов мобильного трафика на уровне от 10 до 100 раз между 2020 и 2030 годами [42], Международный союз электросвязи ИТУ задался целью определить перспективы дальнейшего развития систем ИМТ до 2020 года и далее [43].

Согласно данному видению основное развитие спецификаций ИМТ-2020 планируется к завершению около 2020 года, после чего начнется фаза внедрения соответствующих технологий [44]. Интересно вспомнить, что развитие системы ИМТ-2000 заняло в общей сложности 15 лет (1985–2000) и привело к запуску мобильных сетей третьего поколения 3G (см. рисунок 1.3), тогда как последовавшее за ним развитие системы ИМТ-Advanced и мобильных сетей четвертого поколения 4G уложилось в 12 лет (2000–2012). С учетом дальнейшего повышения темпов разработки и сокращения ее сроков вплоть до 8 лет (2012–2020), ожидается, что перспективное пятое поколение 5G широкополосной мобильной связи значительно усовершенствует все аспекты работы системы, от повышения эффективности использования радиоспектра и механизмов физического уровня до поддержки программно-определяемой конфигурируемости и гибкого управления сетью.

Как было описано выше, ключевые технологии «большой тройки», составляющие основу новейших сетей связи 5G [33], включают в себя дальнейшее уплотнение сети малыми сотами различных размеров на основе разнородных технологий радиодоступа для повышения ее емкости, применение крайне высоких частот (технологии mmWave) для преодоления неизбежного дефицита

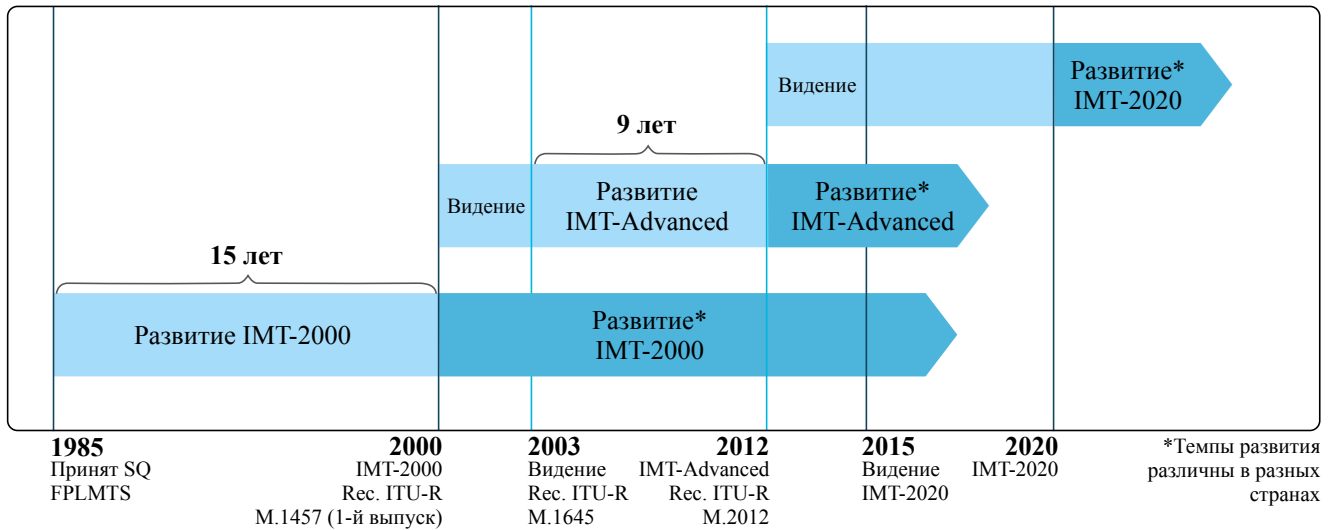


Рисунок 1.3 — Цикл разработки и развертывания систем ИМТ

традиционного беспроводного спектра, а также использование массивных многоантенных схем и технологий в целях повышения спектральной эффективности передачи. Сочетание подобных решений должно обеспечить достижение пиковой скорости передачи данных в 10 (а иногда и в 20) Гбит/с, воспринимаемой пользователем скоростью передачи в 100 Мбит/с на достаточно обширных открытых пространствах и вплоть до 1 Гбит/с внутри помещений.

В конце 2015 года в ответ на эти жесткие требования комитет 3GPP начал активную работу по созданию новой радиотехнологии (New Radio) в рамках системы 5G, которая смогла бы функционировать в диапазоне mmWave без обратной совместимости с предыдущими решениями [45]. Одновременно продолжилось дальнейшее развитие имеющейся системы LTE-Advanced [46]. При помощи разрабатываемой 3GPP сети радиодоступа New Radio перспективные системы 5G получат возможность обслуживать широкий круг ресурсоемких пользовательских приложений [47], включая службы определения местоположения абонентов, облачные и социальные сервисы, мультимедиа-услуги высокого разрешения, приложения виртуальной и дополненной реальности и др.

Однако применение сетей 5G не ограничивается сценариями, ориентированными на обслуживание людей в качестве пользователей, а предполагает также обеспечить поддержку широкого круга сценариев, включающих так называемое межмашинное взаимодействие M2M [48]. Соответственно, дополнительные и более строгие требования возникают относительно задержки, доступности, надежности, масштабируемости, стоимости и энергетической эффективности

связи для приложений М2М [49]. К примеру, технологии 5G должны будут предоставить задержку на уровне радиодоступа не более 1 мс и работать со скоростями движения узлов до 500 км/ч. С другой стороны, системы связи следующего поколения обеспечат высокую плотность подключенных устройств (вплоть до $10^6/\text{км}^2$) с чрезвычайно низким энергопотреблением.

При выполнении ею всех заявленных требований система ИМТ-2020 станет наиболее совершенным средством для обслуживания всевозможных приложений Интернета вещей IoT, построенного на технологиях межмашинного взаимодействия М2М, и, тем самым, позволит подключить к сети огромное многообразие различных машин, функционирующих при минимальном вмешательстве человека [50]. Предполагаемые таким образом сценарии применения гетерогенных сетей связи 5G можно условно разделить на три широких класса (см. рисунок 1.4), два из которых содержат исключительно приложения М2М и Интернета вещей:

- 1. Расширенный мобильный широкополосный доступ.** Мобильный широкополосный доступ нацелен на обслуживание людей в качестве пользователей, получающих доступ к различным мультимедиа-приложениям, услугам и данным. Ожидается, что потребность в поддержке данного вида сервисов продолжит возрастать, а для их обеспечения потребуется расширение сетевых возможностей. В итоге расширенный мобильный широкополосный доступ позволит осуществить сценарии, в которых смогут функционировать новые пользовательские приложения с повышенными требованиями к производительности, предполагающие более высокое качество восприятия услуги. Такие сценарии многочисленны и включают в себя как обеспечение покрытия сетью обширной области пространства (wide-area), так и организацию доступа в отдельных локациях (hotspot), которые, очевидно, должны иметь различные характеристики. Для отдельных локаций, в которых наблюдается высокая плотность размещения абонентов, важны достаточно высокие емкость сети и скорость передачи данных пользователем, тогда как требования по поддержке мобильности являются не такими строгими, как в случае покрытия обширной области. Для такого покрытия необходимо непрерывное соединение даже при перемещении абонента с высокой скоростью.

2. **Массовое межмашинное взаимодействие.** Данный класс сценариев характеризуется очень большим числом подключенных устройств, которые при этом передают сравнительно небольшие объемы устойчивых к задержке данных. Требуется, чтобы подобные устройства имели низкую стоимость, а также длительное время работы без перезарядки или замены своей батареи питания. Такие условия являются типовыми для широкого класса взаимодействующих объектов, от несложных бытовых устройств до более продвинутых приборов с расширенными возможностями. Примерами являются системы наблюдения за интеллектуальными энерго- и сельскохозяйственными системами, подключенные сенсоры и измерители, силовые приводы, носимые устройства, камеры наблюдения, а также автомобили нового поколения.
3. **Надежное взаимодействие с низкой задержкой.** Для этого класса сценариев характерны строгие требования к показателям работы системы, таким как ее пропускная способность, задержка доставки данных и доступность беспроводного соединения. В качестве примера можно привести беспроводные системы контроля промышленного производства и сопутствующих процессов, удаленного осуществления медицинских услуг, распределения нагрузки в интеллектуальных энергосистемах, обеспечения безопасности дорожного транспорта и т.д. Подобные условия функционирования требуют организации межмашинного взаимодействия в режиме реального времени для, например, самоуправляемых автомобилей и беспилотных летательных аппаратов, процессов промышленного производства, оперативного контроля и оптимизации систем транспортной безопасности, работы служб чрезвычайного реагирования и ряда других приложений.

Стремительное развитие в этой связи *конвергентной* экосистемы 5G-IoT [51; 52] предполагает, что разнообразные подключенные устройства получат возможность взаимодействовать друг с другом посредством различных технологий радиодоступа, включая те, которые оперируют с более высокими частотами и более широкой полосой пропускания. Для удовлетворения разнообразных требований ИМТ-2020 понадобится согласованное использование различных радиотехнологий [53], включающее дальнейшее развитие LTE-Advanced и систему New Radio комитета 3GPP совместно с различными протоколами доступа семейства IEEE, такими как широко распространенные сети WiFi и



Рисунок 1.4 — Перспективные сценарии применения систем ИМТ к 2020 году и далее

многочисленные решения для беспроводной связи ближнего радиуса действия. Такие, в высшей степени гетерогенные сети связи 5G неизбежно столкнутся с рядом сложностей, связанных, прежде всего, с поддержкой продвинутых IoT-устройств с повышенными запросами к качеству обслуживания, в том числе беспилотных автомобилей и летательных аппаратов, а также промышленных роботов.

1.3.2 Гетерогенные сети 5G для поддержки Индустриального Интернета

По мере того как технологии 5G готовятся к обслуживанию многочисленных IoT-приложений с повышенными требованиями [54], возникает все более существенное различие между потребительскими и индустриальными сценариями их работы. Действительно, потребительский сегмент Интернета вещей развивался уже значительное время (например, в области автоматизации домашнего хозяйства) с целью улучшения качества жизни людей и общества в целом [55]. При этом индустриальный сегмент Интернета вещей представляет собой относительно недавнее направление, нацеленное на улучшение бизнес-связей между предприятиями [56].

При помощи средств связи 5G промышленные пользователи стремятся достичь повсеместного и надежного взаимодействия, превышающего возможности технологий 4G по обеспечению мобильного широкополосного доступа (см. рисунок 1.5). Однако множеству индустриальных сценариев IoT присущи такие требования, которые не могут быть удовлетворены имеющимися сетями 3G и 4G, например, в системах контроля промышленного производства и обеспечения безопасности дорожного транспорта, при взаимодействии между машинами и их операторами, в сферах материально-технического снабжения и отслеживания перемещения объектов, а также в автомобилестроении, энергетике и т.д. [57].



Рисунок 1.5 — Примеры 5G-приложений Индустриального Интернета для автоматизации производства

Для учета основных требований возникающего таким образом Индустриального Интернета [58] необходимо соблюдать следующие ключевые принципы: совместимость, виртуализация, децентрализация, взаимодействие в реальном времени, ориентация на повышение качества услуг и модульность [59], которые в совокупности могут обеспечить требуемый уровень качества восприятия услуги. Например, в сфере автоматизации производственных процессов при изготовлении продукции, контроле ее качества и осуществлении погрузочно-разгрузочных работ европейскими государствами была запущена многолетняя программа, названная «Industrie 4.0» [60].

Изначально поддержанная правительством Германии, данная инициатива объединяет государственный и частный сегменты, а также академические круги для формирования целостного видения и согласованного подхода к применению беспроводных технологий связи в индустриальном секторе. В ответ на это Китай анонсировал собственную программу «Made in China 2025» для содействия национальной интеграции в сфере цифровых технологий и придания дополнительного импульса промышленному развитию. В России параллельно была запущена программа «Цифровая экономика Российской Федерации», определяющая цели и задачи развития ключевых институтов и основных инфраструктурных элементов цифровой экономики [61].

По мере того как во всем мире ускоряются темпы исследований в области приложений Индустриального Интернета, становится понятно, что технологии 5G станут центральным звеном для обеспечения нужд Industrie 4.0 и схожих программ [62]. В данном контексте особую сложность вызывает поддержка доступности сетевых соединений в удаленных и труднодостижимых областях [63], чего могут потребовать перспективные системы автоматизации производства и службы дистанционного контроля объектов [64]. При этом для организации сервисов виртуальной или дополненной реальности в задачах тактильного управления могут применяться современные мобильные облачные технологии на краю соты [65].

Таким образом, задачи Индустриального Интернета обуславливают переход от традиционных систем распространения информации к надежным сетям передачи навыков [66], которые могут быть недоступны не более нескольких секунд в течение года. К сожалению, инфраструктура современной сети Интернет исторически реализует принцип организации связи «по возможности», и для обеспечения работы надежных сервисов необходимы значительные ее изменения [67]. Помимо крайне надежных беспроводных соединений, некоторые из таких приложений также требуют чрезвычайно низких задержек доставки информации (например, службы непрерывного контроля) [68]. Как следствие, весь стек протоколов системы Индустриального (тактильного) Интернета должен быть переработан в целях достижения жестких гарантий по надежности радиодоступа [69; 70].

Здесь надежность может трактоваться в смысле вероятности получения заданного качества обслуживания на некотором интервале времени в определенных условиях работы. В частности, для улучшения характеристик производи-

тельности радиоканала может быть целесообразно отказаться от использования механизмов повторной передачи сообщений, поскольку они неизбежно вносят дополнительную нежелательную задержку [66]. Альтернативным подходом в целях повышения надежности передачи служит внедрение избыточности в частотной или пространственной областях путем использования множества независимых между собой беспроводных каналов. Тем самым параллельная работа на нескольких радиоканалах, называемая также одновременным подключением или многосвязностью (multi-connectivity), становится важным средством повышения надежности будущих приложений Индустриального Интернета.

С учетом того, что функционал сотовых систем строго регламентирован и изначально предполагает необходимость управления интерференцией, применение операторских сетей мобильной связи становится наиболее привлекательно для организации базового подключения и надежного управления [71]. Однако новейшие приложения Индустриального Интернета требуют более тесного взаимодействия между сотовыми решениями 3GPP LTE и другими технологиями радиодоступа как в общественных, так и в частных сетях. На сегодняшний день консенсус заключается в том, что дальнейшее развитие систем LTE-Advanced будет сопровождаться их интеграцией с уже имеющимися (например, WiFi) и новыми (например, New Radio) радиотехнологиями для достижения более высокоскоростной, надежной и непрерывной связи. Такое видение гетерогенной экосистемы 5G является общепринятым и призвано обеспечить весь спектр ее перспективных приложений.

1.4 Стандартизация сетей связи 5G и анализ ее требований

1.4.1 Возникновение Интернета мобильных и надежных вещей

Число подключенных M2M-устройств может превысить отметку в 28 миллиардов к 2021 году, что значительно превзойдет число беспроводных соединений, устанавливаемых людьми в качестве пользователей [72]. Данное обстоятельство становится катализатором сближения физического и цифрового миров, которое ведет к формированию беспрецедентного рынка приложений и

услуг Интернета вещей, оцениваемого в 19 трлн долларов США в следующее десятилетие. На сегодняшний день большое разнообразие сценариев Интернета вещей включает в себя интеллектуальные транспортные и энергетические системы, дистанционное здравоохранение, организацию работы «умных» счетчиков, поддержку служб промышленной автоматизации, управления и удаленного производства, осуществление видеонаблюдения в целях сохранения общественного порядка и т.д.

В отношении технических требований многочисленные устройства Интернета вещей можно условно разделить на недорогое оборудование, применяющееся в массовых сценариях, а также существенно более сложные решения, которые используются в критических приложениях [73]. С одной стороны, массовые сценарии включают в себя большое число достаточно простых М2М-устройств, таких как интеллектуальные измерители, которые передают свои данные в сеть на регулярной основе. В этом случае стоимость комплексного решения для организации беспроводной связи должна сохраняться достаточно низкой для поддержания экономической целесообразности такой услуги. Помимо невысокой стоимости производства конечных устройств, типовыми требованиями здесь являются длительное время их работы от батареи и широкая область покрытия сетью. С другой стороны, критические приложения предъявляют жесткие требования к надежности, доступности и задержке беспроводного соединения [68]. В этом случае объемы продаж могут быть значительно меньше, но рыночная ценность таких услуг может оказаться даже выше, чем для массовых приложений Интернета вещей.

С учетом стремительных темпов развития различных приложений и услуг Интернета вещей цель опорных радиотехнологий, обеспечивающих межмашинное взаимодействие, заключается в построении адекватных беспроводных соединений для подключения всевозможных стационарных и мобильных М2М-устройств и прочих объектов на обширных площадях. В настоящее время гетерогенные сети связи 5G становятся основным средством создания таких соединений, а стандартизация ключевых технологий 5G приближается к своему завершению. В этой связи особенно важной и непростой в критических приложениях становится задача организации непрерывной связи с подвижными объектами, такими как беспилотные автомобили и летательные аппараты, а также автономные промышленные роботы. Несмотря на то, что разнообразные модели мобильности были всесторонне рассмотрены в контексте обслуживания

людей в качестве пользователей, например, в специализированных («ad-hoc») сетях и сотовых системах связи, влияние мобильности устройств и динамики окружающей среды на характеристики перспективных технологий 5G изучены в сравнительно меньшей степени [74].

Для беспроводных «ad-hoc» сетей воздействие процессов, связанных с мобильностью, на скорость передачи отдельного абонента было обстоятельно изучено в работе [75]. Для приложений с нестрогими требованиями по задержке передачи информации, когда возможно изменение топологии системы в течение времени передачи пакета данных, достижимая скорость передачи может заметно возрасти, если узлы сети начнут перемещаться в пространстве. Следуя подобным рассуждениям, удастся повысить производительность системы в целом, используя различные формы многопользовательских механизмов взаимоподдержки, включая абонентскую кооперацию [76; 77]. Однако этот и аналогичные важные механизмы оказываются неприменимы в сетях 5G, где имеются более строгие ограничения на задержку передачи. В таких сетях некоторый центральный узел управляет доступом к радиоресурсам всей системы. С другой стороны, более жесткий контроль решений 5G за распределением системных ресурсов ведет к улучшению качества беспроводного покрытия, снижению уровня интерференции, а также позволяет достичь более высокой скорости передачи данных.

Другим важным направлением исследований становится рассмотрение согласованной работы нескольких технологий радиодоступа в рамках гетерогенной сети связи 5G для повышения надежности беспроводного соединения, снижения чувствительности к его обрывам, а также для более эффективного использования спектральных ресурсов в среде, которая включает в себя абонентские терминалы с различными возможностями и целевыми требованиями по качеству обслуживания [78]. В этой связи тесная интеграция «ad-hoc» сетей, функционирующих на нелицензированных частотах, с сотовыми системами, работающими в лицензированном спектре [50], призвана обеспечить поддержку наиболее перспективных приложений 5G-IoT, таких как пользовательские и операторские точки доступа, размещенные на беспилотных летательных аппаратах (дронах). Эта относительно новая возможность обусловила интерес к использованию преимуществ одновременного подключения в целях обеспечения высокомобильных приложений и услуг Интернета вещей.

В результате потребность в надлежащем моделировании сетей связи 5G для обслуживания перспективных сценариев Интернета мобильных и надежных вещей в настоящее время весьма высока [79]. Соответственно, актуально изучение гетерогенных систем с возможностью одновременного подключения, в которых устройства могут применять сотовые и прямые соединения, а также пользоваться поддержкой, например, беспилотных автомобилей и летательных аппаратов, предоставляющих непосредственный беспроводной доступ, для организации обмена информацией и обеспечения работы в сети Интернет. В частности, наблюдается дефицит результатов моделирования системы в целом в типовых сценариях 5G-IoT, предполагающих наличие нескольких радиотехнологий, которые пользовательские устройства могут использовать альтернативно или одновременно. При этом важно выяснить запас емкости такой сети и прирост качества обслуживания в ней по сравнению с базовой системой 3GPP LTE, в которой абонентские терминалы связываются при помощи инфраструктурных сотовых соединений.

1.4.2 Цели и задачи стандартизации систем 5G

Широкое распространение всевозможных мобильных и M2M-устройств обусловило необходимость стандартизации новейших технологических решений 5G. Поворотным моментом в этом направлении стало опубликование документа, задающего развитие системы ИМТ к 2020 году и далее [43], которое определило основные цели, а также ключевые требования к системам мобильной связи следующего поколения [80]. Данный документ формально вводит три широких класса сценариев применения систем 5G, с соответствующими крайне различными характеристиками: (1) расширенный мобильный широкополосный доступ; (2) массовое межмашинное взаимодействие и (3) надежное взаимодействие с низкой задержкой. В ответ на эту инициативу комитет 3GPP начал активную разработку новой технологии радиодоступа New Radio без обратной совместимости с предыдущими решениями, функционирующей на крайне высоких частотах, а ее первая коммерческая реализация намечена уже на 2019–2020 годы. Ожидается, что развертывание перспективного решения New Radio будет происходить параллельно с дальнейшей эволюцией сотовых сетей LTE.

Понимая необходимость глобальных стандартов, регламентирующих функционирование перспективных сотовых сетей, комитет 3GPP провел значительную работу по спецификации целого ряда перспективных технологий для интеграции LWA, доступа LAA, связи D2D при организации абонентской кооперации через смартфоны для носимых устройств, а также энергосберегающих механизмов для M2M-устройств, системы oneM2M для обеспечения услуг Интернета вещей и схем передачи M2M с малой скоростью и низкой сложностью. Последние реализованы как новая категория абонентских терминалов LTE (Cat. M1) и впоследствии дополнены новым радиointерфейсом NB-IoT начиная с 3GPP Release 13 и далее [81]. Для Интернета вещей эти наработки означают дальнейшее обособление массовых и критических сценариев, хотя отдельные его приложения могут принадлежать одновременно к обоим типам услуг (например, тревожная сигнализация на производстве).

С учетом поддержки новыми технологиями 3GPP различных требований M2M-устройств, начавшейся уже в 2005 году, а также недавних усилий по стандартизации радиointерфейса NB-IoT, множество сценариев массового межмашинного взаимодействия могут быть с успехом реализованы современной версией технологии LTE. При этом улучшения, на которые нацелено создание технологии New Radio, касаются, прежде всего, широкого развертывания сетей для надежного взаимодействия устройств с низкой задержкой. Соответствующие требования зачастую сопряжены с мобильностью (со скоростями вплоть до 500 км/ч) и временем доставки данных (с круговой задержкой на уровне радиодоступа не более 1 мс) [43]. Их обеспечение особенно актуально для того, чтобы в ближайшее время реализовать богатый рыночный потенциал критических приложений и услуг 5G, и порождает множество важных задач, связанных с разработкой новой системы New Radio.

Соответственно, основной целью быстро развивающейся экосистемы 5G становится организация тесного взаимодействия и интеграции различных технологий радиодоступа с учетом разнообразия целевых сценариев, в результате чего массовые и критические приложения, а также услуги мобильного широкополосного доступа смогут сосуществовать в рамках единой сети связи. Однако требования некоторых приложений и услуг могут различаться настолько значительно, что в ходе разработке системы 5G возникла следующая дилемма. Следовало определиться, стоит ли создавать комплексный набор специализированных сетевых решений или же строить единую универсальную технологию,

которая могла бы удовлетворить запросы всевозможных пользователей. По итогам оживленных дискуссий между мобильными операторами, производителями оборудования, ожидаемыми пользователями и директивными органами была предложена концепция сетевой «нарезки» (network slicing) для эффективной инкапсуляции обслуживания отдельных групп абонентов со схожими свойствами.

Для обеспечения непрерывной работоспособности потребительских и промышленных сценариев 5G [82] необходимо интегрировать все доступные сети связи, включая гетерогенные технологии доступа для осуществления передачи по линии вверх и вниз, а также прямые соединения между устройствами. Применение последних становится особенно привлекательно, поскольку грань различия между пользовательским и операторским оборудованием постепенно стирается. Следовательно, возможно использовать специализированные абонентские терминалы для выполнения некоторых сетевых функций. Это осуществить тем проще, чем более интеллектуальным является взаимодействующее пользовательское оборудование – от сенсоров, носимых устройств и традиционных мобильных терминалов до подключенных автомобилей, подвижных роботов и дронов, работа которых требует более строгих гарантий по обеспечению надежности, поддержке мобильности и организации эффективного управления спектральными ресурсами [83].

В качестве примера стоит привести быстро растущий рынок подключенных и самоуправляемых автомобилей, который накладывает жесткие ограничения на допустимые показатели надежности и задержки доставки данных при движении объектов с высокой скоростью. Их соблюдение особенно затруднительно в связи с тем, что более высокая скорость движения принципиально противоречит задаче повышения надежности передачи. По мере того как становятся понятны перспективные сценарии V2X, исследование механизмов обеспечения более надежной и качественной беспроводной связи посредством гетерогенных сетей 5G становится крайне востребовано. Этого можно достичь при помощи согласованного использования различных радиотехнологий и уплотнения сети доступа, а также за счет применения прямых соединений между устройствами и крайне высоких частот, что и становится предметом рассмотрения двух последующих глав данной работы.

1.4.3 Особенности различных сценариев применения 5G

Как следует из вышеизложенного, адаптивное использование различных технологических возможностей, тесно связанное со свойствами целевого сценария, является ключевым в перспективных сетях связи 5G. С учетом разнообразия требований, предъявляемых множеством их приложений, основным принципом развития экосистемы 5G становится универсальность и гибкость в обслуживании пользователей различных типов. Кроме того, должны быть учтены ограничения, связанные с энергопотреблением подключенных устройств и эффективностью использования спектральных ресурсов. Несмотря на широкий диапазон заявленных возможностей, значение отдельных ключевых параметров 5G может быть различно в зависимости от условий конкретного применения. Ожидаемый уровень их важности для трех введенных классов сценариев 5G приведен на рисунке 1.6 и в соответствии с описанием в [43] представляет собой условную градацию по категориям высокой, средней и низкой значимости.



Рисунок 1.6 — Ключевые особенности сценариев применения технологии 5G

В сценариях расширенного мобильного широкополосного доступа наиболее важными оказываются воспринимаемая пользователем скорость передачи данных, пропускная способность на единицу площади, максимально достижимая скорость передачи данных в системе, поддержка мобильности, а также обеспечение высокой энергетической и спектральной эффективности. Но при этом поддержка мобильности и/или скорости передачи данных имеет различный уровень значимости в зависимости от условий применения. К примеру,

при организации доступа в отдельных локациях (hotspot) более предпочтительна высокая скорость передачи данных, а поддержка мобильности может быть не настолько важна, так как ожидается, что пользователи не будут перемещаться на значительные расстояния. С другой стороны, при реализации покрытия сетью обширной области пространства (wide-area) ситуация может измениться на противоположную.

Напротив, в сценариях, предполагающих надежное взаимодействие устройств с низкой задержкой, оперативность доставки информации приобретает первостепенную значимость, например, для обеспечения работы критически важных приложений. Данное обстоятельство возникает в ситуациях, связанных с взаимодействием объектов, перемещающихся на значительной скорости, в том числе при организации безопасности дорожного движения. При этом возможность передавать данные с высокой скоростью может быть менее важна. Наконец, в сценариях с массовым межмашинным взаимодействием требуется высокая плотность подключенных устройств, что связано с их очень большим числом. Однако сами устройства при этом могут передавать свои данные только периодически, чтобы не превысить общей пропускной способности системы. В этом случае поддержка мобильности не так критична, как длительный срок службы батареи.

В ходе последующего развития гетерогенных сетей связи 5G могут возникнуть дополнительные требования, направленные на обеспечение их более гибкого, надежного и безопасного функционирования при организации всевозможных приложений и услуг, которые будут возникать в дальнейшем.

1.5 Структура последующего изложения

Подводя итоги, следует отметить, что перспективные технологии 5G предлагают набор важных улучшений, направленных на увеличение, прежде всего, емкости системы связи и включающих в себя (1) дальнейшее уплотнение сети малыми сотами различных типов; (2) более продвинутые формы прямого взаимодействия абонентских терминалов для выгрузки трафика; (3) использование крайне высоких частот для увеличения скорости передачи данных и (4) повышение спектральной эффективности за счет использования более сложных

многоантенных и кодовых конструкций. Исследованию большинства этих особенностей с упором на методы управления потоками трафика и посвящено дальнейшее изложение. В частности, изучаются следующие ключевые компоненты экосистемы 5G-IoT, которые становятся предметом рассмотрения последующих глав данной работы.

- Во **второй** главе исследуются интеграция разнородных технологий радиодоступа и особенности уплотнения сети связи.
- В **третьей** главе изучаются использование прямых соединений между устройствами и применение крайне высоких частот.
- В **четвертой** главе рассматривается поддержка приложений Интернета вещей и распространение носимых устройств.
- В **пятой** главе описываются перспективные сценарии применения гетерогенных сетей 5G, а также их перспективные приложения.
- В **заключении** дается краткая характеристика основных результатов диссертационной работы.

Глава 2. Интеграция технологий радиодоступа и уплотнение сети

2.1 Предварительные замечания

Гетерогенные сети связи 5G построены на основе архитектуры, которая включает в себя иерархическую структуру сот различного размера, использующих разнородные технологии радиодоступа. В таких системах абонентские терминалы при передаче своих данных сетевой инфраструктуре имеют возможность использовать несколько радиотехнологий одновременно. С увеличением числа таких терминалов на рынке операторы мобильной связи становятся более заинтересованы в эффективном использовании радиоресурсов нескольких сетей доступа для увеличения емкости интегрированной системы связи и повышения качества пользовательского соединения. Таким образом, взаимодействие между различными технологиями радиодоступа и возможность эффективного управления ими становятся основными особенностями сетей 5G, исследованию которых и посвящена данная глава.

Далее рассматриваются всевозможные варианты интеграции сетей доступа на примере технологий LTE и WiFi. Особое внимание уделяется методам выбора предпочтительной радиотехнологии и схемам планирования радиоресурсов, функционирующим на уровне сети радиодоступа, которые могут быть реализованы как на стороне абонентских терминалов, так и на стороне сетевой инфраструктуры. Проводится сравнение традиционных и новых алгоритмов такого управления в рамках гетерогенной системы связи 5G для выявления преимуществ подхода с так называемым сетевым содействием, при котором абонентские терминалы снабжаются актуальным знанием относительно, например, загрузки сети радиодоступа. Для моделирования подобных методов в этой главе также разрабатывается соответствующий математический подход к анализу гетерогенных систем 5G.

Предложенная пространственно-временная математическая модель учитывает динамику абонентского трафика во времени и размещение узлов сети в пространстве, а также согласуется с результатами имитационного моделирования. Она позволяет изучать продвинутые алгоритмы управления трафиком в гетерогенных сетях 5G, в особенности те, которые применяют сетевое содей-

ствие для организации эффективной работы разнородных радиотехнологий (например, LTE и WiFi) и оперативного выбора предпочтительной сети доступа. Используемый математический аппарат основан на методах стохастической геометрии, которые объединяются здесь с подходами теории массового обслуживания для моделирования пространственных взаимоотношений между узлами совместно с динамикой их трафика на уровне сессий в типовых условиях функционирования систем 5G.

Кроме того, учитывается важная практическая тенденция к уплотнению сетей доступа, связанная с необходимостью передачи все возрастающих объемов данных от мобильных абонентов. В отличие от традиционной выгрузки трафика пользователей из сотовой системы связи в локальную сеть доступа, которая активно применялась в гетерогенных сетях предыдущих поколений, в тесно интегрированных системах 5G возможно совмещение функционала радиотехнологий LTE и WiFi в рамках единого модуля доступа. Работа таких совмещенных систем, особенно для сверхплотных покрытий малыми сотами LTE-WiFi, позволяющих абонентскому терминалу вести передачу данных по нескольким технологиям радиодоступа одновременно, связана со специфической пространственно-временной динамикой, которая также рассматривается в рамках данной главы.

2.2 Варианты интеграции технологий радиодоступа

2.2.1 Необходимость интеграции радиотехнологий

Развитие мобильной связи в эпоху 5G требует принципиальных изменений в сетевой инфраструктуре, а также в архитектуре самих абонентских терминалов UE, которые при этом начинают учитывать контекст их использования [84]. Считается, что наиболее целесообразным методом преодоления дефицита радиоресурсов в сетях 5G по сравнению с сетями предыдущих поколений становится размещение малых сот с более высокой плотностью покрытия. Благодаря относительно коротким каналам радиосвязи использование малых сот позволяет достичь повышенных скоростей передачи данных, а также снизить энерги-

тические затраты на передачу по линии вверх, особенно в условиях городской застройки.

Однако размещение базовых станций с более высокой плотностью для увеличения емкости сети неизбежно ведет к возрастанию сложности управления интерференцией, увеличению стоимости развертывания сети и дополнительным затратам на поддержание ее работоспособности [85]. При этом на сегодняшний день лицензированный радиоспектр является весьма ограниченным и потому дорогостоящим ресурсом, а традиционные методы повышения эффективности его использования приближаются к своим теоретическим пределам. Даже при выделении дополнительного радиоспектра для услуг мобильной связи соответствующие радиочастоты, по всей видимости, останутся сильно фрагментированными и потребуют применения различных технологий радиодоступа RAT. Таким образом, возникает острая необходимость в координировании разнородных технологий радиодоступа с целью повышения емкости сети связи.

Как следствие, становится актуально объединять гетерогенные технологии радиодоступа в более сложную *интегрированную* сеть связи [86], а также разрабатывать методы эффективного использования радиоресурсов, которые доступны в различных спектральных диапазонах и ассоциированы с разнородными технологиями радиодоступа [87]. В частности, ожидается, что существенный прирост емкости можно получить за счет построения гетерогенной архитектуры сети, которая могла бы агрегировать лицензированный и нелицензированный спектральный радиоресурс. К примеру, абонентские терминалы, поддерживающие *одновременную* работу нескольких технологий радиодоступа, могут использовать доступ в нелицензированном диапазоне частот для разгрузки инфраструктурных сотовых соединений [88].

На сегодняшний день, по мере уменьшения зон покрытия базовых станций, эффективные области обслуживания сотовых, локальных и персональных сетей доступа начинают в значительной степени перекрываться. Это создает предпосылки к одновременному использованию нескольких технологий радиодоступа для улучшения качества связи. Предполагается, что такое интеллектуальное использование множества радиотехнологий позволит достичь повышенной эффективности относительно ряда показателей, включающих в себя пространственное, временное и частотное разнесение, координирование интерференции, загрузку сети и т.д. В перспективных сетях пятого поколения радиотехнологии ближнего и дальнего радиуса действия могут быть использованы

совместно в процессе тесного взаимодействия между абонентским терминалом и сетью связи, а также непосредственно между самими терминалами для достижения более высокого качества обслуживания абонентов [89].

Как следствие, возрастает необходимость в координировании совместной работы различных технологий радиодоступа и ожидается, что использование множества радиотехнологий в совокупности с соответствующими средствами координирования на стороне абонентских терминалов и сетевой инфраструктуры станут неотъемлемой характеристикой систем связи нового поколения [90]. В частности, распределенная локальная сеть в нелицензированном диапазоне частот WLAN может получать содействие со стороны *централизованной* функции управления, которая размещена на стороне сотовой системы связи. Это позволяет проводить эффективное установление пользовательского соединения с сетью при наличии альтернативных технологий радиодоступа и, таким образом, повышать производительность системы, а также улучшать надежность соединений.

В соответствии с вышеизложенным в настоящее время наблюдается тенденция к более тесному взаимодействию разнородных технологий радиодоступа. В этой связи в данной работе рассматриваются задачи выбора предпочтительной технологии радиодоступа и планирования радиоресурсов в интегрированной сети связи, а также выявление соответствующего прироста производительности. Последующее изложение сосредоточено на совместной работе различных технологий радиодоступа в рамках гетерогенной системы связи. В качестве примера рассматривается взаимодействие точек доступа локальной сети WLAN с малыми сотами мобильного оператора в контексте организации соответствующей архитектуры системы связи и выяснения преимуществ различных форм интеграции. Цель данного исследования заключается в анализе показателей эффективности, таких как емкость системы связи и качество пользовательского соединения.

Следует отметить, что взаимодействие локальных и сотовых систем доступа неоднократно рассматривалось в исследовательской литературе, но в основном с позиций переключения между ними, называемого вертикальным [91]. Комитет по стандартизации мобильных систем связи 3GPP на протяжении нескольких лет был в значительной степени вовлечен в разработку спецификаций, направленных на обеспечение эффективной совместной работы сотовых и локальных сетей. Создание таких спецификаций проводилось в рамках несколь-

ких направлений работы, которые имели своей целью более тесную интеграцию между двумя типами систем. Они включают в себя решения для доверенного доступа устройств WLAN к услугам 3GPP, организацию беспроводной мобильности между технологиями 3GPP и WLAN, а также разработку функционала для обнаружения и выбора сети доступа ANDSF. В то время как более ранние подходы позволяли обеспечивать неплотное взаимодействие 3GPP и WLAN на уровне опорной сети (ядра сети), начиная с 3GPP Release 12 велась работа по организации более тесного взаимодействия на уровне сети радиодоступа RAN [92].

Данная активность связана с необходимостью повышения качества обслуживания в нелицензированном диапазоне частот в соответствии с требованиями мобильных операторов, стремящихся таким образом увеличить надежность соединений WiFi. В свою очередь, комитет по стандартизации локальных сетей WLAN развернул работу по созданию спецификаций HotSpot 2.0, а также инициировал новое направление по разработке систем WLAN повышенной эффективности в рамках рабочей группы IEEE 802.11. Таким образом, актуальным является исследование решений по интеграции 3GPP и WLAN на уровне RAN, позволяющих достичь более оперативного координирования между радиотехнологиями сотовых и локальных сетей доступа. Ниже описываются возможности обеспечения такого взаимодействия и соответствующие интеллектуальные алгоритмы, а также исследуется их влияние на показатели качества связи абонента и характеристики работы системы в целом. При этом различаются подходы, принимающие решение об управлении информационными потоками, которые функционируют на стороне абонентских терминалов и системы связи соответственно.

2.2.2 Типы архитектуры гетерогенной сети

Как отмечалось выше, ограничения по емкости и качеству соединения в мобильных сетях по сравнению с требованиями 5G обуславливают необходимость более тесной интеграции сотовых и локальных сетей доступа. На рисунке 2.1 схематически показана современная гетерогенная сеть HetNet сотового оператора. Система HetNet включает в себя иерархическую структуру сот различного размера, включая макросоты для обеспечения беспроводного покрытия,

базового доступа и мобильности, а также различные малые соты (пикосоты, фемтосоты, точки доступа WiFi, совмещенные LTE-WiFi модули и др.). Соответствующее оборудование доступа имеет более низкую мощность передачи и стоимость, позволяя повысить емкость системы за счет сокращения расстояния между сетевой инфраструктурой и абонентом в областях с высокой потребностью в обслуживании.

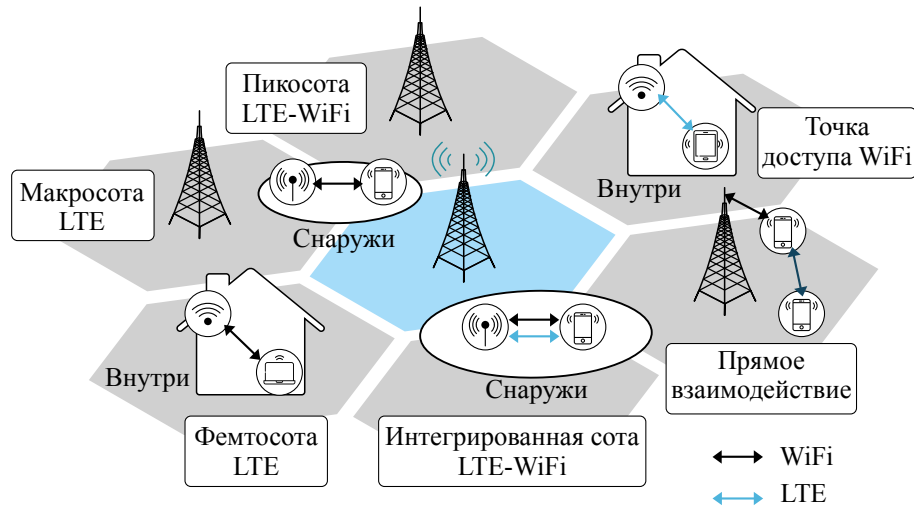


Рисунок 2.1 — Топология типовой гетерогенной сети

Исторически тенденция к интеграции сотовых и локальных систем WLAN возникла в связи с потребностью мобильных операторов в наращивании емкости сети доступа и предотвращении ее перегрузок. Однако ожидается, что использование технологии WiFi для разгрузки сотовых соединений останется частью долгосрочной стратегии развития сотовых операторов. В простейшем случае сети 3GPP и WLAN могут не иметь средств прямого координирования, и тогда абонентские терминалы вынуждены принимать локальные решения о предпочтительной технологии радиодоступа [93]. Если же система WiFi управляется в рамках интегрированной сети радиодоступа мобильного оператора, появляются дополнительные возможности для более тесного взаимодействия технологий 3GPP и WLAN.

Например, возникает новая сетевая архитектура, в которой использование совмещенных малых сот LTE-WiFi позволит достичь более тесного взаимодействия между соответствующими технологиями радиодоступа [94], по сути агрегируя систему WLAN в качестве «виртуальной несущей» сети 3GPP. На сегодняшний день совмещенные малые соты LTE-WiFi являются одной из пе-

редовых инноваций в телекоммуникационной индустрии и позволяют снизить затраты на развертывание системы связи за счет использования гибридной сетевой инфраструктуры. Однако учитывая, что такой вариант размещения сот еще мало распространен на практике, текущие работы по стандартизации направлены также на улучшение сетевой архитектуры, в которой абонентские терминалы принимают локальные решения в условиях ограниченного содействия со стороны сетевой инфраструктуры или же при его отсутствии.

Последующее изложение сосредоточено на рассмотрении различных вариантов интеграции сетей LTE и WiFi как показано на рисунке 2.2. Данные опции отличаются, прежде всего, специфичными механизмами обеспечения взаимодействия технологий радиодоступа, включающими в себя обнаружение радиотехнологии, ее выбор или назначение, управление радиоресурсами RRM, обеспечение мобильности и перенаправления потока данных между технологиями радиодоступа и т.д.

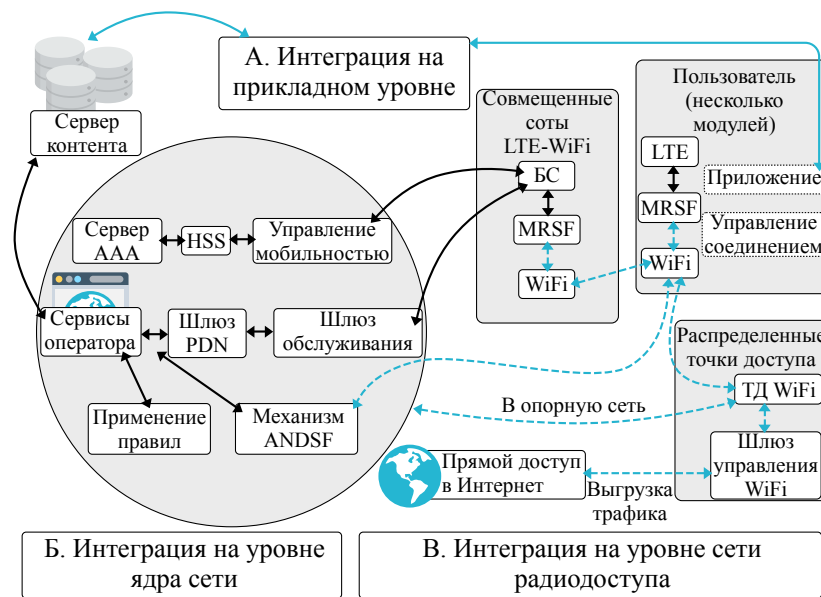


Рисунок 2.2 — Архитектура гетерогенной сети связи

Интеграция на прикладном уровне. На рисунке 2.2 *Вариант А* соответствует архитектуре гетерогенной сети, которая подразумевает интеграцию на прикладном уровне. Соответственно, предполагается наличие специализированного интерфейса верхнего уровня, который позволяет абонентскому терминалу взаимодействовать напрямую с поставщиком услуги, обмениваясь информацией по различным технологиям радиодоступа. Поскольку такая опция не предполагает координирования на сетевом уровне, она является достаточно про-

стой в реализации и уже в значительной степени была исследована в контексте улучшения работы отдельных приложений верхнего уровня. Данный вариант архитектуры в наибольшей степени подходит для оперативного повышения качества восприятия услуги пользователем, но остается существенно зависимым от способа его реализации в конкретном приложении. Он также может не в полной мере учитывать текущие условия работы сети связи, особенно в случаях, когда такие условия меняются динамически.

Интеграция на уровне ядра сети. *Вариант Б* соответствует недавним наработкам в комитете 3GPP, облегчающим интеграцию сотовых и WLAN систем при помощи взаимодействия на уровне опорной сети. Соответственно, механизм обнаружения и выбора сети доступа ANDSF обеспечивает выявление подходящих точек доступа WiFi, а также может задавать рекомендации по выбору сети. Однако окончательное решение об использовании той или иной технологии радиодоступа принимается непосредственно абонентским терминалом. Тем самым UE при принятии своего решения имеет возможность учесть текущее состояние канала радиосвязи, настройки мобильного оператора и предпочтения пользователя, что зачастую позволяет достичь более высокого качества обслуживания.

Данный вариант архитектуры сети обладает рядом преимуществ, поскольку может принять во внимание как настройки оператора, так и предпочтения пользователя. Однако эффективность работы соответствующих функций управления системой по-прежнему остается весьма ограниченной. Это происходит по той причине, что абонентские терминалы располагают частичным знанием текущих условий функционирования сети связи и, таким образом, могут принять только лишь локально оптимальные («жадные») решения, которые зачастую ведут к недостаточно эффективной работе системы в целом.

Несмотря на то, что абонентский терминал имеет возможность сообщить опорной сети о наблюдаемом им состоянии канала связи, обмен такой информацией не может происходить достаточно оперативно во избежание роста нагрузки на управляющие соединения. Следовательно, когда состояние канала меняется динамически, управление радиоресурсами на уровне сети радиодоступа RAN позволит добиться дальнейшего увеличения качества обслуживания. Именно поэтому более продвинутая архитектура, позволяющая интегрировать технологии радиодоступа на нижнем уровне RAN, представляет все возраста-

ющий интерес, поскольку имеет возможность учесть информацию о состоянии каналов радиосвязи в рамках гетерогенной сети в целом.

Интеграция на уровне сети радиодоступа. *Вариант В* описывает передовую архитектуру гетерогенной системы связи, которая интегрирует технологии 3GPP и WLAN на уровне сети радиодоступа RAN. Данная опция предполагает, что содействие со стороны абонентского терминала позволяет упростить обмен информацией между сотовой и локальной сетью доступа, что может потребовать создания соответствующих управляющих интерфейсов. Доступные в этом случае возможности по взаимодействию на уровне RAN ограничены емкостью транзитных (backhaul) каналов между различными сотами и технологиями радиодоступа [95]. В том случае, когда доступны транзитные каналы высокой емкости или же при совмещении функционала LTE и WiFi в рамках единого устройства доступа, может быть достигнуто более оперативное координирование. Это, в свою очередь, позволяет улучшить динамическое управление радиоресурсами и, как следствие, повысить как емкость системы связи, так и качество соединения абонента.

Кроме того, вовлечение сотовой инфраструктуры позволяет улучшить управление мобильностью и трафиком абонентского терминала. Соответственно, UE может использовать протоколы 3GPP для контроля соединения при подключении к малым сотам LTE или же к совмещенным LTE-WiFi модулям, а далее применять локальное переключение для доступа по технологии WLAN и, тем самым, снижать задержку передачи. Преимуществом такого подхода является более гибкая адаптация к динамическим изменениям радиоканала и интерференции в нем и, как следствие, снижение вероятности потери пакета и разрыва соединения. Важно отметить, что настройки мобильного оператора и предпочтения пользователя могут быть учтены через обратную связь со стороны абонентского терминала или же посредством определенным образом сконфигурированной аппаратуры RAN.

В итоге возможности координирования на уровне RAN могут варьироваться от использования простейшей вспомогательной информации (например, загрузка сети радиодоступа) до реализации тесной интеграции с последующим совместным или централизованным управлением радиоресурсами. Далее проводится анализ взаимодействия сотовых и локальных сетей доступа для различных вариантов архитектуры сети и способов ее размещения, а затем описываются некоторые алгоритмы управления информационными потоками. Осо-

бое внимание уделяется практическим вопросам, таким как наличие только лишь частичной вспомогательной информации, доступной абонентскому терминалу в гетерогенной сети, что является распространенной ситуацией в условиях ограничений на емкость управляющих соединений и вычислительный ресурс устройств.

2.3 Анализ интегрированной сотовой и локальной сети доступа

2.3.1 Совместный учет динамики трафика и размещения узлов

В данном разделе рассматривается важная задача выбора предпочтительной радиотехнологии на примере гетерогенной системы связи, включающей в себя сотовую и локальную сеть доступа [14]. Предполагается, что локальная сеть WLAN является частью интегрированной системы HetNet, которая размещена и управляется некоторым мобильным оператором. На основе механизмов, разработанных комитетом по стандартизации 3GPP, сотовая инфраструктура обеспечивает интеллектуальное *codeйствие* в процессе управления информационными потоками. Оно осуществляется на основе вспомогательной информации (например, местоположение абонента, требования к качеству обслуживания, измерения интерференции и загрузки ближайших сетей WLAN и т.д.), которая поступает от абонентских терминалов на специализированный управляющий модуль, расположенный на стороне сотовой инфраструктуры. Данный модуль рекомендует абонентским устройствам ту или иную радиотехнологию, и дальнейшее изложение направлено на изучение типовых взаимодействий в рассматриваемой системе.

Ниже рассматривается несовмещенное размещение малых сот LTE и точек доступа WiFi, а также не предполагается наличие прямого интерфейса взаимодействия между сотовой и локальной сетью доступа. Другой тип размещения, построенный на совмещенных LTE-WiFi модулях и подразумевающий более тесную интеграцию соответствующих сетей [96], будет рассмотрен в следующем разделе. Здесь же предполагается архитектура гетерогенной сети с интеграцией на уровне RAN, располагающая необходимыми динамическими

механизмами выбора сети радиодоступа. Приводится пример анализа системы для передачи данных по линии вверх, что представляет собой более интересный для исследования случай в связи с повышенной сложностью учета радиоинтерференции [97].

Основная новизна предлагаемого подхода состоит в явном учете динамики трафика *на уровне сессий* абонентского устройства. Это в значительной степени расширяет предыдущие результаты для статических (насыщенных) входных потоков. Новая сессия, представляющая собой, например, поток данных реального времени с некоторыми минимальными требованиями по скорости передачи, возникает у абонента случайным образом и заканчивается по истечении своего обслуживания [98]. Соответственно, число активных сессий изменяется со временем, что можно называть динамикой трафика на уровне сессий. Анализ такой приближенной к практике системы представляет собой актуальную задачу, но сопряжен с рядом сложностей. В связи с этим подходы к исследованию подобных динамических сетей связи значительно менее распространены в литературе по сравнению с методами анализа насыщенных систем.

Каждая сессия в рассматриваемой динамической системе связи, вообще говоря, представляет собой поток данных, вызванный отправкой файла, установлением удаленного соединения или же инициированием мультимедиа-передачи. Для моделирования таких сессий, порожденных множеством независимых абонентских терминалов, широко применяются пуассоновские процессы. Исторически проведение моделирования на уровне сессий было характерно для проводных сетей связи. Расширяя область применимости таких постановок для беспроводных систем, необходимо дополнительно учитывать взаимное размещение абонентских терминалов и модулей сетевой инфраструктуры. Подобные подходы, облегчающие совместный учет динамики трафика и размещения узлов, мало изучены в литературе, что обуславливает новизну предлагаемой ниже модели.

На сегодняшний день общеизвестно, что взаимное расположение узлов гетерогенной сети в значительной степени определяет ее итоговую эффективность [99]. Действительно, в связи с нерегулярным размещением терминалов и инфраструктуры наблюдается высокая степень неравномерности характеристик, которая может быть учтена в явном виде при помощи специальным образом подобранного случайного процесса. Например, местоположения абонентских терминалов на плоскости могут быть порождены некоторой реализаци-

ей стохастического потока точек [100]. Объединение моделей случайного размещения узлов и динамики их трафика позволяет принципиальным образом изменить подход к анализу мощности принятого сигнала и радиоинтерференции. При этом могут использоваться методы стохастической геометрии [101], которая широко применялась для анализа общесистемных характеристик случайного размещения абонентских терминалов и модулей сетевой инфраструктуры [102].

Использование стохастической геометрии (т.е. статистического моделирования пространственных взаимоотношений) приобрело в последние десятилетия значительную популярность, что связано с необходимостью анализа общесистемной эффективности, усредненной по результатам множества пространственных реализаций. Сравнительно недавно подобные методы были с успехом применены для исследования современных сотовых технологий связи, начиная с традиционного размещения макросот и вплоть до сверхплотных покрытий малыми (гетерогенными) сотами [90]. Построения стохастической геометрии обычно основаны на рассмотрении некоторого случайного точечного процесса на плоскости или в пространстве для целей статистического моделирования, например, расположения абонентских терминалов и последующего анализа аспектов, связанных с плотностью размещения терминалов, мощностью передачи, потерями при распространении радиоволн и интерференцией.

При отсутствии конкретной информации относительно местоположения абонентских терминалов простейшим способом моделирования их размещения на плоскости является использование равномерного распределения в целевой замкнутой области. Такой характер расположения сопряжен с применением однородного пуассоновского точечного процесса, который предполагает независимое и равномерное размещение некоторого числа точек в рамках данной области. Более сложным и, зачастую, реалистичным подходом к моделированию является использование биномиального процесса, который порождает фиксированное число точек в заданной области, а также пуассоновского кластерного процесса, позволяющего точкам группироваться в определенных локациях. Наконец, также выделяют так называемые процессы твердого ядра, которые получаются просеиванием пуассоновского точечного процесса так, чтобы обеспечить определенное минимальное расстояние между точками.

Последующее изложение ставит своей целью проведение аналитического моделирования гетерогенной системы связи, которое одновременно учитывало

бы динамику трафика во времени и размещение узлов в пространстве [103]. Предлагаемый здесь пространственно-временной подход позволяет провести исследование взаимодействия сотовых и локальных сетей доступа на уровне сессий [104] и, в частности, рассмотреть управление трафиком и выбор (назначение) предпочтительной радиотехнологии. Соответственно, в следующем подразделе вводится новая модель интегрированной гетерогенной сети доступа и затем проводится анализ общесистемных показателей ее эффективности.

2.3.2 Подход к анализу гетерогенных сетей доступа

В данном подразделе рассматривается модель гетерогенной сети доступа, которая включает в себя покрытие макро- и пикосот, а также множество точек доступа (ТД) WLAN [105], которые здесь называются *слоями*. Далее следуют основные допущения предлагаемой модели.

Рассматривается одна (типовая) макросота радиуса R с базовой станцией (БС), размещенной в ее геометрическом центре. Также предполагается наличие нескольких пико- БС и ТД WLAN, как показано на рисунке 2.3. Таким образом, макросота получает название *макрослоя*, тогда как множество пикосот и покрытие ТД WLAN именуются *пикослоем* и *слоем WLAN* соответственно. Каждая ТД/БС рассматриваемой системы связи может обслуживать абонентские терминалы, передающие данные по линии вверх. Рассматриваемый тип трафика соответствует сессиям реального времени с минимальной требуемой скоростью передачи данных r_0 .

На основе приведенных в [106] спецификаций 3GPP предполагается использовать *непересекающиеся* фрагменты беспроводного спектра для размещения трех описанных выше слоев. Как следствие, работа абонентских терминалов в рамках одного слоя не интерферирует с передачей в других слоях. При этом подключенные как к слою WLAN, так и к пикослою терминалы используют радиоресурсы совместно (с учетом соответствующей интерференции), тогда как в макрослое интерференция не рассматривается. Описанная модель системы представлена на рисунке 2.3, где показаны зоны покрытия слоев (макро-, пико- и WLAN), а также отмечены местоположения абонентских терминалов и модулей сетевой инфраструктуры.

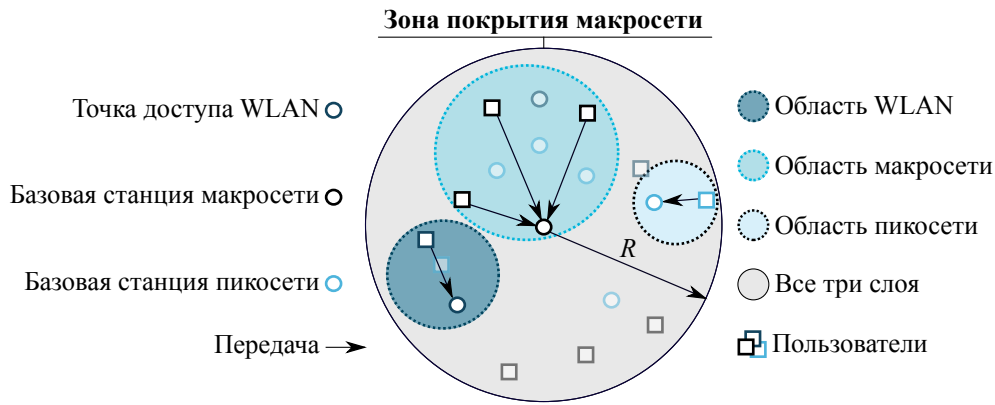


Рисунок 2.3 — Модель гетерогенной сети (слои представлены в виде сечений)

Предполагается, что передающие по линии вверх терминалы инициируют сессии, поступающие на обслуживание в гетерогенную сеть доступа согласно одномерному пуассоновскому процессу интенсивности λ во времени. Таким образом, новая сессия ставится в соответствие терминалу, находящемуся в определенной точке, которая не меняется на протяжении всей сессии. Для сохранения аналитичности также предполагается, что длительности сессий распределены по экспоненциальному закону со средним μ^{-1} , что может соответствовать гологовому или видеозвонку в реальном времени.

Для моделирования размещения узлов в описываемой сети доступа вводятся в рассмотрение ряд случайных процессов, а также формулируется набор базовых допущений.

Допущение (Размещение сетевой инфраструктуры). Местоположения ТД (в слое WLAN) и БС (в пикослое) на плоскости являются независимыми и распределены согласно пуассоновскому точечному процессу (PPP) интенсивности L_w и L_p соответственно.

Здесь значения L_w и L_p могут интерпретироваться как среднее число ТД/БС на единицу площади, а их конкретные местоположения могут, вообще говоря, находиться вне зоны покрытия макросоты (возможно, обслуживая терминалы из других сот).

Допущение (Размещение абонентских терминалов). Для учета динамики трафика принимается следующее допущение относительно местоположения терминалов. Местоположения абонентских терминалов, к которым поступают новые сессии, распределены на плоскости согласно PPP. Соответственно, область моделирования ограничена зоной покрытия рассматриваемой макросо-

ты (т.е. окружностью \mathbf{B} радиуса R), что обеспечивает равномерное размещение терминалов в рамках \mathbf{B} .

На практике ограничение, связанное с размещением абонентских терминалов в рамках определенной области, может быть продиктовано необходимостью учета максимальной мощности передачи и/или потерь в радиоканале. Равномерное распределение предполагается здесь только лишь в качестве примера. Вообще говоря, может быть рассмотрено произвольное распределение взаимного расположения терминалов $f(x,y)$, что приведет к техническому усложнению анализа, но не изменит предлагаемого подхода к моделированию.

Задача совместного учета динамики трафика и размещения абонентских терминалов обуславливает введение *пространственно-временного* процесса с интенсивностью, выражаемой обобщенной функцией $\Lambda(x,t)$, в которой $x \in \mathbb{R}^2$ представляет собой пространственный компонент, а $t \in \mathbb{R}^+$ является компонентом времени и

$$\Lambda(x,t) = \frac{\lambda}{S_R}, \text{ если } x \in \mathbf{B}, \quad (2.1)$$

где $S_R = \pi R^2$ – зона покрытия макросоты. Функция $\Lambda(x,t)$ имеет смысл частоты, с которой возникают новые поступления в пространстве $\mathbb{R}^2 \times \mathbb{R}^+$.

Допущение (Распространение радиосигнала). Коэффициент передачи канала $\gamma_{i,j}$ между терминалом U_i и ТД/БС j зависит от расстояния $d_{i,j}$ между ними. В соответствии с рекомендациями [107] предполагается, что для сессии i коэффициент передачи беспроводного канала связи $\gamma_{i,j}$ находится в степенной зависимости от расстояния:

$$\gamma_{i,j} = \frac{G}{d_{i,j}^{\kappa}}, \quad (2.2)$$

где $d_{i,j}$ – расстояние между ТД/БС и передающим терминалом; κ – экспонента распространения, а G – константа распространения радиосигнала. Параметры κ и G зависят, в частности, от применяемой технологии радиодоступа и задают определенную модель канала связи.

Поясняя приведенное выше допущение, следует отметить, что предложенные в [107] модели состоят из компонентов прямой (LoS) и непрямой (nLoS) видимости, переключаемых в соответствии с распределением Бернулли. Подобные модели можно легко описать степенной функцией, в которой параметры G и κ выбираются в соответствии с целевым сценарием и используемой радиотехнологией.

Поскольку предложенная модель предназначена, прежде всего, для анализа усредненной работы интегрированной системы связи в течение длительного интервала времени и без учета мобильности терминалов [108], влияние эффекта быстрого затухания сигнала может быть усреднено в предположении применения помехоустойчивого кодирования. Соответствующий вклад является в таких условиях минимальным и приблизительно постоянным, что позволяет легко его учесть. Эффект медленного затухания сигнала влияет в первую очередь на пропускную способность канала связи, и предлагаемый подход учитывает его в рамках функции распространения радиосигнала. Дополнительные соображения относительно эффектов затухания приведены в приложении А, где к пропускной способности канала добавляется некоторый случайный компонент. В итоге для учета эффектов затухания в данной модели вводится соответствующий поправочный коэффициент η , который уточняет пропускную способность канала связи.

Для повышения точности моделирования процессов передачи данных можно было бы принять во внимание эффекты затухания сигнала в явном виде. Однако это привело бы к значительному усложнению приведенного анализа, связанному, прежде всего, с необходимостью учета работы таких механизмов системы связи, как организация повторной передачи пакетов (HARQ), выбор модуляционно-кодированной схемы (MCS) и управление мощностью передачи в замкнутом цикле (closed-loop). Такое усложнение еще более затруднило бы наглядный анализ, а также привело бы к появлению дополнительных допущений и ограничений. Таким образом, в рамках данной модели коэффициент передачи канала γ остается ключевой конструкцией, связывающей излучаемую мощность с достижимой скоростью передачи данных в канале между абонентским терминалом и ТД/БС.

Далее предполагается, что скорость передачи данных является непрерывной функцией, а отображение мощность – скорость определяется формулой Шеннона. Такой подход был предложен в [109] и является достаточно точным для современных беспроводных сетей связи.

Допущение (Отображение мощность – скорость). Мощность сигнала p_i , излучаемого абонентским терминалом, и соответствующая скорость передачи данных r_i связаны друг с другом через «обобщенную» формулу Шеннона

$$r_i = B \log (1 + A p_i), \quad (2.3)$$

где p_i – мощность сигнала на выходе радиочастотного усилителя, а A и B представляют собой масштабирующие коэффициенты, которые зависят от выбора определенной технологии радиодоступа. К примеру, они могут быть заданы как

$$A = \frac{\eta \gamma_{i,i}}{N_0 + I}, \quad B = w, \quad (2.4)$$

где $\gamma_{i,i}$ – коэффициент передачи радиоканала между абонентом и ТД/БС для сессии i ; η – поправочный коэффициент; w – полоса пропускания канала; N_0 – уровень шума в канале связи и I – уровень интерференции на приемной стороне.

Для предотвращения неограниченного роста функции скорости передачи данных в зависимости от заданного набора модуляционно-кодовых схем определенной радиотехнологии предполагается, что скорость передачи остается постоянной r_{lim} , когда $d < d_0$. Данное соображение означает, что дальнейшее увеличение соотношения сигнал – шум (SNR) не повлечет за собой соответственного роста скорости передачи после некоторого порогового значения d_0 . Параметр d_0 можно определить как

$$d_0 = \left[\frac{G \cdot p}{(N_0 + I) (e^{r_{\text{lim}}/w} - 1)} \right]^{1/\kappa}. \quad (2.5)$$

В рамках предлагаемой системы допущений анализ случайной топологии сети доступа проводится совместно с исследованием ее динамики на уровне сессий. Такой подход предполагает введение адекватной модели системы массового обслуживания [110; 111], в которую поступают сессии случайной длительности, покидая ее по окончании своего обслуживания. В момент времени, когда новая сессия прибывает или же обслуженная сессия покидает систему, централизованный модуль управления, находящийся на стороне сотовой сети, осуществляет *контроль доступа* и *управление мощностью* во всех слоях. Данные процедуры позволяют определить, будет ли новая сессия допущена на обслуживание в определенный слой (контроль доступа) и какая будет при этом установлена мощность передачи для соответствующего абонентского терминала (управление мощностью).

Допущение (Контроль доступа). Каждая сессия реального времени требует минимальной скорости передачи данных r_0 . Таким образом, сеть связи допускает вновь поступившую сессию, если имеется достаточное количество радиоресурсов для ее обслуживания. Другими словами, каждая текущая сессия i занимает в точности долю r_0/r_i от общего времени работы системы (накладные

расходы на передачу сигнальных сообщений принимаются в расчет позднее), тогда как для всех активных сессий выполняется следующее:

$$\sum_{\text{all sessions}} \left(\frac{r_0}{r_i} \right) \leq \delta, \quad (2.6)$$

где δ – доступные на ТД/БС ресурсы (не включая ресурсы, предназначенные для компенсации затухания); $r_i \leq r_i^{\max}$ – мгновенная скорость передачи данных, зависящая от расстояния между абонентским терминалом и обслуживающей его ТД/БС и r_i^{\max} – предельно достижимая скорость передачи данных на максимальном уровне излучаемой мощности.

Кроме того, процедура контроля доступа должна определить, превышает ли текущий уровень интерференции некоторый заданный порог.

Допущение (Учет уровня интерференции). Предполагается, что слой гетерогенной сети с $n - 1$ активными сессиями принимает на обслуживание новую сессию n , если для множества $\{U_i\}_{i=1}^n$ всех сессий на каждой ТД/БС A_j выполняется следующее условие:

$$r_i \geq r_0 \text{ и } p_i \gamma_{i,j} \leq N_0, \quad \forall j, i \neq j, \quad (2.7)$$

где $\gamma_{i,j}$ – коэффициент передачи радиоканала между абонентом и ТД/БС j , а p_i – установленная соответственным образом излучаемая мощность радиосигнала.

Используя выражение $p_i \gamma_{i,j}$, требуется убедиться, что сессия не вызовет интерференцию выше, чем (модифицированный) уровень шума, допустимый для определенной технологии радиодоступа (см. подробности в [112]). Следовательно, процедура контроля доступа должна удостовериться, что для рассматриваемой сессии будет достигнута минимальная требуемая скорость передачи данных, а также что уровень интерференции на ТД/БС A_i , производимый соответствующим абонентским терминалом U_j , не будет превышать некоторого заданного порога. Этот порог зависит от характеристик используемой радиотехнологии и ниже обсуждается отдельно для слоя WLAN и пикослоя, тогда как в макрослое интерференция не рассматривается.

Допущение (Пороговый уровень интерференции). Дополнительно предполагается, что совокупный уровень шума и интерференции в канале принимает форму $N_0 + I = KN_0$, где значение K представляет собой некоторый масштабирующий коэффициент. Для рассматриваемой сети связи его принимают постоянным и способным учесть интерференцию, возникающую на стороне ТД/БС.

На практике каждая пикосота располагает средствами целенаправленного снижения интерференции, производимой обслуживаемыми ей терминалами на соседние БС. Они включают в себя всевозможные механизмы управления передачей по каналу связи, методы динамического планирования радиоресурсов и т.д. С учетом этого можно применить следующий инженерный подход. Индивидуальные вклады соседних терминалов в общий уровень интерференции могут быть агрегированы в рамках определенного слоя и добавлены к уровню шума в канале связи как некоторая поправка, которую в практических задачах сетевого планирования иногда именуют *коэффициентом интерференции*.

Следует отметить, что рассмотренные здесь оценки интерференции и скорости передачи данных служат в целях предсказания, и в системе связи они предназначены для решения вопроса о допуске той или иной сессии. Если некоторая сессия не может быть принята на обслуживание определенным слоем гетерогенной сети, она становится *заблокированной* с вероятностью $P_{block}^{(1/2/3)}$.

Допущение (Управление мощностью и планирование). Далее введем в рассмотрение два альтернативных подхода к совместному решению задач *управления мощностью и планирования радиоресурсов* [113], которые получают название алгоритм «*Максимальная мощность*» и алгоритм «*Кольцевое планирование*».

1. Алгоритм «Максимальная мощность» вначале устанавливает фиксированную мощность передачи абонентского терминала, равную наибольшему возможному значению в рамках используемой технологии радиодоступа. Затем процедура контроля доступа в сеть проверяет, может ли при данном уровне мощности быть достигнута требуемая минимальная скорость передачи данных.

2. Алгоритм «Кольцевое планирование» вначале удостоверяется, что ресурс сети доступа всегда разделен между передающими терминалами поровну, и затем осуществляет управление мощностью и контроль доступа соответствующим образом. В частности, каждая допущенная к обслуживанию сессия из общего числа активных n получает равную долю общего ресурса системы во времени, т.е. $\frac{r_0}{r_i} = \frac{1}{n}$. После этого абонентские терминалы устанавливают мощность своей передачи таким образом, чтобы достичь минимальной требуемой скорости передачи данных при условии, что необходимая для этого мощность не превышает максимально возможной.

Ниже рассматривается типовой (эвристический) принцип организации контроля доступа и выбора радиосети (см. рисунок 2.4). Соответственно, пред-

полагается *каскадное* обслуживание новых сессий, поступающих в систему. Вначале процедура содействия в выборе сети доступа, функционирующая на стороне сотовой системы связи, предпринимает попытку выгрузки вновь поступившей сессии на обслуживание *ближайшей* ТД WLAN, применяя для этого централизованный алгоритм контроля доступа в сеть WLAN. При этом выбранная ТД может, вообще говоря, находиться за пределами зоны покрытия макросоты R .

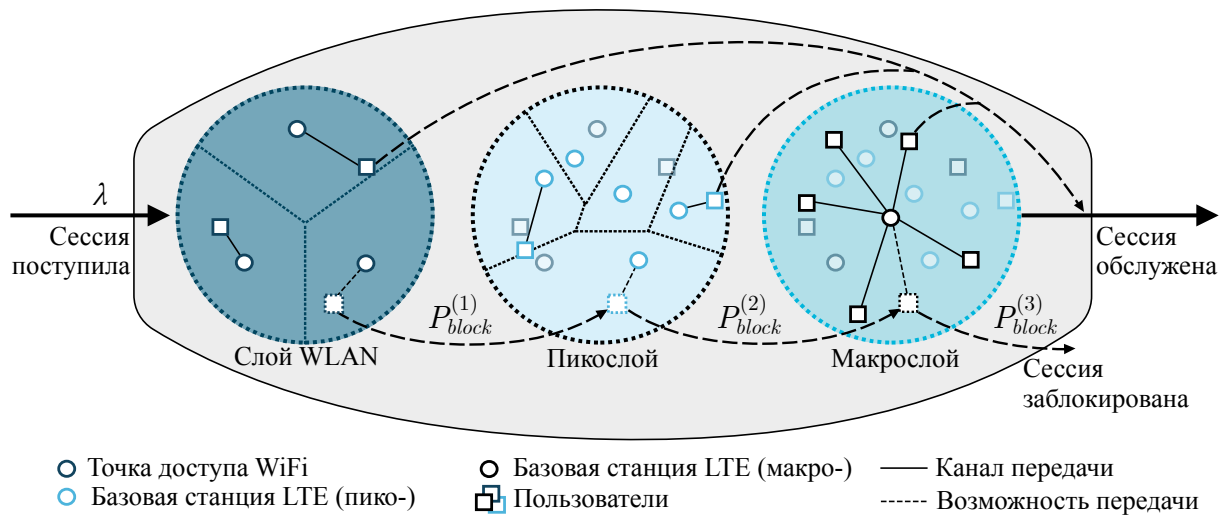


Рисунок 2.4 — Каскадный принцип работы системы со слоями WLAN, пико- и макро-

Если сессия допускается в слой WLAN, она обслуживается в нем без прерываний до момента своего окончания и затем покидает систему. В противном случае, если данная сессия не может быть принята в слой WLAN, запускается процедура контроля доступа в пикослой. Как следствие, сессия либо допускается в пикослой и обслуживается в нем *ближайшей* малой БС, либо же макрослой предпринимает попытку принять эту сессию на обслуживание. В итоге, если некоторая сессия не может быть допущена даже в макрослой, то она считается *окончательно заблокированной* и покидает систему, не получив обслуживания, но не оказывая при этом влияния на последующие поступления новых сессий.

Рассматриваемый здесь последовательный контроль доступа *с сетевым содействием* представляет собой некоторое промежуточное решение между распределенными (некоординированными) и централизованно-управляемыми подходами. Тогда как первые могут работать недостаточно эффективно, применение вторых зачастую влечет за собой значительную вычислительную сложность и

соответствующие накладные расходы в виде управляющих сообщений. Напротив, предлагаемый каскадный принцип организации системы позволяет «упаковать» емкость сети доступа последовательно и, тем самым, становится адекватной опцией для использования в практических гетерогенных системах связи. Другие способы функционирования сети могут быть рассмотрены по аналогии в рамках описанной в данном разделе модели.

Кроме того, выбор представленного на рисунке 2.4 каскадного принципа обслуживания может быть также обусловлен соображениями снижения издержек абонента и оператора. С учетом того, что доступ по технологии WiFi часто возможен без дополнительной оплаты, бывает предпочтительно выгрузить наибольшее возможное количество сессий в слой WLAN, коль скоро соблюдаются требования по качеству обслуживания (которые в данной модели отслеживаются процедурой контроля доступа с сетевым содействием). Когда система WLAN уже достаточно загружена и не может принять на обслуживание дополнительные сессии, вторым предпочтительным кандидатом становится пикослой. Это связано, прежде всего, со стремлением мобильного оператора сбалансировать загрузку размещенных в сети доступа малых сот (особенно для трафика от абонентских терминалов с низкой мобильностью). При таком подходе традиционный способ обслуживания в макрослое сохраняется оператором в качестве резервного и используется, когда остальные становятся неприменимы.

После постановки на обслуживание активный терминал использует долю ресурса системы связи во времени, устанавливая свою излучаемую мощность так, чтобы достичь требуемой скорости передачи данных. При каждом поступлении дополнительной сессии или же когда одна из текущих сессий заканчивается, гетерогенная сеть доступа принимает новое решение относительно планирования ресурсов и управления мощностью передачи для всех активных абонентских терминалов. В каждом слое вводится в рассмотрение свое значение вероятности блокировки $P_{block}^{(i)}$ и вероятности допуска сессии $P_a^{(i)} = 1 - P_{block}^{(i)}$, где $i = 1, 2, 3$ – номера слоев WLAN, пико- и макро- соответственно. При этом исходная интенсивность поступления сессий на (первый) слой WLAN составляет $\lambda_w = \lambda$.

Допущение (Разделение слоев). Для сохранения аналитичности предложенной модели гетерогенной сети предполагается, что три рассматриваемых слоя обслуживают соответствующие абонентские терминалы независимо. Это приво-

дит к случайному просеиванию исходного процесса поступления сессий с соответствующими вероятностями допуска.

Данное допущение представляет собой естественный методический ход в целях разделения сложной для анализа системы слоев на отдельные аналитические компоненты, которые в принципе могут быть заменены на другие или же изъяты из рассмотрения. Например, можно было бы изменить порядок следования слоев в каскадной процедуре контроля доступа.

Ввиду пуассоновского характера просеянного потока, поступления сессий во (второй) пикослой (т.е. не принятых в слой WLAN) происходят в соответствии с пуассоновским процессом интенсивности $\lambda_p = \lambda (1 - P_a^{(1)})$, где $P_a^{(1)}$ – вероятность допуска сессии в слой WLAN. Аналогично поступления в (третий) макрослой подчиняются пуассоновскому процессу интенсивности $\lambda_m = \lambda (1 - P_a^{(1)}) (1 - P_a^{(2)})$, где $P_a^{(2)}$ – вероятность допуска сессии в пикослой.

Абстрагируясь от конкретного местоположения абонентских терминалов, поступления сессий в пико- и макрослой предполагаются равномерными внутри окружности радиуса R . Важно заметить, что это допущение заведомо *не выполняется* на практике. На самом деле существует некоторая зависимость, в соответствии с которой сессии принимаются на обслуживание в слой WLAN и в пикослой. Однако ниже с помощью результатов имитационного моделирования будет показано, что такое предположение о равномерности поступлений является достаточно точным. Данный факт позволяет считать анализ модели, заданной приведенным выше набором допущений, адекватным для аппроксимации функционирования реальной гетерогенной сети доступа.

В итоге *общесистемная вероятность блокировки* P_{block} может быть вычислена как

$$P_{block} = 1 - \left[P_a^{(1)} + (1 - P_a^{(1)}) P_a^{(2)} + (1 - P_a^{(2)}) (1 - P_a^{(1)}) P_a^{(3)} \right], \quad (2.8)$$

где $P_a^{(3)}$ – вероятность допуска сессии в макрослой.

2.3.3 Результаты анализа интегрированной сети доступа

Для проверки точности предложенной аналитической модели и верификации соответствующей системы допущений в данном подразделе приводится на-

бор тестовых примеров. Разработанная система имитационного моделирования в явном виде воспроизводит все случайные процессы, описанные в предыдущем подразделе. Это позволяет учесть «память» стохастической системы и сравнить результаты детального моделирования с предложенными аналитическими аппроксимациями. В частности, по описаниям в рекомендациях ITU/3GPP был воссоздан городской сценарий (см. подробности, например, в [114]), для которого варьировалась плотность БС пикослоя и ТД слоя WLAN. Особое внимание было уделено плотному размещению ТД/БС, поскольку в таких условиях интеллектуальный выбор сети доступа более полезен: сотовая сеть в одиночку испытывает трудности при обслуживании трафика от абонентов.

В качестве типового примера рассматривается сота радиуса R , который составляет 288 метров (согласно рекомендациям в [107]), с переменной интенсивностью поступления новых сессий λ . Каждая вновь прибывшая сессия имеет случайную длительность со средним μ^{-1} , равным 3 секундам, а также минимальную требуемую скорость передачи данных r_0 500 кбит/с. Доступная полоса пропускания для слоев WLAN, пико- и макро- составляет 20, 10 и 10 МГц соответственно. При этом допустимая мощность передачи абонентского терминала ограничена значением в 23 дБм в слое WLAN и в макрослое, тогда как в пикослое это значение принято 20 дБм. Остальные параметры системы установлены в соответствии с доступными спецификациями 3GPP и другими документами стандартизации.

Графики на рисунках 2.5 и 2.6 характеризуют эффективность обслуживания в макрослое с учетом и без учета вклада от эффекта медленного затухания сигнала, который в принципе может оказать влияние на работу каждого из трех рассматриваемых слоев гетерогенной сети доступа. Обычно медленное затухание моделируют случайной величиной, распределенной по логнормальному закону с параметрами $\mu = 0$ и $\sigma = 6$ дБ в соответствии с городским макросценарием в [107]. Сопутствующие вспомогательные распределения могут быть вычислены, как показано в приложении А. Можно заключить, что, несмотря на ряд упрощающих допущений, предложенный подход к анализу гетерогенных систем связи остается достаточно точным на всем интервале рассматриваемых значений интенсивности поступления новых сессий.

Зависимости на рисунках 2.7 и 2.8 иллюстрируют поведение среднего числа сессий в системе как функцию от числа ТД/БС в слое WLAN/пикослое соответственно. Такие результаты подтверждают, что с ростом плотности разме-

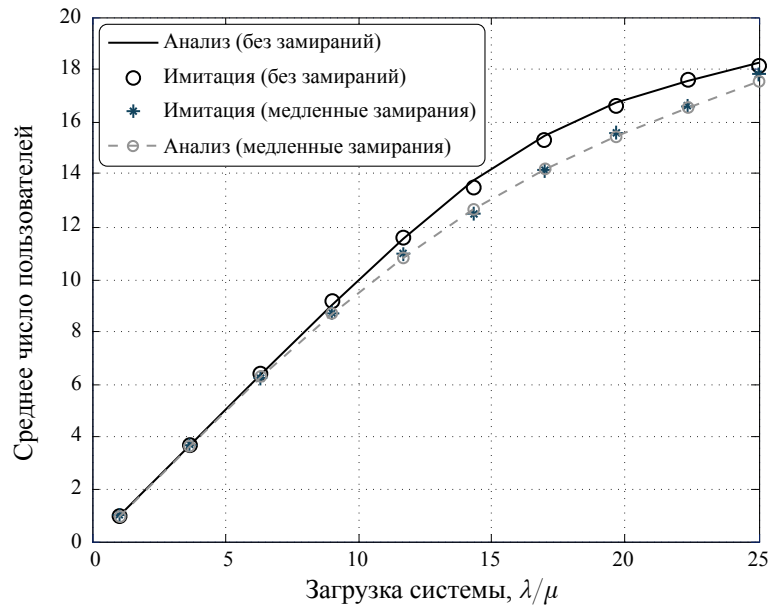


Рисунок 2.5 — Среднее число сессий в макрослое с учетом затухания

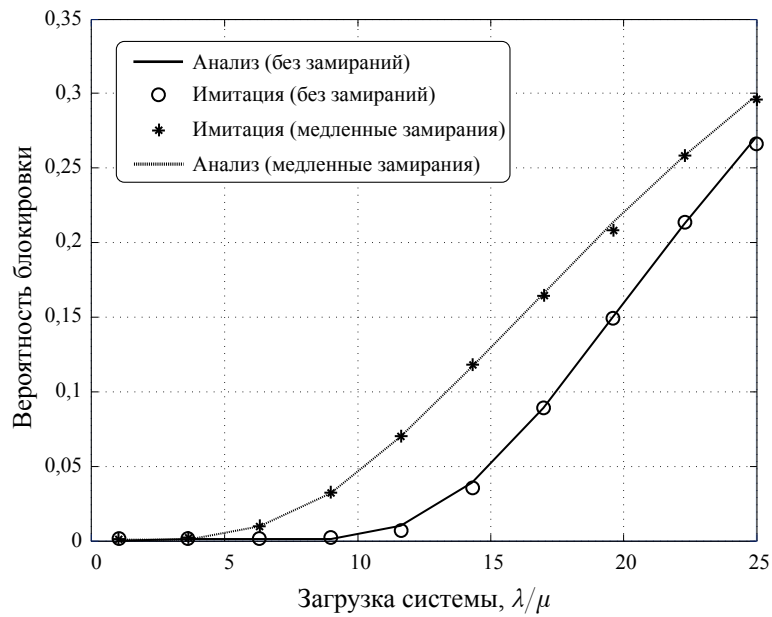


Рисунок 2.6 — Вероятность блокировки в макрослое с учетом затухания

щения модулей сетевой инфраструктуры эффективность обслуживания также существенно возрастает. Наконец, рисунок 2.9 посвящен более глубокому исследованию соответствующих значений вероятности блокировки в рамках рассматриваемой интегрированной сети доступа, а также в каждом из ее слоев по отдельности: макро-, пико- и WLAN. Можно сделать вывод, насколько возрастает эффективность работы гетерогенной сети с двумя дополнительными слоями по сравнению с базовой сотовой системой, включающей только лишь макрослой. Особо стоит отметить значительный прирост производительности

уже для небольшого числа дополнительных модулей инфраструктуры, т.е. четыре ТД WLAN и четыре малые БС в данном примере.

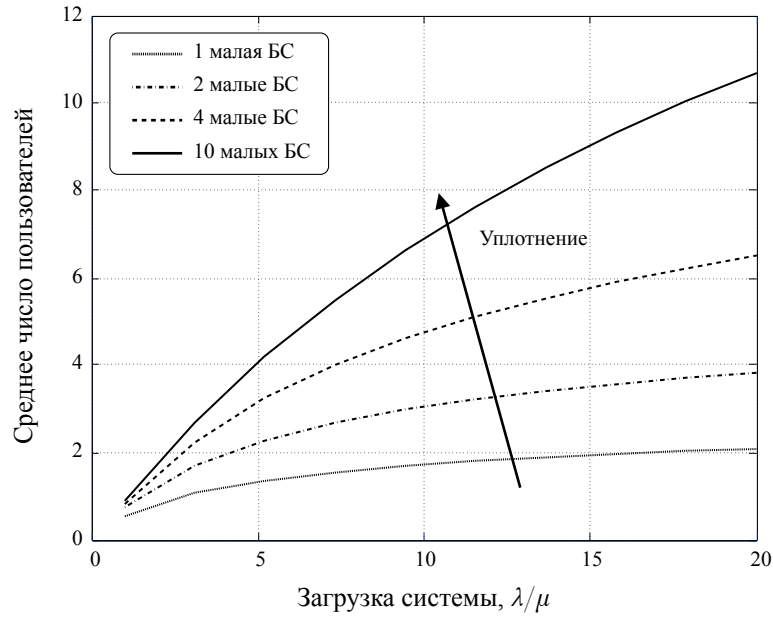


Рисунок 2.7 — Зависимость среднего числа сессий от плотности БС в пиковое

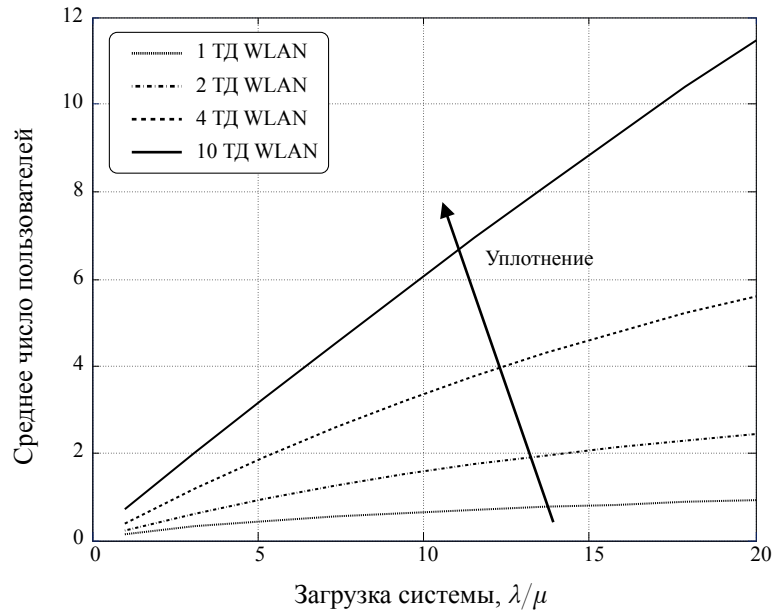


Рисунок 2.8 — Зависимость среднего числа сессий от плотности ТД в слое WLAN

Резюмируя, можно отметить, что различные технологии радиодоступа в совокупности с интеллектуальными процедурами выбора радиосети становятся неотъемлемой частью гетерогенных систем 5G. В частности, согласованное использование множества сетей доступа позволяет достичь значительного прироста

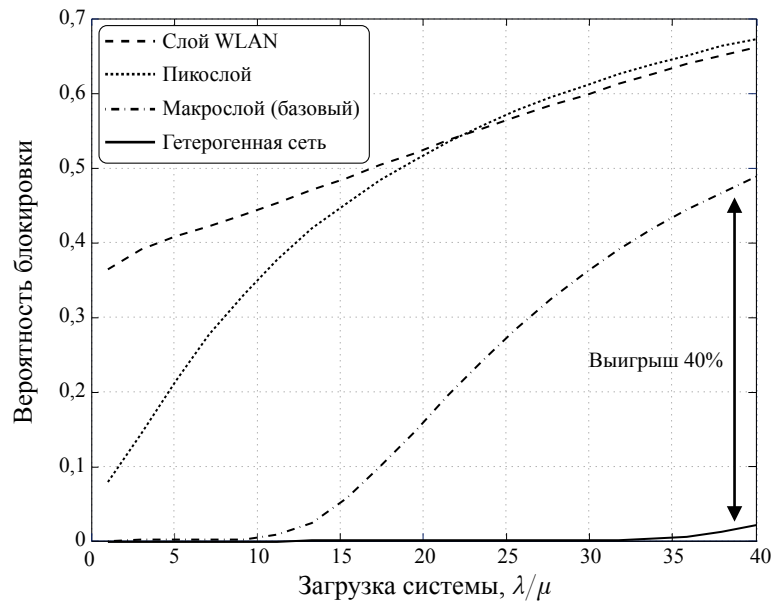


Рисунок 2.9 — Вероятность блокировки в интегрированной гетерогенной сети с тремя слоями

ста эффективности обслуживания абонентов как по величине емкости системы связи, так и в уровне качества пользовательского соединения.

2.4 Анализ уплотнения сетей доступа и учет динамики их загрузки

2.4.1 Совмещенные модули доступа LTE и WiFi

Недорогие малые соты различных типов, имеющие более низкое энергетическое потребление по сравнению с макросотой, являются на сегодняшний день неотъемлемой составляющей типовой сети мобильного оператора, необходимой для обеспечения высоких скоростей передачи данных [115]. Благодаря своим компактным размерам малые БС LTE часто используются для организации непрерывного сетевого покрытия, экономии капитальных средств и операционных расходов, а также снижения энергетического потребления системы доступа [116]. В результате стирается грань между ними и ТД WiFi, которые уже зарекомендовали себя как надежное, проверенное временем сетевое решение для обеспечения локального доступа [91]. Данная тенденция создает предпо-

сылки для все более тесной интеграции технологий радиодоступа LTE и WiFi с целью эффективной разгрузки сотовой сети доступа и гибкого управления трафиком абонентов, что достигается путем совместного использования радиоресурсов обеих систем связи [117].

Как уже отмечалось выше, значительные усилия комитета 3GPP направлены на стандартизацию механизмов интеграции радиотехнологии WiFi как части инфраструктуры сотовой сети LTE. Они включают в себя организацию доверенного доступа WiFi-устройств к услугам сотовой системы связи, обеспечение бесшовного переключения между LTE и WiFi, а также функционал для обнаружения и выбора сети доступа ANDSF. Начиная с 3GPP Release 12 ведется работа по облегчению оперативной интеграции сотовой и локальной систем связи на уровне сети радиодоступа RAN. При этом содействие со стороны сотовой инфраструктуры может привести к ряду улучшений в функционировании интегрированной гетерогенной сети LTE-WiFi [53]: от предоставления простейшей вспомогательной информации (например, о текущей загрузке системы [118]) и вплоть до полноценного (централизованного) управления радиоресурсами.

По мере упрощения взаимодействия между LTE и WiFi становится более выгодно размещать данные технологии радиодоступа совместно. Это связано, прежде всего, с весьма незначительным удорожанием стоимости совмещенных LTE-WiFi модулей по сравнению, например, с малыми БС LTE [119], что является невысокой платой за возможность использования нелицензированных спектральных радиоресурсов WiFi. При применении механизмов тесной интеграции LTE и WiFi в рамках гетерогенной сети доступа абонентский терминал, имеющий два радиомодуля, может вести передачу данных по ним *одновременно*, что существенно повышает эффективность его работы [90]. Однако такие продвинутые режимы передачи для совмещенных систем LTE-WiFi требуют дополнительного функционала, включая соответствующие сетевые интерфейсы [120]. Ожидается, что разработка механизмов управления совмещенными сетями доступа LTE и WiFi приведет к существенному росту показателей качества связи абонента и заметному улучшению характеристик работы системы связи 5G [121].

В итоге размещение малых сот различных типов, функционирующих как в лицензированном (например, LTE), так и в нелицензированном (например, WiFi) спектральном диапазоне, станет основой гетерогенных систем 5G [122]. При этом использование совмещенных LTE-WiFi модулей более выгодно для

целей оперативного уплотнения сети доступа в силу более гибкого управления радиоресурсами [123]. Однако в имеющейся исследовательской литературе подходы к моделированию совмещенного размещения малых БС LTE и ТД WiFi, имеющих прямой интерфейс взаимодействия, еще не получили должного внимания. В следующем подразделе вводится новая модель совмещенной гетерогенной сети LTE-WiFi, которая позволяет провести анализ ее уплотнения и учесть динамику ее загрузки.

Как и в предыдущих разделах этой главы, предлагаемый подход объединяет методы стохастической геометрии для моделирования размещения узлов [102] и методы теории массового обслуживания для моделирования динамики трафика от абонентских терминалов [124]. Учитывая, что загрузка малых сот может значительным образом варьироваться во времени и пространстве [87], ниже рассмотрим ограниченную область моделирования, в которой имеется (сверх)плотное размещение совмещенных LTE-WiFi модулей сетевой инфраструктуры (например, на элементах уличной фурнитуры). Затем проводится сравнительное исследование двух алгоритмов управления функционированием такой совмещенной гетерогенной сети доступа: первый связан с выбором одной (предпочтительной) технологии радиодоступа для осуществления передачи, тогда как второй (предлагаемый) позволяет вести одновременную передачу по обеим радиотехнологиям. Также приводятся результаты имитационного моделирования с применением соответствующего комплекса программ, который был предварительно откалиброван с тестовыми данными 3GPP.

2.4.2 Моделирование сверхплотных совмещенных покрытий

Далее исследуется интегрированная гетерогенная сеть доступа, построенная на сверхплотном покрытии некоторой области совмещенными модулями LTE-WiFi [125], а также вводятся допущения соответствующей математической модели.

Описание целевого сценария. Рассматривается замкнутая *область моделирования* в виде квадрата со стороной D (рисунок 2.10). Кроме того, предполагается наличие множества *интегрированных малых сот*, где каждая совмещенная БС (СБС) оснащена как функционалом LTE, так и функционалом

WiFi. С учетом рекомендаций комитета 3GPP размещать малые соты в выделенном диапазоне радиочастот [126], который не перекрывается с доступными для макросот частотами, все СБС используют общие радиоресурсы LTE. Здесь предполагается, что макрослой служит преимущественно для управления системой и для планирования радиоресурсов, но не участвует в обслуживании трафика, поступающего от абонентских терминалов. При этом сеть WiFi традиционно функционирует в нелицензированном спектральном диапазоне.

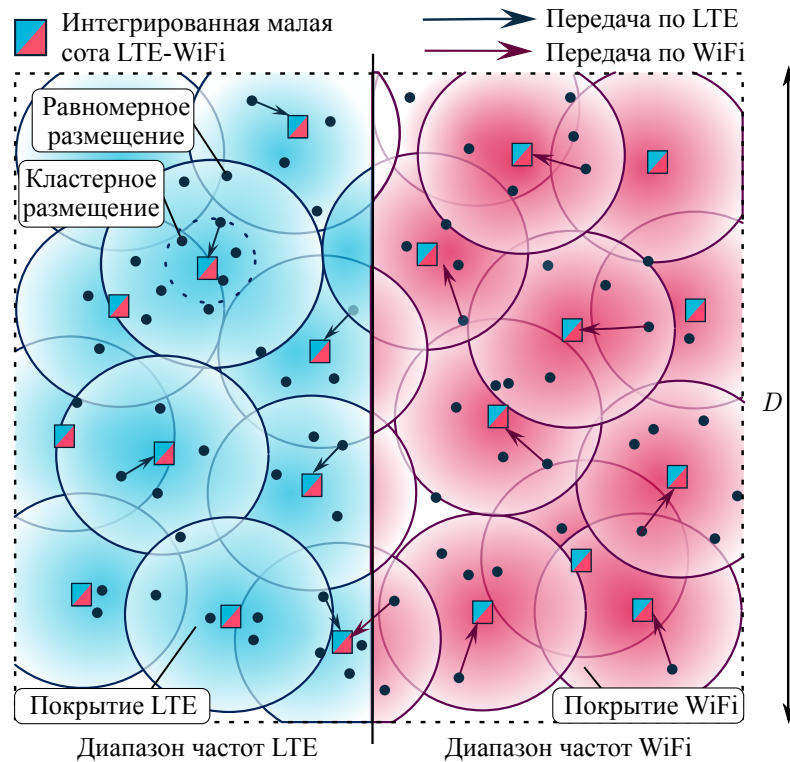


Рисунок 2.10 — Структура совмещенной гетерогенной сети и область моделирования

На практике подобные размещения планируются для обеспечения сетевого доступа в местах большого скопления абонентов, таких как торговые центры. В этих условиях централизованный модуль управления, находящийся на стороне сотовой сети, располагает полной информацией относительно структуры системы связи, включая местоположения СБС и обслуживаемых ими абонентских терминалов. Для последующего моделирования представляет непосредственный интерес только лишь слой СБС, который обслуживает терминалы, передающие свои данные по линии вверх, тогда как управляющий макрослой в явном виде не рассматривается. Такой подход отражает ожидание того, что роль макрослоя в гетерогенных сетях 5G будет постепенно редуцирована ис-

ключительно до функций управления и координирования работы остальных слоев [127].

Последующее изложение сосредоточено на вопросах управления работой абонентских терминалов, находящихся внутри заданной области моделирования и имеющих возможность передавать свои данные по линии вверх с использованием одной или же обеих радиотехнологий: LTE и WiFi [128]. Поскольку в перспективных гетерогенных сетях организация доступа по линиям вверх и вниз фактически обеспечивается двумя различными сетями [90], здесь сужение до анализа исходящей передачи данных от абонентских терминалов связано с более высокой сложностью учета радиоинтерференции по сравнению с доставкой информации в обратном направлении [97]. Как и ранее, предлагаемый подход представляет собой пример пространственно-временного моделирования, проводимого на основе предлагаемой ниже системы допущений.

Размещение абонентских терминалов и сетевой инфраструктуры. Местоположения СБС выбираются согласно однородному пуассоновскому точечному процессу PPP S_{BS} интенсивности μ . Тогда число СБС в целевой замкнутой области представляет собой случайную величину, распределенную по пуассоновскому закону, в то время как число СБС в неперекрывающихся областях – независимые между собой величины. Вообще говоря, ниже предполагается *сверхплотное* размещение совмещенных LTE-WiFi модулей, являющееся более частым по сравнению с типовым сотовым покрытием (шестиугольными сотами) и определяемое здесь формально как $\mu\pi(2R)^2 > 7$, где R – радиус зоны покрытия отдельной СБС. В итоге вероятность того, что покрытие в какой-либо точке пространства отсутствует, достаточно мала.

Из практических соображений такие плотности размещения СБС могут быть оправданы в случаях, когда предпочтительнее использовать множество недорогих станций сравнительно низкой емкости в силу заданных ограничений на минимальную скорость передачи данных или на характеристики покрытия (например, при отсутствии достаточного числа подходящих точек размещения для БС макрослоя). Подобные сценарии наблюдаются на транспортных узлах, в торговых центрах и т.д. Даже при наличии функционала WiFi такие СБС существенно более безопасны и дешевы в эксплуатации по сравнению с традиционными телекоммуникационными технологиями.

Местоположения абонентских терминалов распределены согласно *смеси* процесса PPP и кластерного процесса. При этом определенная доля ε термини-

налов размещена в соответствии с PPP, тогда как остальные терминалы сконцентрированы вокруг обслуживающих их СБС в соответствии с *кластерным* точечным процессом, основанным на процессе S_{BS} . Итоговое распределение расстояний от терминалов до соответствующих узлов инфраструктуры описывается как $f_d(d) = \varepsilon f_u(d) + (1 - \varepsilon) f_c(d)$, $d \leq R$, где $f_c(d)$ – распределение расстояний для терминалов в составе кластеров.

В частности, расположение абонентских терминалов, размещенных в кластерах, подчиняется кластерному процессу Матерна, в котором местоположения точек распределены равномерно внутри окружности радиуса r , описанной вокруг выбранной согласно PPP точки. В рассматриваемом случае такие местоположения сконцентрированы вокруг СБС, а также предполагается, что значение r достаточно мало, чтобы терминалы на соответствующих расстояниях были бы *всегда* ассоциированы с той СБС, которая находится в центре кластера. Данные рассуждения построены на предложениях, содержащихся в актуальной документации 3GPP, которые рекомендуют использование взвешенной суперпозиции равномерного и кластерного распределений для моделирования реального размещения абонентских терминалов [119]. Данные рекомендации соответствуют ряду практических сценариев применения гетерогенных сетей, предназначенных для обслуживания абонентов в магазинах, кафе, ресторанах и прочих общественных местах.

Динамика трафика и модель канала. Новые сессии, связанные с передачей данных по линии вверх, возникают внутри рассматриваемой области моделирования согласно стационарному однородному пуассоновскому процессу интенсивности λ во времени. Далее вновь поступившая сессия ставится в соответствие некоторому абонентскому терминалу на плоскости. Без ограничения общности можно полагать, что смена активным терминалом своего текущего местоположения тождественна появлению «нового» терминала, размещенного в этом изменившемся местоположении. Ввиду отсутствия «памяти» у рассматриваемого процесса, такая интерпретация позволяет в неявном виде учесть мобильность абонентских терминалов.

Для определенности предполагается, что сессия соответствует *передаче файла*, размер которого распределен экспоненциально со средним значением θ бит. По завершении его передачи соответствующий абонентский терминал покидает систему. Фактическое местоположение терминала i , к которому поступает готовый для передачи файл, определяет качество его соединения с обслужи-

вающей СБС. Таким образом, соответствующая скорость передачи данных r_i определяется по формуле Шеннона как

$$r_i = \min\{r_{\text{lim}}, w \log(1 + \text{SINR}_i)\}, \quad (2.9)$$

где $\text{SINR}_i = \frac{\gamma_i}{N_0 + I} p_i$; p_i – уровень излучаемой терминалом мощности; γ_i – коэффициент передачи радиоканала между источником и получателем сессии i ; w – эффективная полоса пропускания канала связи; N_0 – уровень шума в канале; I – уровень интерференции на стороне получателя; r_{lim} – предельно достижимая скорость передачи данных.

Для некоторой сессии коэффициент передачи радиоканала от ее источника до получателя γ_i описывается зависимостью излучаемой мощности от расстояния d_i между ними, т.е. $\gamma_i = \frac{G}{d_i^\kappa}$, где κ – экспонента распространения радиосигнала, а G – константа распространения. Введенное таким образом определение коэффициента передачи радиоканала может быть легко модифицировано для учета большинства типовых случаев, предписанных комитетом 3GPP.

Описываемая здесь модель представляет собой гибкий инструмент для анализа достижимых пределов уплотнения гетерогенной сети путем абстрагирования таких менее значимых факторов, как колебания в уровне излучаемой мощности, связанные с компенсацией эффектов затухания сигнала. Соответственно, детальное исследование эффектов затухания в данной модели не производится в целях сохранения ее аналитичности, но необходимые поправки могут быть введены в виде корректно подобранных коэффициентов затухания или же учтены в виде дополнительных случайных величин, непосредственно участвующих в выражении для коэффициента передачи радиоканала, как это было сделано в предыдущих разделах данной главы.

В зависимости от качества своего соединения активный абонентский терминал использует некоторую долю радиоресурсов LTE и/или WiFi, предоставленную обслуживающей его сессии СБС на основании некоторого *централизованного* алгоритма управления. Применение подобных алгоритмов становится возможным в гетерогенных сетях 5G в силу более тесного координирования (совмещенных) технологий радиодоступа LTE и WiFi, размещенных, как правило, одним и тем же мобильным оператором, а также как следствие концентрации функций управления интегрированной системой связи на стороне сотовой сети доступа.

Планирование и выделение радиоресурсов. Рассматриваемая сеть связи ассоциирует новую сессию с ближайшей к ней СБС в том случае, если число уже обслуживаемых данной СБС абонентских терминалов не превышает некоторого предельного значения N_{\max} . На практике такой максимум фактически существует для любой БС LTE или ТД WiFi, например, в силу ограничений на объем памяти системы планирования. Поэтому предположение о том, что данное число является конечным (пусть и достаточно большим), может быть сделано здесь без потери общности.

В рамках моделируемой системы доступа выделение радиоресурсов LTE происходит согласно алгоритму «Кольцевое планирование», который назначает одинаковый объем радиоресурсов всем активным сессиям данной СБС. При этом радиоресурсы WiFi также достаются абонентским терминалам приблизительно в равных долях благодаря работе протокола доступа к каналу, построенного на принципах случайного множественного доступа. Таким образом, каждая из n текущих сессий, обслуживаемых некоторой СБС, получает свою (равную) долю общего частотно-временного ресурса. Это позволяет выразить предельно достижимую скорость передачи данных как $\tilde{r}_i = \frac{\delta_w/l}{n} r_i$, где r_i – мгновенная скорость передачи, а δ_w и δ_l – доли радиоресурсов, доступные на данной СБС для обслуживания трафика WiFi и LTE соответственно. Новые сессии, получившие скорость передачи \tilde{r}_i в момент своего допуска в систему, обслуживаются в ней без прерываний и после этого немедленно покидают систему.

На практике приведенные выше рассуждения соответствуют случаю, когда число активных сессий меняется во времени незначительно, что очевидно выполняется между моментами времени, совпадающими с принятием на обслуживание новых сессий. Выбор рассматриваемого здесь способа планирования обусловлен отсутствием предварительной информации относительно приоритета между абонентскими терминалами. При необходимости учет такого приоритета может быть проведен с помощью надлежащего выбора весовых коэффициентов и представляет собой отдельную задачу, которая в данной главе не рассматривается. При этом предлагаемый подход к моделированию может быть с легкостью обобщен на случай другой дисциплины обслуживания, а простейший алгоритм «Кольцевое планирование» рассматривается здесь исключительно в качестве примера.

Следует также отметить, что можно было бы, в частности, рассмотреть популярный алгоритм планирования, предоставляющий более «справедливое»

выделение радиоресурсов LTE (proportional-fair), который максимизирует сумму логарифмов скорости передачи с учетом заданных весовых коэффициентов. Для приведенного здесь простейшего случая с равными долями радиоресурсов такие коэффициенты предполагаются одинаковыми в отсутствие дополнительных сведений относительно приоритета трафика. Более того, для так называемого эластичного трафика, рассматриваемого ниже, предполагается, что если доля радиоресурсов, предоставленных терминалу i , составляет ε_i , то можно легко показать, что $\sum_{i=1}^n \log(\varepsilon_i r_i)$ достигает своего максимума, когда $\varepsilon_i = \frac{1}{n}$. Это в том числе означает, что для такого типа трафика анализ «справедливого» способа выделения радиоресурсов тождественен изучению алгоритма «Кольцевое планирование».

В свою очередь, функционирование радиотехнологии WiFi подразумевает вероятностный доступ к каналу с разделением времени между всеми активными терминалами, обслуживаемыми данной СБС. На достаточно длительных промежутках времени такое разделение эквивалентно результату работы «стохастического» алгоритма «Кольцевое планирование». С учетом этих рассуждений последующее изложение предполагает использование алгоритма «Кольцевое планирование» как для системы LTE, так и для системы WiFi.

Управление мощностью передачи и интерференцией. Поскольку изучение вопросов, связанных с радиоинтерференцией в канале связи, является ключевым в сверхплотных гетерогенных сетях 5G [129; 130], далее отдельно вводится набор соответствующих допущений моделирования.

1. Предполагается, что алгоритм управления мощностью передачи абонентского терминала в сети LTE нацелен на поддержание некоторого заданного уровня SINR η [131], который принимается постоянным во всей сотовой системе. В том случае, если целевое значение SINR не может быть достигнуто, терминал передает свои данные на предельно допустимом уровне мощности p_{\max} , пытаясь обеспечить определенное качество обслуживания. Альтернативой могло бы служить изъятие данного абонентского терминала из расписания, но такой вариант здесь не рассматривается.
2. Интерференция в сети LTE поддерживается на некотором постоянном уровне путем явного координирования межсотового расписания передач терминалов [132]. Следовательно, если некоторая сессия, запланированная к обслуживанию на определенных радиоресурсах с некоторой

- требуемой мощностью, вызывает существенный прирост интерференции в другой соте, то соответствующие радиоресурсы *исключаются* из процесса планирования в «пострадавшей» соте (см. рисунок 2.11).
3. Для передачи данных в сети WiFi излучаемая абонентскими терминалами мощность предполагается постоянной и поддерживается на своем предельном уровне p_{\max} , который при этом может отличаться от предельного значения мощности передачи в сети LTE.
 4. Интерференция в сети WiFi регламентируется протоколом доступа к среде, построенным на использовании функции DCF и соответствующих сигнальных сообщений RTS/CTS. Как следствие, каждое соединение WiFi «резервирует» в вероятностном смысле некоторую область вокруг как источника, так и получателя сессии (для передачи данных и подтверждений об их доставке). Новое соединение может быть установлено, если ни ее терминал-источник, ни соответствующий терминал-получатель не находятся в уже «зарезервированных» другими активными WiFi-сессиями областях. При подобном координировании предполагается, что общий уровень интерференции не превышает некоторого известного порога.

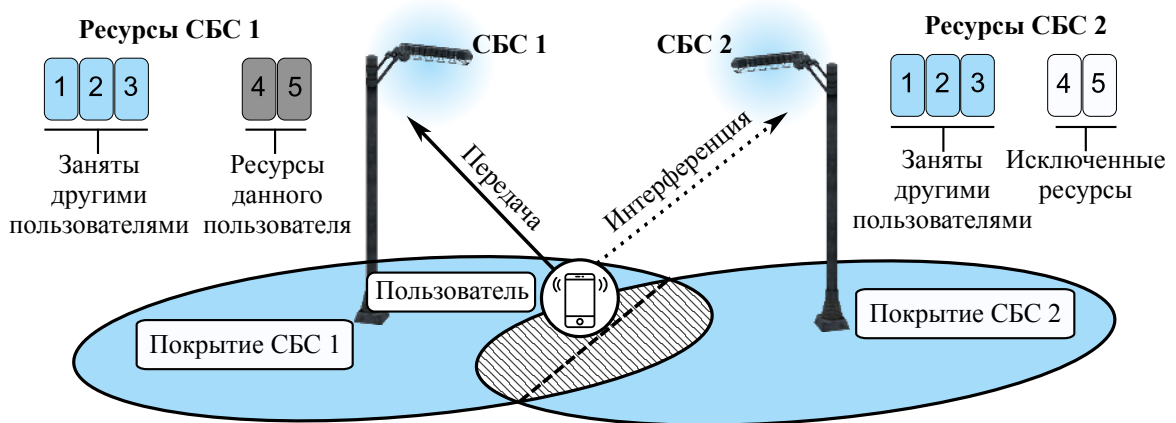


Рисунок 2.11 — Управление интерференцией в сверхплотных сетях LTE

Как отмечалось выше, для координирования межсотовой интерференции в сверхплотной сети доступа СБС LTE может, вообще говоря, использоваться достаточно продвинутое дисциплины планирования [133]. Однако по мере роста зоны покрытия системы связи, использование простых механизмов управления интерференцией может стать предпочтительным за счет их более широкой масштабируемости. В простейшем случае, если абонентский терминал расположен

на пересечении зон покрытия двух СБС и передает свои данные на ближайшую из них, используя набор радиоресурсов X , соседняя СБС исключает то же самое множество радиоресурсов X из своего процесса планирования (рисунок 2.11). Тогда абонентские терминалы из соседних сот не производят достаточно разрушительной интерференции на передачи друг друга.

Если при этом какие-либо радиоресурсы из множества X уже были предварительно исключены в процессе координирования интерференции с другой СБС, никаких дополнительных действий не требуется. Доля исключенных таким образом радиоресурсов составляет $1 - \nu$, где ν представляет собой *коэффициент переиспользования радиоресурсов*, зависящий от параметров системы связи, таких как плотность размещения СБС и радиус зоны их покрытия. Рассмотренный здесь подход к задаче управления интерференцией может быть не самым эффективным, но способен обеспечить выполнение требований по ее целевому уровню, сохраняя при этом приемлемую сложность реализации, что привлекательно для применения на практике.

В некотором смысле аналогичный подход применяется в сети доступа WiFi с тем лишь отличием, что в процессе управления интерференцией исключаются области *фиксированного размера* вне зависимости от требуемого уровня мощности передачи. Это приводит к возникновению в каждый конкретный момент времени случайного набора соединений, которые не интерферируют между собой. В итоге система WiFi может демонстрировать существенно более скромные преимущества от проведения процедуры уплотнения сети по сравнению с сотовой системой связи.

Алгоритмы выбора сети доступа. В заключение введем два альтернативных алгоритма управления функционированием совмещенной гетерогенной сети, связанные с использованием технологий радиодоступа LTE и/или WiFi. Первый *базовый* алгоритм предполагает, что система связи вначале стремится обслужить вновь поступившую сессию с помощью нелицензированных спектральных радиоресурсов сети WiFi, и только если значение SNR в этом случае оказывается ниже некоторого заданного порога η_0 , то эта сессия обслуживается сетью LTE. Такой способ подключения в некотором роде отражает поведение человека при выборе сети доступа на практике, где предпочтительное подключение по WiFi продиктовано соображениями минимизации стоимости соединения [53]. Важно отметить, что предложенная модель совмещенной гетерогенной

сети может служить для анализа и других алгоритмов выбора технологии радиодоступа.

Второй рассматриваемый алгоритм предполагает *одновременную передачу* по обеим радиотехнологиям: LTE и WiFi. Следует заметить, что одновременное использование обеих сетей доступа не обязательно приводит к достижению наибольшей возможной эффективности работы гетерогенной системы связи (например, в силу повышения уровня интерференции). Предлагаемый здесь алгоритм служит скорее для иллюстрации продвинутых возможностей, доступных в гетерогенных сетях 5G. Характеристики двух описанных выше алгоритмов сравниваются в следующем подразделе, а строгий анализ их функционирования производится в приложении Б. В нем содержатся аппроксимации стационарного распределения вероятностей для местоположения абонентских терминалов, для числа активных сессий в совмещенной сети связи, а также для значений вероятности блокировки и среднего времени передачи данных.

2.4.3 Результаты сравнения алгоритмов

При изучении функционирования сверхплотных гетерогенных систем связи важно учитывать особенности соответствующих алгоритмов выбора сети доступа и планирования радиоресурсов [134]. Поскольку такие алгоритмы часто основаны на инженерных соображениях, нежели на формальном решении некоторой оптимизационной задачи, они зачастую неаналитичны и требуют упрощений для достижения исследовательских целей. Ниже приводятся имитационные результаты сравнения работы двух таких алгоритмов, которые ведут к пониманию вопросов, связанных с управлением интерференцией и ее влиянием на эффективность пространственного переиспользования радиоресурсов.

Исследование более сложных процессов, характерных для сверхплотных гетерогенных сетей доступа, обуславливает создание соответствующего комплекса программ имитационного моделирования [135], который был разработан и апробирован в работах [53; 88]. Он позволяет моделировать работу совмещенной системы LTE-WiFi с требуемой степенью детализации и воссоздает сценарий, в котором абонентские терминалы размещены в каждом квадрате 5×5 м внутри рассматриваемой области моделирования. Как уже отмечалось, термини-

налы обслуживаются алгоритмом «Кольцевое планирование» с применением продвинутых методов управления интерференцией. В целях упрощения СБС моделируются как идеализированные получатели сигнала, функционирующие на границе Шеннона, а мощность передачи абонентских терминалов устанавливается так, чтобы поддерживать некоторый целевой уровень SINR. Сами СБС размещены равномерно внутри квадратной области моделирования, и граничные эффекты итоговой зоны покрытия не рассматриваются.

Пример фактического назначения радиоресурсов для абонентских терминалов и СБС, размещенных равномерно, показан на рисунках Б.5 и Б.6 приложения Б. Как подтверждается множеством литературных источников, поделенная таким образом поверхность действительно напоминает диаграмму Вороного практически всюду, кроме областей, в которых отсутствует сетевое покрытие. Важно также отметить, что даже в тех областях, где имеются значительные наложения зон покрытия СБС, удается достичь приемлемого качества обслуживания. Далее исследуется эффективность работы системы, выраженная в терминах пространственного переиспользования радиоресурсов и связанная с характеристиками размещения сети связи. Для рассматриваемого здесь сценария предлагается определенный характер размещения узлов сети поставить в соответствие некоторому коэффициенту ν . Последний отражает степень переиспользования единицы радиоресурсов, которая была исключена из процесса планирования для предотвращения интерференции от другой соты.

Для анализа геометрического размещения узлов на плоскости здесь вводится новая характеристика $D_s = \frac{4 \cdot N_{cells} \cdot S_{cell}}{S_{area}}$, именуемая *фактической плотностью сети*. Она определяет число различных покрытых областей, граничащих с некоторой «усредненной» зоной покрытия соты включая ее саму. К примеру, в традиционном сотовом покрытии $D_s = 7$ представляет собой фактическую плотность сети. Следовательно, соты могут быть условно представлены как шестиугольники, и каждая сота имеет в точности шесть соседних. По аналогии определенное размещение сот можно назвать (сверх)плотным, если $D_s > 7$, так как при этом фактическая плотность сети становится выше, чем в случае традиционного сотового покрытия.

В ходе анализа, описанного в приложении Б, было установлено, что фактическая плотность сети тесно связана с коэффициентом переиспользования радиоресурсов. Соответствующая зависимость показана на рисунке 2.12. Для исследуемого сценария адекватная аппроксимация была получена в виде

функции $\nu = a \cdot x^b + c$ при рассмотрении реальных сверхплотных покрытий ($7 < D_s < 100$). Даже если размещение сот не является сверхплотным, введенный параметр D_s также может быть полезен, поскольку он по-прежнему характеризует коэффициент переиспользования радиоресурсов.

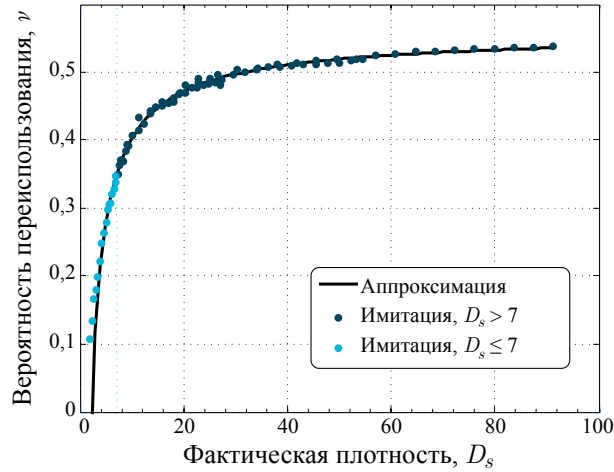


Рисунок 2.12 — Зависимость коэффициента переиспользования радиоресурсов от фактической плотности сети

Далее исследуется эффективность работы сети WiFi, которая значительно отличается от показателей функционирования сотовой системы связи из-за различий в способе управления интерференцией. В отличие от сети LTE, которая координирует свою работу на уровне отдельных СБС, технология WiFi не вводит явного различия между ТД и обслуживаемыми ими терминалами в процессе организации доступа к каналу. В результате все соединения в равной степени состязаются за возможность передачи данных, и те из них, которые испытывают самую высокую конкуренцию, могут быть временно заблокированы процедурой отсрочки передачи (являющейся частью DCF) и/или функцией прослушивания несущей частоты (ССА). Следовательно, для плотной сети WiFi в типовом случае можно наблюдать весьма ограниченное число одновременно обслуживаемых соединений вне зависимости от общего числа активных сессий в системе.

Для изучения функционирования радиотехнологии WiFi также был использован комплекс программ имитационного моделирования [135], который для заданных плотностей размещения СБС воспроизводил описанные выше реальные наборы не интерферирующих между собой соединений. Такие наборы отражают возможные комбинации абонентских терминалов, которые могут

быть активны одновременно, и генерируются как серия «снимков» поведения практической системы WiFi. Пример подобного «снимка» предложен на рисунке 2.13. Для дальнейшего анализа наибольший интерес представляет максимально допустимое число соединений, которое может быть активировано в некоторый момент времени внутри рассматриваемой области моделирования.

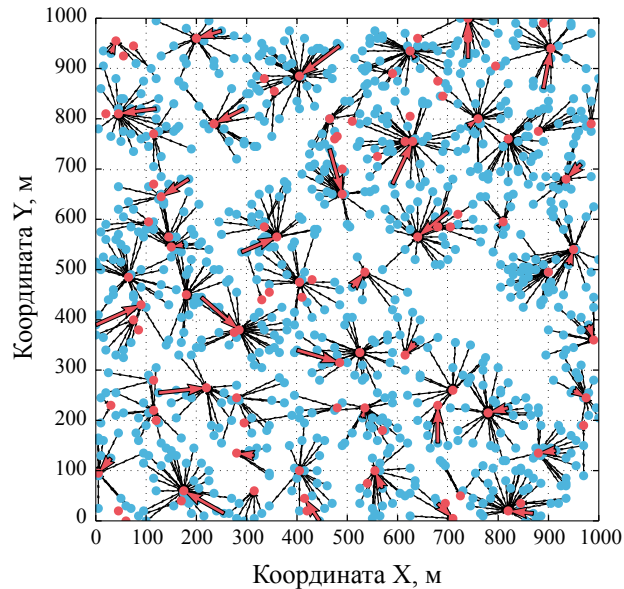


Рисунок 2.13 — Пример активных соединений WiFi

Полученные результаты подтверждают, что в случае, если число активных терминалов существенно выше, чем максимальное число поддерживаемых соединений, то общее число пользователей не оказывает заметного влияния на динамику загрузки сети доступа. Напротив, увеличение числа обслуживающих СБС заметно изменяет эту динамику, поскольку соединения становятся в среднем короче и в меньшей степени влияют на соседние соединения. Несмотря на то, что передача данных по технологии WiFi происходит на предельном значении излучаемой мощности, более короткие соединения здесь также оказываются предпочтительными, как изображено на рисунке 2.14.

Опираясь на полученные результаты, далее можно сравнить эффективность работы двух алгоритмов управления совмещенной гетерогенной сетью LTE-WiFi согласно описанию выше: предпочтительный выбор технологии WiFi и одновременное использование технологий LTE и WiFi. Параметры моделирования для рассматриваемого сценария, а также для приведенных ранее имитационных результатов являются типовыми в перспективных сверхплотных си-

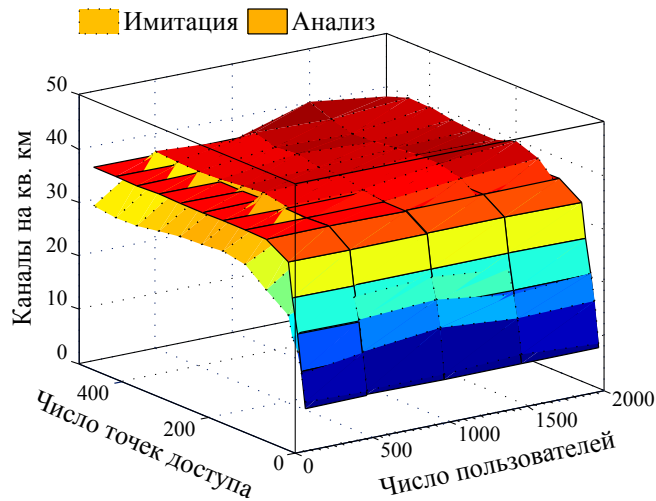


Рисунок 2.14 — Учет переменной плотности размещения

стемах связи с интегрированными малыми сотами и соответствуют актуальным рекомендациям 3GPP (см. таблицу 1).

На рисунках 2.15 и 2.16 изучается влияние интенсивности поступления сессий на эффективность работы совмещенной гетерогенной сети. Маркерами отмечены результаты имитационного моделирования, которые выборочно подтверждают полученные аналитические зависимости, представленные в виде линий. Предлагается набор кривых, соответствующих как работе базового алгоритма (выбор одной технологии), так и функционированию предлагаемого алгоритма (использование обеих технологий) для различных значений плотности покрытия. Можно заключить, что система связи демонстрирует существенное изменение своего поведения при возникновении перегрузки. В частности, если интенсивность поступления сессий достаточно велика, то среднее время передачи данных заметно возрастает. При этом интересное наблюдение можно сделать относительно точки, соответствующей моменту возникновения перегрузки, которая часто находится в обратной зависимости от плотности размещения узлов сети. В сущности, предоставление большего числа СБС, чем требуется для обслуживания текущего числа абонентских терминалов, может оказать отрицательный эффект на итоговую емкость системы.

Для иллюстрации этой важной особенности полезно рассмотреть предельный случай, когда у каждого терминала имеется «своя» персональная СБС, обслуживающая исключительно данные его сессии. С учетом фиксированной зоны покрытия СБС большинство радиоресурсов будет тогда исключено из процесса планирования согласно принципам управления интерференцией. Тем са-

Таблица 1 — Параметры моделирования сверхплотной совмещенной сети

Велич.	Описание/Определение	WiFi	LTE
D	Размер области моделирования	2000 м	
R	Радиус зоны обслуживания	100 м	100 м
μ	Плотность размещения СБС	100 на км ²	
ε	Доля терминалов не в кластерах	0,7	
r	Радиус кластера	10 м	
λ	Интенсивность поступления новых сессий	Переменная	
θ	Средний размер передаваемого файла	1,5 Мбит	
w	Полоса пропускания канала связи	20	10
N_0	Уровень шума	-106 дБм	-106 дБм
r_{lim}	Предельная скорость передачи данных	160 Мбит/с	80 Мбит/с
G	Константа распространения радиосигнала	300	250
κ	Экспонента распространения радиосигнала	6	6
σ	Доступные на СБС радиоресурсы	0,7	1
η	Целевое значение SNR	20 дБ	15 дБ
p_{max}	Максимальная мощность передачи	23 дБм	20 дБм
ν	Коэффициент переиспользования радиоресурсов	0,4	0,4
η_0	Порог SNR для переключения	20 дБм	Нет

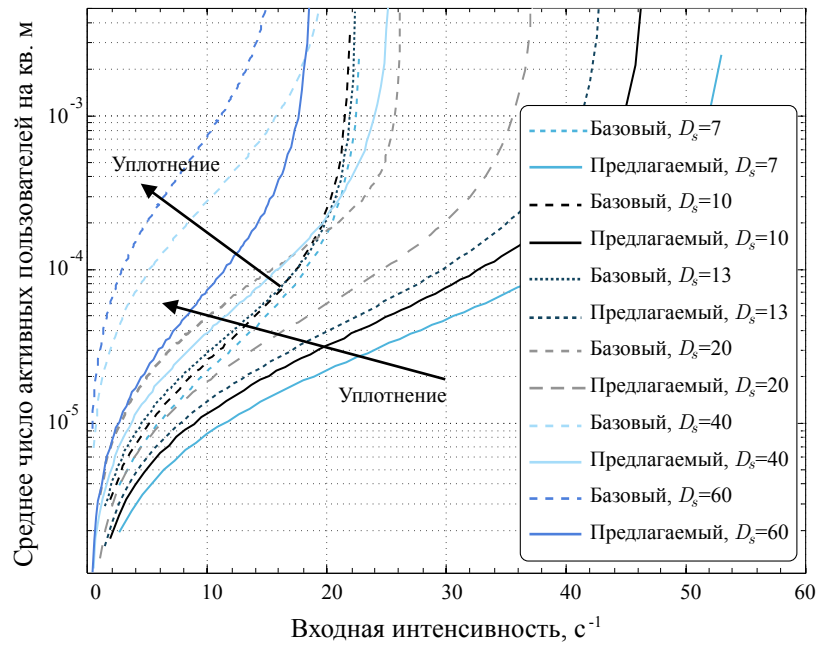


Рисунок 2.15 — Среднее число абонентских терминалов на кв. м

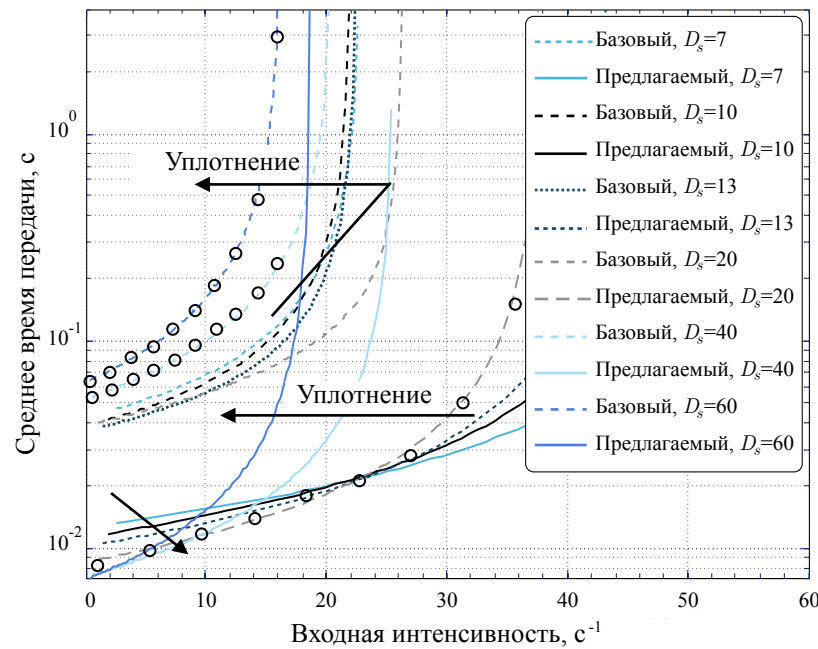


Рисунок 2.16 — Среднее время передачи данных терминалом

мым снижается доля радиоресурсов, которая может быть использована ассоциированным с данной СБС терминалом. С другой стороны, уровень SINR для передачи данных от этого терминала будет чаще всего исключительно высок. Из практических соображений полезное значение SINR фактически ограничено значением приблизительно 25 дБ, что связано с ограниченным набором модуля-

ционно-кодовых схем. Можно заключить, что необходимо жестко контролировать уровень излучаемой мощности при уплотнении сети, иначе в сверхплотных покрытиях можно наблюдать понижение доступной емкости системы связи.

Рисунок 2.17 также отражает обсуждаемый выше эффект, но под другим углом. Легко видеть, что для достаточно высоких плотностей размещения абонентских терминалов продвинутый алгоритм, позволяющий вести одновременную передачу, при фактической плотности сети, равной семи, достигает своей наивысшей эффективности. Стоит напомнить, что такое значение фактической плотности соответствует «почти» регулярному сотовому покрытию. С другой стороны, когда система связи простаивает значительную часть своего времени (например, при наличии всплесков загрузки), использование большого числа малых сот становится предпочтительным. В таких условиях соединения становятся в среднем короче, что приводит к повышению средней скорости передачи данных при сравнительно низкой загрузке. В отличие от продвинутого алгоритма, базовый алгоритм, связанный с выбором предпочтительной радиотехнологии, получает существенно меньший прирост производительности при уплотнении сети: в определенный момент времени большинство терминалов вынуждено использовать технологию WiFi в силу специфики работы алгоритма. Делая вывод, можно отметить, что если абонентский терминал имеет возможность использовать несколько технологий радиодоступа одновременно, то он не должен быть ограничен выбором только лишь одной из них независимо от своего текущего местоположения по отношению к СБС.

2.5 Управление доступом и радиоресурсами в интегрированных сетях

2.5.1 Различные классы алгоритмов управления трафиком

Данный раздел посвящен изучению различных подходов к управлению процессом выделения радиоресурсов в интегрированной гетерогенной сети доступа [53]. С учетом рассмотренной выше архитектуры системы связи здесь исследуются алгоритмы, функционирующие как на стороне абонентских терми-

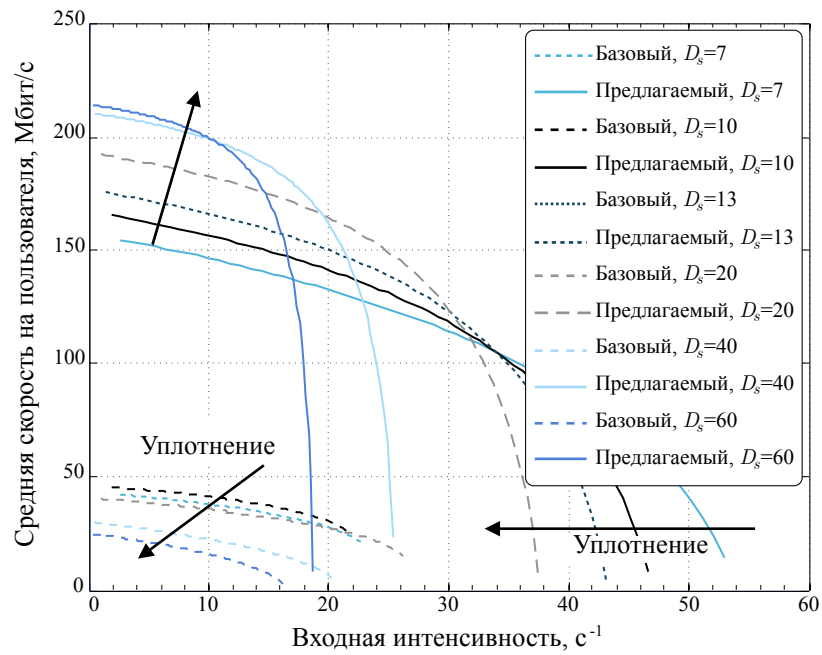


Рисунок 2.17 — Средняя скорость передачи данных терминалом

налов, так и на стороне сетевой инфраструктуры. В случаях интеграции технологий радиодоступа на прикладном уровне или на уровне ядра сети (*Вариант А или В* на рисунке 2.2) могут использоваться только алгоритмы, работающие на стороне абонентских терминалов. Напротив, при интеграции радиотехнологий на уровне сети радиодоступа (*Вариант В* на рисунке 2.2) как правило доступен более широкий набор возможностей, связанных с повышенной оперативностью координирования технологий радиодоступа, составляющих гетерогенную сеть.

Вообще говоря, сеть радиодоступа может выполнять существенную роль при назначении радиоресурсов в рамках интегрированной системы связи. Даже когда она не управляет выделением радиоресурсов напрямую, сеть радиодоступа может предоставлять важную вспомогательную информацию и тем самым повышать эффективность решений, принятых на стороне абонентских терминалов. В системах связи 5G, где контроль над мобильностью и поведением абонентского терминала переходит из опорной сети в сеть радиодоступа, становится доступным более оперативное управление выделением радиоресурсов, обусловленное возможностью быстрого переключения сессий между радиотехнологиями. В совмещенных гетерогенных сетях, а также когда имеется высокоскоростной интерфейс между разнородными технологиями радиодоступа, возможно совместное назначение радиоресурсов, принадлежащих сразу нескольким сетям доступа.

В дальнейшем рассматриваются примеры типовых алгоритмов управления выделением радиоресурсов, начиная с простых традиционных схем, работающих на стороне абонентских терминалов, когда выбирается предпочтительная технология радиодоступа (например, WiFi), и вплоть до сложных интеллектуальных подходов на стороне сетевой инфраструктуры, когда производится совместное координирование радиоресурсов различных сетей доступа.

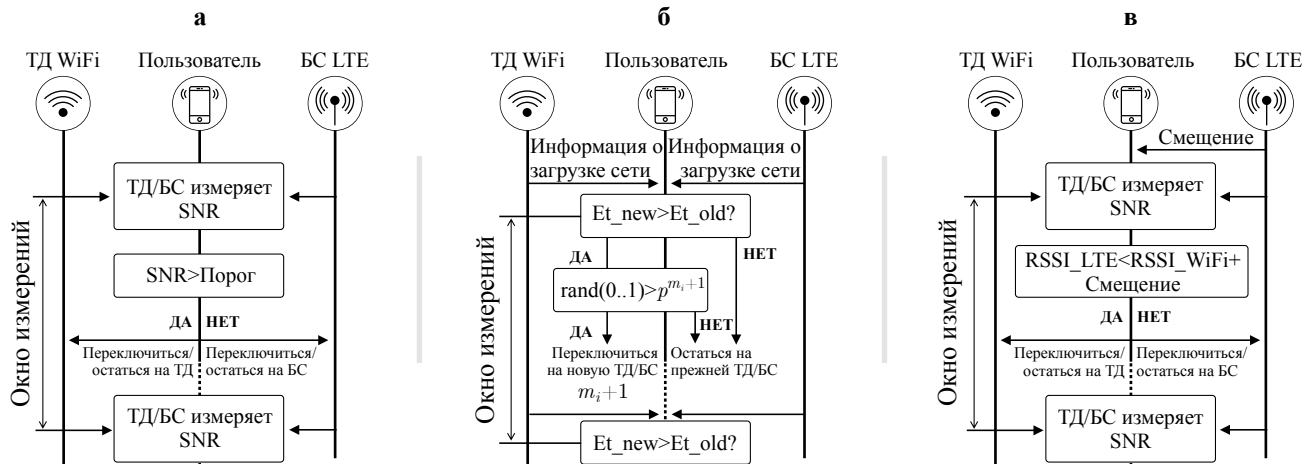


Рисунок 2.18 — Алгоритмы управления трафиком в интегрированной гетерогенной сети

Алгоритмы на абонентских терминалах. Простейшие алгоритмы, основанные на выборе некоторого порогового значения, составляют основу данного класса и служат в качестве базовых для последующего сравнения. Данные схемы предполагают, что абонентский терминал непрерывно отслеживает управляющие сообщения от соседних ТД WiFi для получения актуальной информации относительно уровня SNR в сети. Когда текущее значение SNR превышает заданный порог (который устанавливается здесь равным 40 дБ согласно рекомендациям 3GPP), терминал начинает передавать свои данные на соответствующую ТД WiFi. В противном случае он передает весь имеющийся трафик через сеть LTE (см. рисунок 2.18(а)).

Такие правила алгоритма автоматизируют типовое поведение абонента при выборе предпочтительной радиотехнологии: если доступна ТД WiFi с достаточным уровнем принимаемого сигнала, абонентский терминал переключается на нее и, как правило, получает более высокую скорость передачи данных, снижая при этом расходы абонента, связанные с оплатой сотового трафика. В качестве альтернатив такому функционированию можно предложить подхо-

ды, основывающиеся на выборе технологии WiFi, если удовлетворены некоторые минимальные требования по эффективности работы, качеству соединения, наличию покрытия и т.д. Кроме того, в рассматриваемую категорию можно включить схемы, позволяющие абонентскому терминалу передавать свои данные с использованием нескольких технологий радиодоступа одновременно, но не подразумевающие интеллектуального координирования такой передачи.

Алгоритмы с сетевым содействием. В силу своей простоты, базовый алгоритм с предпочтительным соединением по технологии WiFi, учитывающий пороговое значения SNR, может испытывать существенные ограничения эффективности в плотных покрытиях, подверженных значительному влиянию интерференции, что является распространенной ситуацией в современных городских условиях. К примеру, некоторая ТД WiFi может перейти в состояние перегрузки, когда большое число абонентских терминалов передает через нее свой трафик. Более того, подвижные терминалы, использующие сеть WiFi (например, портативные компьютеры), могут неожиданно потребить значительную долю радиоресурсов локальной сети. Наконец, следует учитывать, что протокол доступа к среде в системе WiFi основан на применении случайного множественного доступа, и это приводит к существенному снижению фактической скорости передачи данных с ростом числа активных пользователей.

Как следствие, использование простейшей схемы, основанной на пороговом значении SNR и не учитывающей текущую загрузку сети, бывает недостаточно эффективным в условиях высокой динамики системы связи. В таком случае абонентские терминалы могут объединить значение актуальной величины SNR с дополнительной информацией относительно загрузки системы, полученной со стороны сетевой инфраструктуры. Учет загрузки сети WiFi может, таким образом, служить важным шагом к улучшению эффективности работы по сравнению с производительностью базового алгоритма, но и его может оказаться недостаточно в ряде практических сценариев. Тогда имеет смысл дополнительно принять во внимание загрузку сети LTE. Алгоритм, реализующий данный подход (см. рисунок 2.18(б)) и учитывающий приведенные ниже соображения, сравнивается далее с базовым алгоритмом, основанным на пороговом значении SNR.

Оценка скорости передачи. Абонентский терминал прослушивает канал связи по обеим технологиям радиодоступа, наблюдая актуальное значение SNR в соседних малых сотах, чтобы оценить свою ожидаемую скорость передачи

данных. Для системы WiFi такая оценка производится на основе предсказания емкости сети и деления ее на число пользователей, обслуживаемых некоторой ТД (такая информация содержится в сигнальных сообщениях, рассылаемых ТД периодически). Кроме того, могут учитываться дополнительные сведения о значении SNR, уровне конкуренции в канале и т.д. В итоге те ТД, которые имеют слишком слабый сигнал (согласно текущему значению SNR), исключаются из рассмотрения. Далее поведение данной прогнозирующей схемы может уточняться с использованием поправок, учитывающих работу протокола случайного множественного доступа в сети WiFi, а также коэффициентов, включающих в себя накладные расходы на передачу сигнальных сообщений. Для системы LTE оценка скорости передачи осуществляется по указанию планировщика радиоресурсов, расположенного на ассоциированной с абонентским терминалом БС, с учетом выбранного способа управления излучаемой мощностью.

Случайный выбор сети. Абонентский терминал имеет возможность выбрать сеть доступа, в которой ожидаемая им скорость передачи данных будет наивысшей. Однако в ряде случаев имеет смысл осуществлять такой выбор с некоторой вероятностью $\text{rand}(0..1) < p^{m_i+1}$, где m_i – число недавних попыток соединения с данной ТД/БС, а p – параметр в диапазоне (0; 1), который представляет собой вероятность повторного соединения. Надлежащее использование параметра p снижает число одновременных соединений с той же самой ТД/БС, что предотвращает неконтролируемое переключение (осцилляцию) между технологиями радиодоступа. Соответственно, если происходит повторное подключение, значение m_i для целевой ТД/БС i увеличивается на единицу. Другие абонентские терминалы также принимают во внимание эту информацию, деля свое ожидаемое значение скорости передачи для данной ТД/БС на $m_i + 1$. Такой подход позволяет управлять динамикой переключений в интегрируемой гетерогенной сети доступа.

Гистерезисное управление. Для дальнейшего снижения числа переключений на границе соты можно применять соответствующие механизмы гистерезисного управления, которые также влияют на оценочное значение скорости передачи данных.

Фильтрация оценок скорости передачи. Дополнительного уточнения оценок скорости передачи данных можно добиться при помощи усреднения. После каждого «окна» измерений фактические значения скорости передачи, наблюдаемые в течение данного периода, могут быть преобразованы, например, с

использованием функции скользящего среднего. Результирующее значение учитывает измеренное и ожидаемое значения скорости передачи данных для уточнения оценки этого параметра на целевой ТД/БС. Подобное усреднение также позволяет повысить надежность проводимого предсказания, предотвращая внезапные флуктуации, связанные с процедурой случайного множественного доступа.

В итоге алгоритмы управления с сетевым содействием имеют возможность использовать дополнительную информацию, получаемую со стороны сетевой инфраструктуры, для повышения эффективности решений, принятых на стороне абонентских терминалов. В принципе сетевое содействие может ограничиваться передачей абонентскому терминалу базовых сведений, таких как загрузка системы, коэффициент ее использования, ожидаемая доля доступных радиоресурсов и т.д. Однако при повышении производительности интерфейса между различными радиотехнологиями (как, например, в совмещенных покрытиях) такое содействие может стать существенно более оперативным и эффективным.

Алгоритмы на стороне сети доступа. Два вышеприведенных способа управления гетерогенной сетью доступа функционируют на стороне абонентских терминалов. Как следствие, они принципиально ограничивают эффективность работы системы связи, которая может быть повышена за счет общесистемной информации, доступной на стороне сетевой инфраструктуры. Применяя централизованный алгоритм управления, непосредственно координирующий работу интегрированной системы связи, можно напрямую назначать абонентским терминалам радиоресурсы каждой из используемых сетей доступа. Такая схема координирования может быть как распределена между малыми сотами, так и реализовываться специализированным модулем, предназначенным для принятия соответствующих решений на стороне сетевой инфраструктуры.

В данном разделе рассматривается традиционный подход, связанный с увеличением эффективной зоны покрытия соты. Он используется в современных сотовых сетях для переключения абонентских терминалов на малые соты путем модификации индикатора принятой мощности сигнала RSSI. На практике значение показателя RSSI можно увеличивать или уменьшать по необходимости, приводя к соответствующему изменению эффективной зоны покрытия ТД WiFi в зависимости от прогнозируемой емкости сети. Ограничением данного метода является тот факт, что добавочная к индикатору RSSI величина долж-

на учитывать общесистемную информацию относительно обслуживаемых абонентских терминалов. Например, можно показать, что целевое значение этой добавочной величины зависит от способа размещения узлов сети, а также от уровня интерференции в ней. Такая информация не всегда доступна на стандартных ТД WiFi, особенно если они не имеют прямого интерфейса с сотовой системой связи. Ниже исследуется алгоритм управления, увеличивающий эффективную зону покрытия малых сот с помощью модификации значения RSSI в рамках некоторого целевого сценария (см. рисунок 2.18(в)). При этом также учитываются необходимые гистерезисные механизмы.

Вообще говоря, реализуемые на стороне сети доступа алгоритмы могут использовать стандартизованные или специализированные интерфейсы между различными сотами и технологиями радиодоступа. Распределенные подходы к управлению со стороны сети доступа были рассмотрены комитетом 3GPP в рамках работы по организации координирования систем WLAN и LTE на уровне сети доступа. Соответственно, в систему связи внедряются механизмы, позволяющие абонентским терминалам сообщать результаты своих локальных измерений состояния радиоканала. Результирующие решения относительно выбора той или иной технологии радиодоступа принимаются затем сотовой БС на основе полученных таким образом сведений. Другие примеры централизованных алгоритмов на стороне сети доступа включают в себя так называемые схемы *двойного соединения*, когда абонентский терминал непрерывно поддерживает управляющее соединение с макрослоем, а также соединение для передачи данных с некоторой малой сотой. При этом система связи организует централизованное управление переключением терминала между различными малыми сотами.

2.5.2 Сравнительный анализ алгоритмов управления трафиком

Далее исследуются различные алгоритмы управления, которые задают способ использования технологий радиодоступа, составляющих интегрированную сеть LTE-WiFi [14]. Предполагается, что система WLAN является составной частью гетерогенной сети доступа, размещенной мобильным оператором. Изучаются типовые практические алгоритмы управления, функционирующие

на стороне абонентских терминалов. С учетом особенностей современных сетевых покрытий рассматриваются слои ТД WiFi и сотовый слой, но не предполагается наличие прямого интерфейса взаимодействия между сетями LTE и WiFi [92]. Кроме того, отдельно обсуждаются преимущества использования совмещенных малых сот LTE-WiFi и приводятся соответствующие показатели эффективности.

В частности, производится сравнительный анализ распределенного алгоритма управления выбором предпочтительной сети доступа, который учитывает информацию о загрузке систем связи LTE и WiFi, и алгоритма, который при таком выборе располагает только измерениями уровня мощности сигнала. Показатели эффективности данных алгоритмов, работающих на стороне абонентских терминалов, сравниваются далее с производительностью алгоритма, функционирующего на стороне сети доступа и непосредственно балансирующего загрузку радиотехнологий. Ожидается, что последний за счет доступности общесистемной информации о состоянии каналов связи продемонстрирует более высокую эффективность по сравнению с алгоритмами, реализуемыми на абонентских терминалах, за счет доступности общесистемной информации о состоянии каналов связи. Однако распределенные схемы, оперирующие без прямого взаимодействия технологий радиодоступа между собой, являются гораздо более простыми в плане реализации и не требуют высоких накладных расходов на передачу сигнальных сообщений. Также такие алгоритмы могут иметь преимущество при учете предпочтений абонентов и требований по качеству обслуживания трафика.

Предлагаемое ниже сравнение построено на основе результатов имитационного моделирования, полученных с применением соответствующего комплекса программ, который воспроизводит работу интегрированной системы LTE-WiFi, а также был откалиброван со спецификациями 3GPP и документами IEEE 802.11. В настоящее время известно незначительное число бесплатных или коммерческих программных продуктов, способных моделировать динамическую гетерогенную сеть доступа. Это связано, прежде всего, с отсутствием необходимых сетевых компонентов, а также с недостаточной масштабируемостью при исследовании сложных зависимостей между радиотехнологиями. В отличие от известных решений используемый здесь комплекс программ представляет собой гибкий инструмент, поддерживающий набор различных способов размещения узлов сети, типов трафика, моделей канала связи и технологий

радиодоступа. Он включает в себя ряд уникальных программных модулей, моделирующих сетевое покрытие и расположение абонентских терминалов, управляющие пакеты для передачи данных и сигнальных сообщений между узлами сети, происходящие в канале связи события, а также механизмы сбора статистической информации для оценки всевозможных показателей качества обслуживания.

Для проведения предлагаемого здесь исследования воссоздается типовой сценарий функционирования интегрированной гетерогенной сети в городских условиях, когда ТД WiFi размещены в той же географической области, что и сотовая сеть LTE [136]. В частности, рассматриваются наружные покрытия малыми сотами согласно рекомендациям 3GPP в [137]. Основные параметры проводимого моделирования собраны в таблице 2. Следует отметить, что предполагается насыщенный абонентский трафик, а ТД размещены при этом равномерно внутри зоны покрытия сотовой сети. Большинство пользователей сгруппированы вокруг ассоциированных с ними ТД в кластеры, что соответствует характеру обслуживания абонентов, генерирующих объемный трафик в аэропортах, кафе, ресторанах, торговых центрах и т.д. Кроме того, около трети терминалов также разместились равномерно в рамках рассматриваемой области моделирования, представляя собой обычных мобильных абонентов. Такой сценарий является типовым и был одобрен комитетом 3GPP для изучения показателей эффективности интегрированных гетерогенных сетей.

Важное преимущество, ожидаемое от использования ТД WiFi в рамках гетерогенной сети, связано с эффективной выгрузкой абонентского трафика из сотовой системы связи. По этой причине основным изучаемым здесь показателем эффективности становится скорость передачи данных абонентским терминалом *по линии вверх*. Такой подход отличается от множества уже имеющихся исследовательских результатов относительно передачи данных по линии вниз и позволяет охарактеризовать доступную емкость системы. Функция распределения CDF значений индивидуальной скорости передачи данных служит для сравнения функционирования алгоритма, основанного на выборе порогового значения SNR (предпочтительное подключение по WiFi), с алгоритмом, использующим информацию о загрузке сети (см. рисунок 2.19). Результаты сравнения подтверждают, что последний предлагает существенные преимущества, особенно на краю зоны покрытия (например, более 75% улучшения на уровне 5% квантили). Кроме того, наблюдается некоторое повышение средней скорости

Таблица 2 — Параметры моделирования алгоритмов управления трафиком

Параметр	Значение
Конфигурация LTE/WiFi	10 МГц (частотное разделение) / 20 МГц
Макрослой	7 сот, 3 сектора в каждой
Режим сигнализации LTE	2 из 20 специальных подкадров, короткий префикс, кадр 10 мс
Расстояние между сотами	500 м
Конфигурация антенн макро-	1x2 (разнесенный прием)
Затухание в канале терм.-сеть	Рекомендации ITU
Коэффициент усиления БС	14 дБ
Макс. излучаемая мощность БС/ТД/терм.	43/20/(23/20 LTE/WiFi) дБм
Высота антенны терминала/БС/ТД	1,5/25/10 м
Модель трафика	Насыщенная
Число терминалов/ТД	30/4 на макросоту (3 сектора)
Способ размещения ТД/терминалов	Равномерное/Кластерное (случай 4b в [137])
Расстояние ТД/терм.-БС; ТД/терм.-терм.	> 75/35 м; 40/10 м
Размер фрагмента WiFi	1500 байт
Время моделирования	3 с
Количество запусков на точку	30

передачи данных в совмещенных покрытиях (т.е. построенных на совмещенных LTE-WiFi модулях).

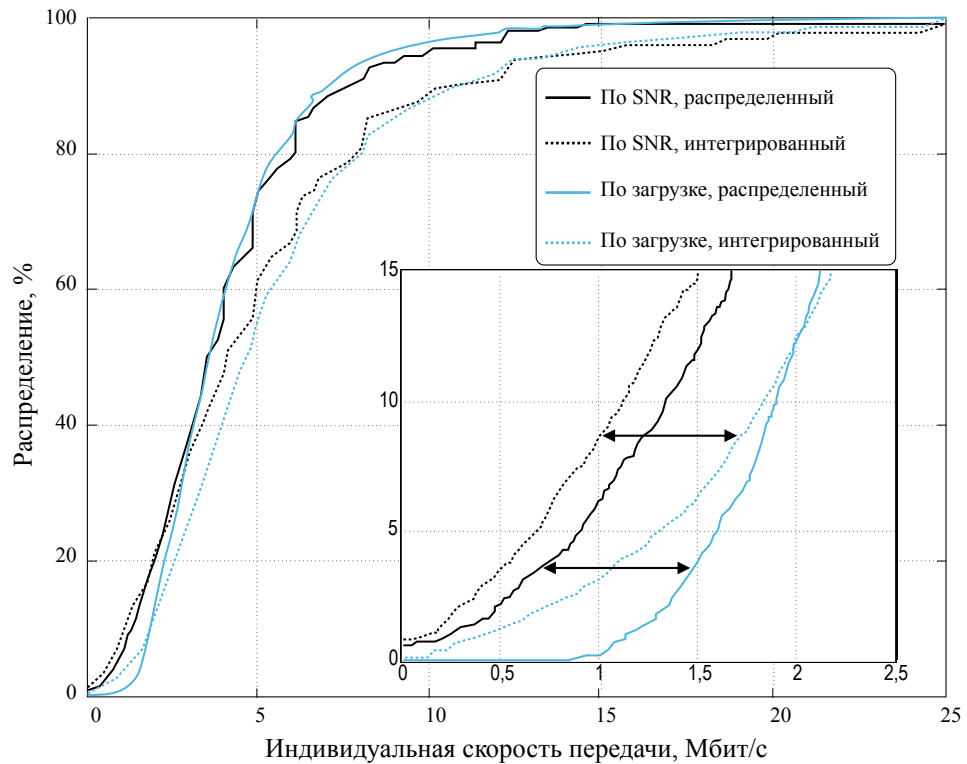


Рисунок 2.19 — Скорость передачи для различных алгоритмов управления трафиком

Кроме того, энергетическая эффективность становится важной характеристикой производительности гетерогенных сетей 5G в силу ограниченных ресурсов аккумуляторной батареи абонентских терминалов [138]. Соответственно, наблюдается значительный прирост в показателе бит/Дж как при распределенном (19%), так и при совмещенном (29%) размещении узлов инфраструктуры, которые исследуются на рисунке 2.20. Поскольку уровень качества обслуживания также представляет немалый интерес, здесь учитывается справедливость распределения радиоресурсов между абонентскими терминалами на основе показателей индивидуальной скорости передачи данных и ее усредненного значения в сети. Для оценки такой справедливости применяется так называемый индекс Джейна (см. таблицу на рисунке 2.20), который для алгоритма с учетом загрузки сети оказывается лучше (0,72/0,63), чем данный показатель для алгоритма с порогом по SNR (0,65/0,54).

Стабильность работы алгоритмов на абонентских терминалах является ключевым практическим показателем в распределенных схемах управления,



Рисунок 2.20 — Сравнительный анализ алгоритмов управления трафиком

поскольку нежелательные переключения между различными технологиями радиодоступа могут вызвать дополнительные накладные расходы, привести к росту задержки при передаче данных, а также снизить энергетическую эффективность терминала. На рисунке 2.20 дополнительно приводится число переключений между радиотехнологиями LTE и WiFi (выраженное в числе подключений в единицу времени). Для улучшения рассматриваемого показателя рекомендуется применять соответствующие механизмы гистерезисного управления.

Приведенное здесь сравнение алгоритмов управления также включает в себя рассмотрение работы схемы с модификацией RSSI на стороне сетевой инфраструктуры, подстраивающей радиус зоны покрытия под особенности конкретного размещения узлов. При этом наличие общесистемной информации может улучшить качество обслуживания в сети доступа за счет подбора подходящего значения добавки к RSSI, что имеет своей целью оптимизацию эффективной зоны покрытия БС LTE и ТД WiFi. Такая оптимизация улучшает показатели эффективности в сценарии с равномерным размещением абонентских терминалов, но в случае кластерного размещения характеристики ухудшаются из-за влияния интерференции. Для учета такого влияния централизованный алгоритм адаптирует радиус зоны покрытия, основываясь на общесистемном

знании относительно способа размещения абонентских терминалов (см. рисунок 2.21).

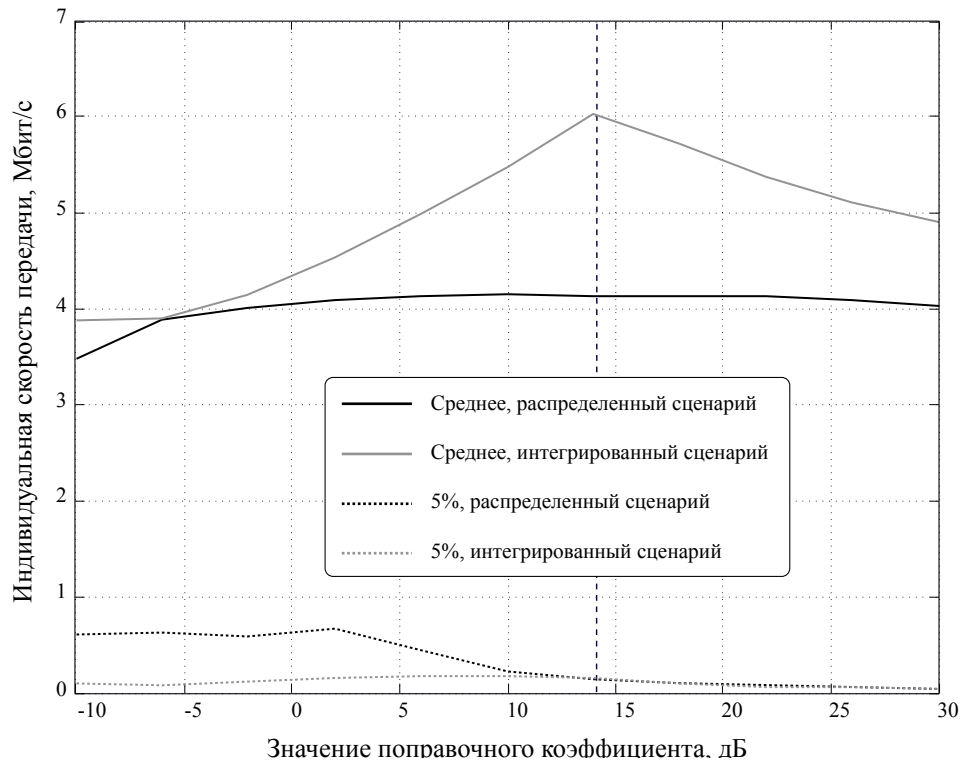


Рисунок 2.21 — Индивидуальная скорость передачи абонентским терминалом

Однако рисунок 2.22 демонстрирует, что даже дополнительная информация на стороне сетевой инфраструктуры (здесь было выбрано модифицированное значение в 14 дБ) не приводит к значительному росту скорости передачи пользовательских данных, которая мало отличается от аналогичного показателя для распределенного алгоритма, учитывающего загрузку сети. В частности, на рисунке 2.22 видно (см. гистограмму), что средняя доля времени, проводимая терминалом на каждом из радиointерфейсов распределена следующим образом: например, для алгоритма с учетом загрузки сети технология радиодоступа WiFi используется достаточно эффективно, тогда как передача по технологии LTE сбалансирована между пико- и макрослоями.

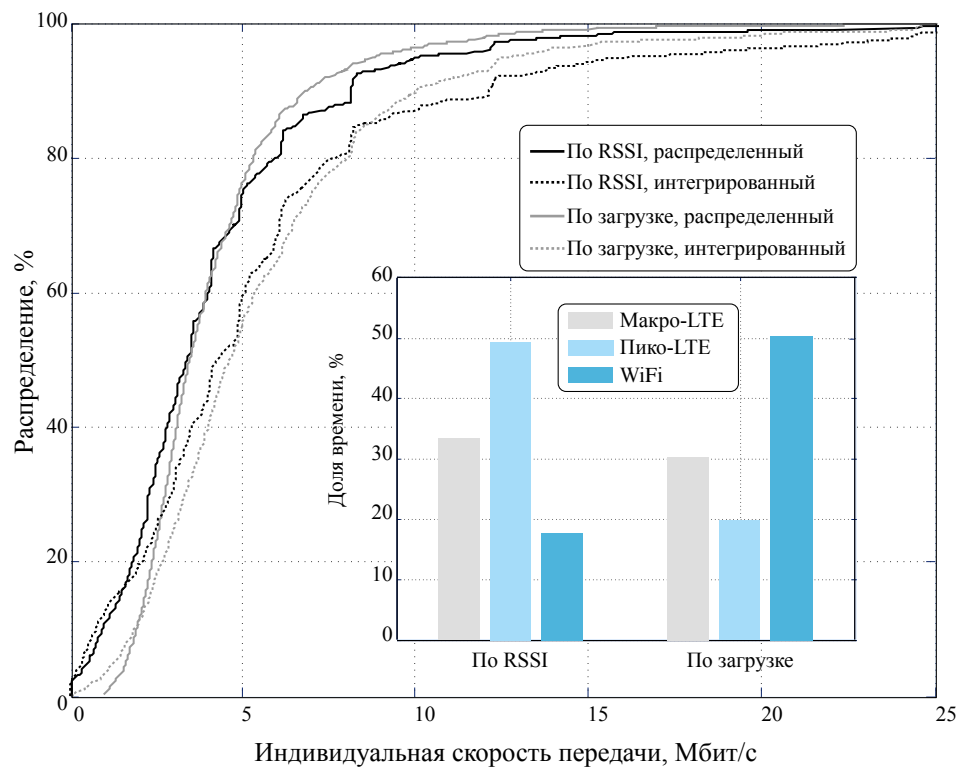


Рисунок 2.22 — Распределение индивидуальной скорости передачи данных

2.6 Выводы

В результате разработанная модель гетерогенной сети с набором интегрированных между собой технологий радиодоступа, учитывающая обслуживание потоков трафика (сессий) от множества пользователей и высокую плотность размещения узлов сетевой инфраструктуры, позволяет оценивать параметры качества обслуживания, такие как среднее число одновременно передающих пользователей и вероятность блокировки их сессий. Кроме того, предложенный алгоритм с разделением пользовательской сессии применим для управления потоками трафика в гетерогенной сети с двумя совмещенными технологиями радиодоступа, которая характеризуется высокой плотностью размещения узлов сетевой инфраструктуры, и позволяет повысить скорость отправки данных пользователями, а также снизить среднее время их передачи.

Глава 3. Использование прямых соединений между устройствами и крайне высоких частот

3.1 Предварительные замечания

Развернутые в настоящее время системы связи могут не справляться со все возрастающей загрузкой от своих абонентов. В качестве альтернативы уплотнению сетевого покрытия для повышения емкости сети, предлагается выгрузка сотового трафика на прямые соединения «устройство – устройство» D2D между абонентскими терминалами, находящимися в непосредственной близости друг к другу. Учитывая наличие поддержки нескольких технологий радиодоступа, например, LTE и WiFi, в современных устройствах связи актуально рассмотреть применение связи D2D в нелицензированном спектре. Данная глава посвящена использованию прямых соединений на нелицензированных частотах в составе системы связи 5G, обеспечивающей сетевое содействие при управлении D2D-соединениями.

Далее приводятся различные варианты организации прямых соединений и альтернативные технологии D2D, которые могут применяться на уровне сети радиодоступа. Это обуславливает возможность исследования как характеристик прямых соединений «устройство – устройство» в контексте систем связи 5G, так и способов интеграции соответствующего функционала для выгрузки абонентского трафика с существующей сотовой архитектурой 3GPP. Особое внимание уделяется методам управления выгрузкой данных с сетевым содействием, для чего формулируется соответствующая модель множественного доступа, проводится ее анализ и предлагаются результаты имитационного моделирования. Разработанный математический подход учитывает динамику абонентского трафика во времени и размещение узлов сети в пространстве, что также подтверждается при помощи имитационных результатов.

Кроме того, рассматривается перспективная возможность передачи данных на крайне высоких частотах, подверженная блокированию радиосигнала небольшими объектами из-за более коротких длин волн. Таким образом, приобретает важность изучение характеристик миллиметровых волн mmWave в реальной среде их распространения, а также анализ динамики процесса блоки-

рования передачи mmWave во времени и пространстве. Для предотвращения блокирования mmWave-соединения подвижными объектами, такими как тело человека, а также поддержания непрерывности сессии применяется подход с одновременным подключением абонентского терминала к двум и более сотам mmWave. Ряд алгоритмов управления таким одновременным подключением и исследование характеристик их работы также приводятся в рамках данной главы.

Для понимания прироста производительности от использования различных алгоритмов управления одновременным подключением в типовых условиях применения связи mmWave в этой главе проводится соответствующее моделирование. В частности, рассматриваются основные показатели эффективности функционирования как сети в целом, так и отдельного абонентского терминала. Для этого применяется метод, основанный на использовании имитатора распространения радиосигнала в реальных условиях городской застройки. Он учитывает динамическое блокирование соединений mmWave различными препятствиями, а также особенности работы альтернативных алгоритмов управления одновременным подключением. Последние служат для оперативного перенаправления абонентской сессии в другую соту mmWave при блокировании текущего соединения на крайне высоких частотах.

3.2 Целесообразность и варианты использования прямых соединений

3.2.1 Обоснование и применение связи «устройство – устройство»

Развитие систем связи 5G затрагивает различные компоненты сотовой архитектуры, в том числе и те, которые отвечают за обслуживание трафика от всевозможных мобильных устройств. По мере роста объемов передаваемых абонентских данных, сетевая инфраструктура уплотняется, что связано с уменьшением зон покрытия малых сот и увеличением их числа в областях значительной загрузки. Уплотнение сети, рассмотренное в главе 2, позволяет повысить емкость системы связи и качество пользовательского соединения за счет сокраще-

ния его длины [84]. Однако увеличение плотности сетевого покрытия влечет за собой существенный рост издержек со стороны мобильного оператора, а также усложняет управление интерференцией. Как следствие, необходимы альтернативные подходы к улучшению качества связи, свободные от этих недостатков.

Известно, что значительную долю абонентского трафика составляют данные пиринговых (P2P) приложений [17], пользователи которых находятся в непосредственной близости друг к другу. Такая структура трафика предоставляет уникальную возможность выгружать соответствующие данные из сотовой системы связи на прямые соединения «устройство – устройство», которые обычно короче, а соответствующие абонентские терминалы размещены ниже по сравнению с традиционными инфраструктурными узлами. Эффективные механизмы для обнаружения P2P-соседей, установления связи D2D и поддержания непрерывного соединения позволяют мобильным операторам применять данную форму выгрузки трафика для снижения загрузки сотовой сети доступа без привлечения дополнительных затрат или даже с получением прибыли.

С точки зрения абонентов системы связи, преимущества технологии D2D очевидны: она позволяет увеличить скорость передачи данных, снизить задержку при их доставке и повысить энергетическую эффективность соединения [5]. Такие возможности особенно привлекательны в связи с бурным ростом числа приложений, которые могут использовать географическую близость абонентов для улучшения качества передачи данных. Это обуславливает активные исследования технологии D2D как в академической, так и в индустриальной сфере с последующей быстрой стандартизацией. Связь D2D имеет множество применений, таких как локальные голосовые услуги (выгрузка звонков между близко расположенными абонентами), обмен мультимедийной информацией, игровые приложения, групповое вещание, контекстные услуги и службы общественной безопасности [18].

Однако в зависимости от характеристик движения абонентов определенные услуги могут быть более или менее пригодны для выгрузки на прямые соединения «устройство – устройство» [13]. К примеру, если терминалы D2D в значительной степени подвижны, то состояние канала связи может сильно меняться во времени, усложняя тем самым поддержание качества работы приложения. Тогда предпочтительными кандидатами для выгрузки являются устойчивые к задержке услуги, которые могут использовать буферизацию в случае разрыва D2D-соединения или же переключаться обратно на инфраструктурную связь

(например, видео по запросу или передача файлов). Напротив, если терминалы D2D малоподвижны, то множество пригодных к выгрузке услуг расширяется и включает в себя групповые потоковые приложения, а также многопользовательские игры.

3.2.2 Альтернативные радиотехнологии для соединений D2D

Ввиду широкого спектра потенциальных приложений и услуг связи, которые могут использовать соединения D2D, данная технология была предложена и исследована в [139], а также множестве последовавших за этой работ. Изначально задумывалось, что слой D2D будет функционировать как «подложка» сотовой системы, работающая в том же диапазоне радиочастот, а терминалы D2D станут подстраивать свою мощность передачи так, чтобы снизить интерференцию с пользователями сотовой сети. На сегодняшний день возможность установления соединений D2D в лицензированном спектральном диапазоне хорошо изучена теоретически, однако стандартизация этой концепции столкнулась с рядом трудностей, например, высокая сложность соответствующего управления системой связи, и это затрудняет реализацию данного подхода на практике.

В отличие от «подложки» D2D в лицензированном спектре, протоколы установления прямого соединения на нелицензированных частотах уже доступны на рынке, и они также могут быть использованы в системах D2D [88]. К сожалению, в большинстве имеющихся решений процесс обнаружения соседей слишком энергозатратен, установление прямого соединения затруднено, отсутствуют надежные средства поддержания непрерывности связи, а управление радиоресурсами зачастую неэффективно [140]. Однако все перечисленные недостатки и ограничения могут быть частично или полностью устранены применением сетевого содействия [141], которое снабдит абонентские терминалы дополнительными сведениями, доступными лицензированной сети мобильного оператора и полезными для обеспечения их эффективной работы в нелицензированном спектральном диапазоне.

Связь D2D в лицензированном спектре. Лицензированный спектр представляет собой ограниченный и дорогой ресурс, и несмотря на непрерывные попытки высвободить дополнительные частоты для нужд мобильной свя-

зи, их оказывается недостаточно для удовлетворения текущей потребности. Таким образом, сотовые операторы заинтересованы в поиске методов, позволяющих повысить емкость уже имеющихся покрытий. Технология D2D с сетевым содействием является при этом многообещающим перспективным направлением, что подтверждается быстро растущим количеством работ в данной области, которые предполагают различные уровни инфраструктурной поддержки. С одной стороны, рассматривается простейшая форма сетевого содействия для синхронизации взаимодействующих терминалов (например, в системе Aura-net [142]). С другой стороны, сотовая сеть может непосредственно управлять D2D-соединениями в виде «подложки» [139].

Управление интерференцией, включая контроль доступа в сеть и назначение мощности передачи [143], требуется в этом случае для поддержания множества соседних соединений D2D. Исследовательские работы [144–146] предлагают способы снижения интерференции, учитывающие состояние канала связи CSI, местоположение абонента и т.д. Такие подходы подразумевают, что D2D-соединения (1) используют лицензированные спектральные ресурсы совместно с традиционными сотовыми соединениями; (2) используют выделенные радиоресурсы или (3) не устанавливаются и данные передаются сотовой инфраструктурой. Процесс принятия решения в пользу какого-либо из этих вариантов получил название выбора *режима передачи* и был изучен в большом числе работ, ставящих различные оптимизационные задачи в терминах SINR [147], пропускной способности [145], энергетической эффективности [148], задержки передачи [149], справедливости планирования радиоресурсов и вероятности блокировки [150].

Вообще говоря, опубликованные результаты исследований технологии D2D различаются по числу рассматриваемых взаимодействующих узлов (базовых станций, сотовых терминалов и D2D-пользователей), по предположению о передаче данных по линии вниз или вверх, по способу выделения радиоресурсов (ортогональное или неортогональное), по объему доступного сетевого содействия, а также по типу дуплексирования сотовых и D2D-соединений. Большинство имеющихся работ подразумевает интеграцию соединений D2D с сотовой системой связи LTE комитета 3GPP [151; 152]. Однако ряд статей также уделяет внимание сетям предыдущих поколений [153]. На практике технология FlashLinQ задает распределенную систему связи D2D [154], которая использует сотовую сеть исключительно для целей синхронизации.

При использовании связи D2D в лицензированном спектре для разгрузки системы LTE разработано множество различных методов повышения ее производительности, таких как схемы MIMO, пригодные для работы с соединениями D2D [155], соответствующие техники сетевого кодирования [156], подходы к последовательному погашению интерференции [157] и способы доставки видеоданных поверх D2D [158]. Можно заключить, что с учетом систематических результатов исследования систем D2D, представленных в работах [145;146], данное направление является хорошо изученным. В ответ на это комитет 3GPP уже несколько лет назад начал стандартизацию технологии LTE-Direct [159], которая представляет собой синхронную систему связи в лицензированном спектре под полным контролем мобильного оператора. Однако множество технических трудностей и разногласий при разработке такого решения задержали выпуск соответствующих рыночных продуктов, что в итоге сместило внимание в сторону реализации соединений D2D на нелицензированных частотах.

Связь D2D в нелицензированном спектре. В нелицензированном спектре, например, в частотном диапазоне для промышленного, научного и медицинского применения ISM, запрещено эксклюзивное использование радиочастот. В результате на этих частотах возможно возникновение значительной неконтролируемой интерференции, которая, в свою очередь, требует надежной технологии радиодоступа, устойчивой к случайным флуктуациям помех. Современные системы Bluetooth и WiFi были разработаны с учетом этого факта и, тем самым, стали широко распространяться в беспроводных персональных и локальных сетях доступа (WPANs/WLANs). Основанная на семействе стандартов IEEE 802.11 технология WiFi представляет собой наиболее популярное решение для связи на нелицензированных частотах как в инфраструктурном режиме, так и при прямом взаимодействии устройств [140;160].

К сожалению, в традиционных сетях WLAN инфраструктурные ТД (точки доступа) не имеют механизмов непосредственного управления радиоресурсами обслуживаемых ими пользователей, которые вынуждены конкурировать за возможность передачи в распределенном режиме. По этой причине системы WiFi часто критикуют за отсутствие строгих гарантий по качеству обслуживания [13]. При этом WiFi-соединения часто получают более высокие скорости передачи данных и энергетическую эффективность по сравнению с альтернативными технологиями [140], поэтому целый ряд исследовательских работ направлен на повышение стабильности при передаче данных без непосредственно-

го участия ТД [161]. Сравнительно недавно была разработана технология WiFi-Direct (WFD) [162], которая позволяет абонентским терминалам соединяться и обмениваться информацией даже при отсутствии инфраструктуры ТД WiFi путем назначения одного из устройств «владельцем группы», тогда как остальные узлы становятся его «клиентами». Каждый «владелец группы» WFD предоставляет синхронизацию всем подключенным к нему «WFD-клиентам», помогая им обнаруживать соседей и подключаться друг к другу.

Поскольку продвинутые абонентские терминалы уже поддерживают технологию WFD, а ТД WLAN продолжают распространяться повсеместно, уровень интерференции в нелицензированном спектре ISM неуклонно возрастает. В этих условиях пользователи услуг WLAN могут получить существенные преимущества от передачи контроля над назначением радиоресурсов, а также ряда других процедур некоторой централизованной функции управления, которая размещена на стороне сотовой системы связи [5]. Централизованная сеть может облегчить решение задач по синхронизации узлов, управлению радиоресурсами и обнаружению других устройств/услуг. Если абонентские терминалы непрерывно ассоциированы с некоторой сотой, она также может содействовать при назначении предпочтительной радиотехнологии (LTE/WFD), управлении мощностью, контроле доступа к среде и выборе формата передачи (модуляционно-кодированная схема, режим использования MIMO и т.д.). Кроме того, при помощи сотовой сети упрощается процесс аутентификации устройств и обеспечения безопасности соединений D2D.

Как отмечалось выше, многие современные устройства связи реализуют поддержку WFD и их доля быстро растет. Это обстоятельство обуславливает необходимость снижения накладных расходов на передачу сигнальных сообщений между ними. Следовательно, снабжение перспективных систем D2D некоторой формой сетевого содействия при работе в нелицензированном спектре может быть весьма предпочтительной [18] в качестве альтернативы непосредственному управлению D2D-соединениями на лицензированных частотах. Как архитектура, так и функции управления традиционными системами связи должны быть в таком случае пересмотрены и дополнены функционалом по обнаружению требуемых устройств/услуг, а также отслеживанию их перемещения. При этом предполагается, что абонентские терминалы по-прежнему ответственны за поддержание требуемого качества связи и переключаются обратно на сотовую инфраструктуру при недостаточной скорости соединения D2D.

3.2.3 Обнаружение соседей и установление соединения D2D

Для обеспечения работы системы D2D требуется выполнить два основных шага: обнаружить соседей и установить прямое соединение. В принципе, оба эти действия могут быть организованы распределенно, но сетевое содействие предоставляет в этой связи ряд важных преимуществ. Поскольку сотовая сеть способна отслеживать местоположение обслуживаемых ею абонентских терминалов, она может в значительной степени снизить долю времени, которую терминал затрачивает на обнаружение других устройств или услуг, информируя его тогда, когда они оказываются поблизости. Такая поддержка позволяет терминалам не использовать радиотехнологию D2D для обнаружения и переключаться на нее только в момент непосредственного установления соединения, что экономит как ресурсы аккумуляторной батареи, так и радиоресурсы.

Более того, так как сотовая система связи располагает средствами безопасного доступа ко всем своим пользователям (а также к абонентам сетей других операторов через соответствующие протоколы взаимодействия), она может произвести безопасное обнаружение «незнакомых» устройств, что повышает доступность данных и/или услуг. Другими словами, снимаются ограничения традиционных способов обнаружения, которые предлагают поиск только среди «знакомых» устройств, предоставляющих данные/услуги, или же требуют установления серии безопасных соединений с «незнакомыми» устройствами в поиске требуемых данных/услуг. При наличии сетевого содействия абонентский терминал получает возможность непосредственного доступа к такой информации, а сотовая система связи информирует его, когда доверенное устройство, предоставляющее целевые данные/услугу по технологии D2D, оказывается достаточно близко для установления прямого соединения.

Другое важное преимущество сетевого содействия состоит в том, что оно позволяет сохранить анонимность абонентских терминалов в процессе обнаружения соседей и установления соединения D2D путем маскирования их исходных идентификаторов. Таким образом, терминалы идентифицируют себя на уровне приложений и далее используют временные идентификаторы канального уровня для работы с соединением D2D, что позволяет им оставаться анонимными для всех устройств сети, кроме своих непосредственных D2D-партнеров. Обнаружение соседей при помощи механизмов сетевого содействия как при вы-

боре партнера, так и при поиске требуемых данных/услуг может быть реализовано различными способами. Наиболее последовательным и эффективным решением является размещение функционала по управлению доступными данными/услугами на сторонних серверах P2P, тогда как контроль над обнаружением и соединением D2D производится в 3GPP-сети мобильного оператора.

На рисунке 3.1 представлены участники типовой сессии P2P. Большинство имеющихся P2P-сервисов предоставляют функционал отслеживания данных/услуг (контента) для обмена [163], который, как правило, располагается на стороннем *сервере приложений*. Последний представляет собой доверенный сервер, размещенный в сети Интернет, к которому получают доступ все зарегистрированные абоненты. Специализированный модуль отслеживания контента протоколирует все доступные данные/услуги, которыми обмениваются пользователи, аутентифицирует их и авторизует доступ к контенту. В традиционных облачных сервисах такая функция отслеживания контента обычно совмещена с системой для его доставки, которая регламентирует процесс обмена. Данная структура характерна для большинства современных социальных сетей, таких как Facebook (использующий Akamai для доставки контента), YouTube (использующий выделенный набор серверов для хранения данных и взаимодействия с пользователем) и множество других.

Для упрощения работы приложений P2P поверх соединений D2D модуль отслеживания контента должен также хранить информацию о местоположении всех данных/услуг, предоставляемых зарегистрированными пользователями. Например, модуль отслеживания контента для YouTube будет сохранять адреса всех Интернет-серверов, которые содержат искомое видео, а также запоминать идентификаторы абонентов, которые предлагают скачать это видео через инфраструктурные или D2D-соединения. В таком случае рассматриваемый модуль передает пользователю, запрашивающему данное видео, список альтернативных источников для его загрузки. К сожалению, пока соединение D2D не установлено, оно не может управляться традиционными методами. Для этого в систему вводится новый доверенный сервер, называемый *сервером D2D*. Он располагается в опорной сети сотового оператора и ведает управлением соединений в процессе обнаружения соседей и установления связи D2D (см. рисунок 3.2).

Предлагаемое здесь размещение сервера D2D в опорной сети мобильного оператора (при рассмотрении сотовой системы 3GPP) обладает рядом дополнительных преимуществ. Во-первых, при нарушении каким-либо абонентским

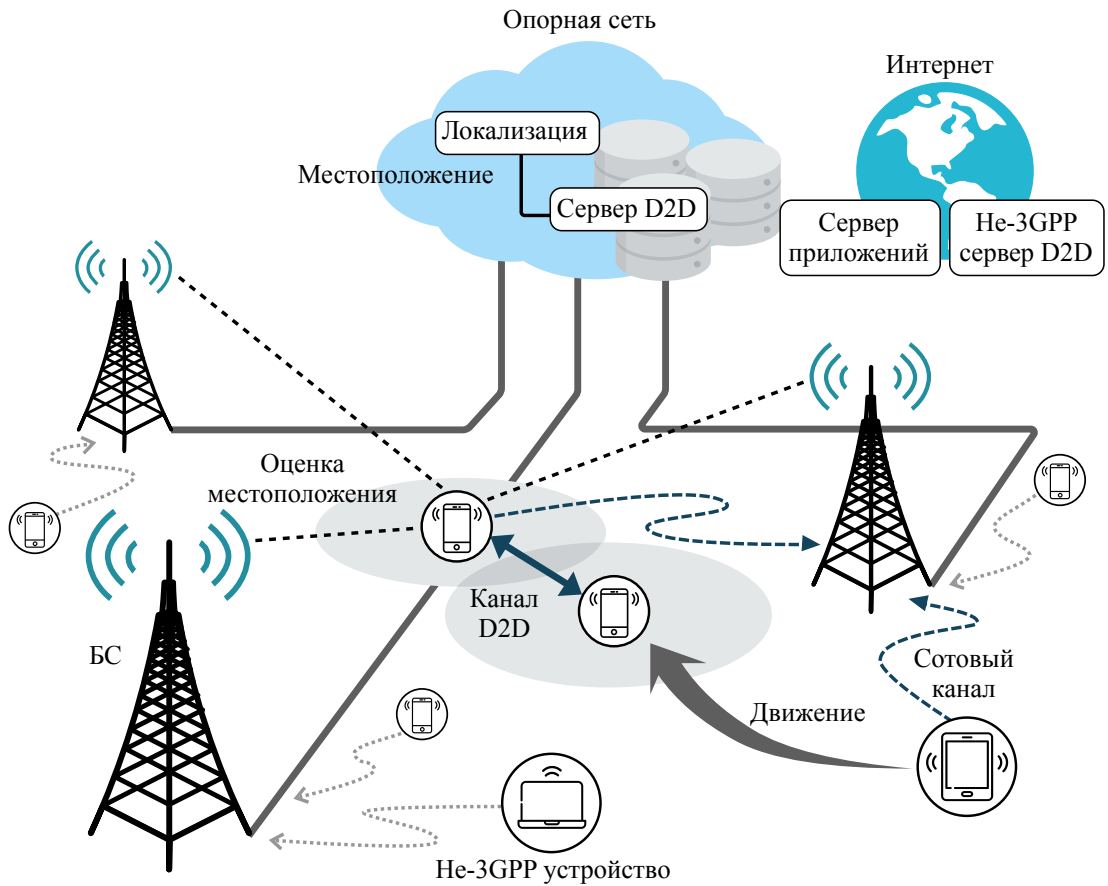


Рисунок 3.1 — Пример работы системы выгрузки на основе соединений D2D

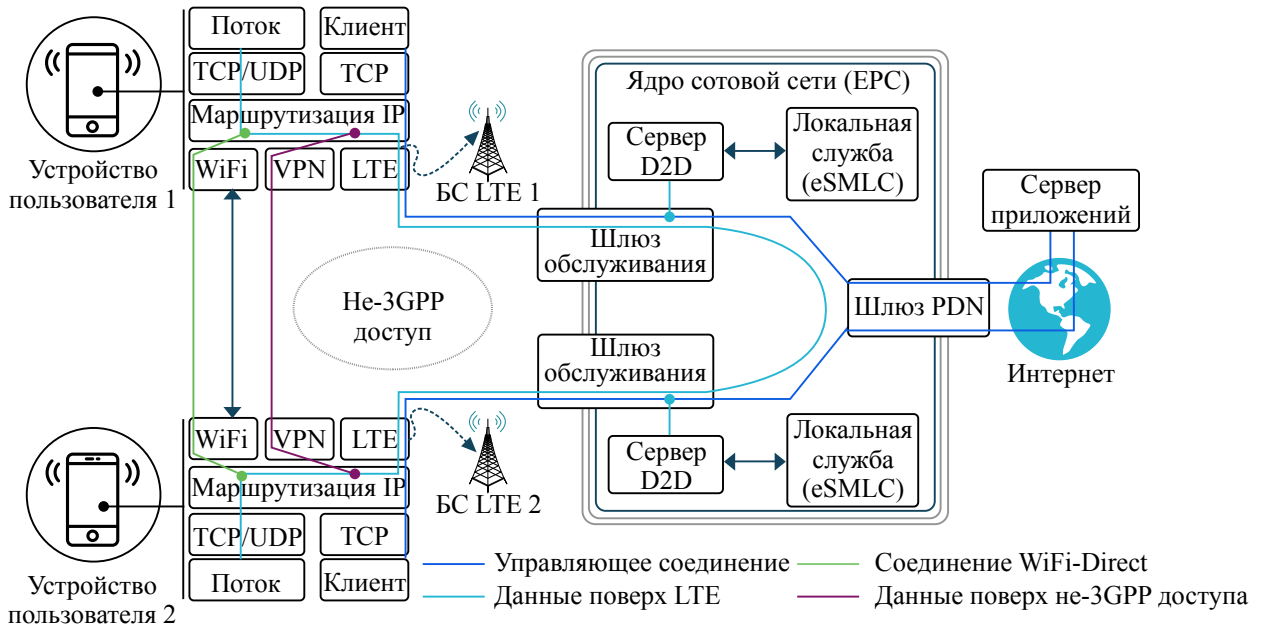


Рисунок 3.2 — Архитектура системы выгрузки на основе соединений D2D

терминалом правил работы, установленных в сети D2D, он может быть окончательно исключен из системы на основании своего аппаратного адреса. Во-вторых, сетевое содействие может быть реализовано через управляющие соединения 3GPP, что гарантирует низкую круговую задержку между сервером D2D и абонентским терминалом даже при высокой загрузке системы связи. Наконец, находясь в ядре сети, сервер D2D получает непосредственный доступ к функциям локализации пользователей для отслеживания их текущего местоположения.

3.3 Использование прямых соединений на нелицензированных частотах

3.3.1 Изучение связи D2D в нелицензированном спектре

Радиоресурсы современных систем связи могут использоваться недостаточно [98], кроме ситуаций повышенной загрузки в определенных областях (часы пик, массовые мероприятия и т.д.). Как следствие, характеристики работы сетей 5G, в том числе поддерживающих прямые соединения D2D, в значительной степени зависят как от интенсивности поступления на обслуживание новых абонентских сессий, так и от географического местоположения соответствующих терминалов. Таким образом, размещение узлов сети в пространстве важно учитывать совместно с динамикой ее трафика во времени. В этой связи многообещающим направлением становится использование пространственно-временного принципа, который комбинирует методы стохастической геометрии [101] с подходами теории массового обслуживания [110; 111; 164]. Этот принцип был применен в предыдущей главе для анализа интегрированных гетерогенных систем связи.

Применение теории массового обслуживания позволяет моделировать динамику системы на уровне сессий, которые поступают в нее на обслуживание и покидают ее по мере его завершения [165]. Для соединений D2D сессия представляет собой поток данных реального времени от одного абонента сети к другому. В данной главе рассматривается возникновение новых сессий согласно

пуассоновскому точечному процессу PPP, который часто применяется в исследовательских работах для анализа взаимодействия сотовых и локальных сетей [166], для изучения способов обнаружения соседей при работе технологии FlashLinQ [167], для учета производительности многослойных гетерогенных систем [97], а также для описания распределений излучаемой мощности и значения SINR в сетях D2D [168]. Однако в большинстве случаев использование стохастической геометрии не позволяет учесть динамику системы связи во времени. Более того, в имеющейся технической литературе отсутствуют подходы, позволяющие проводить пространственно-временное моделирование систем D2D с сетевым содействием.

В данном разделе предлагается новый подход к анализу взаимодействия прямых соединений «устройство – устройство» в нелицензированном спектре и инфраструктурных сотовых соединений в лицензированном спектре, когда сеть оператора имеет возможность выгружать абонентские сессии в слой D2D. Представляют интерес различные практические сценарии такой выгрузки, алгоритмы управления потоками трафика (предпочтительной технологии радиодоступа, LTE или WFD) и характеристики работы системы связи в целом [169]. Например, рассматриваются вероятности блокировки (отказа), когда абонентская сессия не принимается на обслуживание слоем D2D, сотовым слоем или же ими обоими. При возрастающем внимании к энергетической эффективности мобильных устройств с ограниченным объемом аккумуляторной батареи [26], также важно оценить энергетическое потребление типовой абонентской сессии с учетом модели потребления мощности, приведенной в [98].

Последующее изложение нацелено на построение математической модели для анализа выгрузки сессий сотовых абонентов на прямые соединения D2D в системах 5G, предоставляющих сетевое содействие. В отличие от имеющихся результатов со статическим размещением узлов, проведенное далее моделирование позволяет учесть динамику системы связи во времени. В частности, производится всестороннее изучение соединений D2D, соседствующих одновременно в некоторой географической области. Поскольку связь в нелицензированном спектре неизбежно подвержена влиянию интерференции, сформулированный здесь подход позволяет в явном виде описывать взаимодействие абонентских терминалов, находящихся в непосредственной близости друг к другу. При этом предполагается, что сотовая сеть содействует системе D2D на этапе контроля доступа новых сессий.

Для выявления преимущества связи D2D также рассматривается модель традиционной сотовой сети, которая в явном виде управляет уровнем возникающей интерференции, а также решает задачи контроля доступа, назначения мощности передачи и планирования радиоресурсов для обслуживаемых ею абонентских терминалов. Передавая данные реального времени, терминалы устанавливают уровень своей излучаемой мощности в соответствии с подходом в [98], и при этом предлагаются альтернативные методы управления потоками трафика. На этой основе проводится имитационное моделирование сети 5G с соединениями D2D для подтверждения правомерности используемых методов аналитического моделирования, а также в целях выявления прироста производительности от разгрузки сотовой системы связи с помощью прямых соединений «устройство – устройство» при наличии сетевого содействия.

3.3.2 Подход к анализу систем D2D

Данный подраздел вводит в рассмотрение модель системы D2D, включающую в себя множество прямых соединений и традиционное сотовое покрытие, которые здесь называются слоями. Далее следуют основные допущения предлагаемой модели.

Общие соображения. Предлагаемый здесь подход к анализу рассматривает сотовую систему связи на лицензированных частотах, которая взаимодействует с множеством соединений D2D в нелицензированном спектре для обеспечения более эффективного обслуживания данных от абонентских терминалов [170]. Предполагается, что передаваемый трафик соответствует сессиям реального времени с минимальной требуемой скоростью передачи данных r . Для каждой сессии i вводится *передающий терминал* T_i , который становится источником данных, и *принимающий терминал* R_i , являющийся соответствующим получателем. В силу использования неперекрывающихся частотных диапазонов, передачи абонентов в двух введенных слоях не интерферируют между собой, тогда как каждый T_i может отправлять свои данные R_i либо через сотовый слой (*инфраструктурное соединение*), либо через слой D2D (*прямое соединение*), как показано на рисунке 3.3. Для упрощения последующих рассуждений

здесь рассматривается только передача данных, которая предназначена для получения непосредственными D2D-партнерами.

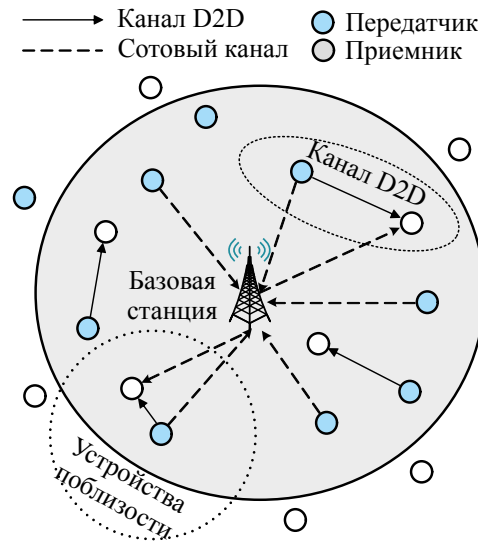


Рисунок 3.3 — Модель выгрузки сотового трафика на соединения D2D

Для моделирования пространственного расположения узлов системы в явном виде вводится следующее основное допущение.

Допущение. Передающие терминалы размещены согласно однородному пуассоновскому точечному процессу PPP в трехмерном пространстве (как точки на двумерной плоскости, а также на оси времени) с интенсивностью $\lambda f(x)$, которая не зависит от времени, где $f(x)$ задает распределение местоположения абонентов на плоскости. Далее для упрощения анализа такое распределение выбирается равномерным внутри окружности заданного радиуса R .

Для некоторого передающего терминала T_i соответствующий принимающий терминал R_i поступает в систему одновременно с T_i так, что местоположение R_i также распределено равномерно внутри той же окружности радиуса R . Дополнительно предполагается, что длительность сессии каждого терминала T_i распределена по экспоненциальному закону со средним $\frac{1}{\mu}$.

Модель сотового слоя. Здесь изучается изолированная сота централизованной сети мобильного оператора, для которой не учитывается межсотовая интерференция. Это упрощение приводит к рассмотрению системы связи, не подверженной влиянию какой-либо интерференции, поскольку передачи данных от различных абонентских терминалов разнесены и не влияют друг на друга. В предположении, что отправка информации по линии вниз, как правило, более надежна и располагает большим количеством радиоресурсов, далее

исследуется только передача данных по линии вверх, т.е. часть инфраструктурного соединения от пользователя до обслуживающей его сотовой БС (базовой станции). Таким образом, анализ сужается до рассмотрения исключительно передающих терминалов.

Распространение радиосигнала в канале связи моделируется в соответствии с документами стандартизации [171] и в целях сохранения аналитичности предполагается, что для сессии i величина SNR на единицу мощности γ_i принимает вид

$$\gamma_i = \frac{G^\kappa}{d_i^\kappa}, \quad (3.1)$$

где d_i – расстояние между БС и передающим терминалом T_i ; κ – экспонента распространения, а G – константа распространения радиосигнала.

Без потери общности также предполагается, что скорость передачи данных является непрерывной функцией, а отображение мощность – скорость определяется формулой Шеннона.

Допущение. Мощность сигнала p_i , излучаемого терминалом i , и соответствующая скорость передачи данных r_i связаны друг с другом через формулу Шеннона, не учитывающую вклад интерференции,

$$r_i = w \log(1 + \gamma_i p_i), \quad (3.2)$$

где p_i – мощность сигнала на выходе радиочастотного усилителя; γ_i – значение SNR на единицу мощности и w – полоса пропускания канала связи.

Здесь используется реалистичная и в то же время достаточно простая модель потребляемой мощности (см. рисунок 3.4), задающая различные *уровни мощности* для каждого передающего терминала, а также учитывающая эффективность антенны η . Данный подход в принципе следует работе [98] и включает в себя (1) переменную мощность передачи $p_{tx} = \frac{1}{\gamma_i \eta} (e^{r_i/w} - 1) + p_c$, которая потребляется в процессе отправки пользователем своих данных ($r_i > 0$), где p_c – некоторое постоянное значение, связанное с потреблением электроцепи, а также (2) постоянную мощность простоя p_a , называемую здесь активной, которая потребляется в тех случаях, когда терминал не отправляет информацию, активно ожидая возможности передачи.

Сотовая БС управляет работой сети при помощи двух альтернативных подходов, называемых здесь *алгоритмами передачи*. Каждый алгоритм решает задачи допуска в сеть, управления мощностью и планирования радиоресурсов.

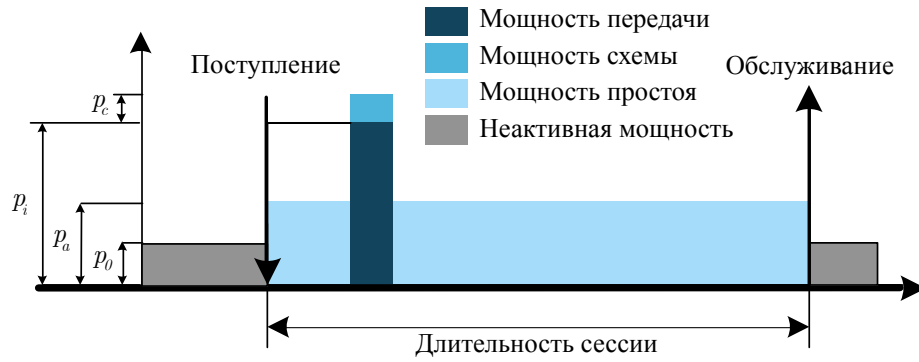


Рисунок 3.4 — Уровни потребляемой передающим терминалом мощности

При принятии на обслуживание передающий терминал занимает долю общего ресурса системы во времени и устанавливает свою излучаемую мощность в соответствии с указаниями БС для достижения требуемой скорости передачи данных. БС пересматривает принятые решения относительно назначения радиоресурсов и мощности передачи для всех активных терминалов каждый раз, когда в систему поступает новая сессия или же текущая сессия завершает свое обслуживание и покидает ее.

Для алгоритма передачи на *максимальной скорости* (MR) предполагается, что терминалы отправляют свои данные на предельно допустимом уровне мощности. В условиях отсутствия интерференции это соображение приводит к тому, что каждый абонент получает скорость передачи, равную своему наибольшему возможному значению. С учетом заданного здесь соотношения мгновенная скорость передачи данных для сессии i устанавливается исходя из значения максимальной мощности передачи p_{\max} :

$$r_i^{\max} = w \log(1 + \gamma_i p_{\max}). \quad (3.3)$$

Следовательно, сотовый слой допускает вновь поступившую сессию, если имеется достаточное количество радиоресурсов для ее обслуживания. Другими словами, каждая текущая сессия i занимает в точности долю r/r_i^{\max} от общего времени работы системы, тогда как для всех активных сессий выполняется следующее:

$$\sum_{all\ sessions} \left(\frac{r}{r_i^{\max}} \right) \leq 1. \quad (3.4)$$

Для алгоритма передачи MR, даже в случае достаточно высокой интенсивности входного потока сессий, радиоресурсы сети связи могут использоваться

недостаточно в том смысле, что всегда имеется (стремящаяся к нулю) доля времени работы системы, когда абонентские терминалы не осуществляют передачу своих данных (т.е. система простаивает).

В качестве альтернативы вводится алгоритм передачи с *полным использованием* (FU), который контролирует, чтобы время работы системы связи всегда распределялось полностью. Соответственно, каждая текущая сессия получает равную долю общего ресурса системы во времени, т.е. $\frac{r}{r_i} = \frac{1}{n}$, а передающие терминалы подстраивают свою излучаемую мощность так, чтобы достичь минимальной требуемой скорости передачи данных. Таким образом, в случае n одновременно активных сессий выполняется следующее:

$$\frac{r}{r_i} = \frac{1}{n}, \quad r_i = rn, \quad \forall i = \overline{1, n}. \quad (3.5)$$

Как следствие, для приема на обслуживание новой сессии, БС должна повысить мощность передачи других активных абонентских терминалов так, чтобы они «уложились» в меньшие интервалы времени после переупаковки радиоресурсов. Если такое повышение невозможно хотя бы для одной из n активных сессий, т.е. $r_i^{\max} = w \log(1 + \gamma_i p_{\max}) < nr$, тогда рассматриваемой сессии будет отказано в доступе в сотовый слой. В противном случае система перераспределит время своей работы для активных сессий, и абонентским терминалам будет установлен более высокий уровень излучаемой мощности

$$p_i = \frac{1}{\gamma_i} \left(e^{nr/w} - 1 \right). \quad (3.6)$$

В результате алгоритмы передачи MR и FU предоставляют возможность выбора между повышением емкости сети связи (приводящим к росту энергетического потребления) или же увеличением эффективности использования радиоресурсов (позволяющим снизить излучаемую мощность). Анализ обоих альтернативных алгоритмов, приведенный в приложении В, позволяет в дальнейшем подобрать подходящее решение в зависимости от целей исследования.

Модель слоя D2D. Поскольку сеть D2D работает на нелицензированных частотах, одновременные сессии передачи данных от абонентских терминалов могут интерферировать между собой. При проведении аналитического моделирования такая интерференция должна быть учтена в явном виде. При этом коэффициент передачи радиоканала $\gamma_{i,j}$ между передающим T_i и принимающим R_j терминалами зависит от расстояния $d_{i,j}$ между ними. В отличие от

модели сотового слоя, для зависимости между мощностью сигнала и скоростью передачи данных делается следующее допущение.

Допущение. Мощность сигнала p_i , излучаемого абонентским терминалом, и соответствующая скорость передачи данных r_i связаны друг с другом через формулу Шеннона, учитывающую вклад интерференции:

$$r_i = w \log(1 + \text{SINR}_i) = w \log\left(1 + \frac{p_i \gamma_{i,i}}{N_0 + I}\right), \quad (3.7)$$

где p_i – мощность сигнала на выходе радиочастотного усилителя; $\gamma_{i,i}$ – коэффициент передачи радиоканала между передающим и принимающим терминалами для сессии i ; w – полоса пропускания канала; N_0 – уровень шума в канале связи и I – уровень интерференции на приемной стороне.

Здесь также рассматриваются два различных уровня потребляемой мощности для каждого абонентского терминала в системе: (1) мощность передачи, включающая компоненту p_c , связанную с потреблением электроцепи, и (2) мощность простоя. В D2D-слое мощность передачи полагается постоянной и равной своему максимальному значению p_{\max} . Кроме того, делается допущение о том, что совокупный уровень шума и интерференции в канале не превышает некоторого общесистемного порога $N_0 + I \leqslant K N_0$. Наконец, предполагается, что слой D2D, уже обслуживающий $n - 1$ активную сессию, допускает вновь поступившую сессию n , если для множества $\{T_i\}_{i=1}^n$ передающих терминалов выполняются следующие условия на каждом из принимающих терминалов R_j , $j = \overline{1, n}$:

$$\frac{p_{\max} \gamma_{i,i}}{K N_0} \geqslant e^{\frac{r}{w} - 1} \text{ и } p_{\max} \gamma_{i,j} \leqslant N_0, \quad \forall i, \quad j \neq i, \quad (3.8)$$

где r – минимальная требуемая скорость передачи данных, а значение постоянной K зафиксировано для всего слоя D2D. Эти условия означают, что целевая скорость передачи r может быть достигнута для каждого соединения i (см. левую часть данного выражения) и одновременно интерференция на стороне R_j , произведенная T_i , не превышает заданного порога N_0 (см. правую часть данного выражения).

Работа системы и ее характеристики. Общий принцип работы рассматриваемой системы связи пояснен на рисунке 3.5. Соответственно, предполагается последовательный контроль доступа с сетевым содействием для новых сессий, поступающих в систему. Вначале процедура содействия в выборе сети доступа, функционирующая в сотовом слое, предпринимает попытку выгрузки

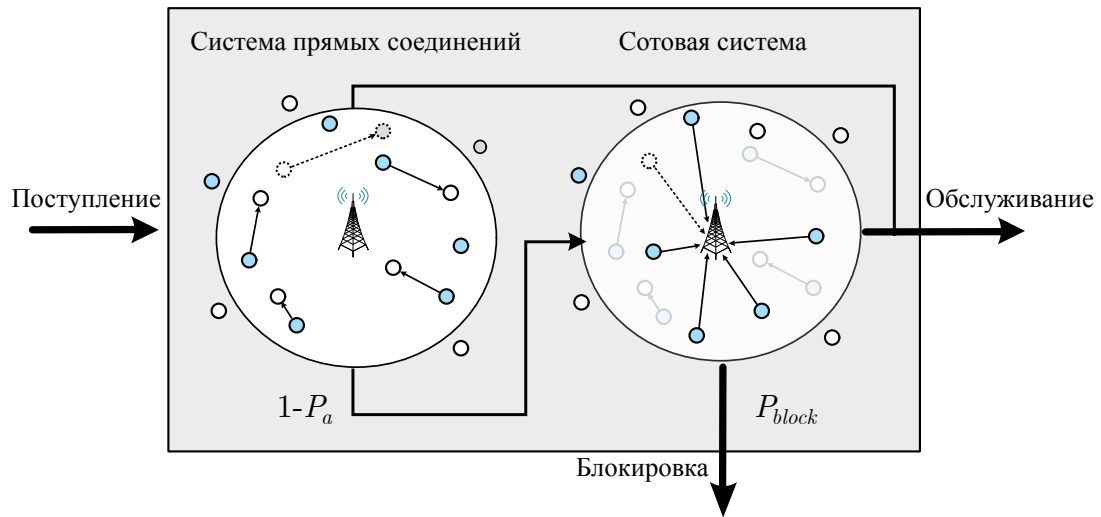


Рисунок 3.5 — Организация работы рассматриваемой системы связи

вновь прибывшей сессии на обслуживание в слой D2D. В случае успеха выгруженная в слой D2D сессия обслуживается в нем без прерываний до момента своего окончания и затем покидает систему. В противном случае запускается процедура контроля доступа в сотовый слой согласно алгоритму MR или FU. В итоге, если некоторая сессия не может быть допущена даже в сотовый слой, то она считается окончательно заблокированной и покидает систему, не получив обслуживания.

Следует напомнить, что исходная интенсивность поступления сессий на (первый) слой D2D составляет λ . Ввиду пуассоновского характера просеянного потока, поступления сессий во (второй) сотовый слой (т.е. не принятых в слой D2D) происходят в соответствии с пуассоновским процессом интенсивности $\lambda(1 - P_a)$, где P_a – вероятность допуска сессии в слой D2D. Абстрагируясь от конкретного местоположения абонентских терминалов в целях сохранения аналитичности, можно предположить, что поступления сессий в сотовый слой также происходят равномерно внутри окружности радиуса R . В результате *общесистемная вероятность блокировки* P_{block} может быть вычислена как

$$P_{block} = 1 - [P_a + (1 - P_b)(1 - P_a)], \quad (3.9)$$

где P_a – вероятность допуска сессии в слой D2D; P_b – вероятность блокировки (отказа) сессии сотовым слоем.

Другой важной характеристикой работы системы связи, рассматриваемой в данном разделе, является энергетическое потребление ϵ для типовой абонентской сессии при условии, что она удовлетворяет требованию по минимальной

скорости передачи данных. Эта величина следует из закона Литтла может быть найдена как

$$E[\epsilon] = \frac{E[P]}{\lambda P_a}. \quad (3.10)$$

Здесь вероятность P_a для слоя D2D может быть заменена на вероятность $1 - P_b$ для сотового слоя.

Применимость предложенной модели. В отличие от централизованных сотовых (3GPP LTE) сетей, в которых все соединения устанавливаются через БС, распределенные системы WLAN (IEEE 802.11 или WiFi), как правило, предоставляют возможность установления прямых соединений между абонентскими терминалами. Однако отсутствие удовлетворительного решения для ряда ключевых задач являлось долгое время сдерживающим фактором развития существующих технологий связи «устройство – устройство». Основной проблемой становится отсутствие эффективных методов обнаружения соседей и установления/поддержания соединения, поскольку в распределенных сетях реализация быстрых и надежных механизмов для достижения этих целей представляет собой значительную сложность. При этом современные абонентские терминалы имеют возможность использовать несколько технологий радиодоступа одновременно, и сотовая сеть может тем самым содействовать системе WLAN при установлении соединений D2D.

Вообще говоря, потенциальное соединение D2D, которое еще не было установлено, не может управляться традиционными способами. Соответственно, в состав архитектуры сотовой сети должны быть введены дополнительные модули для организации такого управления, в частности, для отслеживания контента, обеспечения безопасности, а также контроля D2D-соединений. В этой связи необходимо провести надлежащую изоляцию функционала этих новых модулей, и в данной главе было предложено разделить функции управления соединением (реализуемые сервером D2D) и функции отслеживания контента (реализуемые сервером приложений). Преимуществом такого подхода является простота интеграции технологии D2D с существующей архитектурой системы 3GPP LTE [88]. Далее обсуждаются вопросы соответствия предложенной выше модели особенностям современных радиотехнологий.

В соответствии с допущениями модели системы, предложенный подход к анализу соединений D2D позволяет моделировать динамическое взаимодействие сетей 3GPP LTE и WFD (WiFi-Direct). Однако основные его положения

могут быть обобщены на случай альтернативных способов реализации технологии связи D2D, например, в лицензированном спектре. В частности, допущение относительно экспоненциальной длительности сессий сделано здесь исключительно для наглядности. Все соответствующие рассуждения могут быть проведены для произвольного распределения длительности сессий. При моделировании взаимодействия систем LTE и WFD в явном виде важно предположить, что все обслуживаемые абонентские терминалы имеют хотя бы два соответствующих радиомодуля, позволяющие им использовать обе технологии радиодоступа одновременно. Также требуется, чтобы терминалы постоянно поддерживали активным некоторое управляющее соединение с сотовой БС, через которое осуществляется сетевое содействие.

Согласно допущениям модели системы отображение мощность – скорость определяется формулой Шеннона. Было наглядно показано, что такой подход может служить удовлетворительной аппроксимацией для современных радиотехнологий [109]. Для дальнейшего приближения предложенного здесь математического подхода к реальности можно применить ряд дополнительных ограничений, наложенных использованием практических модуляционно-кодовых схем в системе LTE и/или WFD. В частности, передающие терминалы не могут превысить некоторое максимальное значение скорости передачи данных r_{\max} (составляющее, например, около 60 Мбит/с в сети LTE и 56 Мбит/с в сети WFD), ограничивая, таким образом, наибольшую эффективную величину SINR на приемной стороне как

$$r_{\max} = w \log(1 + \text{SINR}_{\max}). \quad (3.11)$$

Тем самым значения γ_i и $\gamma_{i,j}$ не могут возрасти неограниченно, по мере того как d_i или $d_{i,j} \rightarrow 0$, и после некоторого $\gamma_{\max} = \frac{1}{P_{\max}} (e^{\frac{r_{\max}}{w}} - 1)$ скорость передачи данных далее не увеличивается.

В системе связи D2D процедура контроля доступа определяет, может ли некоторая сессия быть принята на обслуживание. При идеализированном планировании работы слоя D2D уровень излучаемой мощности для каждого передающего терминала будет установлен индивидуально, в целях максимизации, например, общей пропускной способности сети связи. Однако реализация такого планирования на практике не представляется возможной в силу слишком высоких накладных расходов на передачу сигнальных сообщений. Как следствие, целесообразно использование упрощенных решений, которые отражают аналогичные подходы в системах стандарта IEEE 802.11. Соответственно пред-

полагается, что (1) мощность передачи является постоянной и (2) уровень шума в канале связи не превышает некоторого фиксированного значения N_0 . Каждый раз, когда данное условие не выполняется на приемной стороне (или же не удовлетворяется на одном из других принимающих терминалов), передающий терминал отказывается от отправки данных и соответствующий пользователь покидает слой D2D.

Рассматриваемая здесь процедура в значительной степени повторяет механизм прослушивания несущей частоты при работе технологии WFD. Он удостоверяется, что уровень интерференции, произведенной некоторой передачей на каждый из принимающих терминалов, никогда не превысит значения N_0 . При этом, тем не менее, нет гарантии, что сумма уровней интерференции, произведенной всеми активными сессиями в совокупности, не превысит N_0 . Учитывая такую кумулятивную интерференцию, можно ввести дополнительный коэффициент K . Его значение на практике может быть установлено как наибольшее число потенциально интерферирующих соединений в непосредственной близости от рассматриваемого принимающего терминала.

3.3.3 Имитационное моделирование и его результаты

Для проверки результатов, полученных в рамках предложенной аналитической модели, был использован комплекс программ имитационного моделирования, разработанный с учетом актуальных спецификаций 3GPP LTE и IEEE 802.11 (с поддержкой функционала WiFi-Direct). С одной стороны, на сегодняшний день отсутствует как бесплатное, так и коммерчески распространяемое программное обеспечение, которое позволяет проводить имитационное моделирование полноценной системы связи D2D. Это происходит ввиду нехватки соответствующих сетевых функций или же недостаточного уровня масштабируемости имитаторов для изучения зависимостей между характеристиками D2D-соединений. С другой стороны, применяемый здесь комплекс программ представляет собой гибкий инструмент, созданный для исследования различных способов размещения узлов сети, моделей трафика, параметров радиоканала и алгоритмов управления системой (см. подробности по ссылке [135]).

Разработанная система имитационного моделирования воссоздает традиционную архитектуру сетей LTE и WiFi, включая всевозможные варианты размещения абонентских терминалов и модулей сетевой инфраструктуры (шестиугольные или квадратные соты, среда с «оберткой» сотами или без нее, равномерное или кластерное распределение узлов и т.д.). В целях улучшения производительности системы применяется ряд упрощений в моделировании реальных технологий радиодоступа, но при этом основные механизмы их функционирования воссоздаются достаточно точно. С помощью данного инструмента была продемонстрирована целесообразность установления соединений D2D в нелицензированном спектре с сетевым содействием со стороны сотовой системы связи, а также выявлены их возможности по повышению качества пользовательского соединения [172; 173].

Структура и работа системы моделирования. Согласно введенным ранее предположениям относительно функционирования исследуемой системы связи, рассматривается сотовая сеть 3GPP LTE для организации инфраструктурных соединений и радиотехнология IEEE 802.11 (WiFi-Direct) для поддержки соединений D2D. В рамках технологии LTE используемая система имитационного моделирования (в отличие от аналитического подхода, рассмотренного выше) реализует ее следующие практические особенности: структуру кадров для передачи данных, запросы на резервирование полосы пропускания канала связи и планирование радиоресурсов со стороны БС. Для сети D2D имитационное моделирование следует правилам работы процедуры контроля доступа в сеть с прослушиванием несущей частоты согласно стандарту IEEE 802.11. Однако для соотнесения данного функционала с возможностями введенной системы допущений аналитического моделирования применяются следующие упрощения работы практических технологий радиодоступа.

1. Процедура контроля доступа в сеть предполагает, что качество канала связи между всеми абонентскими терминалами известно заранее. Это допущение правомерно при наличии сетевого содействия, когда БС отслеживает, хранит и сообщает данную информацию терминалам.
2. В целях упрощения предполагается, что сессия данных от абонента получает доступ к каналу на протяжении всей своей длительности в отличие от реальных протоколов, которые резервируют радиоресурсы по факту наличия данных для передачи. Такое поведение алгоритма резервирования может служить оценкой для характеристик реальной

системы. При этом имитатор контролирует, что в случае установления D2D-соединения оно гарантировано получает минимальную требуемую скорость передачи данных при наличии свободных радиоресурсов.

3. Когда соединение уже установлено, передача данных от абонентского терминала фрагментируется на пакеты одинаковой длины (здесь предполагается размер в 1000 байт), которые отправляются через равные промежутки времени, подбираемые так, чтобы достичь требуемой скорости передачи данных на протяжении всей длительности сессии.

Пример моделируемого сценария. Далее описывается пример сценария для проведения имитационного моделирования, который демонстрирует выгрузку абонентских сессий из системы LTE на соединения WiFi-Direct при наличии сетевого содействия. Соответственно, рассматривается ограниченная *область моделирования* [174], в которой сотовое покрытие совмещено с зоной обслуживания сети D2D (например, торговый центр или транспортный узел). В частности, изучается работа изолированной круглой соты радиуса $R = 100$ м, на которую не воздействует интерференция со стороны соседних сот. В данной области абоненты обмениваются небольшими фрагментами мультимедиа-данных с минимальной требуемой скоростью передачи $r = 4,8$ Мбит/с. Поскольку длительность сессии распределена экспоненциально со средним $s^{-1} = 3$ с, типовая посылка данных от абонента содержит 2 Мбита информации.

Предполагается, что интервалы времени между поступлениями новых сессий распределены экспоненциально так, что в среднем λ новых сессий прибывает в систему каждую секунду и требует своего обслуживания. Каждый передающий терминал отправляет свои данные соответствующему принимающему терминалу, также находящемуся внутри рассматриваемой области моделирования. При этом некоторый передающий терминал может быть или допущен в слой D2D, или же ему может быть отказано в доступе, и он будет вынужден запрашивать обслуживания у БС LTE. Если свободных радиоресурсов сотового слоя недостаточно для обеспечения потребности этого пользователя, то он получает окончательный отказ и блокируется.

Далее рассматривается пример работы алгоритма передачи на максимальной скорости MR, когда все терминалы передают с наибольшим возможным значением излучаемой мощности. Дополнительные параметры, связанные с мощностью передачи, описаны в [175] для WiFi и в [176] для LTE, тогда как остальные настройки системы связи приведены в таблице 3.

Таблица 3 — Параметры моделирования системы прямых соединений

Параметр	Сотовая сеть	Сеть D2D
R , м	100	100
r , Мбит/с	4,8	4,8
μ^{-1} , с	3	3
k	5	6,5
G	197,43	2,4
w , МГц	10 (ITU-R M.2135)	20 (IEEE 802.11)
η	0,5	0,5
N_0 , дБ	-60	-70
r_{\max} , Мбит/с	60	56
p_a , мВт	0,05	0,005

Имитационные результаты и выводы. На основании описания рассматриваемой сети связи далее проводится имитационное моделирование работы системы D2D, изображенной на рисунке 3.5, а также предлагается сравнение полученных результатов с аналогичными величинами для сотовой сети, не поддерживающей прямых соединений.

Одной из основных характеристик функционирования изучаемой системы является ее емкость, выражаемая в числе сессий, которые могут обслуживаться в ней одновременно. Соответственно, на рисунке 3.6 данная величина сравнивается для базовой сети LTE и для системы с поддержкой связи D2D, что позволяет сделать вывод о значительных преимуществах (улучшение порядка 20%) при использовании прямых соединений. Здесь непрерывные линии отображают результаты имитационного моделирования (И), тогда как символы

соответствуют значениям, полученным при помощи математического анализа (А). Легко видеть, что общий характер зависимостей подтверждает рост среднего числа активных сессий вплоть до момента насыщения, который зависит от типа размещения узлов, способа планирования радиоресурсов и метода интеграции прямых соединений.

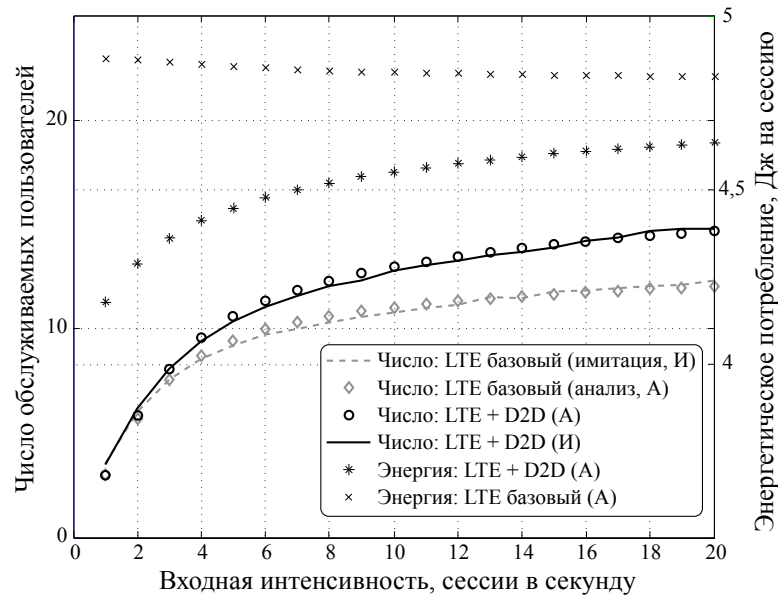


Рисунок 3.6 — Емкость системы и энергопотребление абонентских терминалов

Вероятность блокировки в рассматриваемой системе (см. рисунок 3.7) связана с ее емкостью и определяется долей запросов на обслуживание, которые не могут быть удовлетворены сетью связи. Исследуются значения общесистемной вероятности блокировки P_{block} , вероятности отклонения сессии слоем D2D и вероятности блокировки сетью LTE в зависимости от загрузки системы связи. Следует напомнить, что сотовая сессия будет заблокирована, если она не может быть включена в расписание передач на момент своего поступления, тогда как в сети D2D различаются (1) отклонение сессии по причине превышения порога интерференции от одновременно передающих терминалов и (2) отклонение по причине превышения эффективной длительности соединения в плане минимальных требований по скорости передачи (при условии, что ограничение по интерференции отсутствует).

Важно проанализировать структуру процессов блокировки в обеих рассматриваемых системах. В сети D2D при низкой загрузке отказы связаны в основном с превышением эффективной длины соединения, а при увеличении загрузки возрастает доля отказов, связанных с превышением порога по интер-

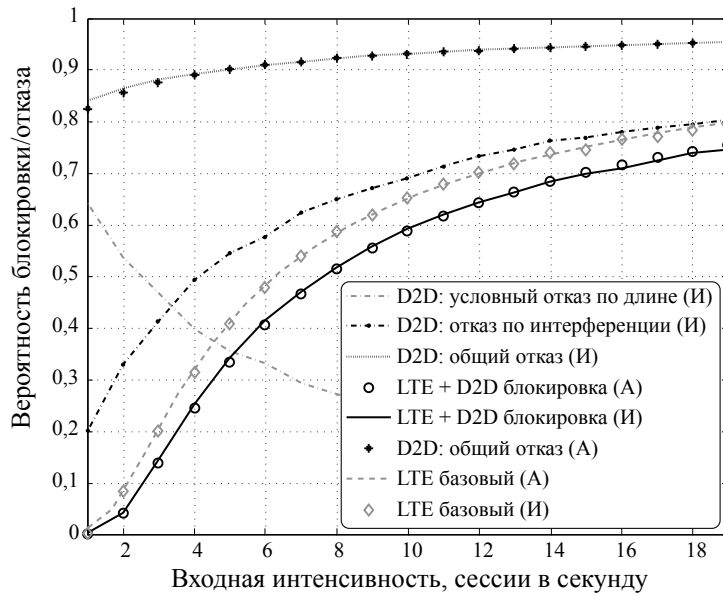


Рисунок 3.7 — Вероятность блокировки (отказа) для сессии

ференции. Интересно отметить, что в сети LTE блокировка сессий не происходит по порогу, несмотря на работу алгоритма планирования. На практике в исследуемых сценариях сотовая система связи никогда не достигает полной загрузки, которая привела бы к стопроцентной блокировке. Это можно объяснить тем фактом, что сеть продолжает допускать новые сессии при условии, что их обслуживание может быть включено в расписание передач. При этом в случае высокой загрузки неявно получают приоритет те сессии, которые имеют более высокие скорости передачи (см. рисунки 3.8 и 3.9).

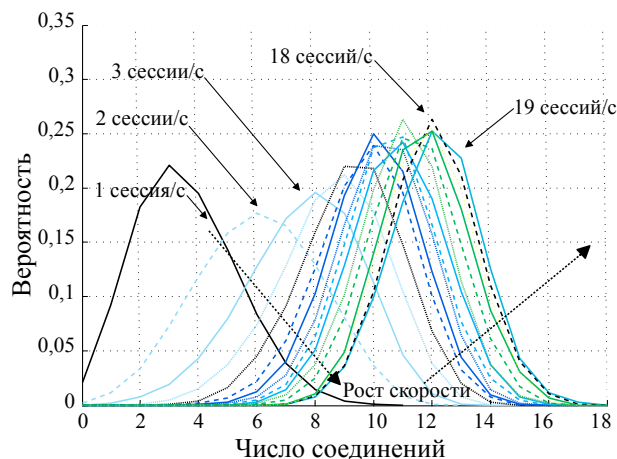


Рисунок 3.8 — Распределение числа соединений LTE

Кроме того, изучается энергетическое потребление абонентских терминалов в сети LTE, а также в интегрированной системе LTE-D2D, которое также

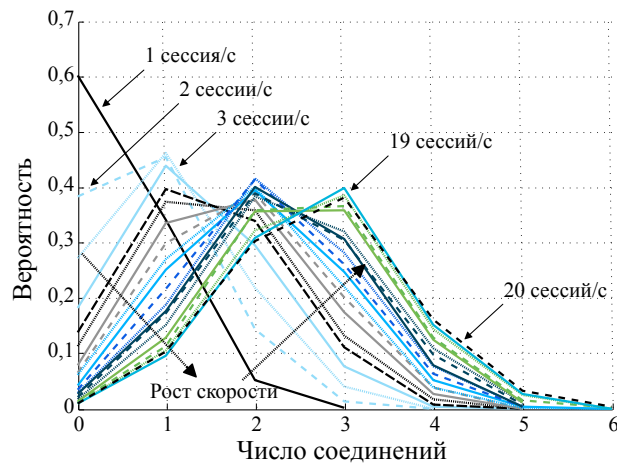


Рисунок 3.9 — Распределение числа соединений D2D

приведено на рисунке 3.6. Можно заметить, что на низких интенсивностях входного потока сессий соединения D2D существенно влияют на значение энергетической эффективности в интегрированной системе, улучшая ее вплоть до 14%. Однако с увеличением загрузки, когда слой D2D более не может принимать на обслуживание дополнительные сессии, прирост энергетической эффективности становится менее значительным.

Разумеется, наблюдаемый здесь эффект снижения энергетического потребления во многом зависит от конкретных параметров передающих терминалов. Поэтому предложенный в данной главе подход к анализу системы связи D2D становится полезен при проведении расчета энергетической эффективности для требуемой модели потребляемой мощности в заданном диапазоне значений интенсивности входного потока сессий. Можно сделать вывод о том, что выгрузка пользовательского трафика из сети LTE на прямые соединения D2D, функционирующие по технологии WiFi-Direct, при содействии со стороны сотовой инфраструктуры существенно снижает возможность блокировки сессии и в то же время повышает энергетическую эффективность передачи данных.

3.4 Анализ использования крайне высоких частот в сетях нового поколения

3.4.1 Перспективы связи на крайне высоких частотах

Множество практических результатов подтвердило целесообразность использования крайне высоких частот, т.е. миллиметровых волн mmWave для организации соединений в сетях 5G [35]. Как следствие, возможность построения беспроводной связи на основе mmWave привлекла внимание как академического, так и индустриального сообществ. Действительно, диапазон с длиной волны от 10 до 1 мм, соответствующий частотам от 30 до 300 ГГц, содержит значительный спектральный запас, обещающий закрыть существенную долю потребности сетей 5G в радиоресурсах. В этой связи ряд организаций, включая Комитет по стандартизации мобильных систем 3GPP, активно разрабатывает спецификации новых радиотехнологий с использованием миллиметровых волн. В частности, ожидается, что сотовая связь mmWave будет работать вначале на частоте около 28 ГГц, а потом и на более высоких частотах [177].

Дополнительная емкость сети, которая станет доступна после размещения малых сот mmWave с высокой плотностью покрытия, сделает возможной поддержку критически важных приложений мобильной связи, таких как продвинутое пользовательские услуги и новые сервисы Интернета вещей, функционирующие без непосредственного участия человека [178]. Для своей оперативной и надежной работы эти приложения требуют значительных радиоресурсов [179], что обуславливает создание новой радиотехнологии в рамках архитектуры 5G, которая получила название «Новое Радио» (NR). Поскольку система NR в том числе обеспечивает поддержку связи mmWave, необходимо всестороннее исследование особенностей организации соединений на крайне высоких частотах [180], включая *блокирование* передачи небольшими объектами из-за более коротких длин волн, повышенный уровень затухания сигнала в канале связи и т.д.

Особенности связи mmWave послужили предметом большого числа исследовательских статей, посвященных обеспечению надежной связи на крайне высоких частотах, а также поддержке перспективных приложений 5G с повышен-

ными требованиями к качеству обслуживания [181]. В итоге многие результаты нашли отражение в работе комитета 3GPP по созданию спецификаций технологии NR, которая в 2018 году перешла в фазу практической апробации [182]. По мере того как накапливается объем знаний об использовании крайне высоких частот в сетях 5G и завершается соответствующая стандартизация, особый интерес представляет анализ и управление связью mmWave в типовых сценариях ее применения. Здесь учитывается множество практических факторов, таких как параметры и особенности среды передачи, а также динамика и качество соединений mmWave.

Помимо явных преимуществ, например, относительно широкая полоса пропускания, перспективная технология связи mmWave имеет ряд специфических особенностей, которые требуют детального изучения. Во-первых, возникает необходимость в разработке принципиально новых моделей распространения радиосигнала, которые отличаются от известных подходов для современных микроволновых сетей доступа LTE и WiFi. Комитетом 3GPP при поддержке ведущих специалистов в данной области были проведены измерения и предложены соответствующие модели [177; 183–186]. В частности, работа [187] вводит новую модель потерь при распространении радиоволны mmWave в городских сценариях, которая получается путем параметризации известных моделей ITU-RM.2135 «urban microcellular» («городские малые соты», UMi) для прямой видимости (LoS) и «Manhattan grid» для непрямой видимости (nLoS) на основе практических измерений.

Кроме того, детальное изучение точности различных моделей распространения, предложенных для систем mmWave, произведено с применением практических измерений в работе [188]. В статье [189] проведен анализ вероятности блокирования передачи mmWave с учетом более коротких длин волн, из которого следует, что относительно небольшие объекты, такие как автомобили, люди и элементы уличной фурнитуры, становятся препятствиями для распространения сигнала в условиях LoS. Повышенный уровень затухания при установлении соединений привел к необходимости разработки новых методов формирования диаграммы направленности антенн для оборудования mmWave [178; 190], что свою очередь потребовало создания соответствующих моделей интерференции [191–193].

В перспективных сетях mmWave особое место занимает применение массивных многоантенных систем MIMO, поскольку они позволяют достичь бо-

лее высоких скоростей передачи данных [194]. Ряд исследований протекает в этом направлении [195]. Например, в статье [196] предложен имеющий высокую производительность метод mmWave MIMO с погашением интерференции. В работе [197] описывается новый математический подход к анализу спектральной эффективности двухслойной гетерогенной сети с БС, поддерживающими технологию MIMO, а также с покрытием малыми сотами. Исследование [198] вводит способ формирования диаграммы направленности для mmWave-систем с множеством сот и обслуживаемых абонентских терминалов, а также осуществляет анализ спектральной и энергетической эффективности.

Теоретическое изучение сотовых сетей с возможностью установления соединений mmWave зачастую проводится методами стохастической геометрии. Подход в работе [199] предлагает математическую модель для анализа величины SINR и производных от нее характеристик системы связи при передаче как по линии вниз, так и по линии вверх. При этом предполагается множество равномерно распределенных на плоскости БС и учитываются основные особенности соединений mmWave, такие как возможность блокирования передачи объектами небольшого размера и использование антенн с острой диаграммой направленности. Для снижения вероятности отсутствия покрытия mmWave и повышения достижимой скорости передачи данных также разработана двухшаговая полнодуплексная схема с ретрансляцией, пригодная для применения в сетях 5G [200].

Большинство имеющихся исследований в области связи mmWave сходятся к заключению, что ввиду более коротких длин волн передача на крайне высоких частотах подвержена блокированию радиосигнала сравнительно небольшими объектами. Этот эффект приводит к падению уровня мощности на принимающей стороне и, тем самым, делает соединения mmWave менее надежными [201; 202]. Как следствие, уплотнение сотовой инфраструктуры статическими малыми сотами mmWave может быть дополнено подвижными узлами мобильной связи, что приводит к увеличению надежности установления соединения [203]. Для организации работы сети mmWave с повышенной надежностью возможно использование *одновременного подключения* или многосвязности (multi-connectivity) абонентского терминала, когда он соединен с двумя и более сотами одновременно [204; 205]. В статье [206] такая связь с возможностью одновременного подключения была изучена в приложении к облачным системам радиодоступа для повышения эффективности обслуживания терминалов, расположенных на краю соты.

С учетом имеющихся требований к сетям 5G [43] ожидается, что высоконадежные соединения mmWave станут неотъемлемым компонентом систем связи следующего поколения [207]. При этом, несмотря на множество результатов исследований отдельных особенностей связи на миллиметровых волнах, общесистемный анализ применения соединений mmWave в 5G в технической литературе практически отсутствует. Особый интерес в данном контексте представляет изучение динамического управления соединением mmWave с возможностью одновременного подключения к нескольким сотам для повышения надежности связи, подверженной случайному блокированию передачи подвижными объектами, например, телом человека.

3.4.2 Подход к моделированию связи mmWave

Реалистичное моделирование сотовой системы с возможностью установления соединений mmWave в динамике представляет собой сложную задачу [208], решение которой требует учета множества факторов. Это связано с существенным отличием связи mmWave от передачи данных в традиционных микроволновых сетях доступа и, в частности, со значительным влиянием разнообразных объектов на распространение радиосигнала mmWave. Для проведения соответствующего анализа в данном разделе предлагается новый подход, который включает в себя методы как аналитического, так и имитационного моделирования (см. рисунок 3.10). Он подразумевает создание детальной трехмерной карты зоны покрытия сети mmWave, содержащей всевозможные объекты, которые оказывают заметное влияние на распространение радиосигнала mmWave (т.е. с линейными размерами от нескольких сантиметров и выше). Также моделируется размещение БС малых сот mmWave, которое затем исследуется с учетом излучаемой мощности, несущей частоты, полосы пропускания и других важных параметров. Наконец, рассматриваются местоположения абонентских терминалов.

По завершении начального этапа моделирования, построенный в деталях трехмерный сценарий моделируется в имитаторе распространения радиосигнала, который специальным образом сконструирован для изучения особенностей передачи радиоволн на крайне высоких частотах и наблюдения реалистичной

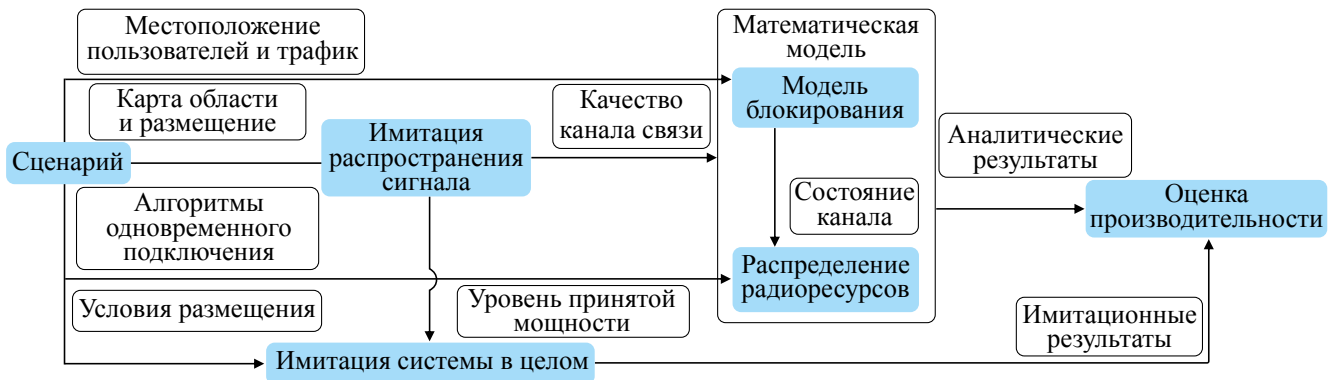


Рисунок 3.10 — Структура предлагаемого подхода к моделированию системы mmWave-связи

картины радиопокрытия внутри некоторой области моделирования. Имитатор был откалиброван с практическими измерениями из работы [35], а также поставлен с актуальными моделями 3GPP, приведенными в документе [177]. Используемый инструмент воссоздает различные эффекты, связанные с передачей миллиметровых волн в практических условиях, например, потери при распространении радиоволны, ее отражения и *диффузное рассеяние* лучей при столкновении с различными объектами для типового взаимного расположения БС и пользователя.

В качестве следующего шага моделирования результаты, полученные с помощью имитатора распространения радиосигнала, конвертируются в значения SNR. Затем последние преобразуются в соответствующие масштабирующие коэффициенты (в бит/с/Гц), которые назначаются абонентским терминалам в зависимости от состояния их канала связи. Такие масштабирующие коэффициенты задают нижнюю границу доли накладных расходов, которые передающие терминалы должны принять во внимание при пересчете своих требований по обслуживанию трафика на уровне приложений в фактическую скорость передачи данных в канале. На основании этих сведений область моделирования сегментируется на ряд непересекающихся зон обслуживания, в которых соединения mmWave получают различную скорость передачи.

Набор полученных таким образом зон, местоположений абонентов, их требований по качеству обслуживания, а также используемых алгоритмов, управляющих возможностью одновременного подключения, служит в качестве входных данных применяющегося далее аналитического подхода состоящего из двух

компонентов [209]. Первый из них представляет собой трехмерную модель блокирования передачи mmWave подвижными объектами, такими как тело человека, которая отслеживает события, связанные с ухудшением качества соединения из-за возникновения препятствий. Данная модель позволяет изучать динамику рассматриваемой системы, когда какое-либо соединение mmWave становится заблокированным или же разблокированным с течением времени. При этом учитываются состояния отдельных mmWave-соединений внутри каждой зоны обслуживания в заданные моменты времени.

Второй компонент применяемого математического подхода основан на использовании методов теории массового обслуживания и в некотором смысле объединяет результаты всех остальных шагов моделирования [210]. Он позволяет учесть различные характеристики передаваемого трафика, способы планирования радиоресурсов на БС mmWave, а также алгоритмы одновременного подключения и, тем самым, проанализировать основные показатели эффективности работы как сети связи в целом, так и отдельного абонентского терминала. Основными показателями эффективности, исследуемыми в данной главе, становятся, таким образом, вероятность блокировки (отказа) в обслуживании, вероятность разрыва соединения mmWave и среднее число активных абонентских сессий в системе.

Наконец, для проверки точности проведенного анализа приводятся данные имитационного моделирования системы в целом. Такое моделирование производится с помощью специализированного программного инструмента, который отражает особенности поведения сети mmWave при обслуживании пользователей с требуемой степенью детализации. В итоге результаты математического анализа подтверждаются данными имитационного моделирования, в процессе чего выявляются возможные несоответствия. Последующее изложение сосредоточивается на описании основных положений и результатов применения описанного здесь подхода к исследованию систем связи mmWave.

3.4.3 Динамика процесса блокирования передачи

Центральным звеном предложенного здесь подхода к анализу системы связи mmWave является изучение процесса блокирования канала прямой ви-

димости LoS при установлении соединения между абонентским терминалом и обслуживающей его БС mmWave [211]. Основные обозначения, используемые в рамках данного исследования, собраны в таблице 4. Рассматривается произвольным образом выбранная зона обслуживания U_i и предполагается, что P_i представляет собой ее геометрический центр. При этом интенсивность появления новых абонентов, требующих своего обслуживания в этой зоне, составляет λ_i , тогда как интенсивность возникновения блокирующих объектов (других абонентов, случайных прохожих и т.д., называемых ниже препятствиями) составляет λ_{iB} .

Далее случайным образом выбирается некоторый момент времени t . Графическая иллюстрация блокирования канала LoS между абонентским терминалом, расположенным в точке P_i , и БС mmWave, размещенной на расстоянии r от терминала, приведена на рисунке 3.11. Сигнал LoS, например, от БС mmWave может быть закрыт препятствием (телом человека), которое расположено внутри так называемой *зоны блокирования прямой видимости* LoS [211]. Длина такой зоны составляет

$$d = \frac{r(h_P - h_{P,T})}{h_T - h_{P,T}} + w_P/2, \quad (3.12)$$

где h_P – высота блокирующих объектов; $h_{P,T}$ – высота абонентского терминала; h_T – высота размещения БС mmWave, а r – расстояние между БС mmWave и терминалом.

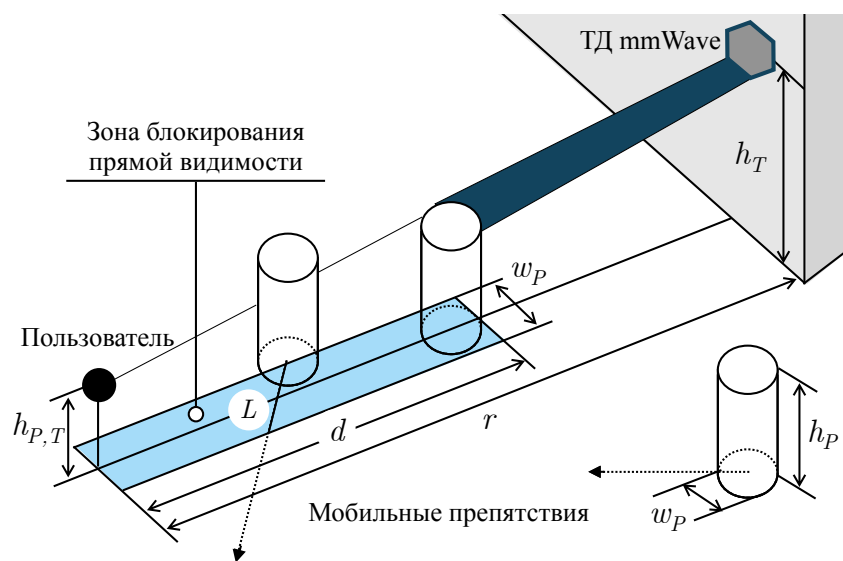


Рисунок 3.11 — Динамическое блокирование канала прямой видимости

Таблица 4 — Параметры моделирования связи на крайне высоких частотах

Обозн.	Описание
λ_i	Интенсивность абонентских терминалов в зоне i
λ_{iB}	Интенсивность блокирующих объектов в зоне i
w_P	Диаметр препятствия
h_P	Высота препятствия
$h_{P,T}$	Высота абонентского терминала
h_T	Высота размещения БС mmWave
r	Расстояние между БС mmWave и терминалом
d	Длина зоны, где препятствия могут блокировать канал LoS
S_B	Площадь зоны блокирования LoS
p_L	Вероятность сохранения канала LoS
λ_{iT}	Интенсивность препятствий, появляющихся в зоне блокирования
V	Скорость перемещения подвижных препятствий
ξ	Среднее время между моментами встреч двух объектов
$f(x,y)$	Стационарное распределение вероятностей для модели RDM
S_{U_i}	Площадь зоны i
$1/\lambda_{iT}$	Среднее время между последовательными моментами появления двух препятствий
L	Расстояние, пройденное препятствием в зоне блокирования

Важно отметить, что для реальных расстояний между абонентским терминалом и обслуживающей его БС mmWave площадь зоны блокирования LoS может быть приближенно представлена в виде прямоугольника [202]. Тогда LoS можно выразить как

$$S_B = w_P \left(\frac{r(h_P - h_{P,T})}{h_T - h_{P,T}} + w_P/2 \right), \quad (3.13)$$

где w_P – диаметр препятствия.

Для дальнейшего учета динамики процесса блокирования [212] прежде всего требуется определить интенсивность появления препятствий λ_{iT} , возникающих в зоне U_i с течением времени. Для получения аппроксимации λ_{iT} можно предположить, что как препятствия, так и абоненты перемещаются внутри U_i в соответствии с моделью случайного движения на плоскости RDM [213] и имеют постоянную скорость V . В работе [214] было показано, что интервалы времени между последовательными моментами встреч двух объектов, которые движутся в соответствии с RDM со случайными скоростями V_1 и V_2 внутри квадратной области со стороной L (и имеют круговые зоны покрытия с радиусом R), могут быть приближенно описаны экспоненциальным распределением с параметром

$$\zeta = 2RE[V] \int_0^L \int_0^L f^2(x,y) dx dy, \quad (3.14)$$

где $E[V]$ – относительная скорость двух объектов; $f(x,y)$ – стационарное распределение вероятностей для модели мобильности RDM, которое является равномерным, т.е. $f(x,y) = 1/L^2$ [214].

Рассматривая зону U_i произвольной формы и предполагая, что абонентский терминал остается неподвижным, тогда как препятствия могут перемещаться вокруг него, можно аппроксимировать время между последовательными моментами встреч внутри зоны блокирования LoS в виде

$$\zeta = \left(2w_P V + w_P \left(\frac{r(h_P - h_{P,T})}{h_T - h_{P,T}} + w_P \right) \right) \iint_{S_{U_i}} f^2(s) ds, \quad (3.15)$$

где V – средняя скорость перемещения препятствия; $f(s)$ – стационарное распределение вероятностей для модели мобильности RDM.

Основываясь на результатах из статьи [211], для анализа динамического блокирования можно применить альтернирующий процесс восстановления, согласно которому интервалы наличия и отсутствия сигнала LoS чередуются

друг с другом. Пусть B и A представляют собой случайные величины, обозначающие периоды времени, когда канал LoS становится заблокированным или же разблокированным соответственно. Поскольку препятствия в рассматриваемой зоне возникают согласно пуассоновскому процессу с интенсивностью λ_{iT} , то интервал времени A распределен экспоненциально с параметром λ_{iT} , т.е. $F_A(x) = 1 - e^{-\lambda_{iT}x}$, и имеет среднее значение $E[A] = 1/\lambda_{iT}$.

Кроме того, важно изучить распределение интервала времени B , который может состоять из нескольких перекрывающихся периодов, соответствующих блокированим, произведенным различными препятствиями. При этом интервал T , возникающий под воздействием некоторого определенного препятствия, называется здесь временем его пребывания в зоне блокирования LoS. Пусть тогда $F_T(x)$ и $F_B(x)$ обозначают функции распределения для интервала времени пребывания в зоне блокирования LoS и для совокупного интервала отсутствия канала LoS соответственно. Вид функции $F_B(x)$ является известным результатом [215] и приведен в выражении ниже, которое может быть решено численно для нахождения искомой функции распределения времени пребывания в зоне блокирования LoS $F_T(x)$.

$$F_B(x) = 1 - \left([1 - F_T(x)] \left(1 - \int_0^x (1 - F_B(x-z)) \exp(-\lambda_{iT}F_T(z)) \lambda_{iT} dz \right) + \int_0^x (1 - F_B(x-z)) |de^{-\lambda_{iT}F_T(z)}| \right), \quad x > 0. \quad (3.16)$$

Для завершения вычислений осталось описать функцию распределения для времени пребывания в зоне блокирования LoS $F_T(x)$. Вначале целесообразно выразить функцию распределения для расстояния, пройденного препятствием внутри зоны блокирования LoS, а затем учесть постоянную скорость его движения V . Таким образом, распределение расстояния L , пройденного препятствием внутри зоны блокирования LoS, может быть получено с использованием методики преобразования случайных величин следующим образом:

$$F_L(x) = \begin{cases} 0, & x \leq 0, \\ w_1 F_{L_1^*}(x) + w_2 F_{L_2^*}(x), & 0 < x \leq \sqrt{d^2 + w_P^2}, \\ 1, & x > \sqrt{d^2 + w_P^2}, \end{cases} \quad (3.17)$$

где веса w_1 и w_2 обозначают (1) вероятность события, что препятствие начинает свое движение от стороны с длиной d и заканчивает его у стороны с длиной w_P (или наоборот), а также (2) вероятность события, что препятствие начинает свое движение от стороны с длиной d и заканчивает его у стороны с длиной d соответственно. Эти вероятности могут быть получены в виде

$$w_1 = \frac{d^2 + 3dw_P}{d^2 + 3dw_P + 2w_P^2}, \quad w_2 = \frac{2w_P^2}{d^2 + 3dw_P + 2w_P^2}. \quad (3.18)$$

Тогда распределение времени блокирования может быть найдено путем линейного преобразования L/V , где V – постоянная скорость перемещения препятствий. Важно отметить, что плотность вероятности для линейного преобразования $Y = a + bX$ задается в виде

$$f_Y(y) = f_X[g^{-1}(y)] \left| \frac{dx}{dy} \right| = f_X \left(\frac{y - a}{b} \right) \frac{1}{|b|}. \quad (3.19)$$

В результате функции распределения представляют собой

$$F_{L_1^*}(x) = \begin{cases} 0, & x \leq 0; \\ \frac{\pi x^2}{4w_P d}, & 0 < x \leq \min(w_P, d); \\ \frac{1}{2w_P d} (\min(w_P, d) \sqrt{x^2 - \min(w_P, d)^2} + \\ + x^2 \arcsin(\frac{\min(w_P, d)}{x})), & \min(w_P, d) < x \leq \max(w_P, d); \\ \frac{1}{2w_P d} (\min(w_P, d) \sqrt{\max(w_P, d)^2 - \min(w_P, d)^2} + \\ + d(\sqrt{x^2 - d^2} - \sqrt{\max(w_P, d)^2 - d^2}) + \\ + w_P(\sqrt{x^2 - w_P^2} - \sqrt{\max(w_P, d)^2 - w_P^2}) + \\ + \max(w_P, d)^2 (\arccos(\frac{w_P}{\max(w_P, d)}) + \\ + \arcsin(\frac{\min(w_P, d)}{\max(w_P, d)}) - \arcsin(\frac{d}{\max(w_P, d)})) + \\ + x^2 (\arcsin(\frac{d}{x}) - \arccos(\frac{w_P}{x})), \\ \max(w_P, d) < x \leq \sqrt{d^2 + w_P^2}; \\ 1, & x > \sqrt{d^2 + w_P^2}, \end{cases}$$

$$F_{L_2^*}(x) = \begin{cases} 0, & x \leq d; \\ \frac{d^2 - x^2 + 2w_P \sqrt{x^2 - d^2}}{w_P^2}, & d < x \leq \sqrt{d^2 + w_P^2}; \\ 1, & x > \sqrt{d^2 + w_P^2}. \end{cases} \quad (3.20)$$

Таким образом, можно вычислить функцию плотности вероятности для случайной величины $T = L/V$ как

$$f_T(x) = f_L(Vx)V, \quad (3.21)$$

где $f_L(x) = F'_L(x)$.

Можно сделать вывод о том, что блокирование сигнала прямой видимости для абонентского терминала, расположенного в геометрическом центре U_i , представляет собой процесс восстановления, в котором длительность интервала, когда канал LoS становится разблокированным, распределена по экспоненциальному закону, а длительность интервала, когда канал LoS становится заблокированным, подчиняется общему распределению [211]. Применяя это знание в контексте использования систем массового обслуживания, представляется возможным охарактеризовать процесс планирования радиоресурсов в сети mmWave [209]. Для этого следует принять во внимание следующие вспомогательные допущения: (1) все терминалы внутри U_i наблюдают идентичные характеристики процесса блокирования и (2) распределение заблокированных интервалов возможно получить как описано ранее. С учетом данных допущений процесс поступления пользователей на обслуживание некоторой БС mmWave становится пуассоновским.

3.5 Перспективные сценарии применения связи на крайне высоких частотах

3.5.1 Моделирование распространения миллиметровых волн

Поскольку длина радиоволн при использовании связи mmWave составляет менее сантиметра, она оказывается меньше множества объектов в реальной среде распространения. Соответственно, передача миллиметровых волн отличается от излучения радиоволн на других частотах, которые применяются в современных сетях доступа. Например, сигнал mmWave подвержен более сильному затуханию, ослабляется атмосферой, а также испытывает значительные дифракционные потери по сравнению с микроволновым излучением [216; 217].

Несмотря на то, что основные свойства передачи mmWave-сигналов хорошо изучены, остается ряд нерешенных исследовательских задач, связанных с распространением миллиметровых волн в городских условиях. Таким образом, цель последующего изложения состоит в решении одной из таких задач, а именно, в изучении эффекта диффузного рассеяния для типового сценария применения связи на крайне высоких частотах.

В частности, представляет интерес тот вклад, который диффузное рассеяние привносит в общую мощность принятого сигнала, распространяющегося в каналах прямой (LoS) и не прямой (nLoS) видимости. Этот физический эффект является специфичным для связи mmWave и только лишь незначительно проявляется при передаче на более низких частотах. Для его исследования был создан специализированный имитатор распространения радиосигнала [218], который позволяет разложить мощность принятого сигнала mmWave на отдельные компоненты, связанные с влиянием различных физических эффектов, в том числе и диффузного рассеяния. Для этих целей пригодны методы имитации распространения луча, такие как «ray launching» (RL) и «ray tracing» (RT), позволяющие эффективно моделировать физические особенности сигналов mmWave. Основанные на этих методах имитаторы распространения радиосигнала позволяют детально изучать характеристики миллиметровых волн в различных условиях передачи.

Однако учет диффузного рассеяния при использовании методов имитации RL и RT считается сложной задачей [219]. Как следствие, данный эффект не поддерживается большинством современных программных инструментов, реализующих методы RL и RT. Вместо этого, как правило, предполагается, что влияние его незначительно, и он далее не рассматривается. Примером такого подхода является распространенный коммерческий имитатор Wireless InSite, который ограничивается моделированием зеркального отражения от шероховатой поверхности с последующим ослаблением сигнала [220]. Поэтому для получения приведенных ниже результатов был создан новый имитатор распространения радиосигнала, который реализует методы, предложенные в работе [221], усовершенствует их с точки зрения вычислительной сложности, а также внедряет поддержку эффекта диффузного рассеяния.

Основной вопрос, освещаемый в данном разделе, состоит в совместном изучении эффектов диффузного рассеяния и ослабления сигнала mmWave (как в канале LoS, так и в канале nLoS) при взаимодействии его с шероховатой

поверхностью в реалистичных условиях распространения. В частности, исследуется распределение мощности принятого сигнала для предсказания фактической зоны покрытия БС mmWave. Для этого рассматривается система связи, состоящая из одного передающего узла (БС mmWave, обозначаемая как TX) и большого количества принимающих узлов (абонентских терминалов, обозначаемых как RX). Работа такой системы моделируется при помощи имитатора распространения радиосигнала, реализующего метод RL, который является здесь предпочтительным по отношению к методу RT в силу экспоненциального роста вычислительной сложности последнего с увеличением числа принимающих узлов [222].

Для учета особенностей среды распространения сигнала mmWave в условиях городской застройки, далее предлагается стандартный способ размещения узлов сети, называемый «Manhattan grid» и изображенный на рисунке 3.12. Соответственно, большое число узлов RX располагается на и вокруг открытой площади в центре. Такая топология размещения является типовой для современных городских условий, в которых планируется использование средств связи mmWave, и позволяет сделать обоснованные выводы об ожидаемых характеристиках распространения радиосигнала. В частности, подобный способ расположения узлов был выбран при получении результатов моделирования в рамках крупного исследовательского проекта WINNER-II [223].

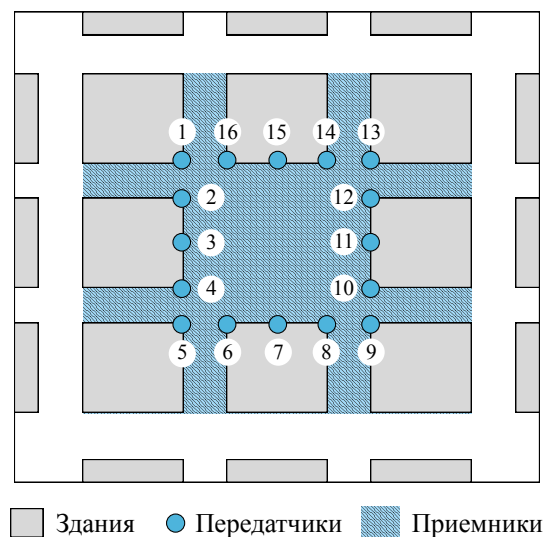


Рисунок 3.12 — Рассматриваемая модель городского сценария (вид сверху)

Поскольку в рассматриваемом сценарии узлы RX расположены на и вокруг центральной площади, далее предполагается, что узел TX, представляю-

щий собой БС mmWave, может быть установлен на углу здания в одной из 16 локаций на краях площади (отмеченных точками на рисунке 3.12). В процессе моделирования каждый узел TX и RX был представлен в виде изотропной сферы, взаимодействующей с синусоидальными волнами на частоте 28 ГГц. Мощность, излучаемая узлом TX, составляла 25 дБм. Остальные параметры передачи были выбраны так, чтобы исключить рассмотрение эффектов излучения ближнего поля. Кроме того, антенны как узла TX, так и узлов RX были поляризованы вертикально во избежание несовпадения поляризации.

Полученные таким образом результаты имитационного моделирования представлены на рисунке 3.13 отдельно для каждой из двух групп узлов RX (имеющих канал LoS или nLoS соответственно). Можно заметить, что эффект диффузного рассеяния приносит вклад в общую мощность принятого сигнала порядка 8 дБ для группы nLoS, подтверждая тем самым предположение о его важности в системе mmWave из-за рассеяния лучей шероховатой поверхностью. Следует отметить, что в рассматриваемом сценарии повышение общей мощности принятого сигнала на 8 дБ благодаря наличию диффузного рассеяния соответствует увеличению числа обслуживаемых узлов, принадлежащих группе nLoS, на 20%. Также на рисунке видно, что вклад компонента диффузного рассеяния в общую мощность сигнала, принятого группой узлов LoS, пренебрежимо мал, как это и ожидалось.

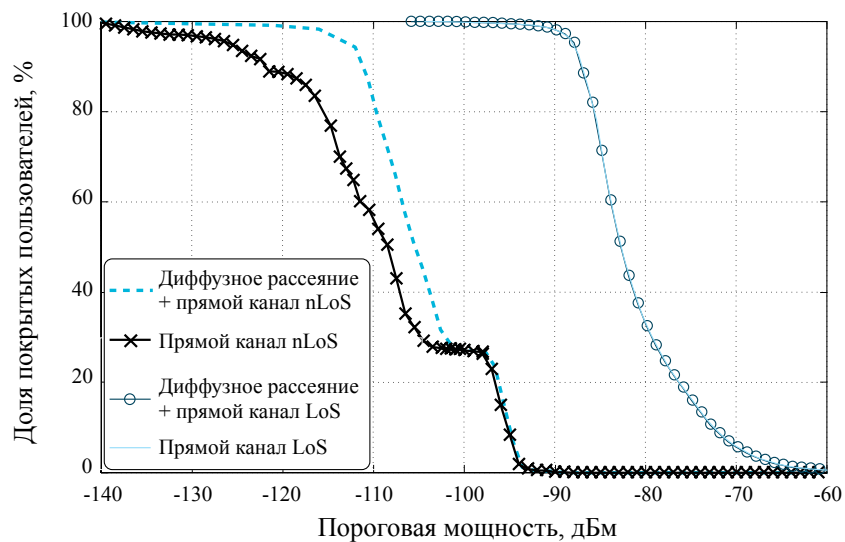


Рисунок 3.13 — Доля покрытых абонентских терминалов на частоте 28 ГГц

Детальный разбор графика на рисунке 3.13 показывает, что для группы узлов nLoS соответствующая функция распределения имеет некомпенсирован-

ное «плато» (т.е. когда не все отраженные лучи mmWave принимаются во внимание при сборе приведенной статистики). Это плато частично компенсируется при учете лучей, которые испытывают несколько отражений. Теоретически некомпенсированная часть данного плато может быть «выровнена» после рассмотрения всех лучей, в том числе и тех, которые получают большое количество отражений. Используемый имитатор распространения радиосигнала в принципе имеет возможность моделирования лучей, испытывающих множество отражений, но требуемое для этого время расчета неприемлемо велико для исследуемого здесь сценария.

3.5.2 Изучение реалистичного сценария использования mmWave

Для демонстрации предлагаемого подхода к моделированию соединений на крайне высоких частотах с возможностью одновременного подключения абонентского терминала к двум и более сотам в городских условиях далее проводится изучение характеристик работы сети mmWave. Данное исследование основывается на применении имитатора распространения радиосигнала, обсуждаемого в предыдущем подразделе, а также использует методы имитационного моделирования системы в целом при помощи инструмента, аналогичного описываемому в главе 2, и математического анализа [209]. Полученные таким образом результаты восполняют имеющийся в технической литературе пробел, связанный с недостатком сведений о функционировании системы mmWave в реалистичных сценариях, где пользователи имеют возможность одновременного подключения для увеличения надежности передачи своих данных (см. рисунок 3.14).

Для последующего моделирования выбираются типовые городские условия, в частности, площадь Таймс-сквер в Нью-Йорке, США. Обоснование такого выбора заключается в том, что Таймс-сквер представляет собой одну из наиболее часто посещаемых достопримечательностей и, таким образом, на ней и в ее окрестностях зачастую находится множество пользователей услуг мобильной связи. Рассматриваемая область моделирования (1) включает в себя обширную пешеходную зону посередине, позволяющую изучать эффекты, связанные с движением абонентов, на работу связи mmWave, а также (2) окружена множеством высотных зданий сложной формы. Кроме того, на прилежащих к

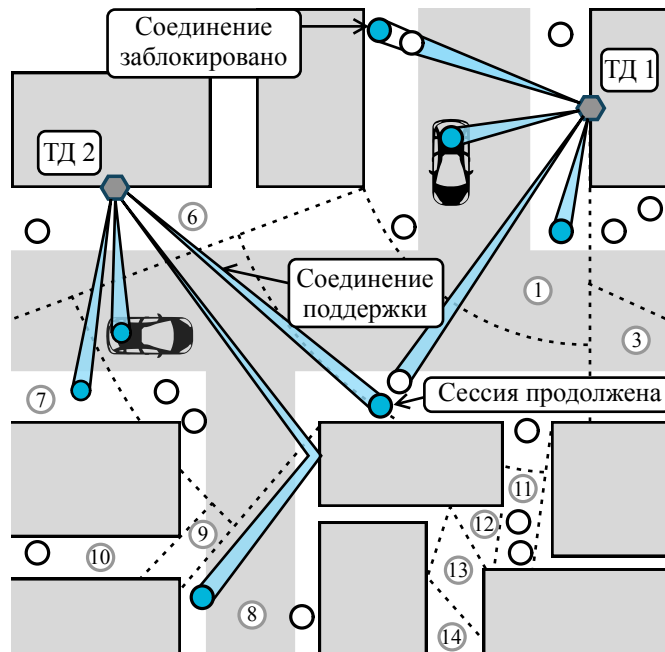


Рисунок 3.14 — Пример использования связи mmWave в реалистичном сценарии

ней улицах имеется большое количество автомобилей, что неоднозначным образом влияет на распространение радиосигнала mmWave и должно быть учтено при соответствующем исследовании.

Так как известно, что имитаторы распространения радиосигнала крайне чувствительны к точности входной информации [224], была воссоздана детальная карта рассматриваемой области моделирования с использованием реальных данных из открытого ресурса OSM. С помощью разработанного программного инструментария, включающего в себя модуль перевода координат объектов из геодезической системы координат в декартовскую (позволяющую проводить более точное моделирование), данные ресурса OSM были переработаны в точную трехмерную карту среды передачи [225]. Для параметризации созданной карты использовалась реальная статистическая информация о плотности транспортного потока, приведенная в отчете [226], а неподвижные объекты и элементы уличной мебели были добавлены при помощи данных системы Google Street View. Построенная таким образом модель сценария была представлена в виде, пригодном для чтения стандартным обработчиком трехмерных объектов, т.е. разбита на плоскости, ребра, ограничивающие рамки и т.д. (см. рисунок 3.15).

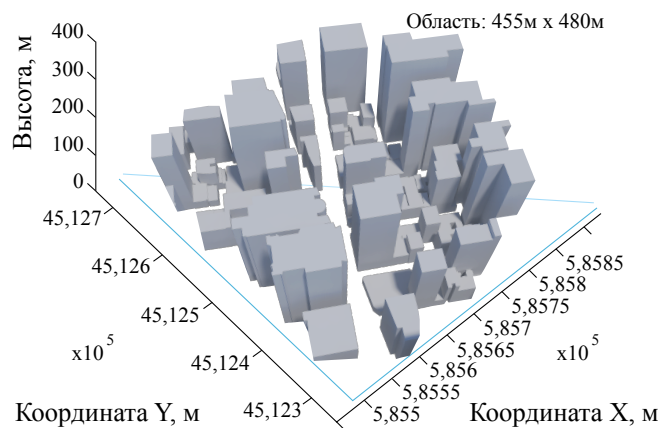


Рисунок 3.15 — Детализированная трехмерная модель области моделирования

На основании этих практических данных далее были рассмотрены возможные локации БС mmWave и абонентских терминалов. В отсутствие какой-либо информации о местонахождении узлов сети mmWave на практике, было выбрано несколько альтернативных конфигураций для размещения обслуживающих БС. В итоге оказалось отобрано 12 вариантов расположения БС mmWave (отмеченных треугольниками на рисунке 3.16(слева)), которые можно условно разделить на две группы: угловые расположения и расположения у стен. Предварительное исследование показало, что прочие местоположения БС не оказывают существенно различающегося влияния на характеристики работы сети по сравнению с уже рассмотренными двумя группами. Поскольку пользователи могут в принципе запросить обслуживание в произвольной точке области моделирования, возможные локации абонентских терминалов были представлены в виде сетки с шагом 1 м внутри всей пешеходной зоны сценария (отмечены точками на рисунке 3.16(слева)). Всего было рассмотрено 12x6031 пар БС–терминал, что составляет свыше 72 тысяч потенциальных соединений mmWave.

Последующее моделирование в имитаторе распространения радиосигнала проводилось для заданных локаций БС mmWave, изображенных на рисунке 3.16(слева). Вначале было получено распределение мощности принятого сигнала во всех точках на сетке размещения терминалов и для каждой из возможных локаций БС. Соответственно, рисунок 3.16(в центре) демонстрирует дополнительную функцию распределения для значения принятой мощности. Видно, что рассматриваемые местоположения различаются вплоть до 10 дБ, тогда как локации 2–5 и 9–11 являются наиболее предпочтительными. Данные результаты сопоставлялись с предполагаемыми практическими требованиями

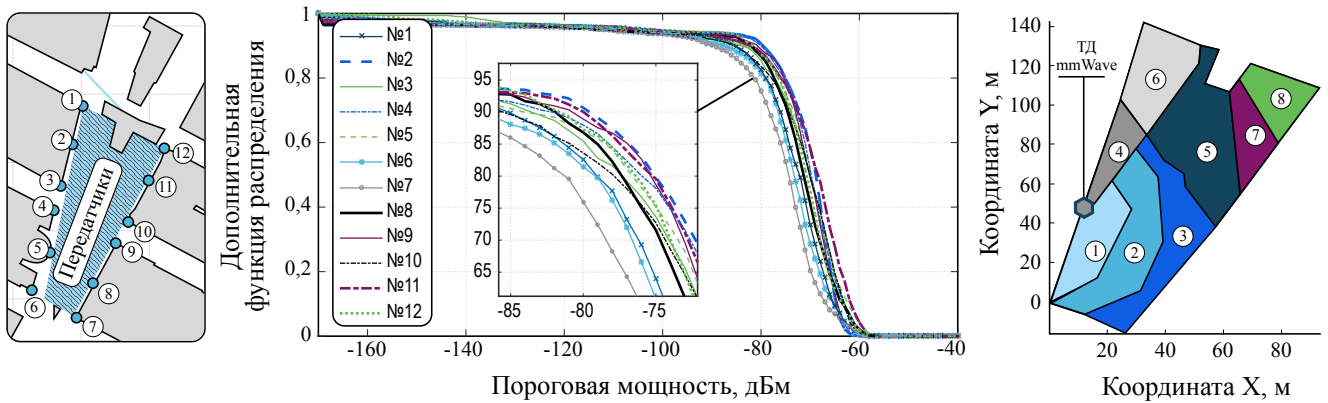


Рисунок 3.16 — Рассматриваемые локации БС (слева); дополнительная функция распределения для значения принятой мощности (в центре); пример выбора MCS для БС номер 5 в условиях nLoS (справа)

к размещению оборудования mmWave с точки зрения максимального расстояния между БС и далее выбирались местоположения под номерами 2, 5 и 11 в качестве тестового множества для проведения последующего изучения.

Каждой из таких отобранных локаций ставится в соответствие подробная карта мощности принятого сигнала как для связи в условиях LoS (при отсутствии блокирования прямой видимости), так и для условий nLoS (когда прямое соединение заблокировано телом человека). Полученные таким образом значения принятой мощности затем пересчитываются в величину SNR с учетом предположения, что уровень шума в канале связи равен -84 дБм и полоса его пропускания составляет 1 ГГц. На завершающем этапе имеющиеся значения SNR отображаются на практические модуляционно-кодовые схемы MCS с учетом фактического качества канала связи. Для наглядности здесь предполагаются общеизвестные схемы, основанные на мультиплексировании с ортогональным частотным разделением OFDM и используемые в радиотехнологии IEEE 802.11ad под номерами от 13 до 24.

В результате проведенных преобразований для каждого из шести режимов (три локации БС mmWave в условиях LoS и nLoS) исходная сетка размещения пользователей была разбита на области, соответствующие выбору абонентскими терминалами определенной MCS для передачи своих данных на БС mmWave. Соответственно, рисунок 3.16(справа) иллюстрирует пример такого выбора для БС номер 5 в условиях nLoS. Для большей наглядности результаты прошли фильтрацию, которая отсеяла области площадью менее 5 м^2 . Получен-

ные номера MCS в итоге были пересчитаны в значения фактической спектральной эффективности (в Гц на бит/с) для каждой из рассматриваемых областей, которые и были использованы в качестве входных данных анализа, проведенного в работе [209]. Результаты произведенного моделирования системы обсуждаются далее.

Для подтверждения результатов анализа сети связи на крайне высоких частотах использовалась среда имитационного моделирования системы в целом, которая включает в себя необходимые модули, задающие работу реалистичной технологии mmWave. Для реализации данного программного инструмента использовались рекомендации комитета 3GPP, а также была проведена предварительная калибровка с тестовыми данными из [227] для микроволновых сетей доступа. В целях интеграции радиотехнологии mmWave в состав данного комплекса имитационного моделирования были применены актуальные спецификации физического уровня, разработанные рабочей группой RAN 1 3GPP и собранные в документе [228].

3.5.3 Полученные результаты и их обсуждение

Параметры сценария и характеристики работы сети. Для изучения характеристик работы сети mmWave в реалистичных городских условиях следовало настроить рассматриваемый сценарий моделирования. В начале, с использованием фотоснимков из системы Google Street View, были получены относительные плотности абонентов и блокирующих соединений mmWave объектов, которые составили от 0,3 до 2 при мысленном движении от краев пешеходной области к центру рассматриваемой площади. Как следствие, средняя плотность препятствий внутри области моделирования составила $\lambda_B = 0,5$, что включает в себя $\lambda_{iB} = 3$ для участка в центре площади и $\lambda_{iB} = 0,2$ для участков на краю площади, смежных с прилегающими улицами.

Кроме того, было рассмотрено несколько типов передаваемого трафика на примере потока данных в 50 Мбит/с, соответствующего передаче видеoinформации с разрешением 4К, и в 10 Мбит/с, соответствующего видеозвонку высокого разрешения. Соотношение между этими двумя типами трафика внутри области моделирования было выбрано как 1:4, т.е. каждая вновь возника-

ющая абонентская сессия требовала для своего обслуживания 50 Мбит/с с вероятностью 0,2 или же 10 Мбит/с с вероятностью 0,8. Средняя длительность сессий обоих типов составляла 60 с. Активные абонентские терминалы, к которым поступали новые сессии, вначале устанавливали mmWave-соединение с БС, которая обеспечивала им максимальное значение мощности принятого сигнала.

В рамках рассмотренного выше сценария представляет интерес нахождение следующих характеристик работы системы связи mmWave.

- *Вероятность отказа в обслуживании.* Вероятность того, что вновь поступившей в систему абонентской сессии будет отказано в обслуживании со стороны некоторой БС mmWave (т.е. сессия не будет обслужена).
- *Вероятность разрыва сессии.* Вероятность того, что абонентская сессия, уже принятая на обслуживание некоторой БС mmWave какое-то время назад, была прервана до своего успешного завершения либо на текущей БС, либо на другой БС, куда эта сессия была перенаправлена алгоритмом управления одновременным подключением.
- *Среднее число активных сессий.* Среднее число активных абонентских сессий, которое обслуживается в некоторый момент времени всеми БС mmWave в системе.

Каждая из введенных выше вероятностей представляет самостоятельный интерес для изучения. В частности, прерывание текущей сессии (например, видеозвонка) непосредственно ухудшает качество восприятия услуги связи. При этом разрыв сессии в ряде случаев может быть даже менее желателен для пользователей, чем отказ в обслуживании, поскольку не ведет к формированию у них ожиданий, что устанавливаемая сессия будет успешно завершена. Описанные здесь характеристики функционирования сети mmWave исследуются далее для различных алгоритмов управления одновременным подключением абонентского терминала к двум и более сотам mmWave.

Алгоритмы управления одновременным подключением. Важно заметить, что блокирование соединения mmWave представляет собой случайное событие, которое требует немедленной реакции со стороны системы связи [229] во избежание разрыва активной сессии (например, видеозвонка высокого разрешения). Следовательно, необходимо управление одновременным подключением абонентского терминала к двум и более сотам mmWave для оперативного перенаправления данной сессии на другую БС при блокировании текущего соединения. Такое управление является новым функционалом сети доступа и

отличается от известных подходов, связанных с выгрузкой абонентского трафика на другие технологии радиодоступа, рассмотренные в начале данной главы. Ниже вводится в рассмотрение ряд соответствующих алгоритмов управления, которые предлагаются в порядке увеличения сложности своей практической реализации.

Далее рассматриваются исключительно алгоритмы управления одновременным подключением, реализуемые на абонентских терминалах, где решение относительно перенаправления текущей сессии принимается локально и, возможно, учитывает при этом дополнительную информацию сетевого содействия со стороны сотовой инфраструктуры. Алгоритмы, функционирующие на стороне сети доступа, являются значительно более сложными в реализации, и их работа может привести к дополнительным задержкам переключения, поскольку решения принимаются удаленно. Далее будет показано, что даже сравнительно простые алгоритмы на абонентских терминалах существенным образом улучшают характеристики функционирования системы связи mmWave. Это соображение может потребовать усилий со стороны комитета 3GPP по внедрению соответствующих механизмов для их поддержки.

Без перенаправления (Базовый, NORECON). Данный алгоритм является простейшим и выбирается здесь в качестве базового. Он не предполагает в системе mmWave возможности одновременного подключения абонентского терминала к двум и более сотам. Как следствие, терминал всегда ассоциирован с единственной обслуживающей его БС. Соответственно, абонентская сессия будет прервана, если у данной БС не имеется достаточного количества радиоресурсов для ее дальнейшего обслуживания.

Двойное подключение (DCON). Данный алгоритм нацелен на снижение доли времени, когда абонентский терминал остается ассоциирован с БС mmWave, соединение с которой заблокировано. В этой связи он аналогичен соответствующим разработкам комитета 3GPP по созданию схем двойного подключения для микроволновых систем доступа [205]. В случае перехода mmWave-соединения с текущей БС в заблокированное состояние обслуживание терминала перенаправляется на другую (резервную) БС, имеющую достаточно низкую вероятность блокирования соединения в текущей зоне нахождения абонента. Тем самым алгоритм DCON косвенно снижает вероятность разрыва сессии из-за недостаточного качества соединения, вызванного его блокированием. Для работы этого алгоритма на абонентском терминале необходима информация о том, какая из

потенциальных БС mmWave будет в наименьшей степени заблокирована для него в текущей зоне обслуживания.

Множественное подключение, учет блокирования (MCON, блок.). Данный алгоритм представляет собой расширение предыдущего подхода, направленное на снятие следующего ограничения, присущего алгоритму DCON: множество абонентских терминалов в некоторой зоне обслуживания может быть перенаправлено на одну и ту же БС, что приведет к исчерпанию имеющихся у нее свободных радиоресурсов. Напротив, алгоритм MCON, блок. перенаправляет терминалы, mmWave-соединение которых заблокировано, на другие БС случайным образом, но пропорционально вероятностям сохранения для них незаблокированного соединения с этими БС. Объем необходимой вспомогательной информации для организации работы этого алгоритма выше, чем в предыдущем случае, поскольку должна быть известна вероятность блокирования для каждой БС.

Множественное подключение, учет скорости передачи (MCON, скор.). Данный алгоритм дополнительно принимает во внимание, что пользовательские сессии не обязательно прерываются в случае блокирования. Действительно, текущая БС может продолжить обслуживание заблокированного соединения mmWave, но на более надежной и менее эффективной MCS. Таким образом, имеет смысл не избегать блокирования как такового, а максимизировать номер текущей MCS (в порядке увеличения эффективности) для снижения вероятности разрыва сессии. Согласно предлагаемому здесь алгоритму, в случае блокирования соединения с обслуживающей его БС mmWave абонентский терминал сравнивает текущий номер MCS, назначенный ему этой БС в условиях nLoS, с ожидаемыми номерами MCS, которые предлагают другие БС. После этого терминал выбирает ту БС, которая «обещает» ему обслуживание на самой высокой скорости и предпринимает попытку перенаправления своей сессии на нее. Такой алгоритм требует дополнительной информации на стороне абонентского терминала, которая включает в себя ожидаемые номера MCS от всех соседних БС.

Множественное подключение, учет блокирования и скорости передачи (MCON, блок. + скор.). Данный алгоритм представляет собой модифицированную версию предыдущего подхода, при работе которой абонентский терминал в заблокированном состоянии сначала в явном виде проверяет мощность принятого сигнала на соединениях со всеми возможными БС mmWave, а

затем выбирает то из них, которое соответствует максимальному *фактическому* номеру MCS. По сравнению с алгоритмом MCON, скор. такой способ управления снижает «степень» случайности, поскольку наилучшее в отношении скорости передачи данных mmWave-соединение может при этом иметь низкий номер MCS в условиях nLoS, а также бывает заблокировано в момент перенаправления. В то же время предлагаемый здесь алгоритм является более сложным при своей практической реализации на стороне абонентского терминала, так как он требует не только знания об актуальном состоянии системы mmWave, но также наличия дополнительных механизмов для выяснения качества канала связи при каждой попытке перенаправления.

Далее проводится сравнительный анализ характеристик работы описанных выше алгоритмов в рамках рассматриваемого городского сценария.

Исследование характеристик работы алгоритмов. В данном подразделе предложенные алгоритмы управления одновременным подключением изучаются при помощи методов имитационного моделирования, а также математического анализа из работы [209].

Вначале исследуется вероятность отказа в обслуживании в зависимости от интенсивности входного потока в плотной городской среде ($\lambda_B = 0,5$), см. рисунок 3.17. Для базового алгоритма NORECON можно сделать вывод о том, что существуют три отчетливо различимых участка. Первый соответствует интервалу от $\lambda_I = 0$ до $\lambda_I \approx 2,27$, на котором вероятность отказа в обслуживании сохраняется достаточно низкой, поскольку система слабо загружена (т.е. все еще может принимать на обслуживание новые сессии) вне зависимости от текущего состояния канала связи. Затем, в интервале от $\lambda_I \approx 2,27$ до $\lambda_I \approx 5,92$, производительность алгоритма NORECON значительно падает по мере того как количество активных сессий в системе растет, что снижает долю радиоресурсов, доступных для вновь прибывающих сессий. Наконец, на третьем интервале, начиная с $\lambda_I \approx 5,92$ и выше, большинство радиоресурсов оказывается занято. В этом случае вероятность отказа в обслуживании стремится к 1 при $\lambda_I \rightarrow \infty$.

Продолжая изучение алгоритмов управления одновременным подключением, следует отметить, что они демонстрируют схожее качественное поведение своих характеристик, но различаются количественно в отношении наибольшей поддерживаемой ими загрузки трафиком: от приблизительно 25% для простейшего DCON до примерно 55% для более сложного MCON, блок. + скор. (при

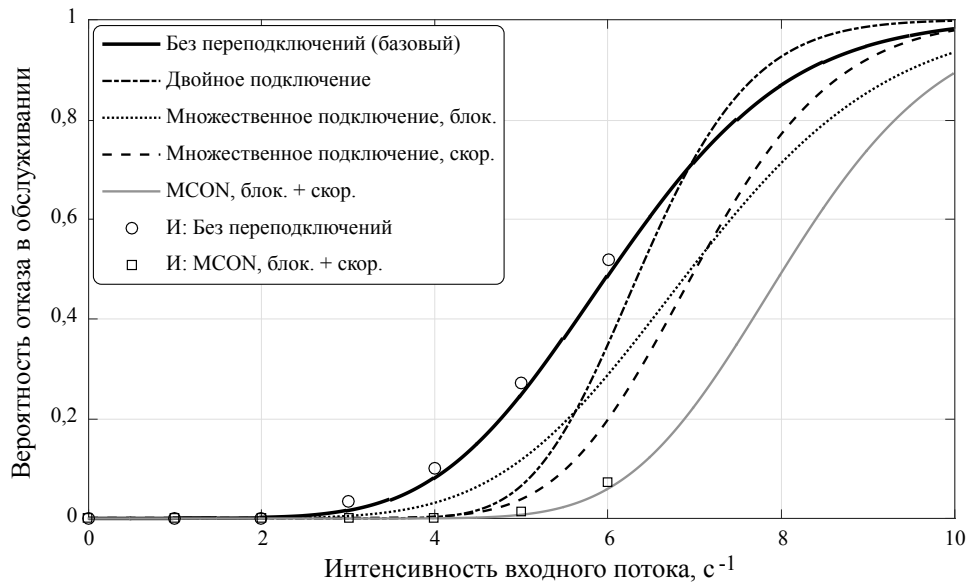


Рисунок 3.17 — Зависимость вероятности отказа в обслуживании от интенсивности входного потока

максимально допустимом уровне вероятности отказа в обслуживании, выбранном как 0,1). Особый интерес представляет участок, начиная с $\lambda_I \approx 7,02$, на котором алгоритм DCON показывает значения вероятности отказа выше, чем базовый алгоритм NORECON. Это можно объяснить тем соображением, что алгоритм DCON, стремящийся поддержать непрерывность соединений mmWave в среднем, в итоге принимает на обслуживание большее число соединений для того же значения интенсивности входного потока. Как следствие, доля доступных на БС mmWave радиоресурсов для приема новых сессий в таком случае снижается, и некоторым из них может быть отказано в обслуживании.

На рисунке 3.18 приведены значения вероятности разрыва сессии в зависимости от интенсивности входного потока в плотной городской среде ($\lambda_B = 0,5$). Тогда как вид наблюдаемых зависимостей схож с кривыми вероятности отказа в обслуживании, сами значения оказываются существенно ниже ввиду особенностей работы рассматриваемой системы: по мере роста интенсивности входного потока сеть mmWave вначале воздерживается от приема на обслуживание новых сессий и прерывает уже текущие только в том случае, когда они начинают требовать чрезмерного количества радиоресурсов из-за блокирования. Кроме того, вероятность разрыва сессии стабилизируется при достижении некоторого значения, которое зависит от плотности препятствий, характеристик пользовательского трафика и объема доступных радиоресурсов. Сессия, которая уже

принята на обслуживание, может быть прервана только в случае блокирования. Сравнивая различные алгоритмы управления одновременным подключением, видно, что они дают выигрыш от $\approx 15\%$ для MCON, блок. до $\approx 40\%$ для MCON, блок. + скор. с точки зрения значения наибольшей поддерживаемой загрузки.

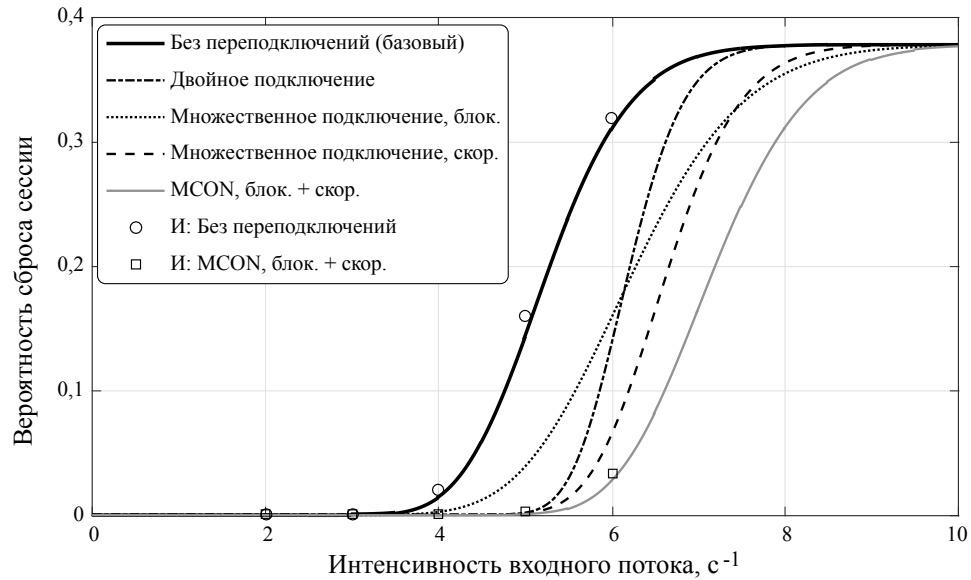


Рисунок 3.18 — Зависимость вероятности разрыва сессии от интенсивности входного потока

В продолжение анализа системы mmWave рассматривается ее функционирование в зависимости от плотности препятствий. Вероятность отказа в обслуживании λ_B на всех участках приводится на рисунке 3.19 для $\lambda_I = 5$. Можно заметить, что выбор алгоритма управления одновременным подключением оказывает существенное влияние на значение рассматриваемой вероятности. В частности, наблюдается выигрыш вплоть до $\approx 70\%$ для $\lambda_B \in [0,5; 1]$ (т.е. в плотной городской среде). При этом данный прирост падает практически до нуля при повышении плотности препятствий, которые блокируют большинство активных соединений mmWave. Однако на практике плотности, превышающие 3 чел/м², возникают крайне редко в типовых городских условиях. Следовательно, наибольшее значение вероятности отказа в обслуживании снижается и зависит от значения λ_I , а также от доли радиоресурсов, используемых системой в среднем (рассматриваемое здесь значение $\lambda_I = 5$ приблизительно соответствует 90% занятых радиоресурсов). Наконец, важно отметить, что алгоритм MCON, блок. оказывается более надежным не только при значительных интенсивностях входного потока, но и при высоких плотностях размещения препятствий.

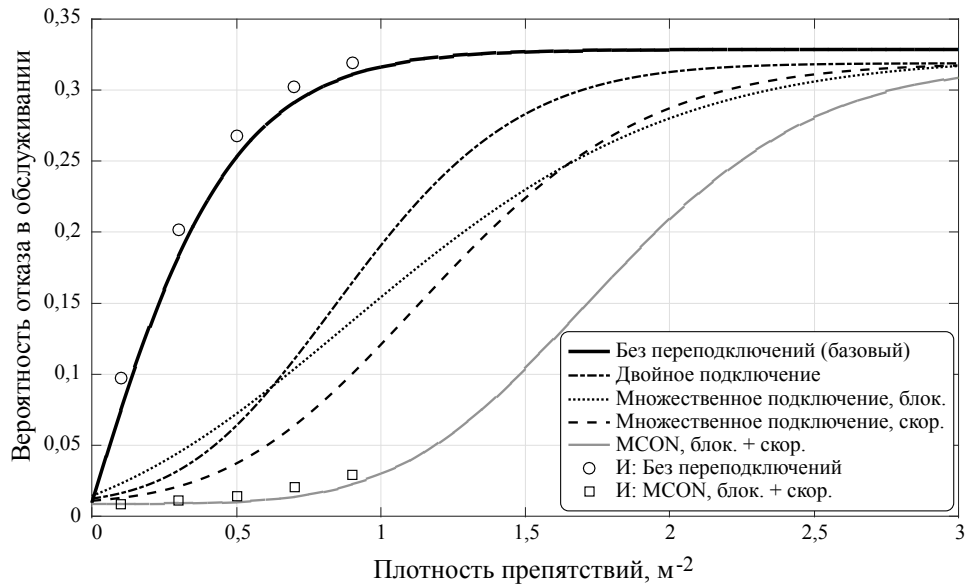


Рисунок 3.19 — Зависимость вероятности отказа в обслуживании от плотности препятствий

Углубляя данное исследование, полезно также понять изменение вероятности разрыва сессии в зависимости от плотности препятствий, что приведено на рисунке 3.20. Для изучения реакции системы, когда более 95% ее радиоресурсов заняты, предполагается более высокое значение интенсивности входного потока ($\lambda_I = 6$). Во-первых, видно, что для низких интенсивностей λ_B вероятность разрыва сессии близка к нулю, как и ожидалось. При этом ее значение возрастает вплоть до своего максимума при $\lambda_B \approx 0,5$, когда множество сессий, изначально принятых в условиях LoS, оказываются прерванными вследствие блокирования по мере того как заканчиваются свободные радиоресурсы, используемые соединениями в условиях nLoS. После этой точки система начинает принимать больше сессий в состоянии nLoS, поскольку возможность возникновения новой сессии в состоянии LoS в точности за моментом окончания обслуживания какой-либо из активных сессий (которая высвободит часть радиоресурсов) снижается с ростом плотности препятствий, становясь пренебрежимо малой после $\lambda_B \approx 2,8$.

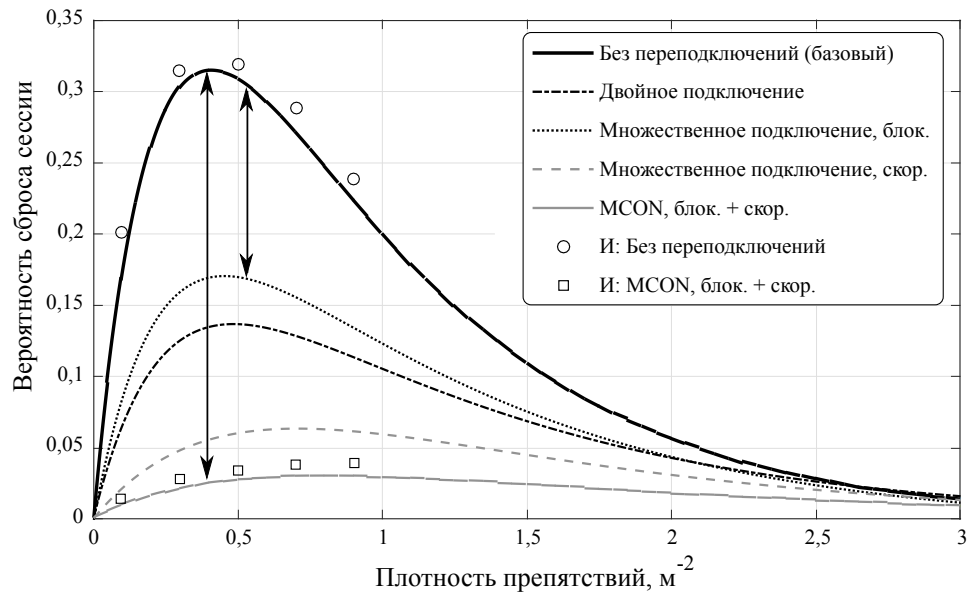


Рисунок 3.20 — Зависимость вероятности разрыва сессии от плотности препятствий

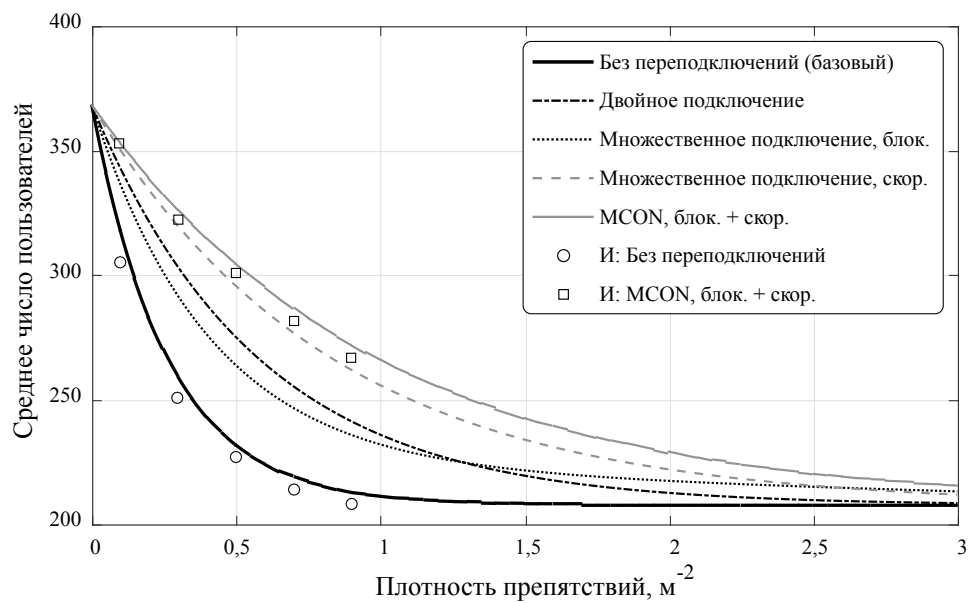


Рисунок 3.21 — Зависимость среднего числа активных сессий от плотности препятствий

При таких значениях интенсивности λ_B практически все принятые сессии сразу же оказываются в состоянии nLoS и обычно не возвращаются в состояние LoS до момента окончания своего обслуживания из-за частого блокирования. Снова можно отметить различие примерно 50% между алгоритмом NORECON и простыми алгоритмами управления одновременным подключением DCON и

MCON, блок., тогда как более сложные алгоритмы MCON, скор. и MCON, блок. + скор. оказываются в 7–10 раз лучше, чем базовый для $\lambda_B \approx 0,5$. Данное наблюдение позволяет сделать вывод, что использование алгоритмов управления одновременным подключением особенно полезно при высокой загрузке сети mmWave, когда при этом плотность препятствий остается средней. Эта ситуация приводит к частым прерываниям активных сессий ввиду множества переходов из состояния LoS в состояние nLoS. Поскольку текущие сессии терминалов без возможности одновременного подключения имеют достаточно высокую вероятность разрыва уже при первом переходе из состояния LoS в состояние nLoS, рассматриваемые здесь алгоритмы управления (в особенности MCON, скор. и MCON, блок. + скор.) предоставляют им возможность продолжить свое обслуживание в условиях LoS, улучшая таким образом качество пользовательского соединения.

Наконец, после изучения характеристик работы абонентских терминалов полезно рассмотреть производительность системы в целом, выраженную средним числом активных сессий в зависимости от плотности препятствий, что представлено на рисунке 3.21 для $\lambda_I = 6$. В случае работы алгоритма NORECON возможно наблюдать более чем 35%-ное сокращение среднего числа обслуживаемых сессий для $\lambda_B = 0,5$. При том же значении λ_B снижение числа активных сессий для более продвинутого алгоритма управления одновременным подключением MCON, блок. + скор. составляет уже 18%. Анализируя весь диапазон значений λ_B , можно заметить, что простой алгоритм DCON обслуживает до 50 дополнительных сессий по сравнению с базовым алгоритмом, тогда как наиболее продвинутый алгоритм MCON, блок. + скор. увеличивает это число еще на 70 сессий.

3.6 Выводы

В результате построенная модель гетерогенной сети с двумя одновременно функционирующими технологиями радиодоступа и возможностью установления прямых соединений между соседними устройствами при высокой плотности их размещения применима для оценки параметров качества обслуживания, включающих в себя среднее число одновременно передающих пользователей и

вероятность блокировки их сессий. Кроме того, сформулированный алгоритм с одновременным подключением пользователя к нескольким узлам сетевой инфраструктуры в условиях их плотного размещения, учитывающий блокирование канала прямой видимости препятствиями при передаче на крайне высоких частотах, применим для улучшения параметров качества обслуживания, таких как среднее число одновременно передающих пользователей и вероятность блокировки их сессий.

Глава 4. Поддержка приложений Интернета вещей и распространение носимых устройств

4.1 Предварительные замечания

На сегодняшний день Интернет вещей сделал возможной интеграцию разнообразных устройств, осуществляющих самые различные приложения и услуги. Поскольку множество разнородных технологий радиодоступа для обеспечения такого межмашинного взаимодействия M2M крайне широко, в данной главе вначале проводится актуальный обзор наиболее многообещающих решений с упором на особенности их применения и соответствующие характеристики производительности при организации связи M2M. Особое внимание уделяется сотовым системам M2M, которые используют технологию 3GPP LTE с необходимыми дополнениями. Кроме того, обсуждаются соответствующие наработки комитета IEEE 802.11 для обеспечения межмашинного взаимодействия, а также так называемые энергоэффективные сети дальнего радиуса действия.

Поскольку сотовые системы M2M являются наиболее перспективными для организации межмашинного взаимодействия, далее они изучаются более подробно в рамках типового сценария (например, множество сенсоров, счетчиков и измерителей в составе интеллектуальной энергосистемы). Предполагается, что большое число таких устройств стремится установить соединение с сетью практически одновременно для осуществления некоторого действия, что может вызвать всплеск нагрузки на канал связи. Поскольку такой канал управляется методами случайного доступа, в данной главе производится анализ возникающей при этом перегрузки с целью последующего ее контроля. Особенностью такого исследования становится рассмотрение различных параметров сети доступа с большим количеством устройств, включая энергетическую эффективность ее узлов.

Далее в этой главе рассматриваются особенности межмашинного взаимодействия для носимых устройств, которые в настоящее время активно развиваются и распространяются на практике. С учетом все возрастающего числа подключенных носимых устройств, ряд которых может, вообще говоря, порождать интенсивные потоки данных (например, очки дополненной или виртуаль-

ной реальности), представляет интерес изучение перспективных сценариев их использования. В частности, здесь рассматриваются условия работы носимого оборудования, размещенного с высокой плотностью, и проводится обзор современных решений беспроводной связи, пригодных для обслуживания подобных массовых сценариев. Можно сделать вывод о том, что большинство имеющихся радиотехнологий ближнего радиуса действия не способно обеспечить приемлемые характеристики функционирования.

Однако наиболее современные беспроводные решения, применяющие крайне высокие частоты (миллиметровые волны, или mmWave) для обеспечения связи, позволяют справиться с большими объемами передаваемого трафика в массовых сценариях. В данной главе производится моделирование mmWave-системы носимых устройств с целью повышения энергетической эффективности их доступа, которое учитывает форму диаграммы направленности антенны и интерференцию, возникающую между персональными сетями соседних пользователей. Кроме того, предлагается удобная для математического описания абстракция протокола mmWave-доступа, а также даются рекомендации по управлению параметрами такой системы. Результаты, полученные в этой главе, могут быть использованы для повышения энергетической эффективности работы практических радиотехнологий, таких как IEEE 802.11ad/ay.

4.2 Способы организации беспроводного доступа вещей в современных сетях

4.2.1 Особенности возникновения и развития Интернета вещей

Происхождение технологий М2М. По мнению экспертов, человечество находится на пороге новой промышленной революции, которая сулит колоссальные экономические и социальные блага, в результате повсеместного подключения к сети Интернет регионов, объектов и даже процессов. Центральная доктрина Интернета вещей предполагает соединение всевозможных окружающих нас сущностей, называемых «вещи», их обнаружение и именование, а также интеграцию с инфраструктурой сети Интернет [230]. Данная область знаний во-

брала в себя целый ряд различных сенсорных, телекоммуникационных, сетевых и информационных технологий, а также методов хранения и обработки данных и в настоящее время имеет значительный потенциал для создания технологических инноваций, использующихся множеством компаний для получения своего дохода.

Основополагающие технологии в рамках Интернета вещей базируются на концепции межмашинного взаимодействия M2M, предоставляя автономным и автоматическим устройствам (таким как сенсоры, силовые приводы, промышленные роботы, интеллектуальные измерители и т.д.) возможность взаимодействовать друг с другом, а также с сетью Интернет при минимальном вовлечении пользователя системы. По мере сокращения прибыли от традиционных голосовых услуг, операторы мобильной связи более заинтересованы в продвижении всевозможных M2M-сервисов. В этой связи технологии M2M позволяют устройству передавать имеющиеся у него сообщения через другое устройство или же непосредственно через сеть некоторому удаленному приложению, которое затем интерпретирует эти данные и может произвести какие-либо действия на их основе.

Ранние формы M2M-соединений относят к промышленным системам диспетчерского управления и сбора данных SCADA, которые активно использовались с 1980-х годов, но, по сути, тогда представляли собой изолированные и единичные решения. На пути своего дальнейшего развития технологии M2M абсорбировали методы радиочастотной идентификации RFID, а также последовавшие за ними средства организации беспроводных сенсорных сетей WSN. Предпосылками активного распространения систем M2M в недавние годы послужили более дешевая и доступная беспроводная связь, эффективные облачные системы хранения и обработки данных, а также большое разнообразие устройств различной производительности, стоимости и потребляемой мощности. Ожидается, что последующая эволюция M2M и IoT будет тесно связана с киберфизическими системами, которые объединяют вычислительные и физические процессы, а также с кибербиологическими системами, дополняющими сетевое взаимодействие информацией от живых организмов.

На сегодняшний день сервисы M2M включают в себя широкий спектр приложений – от носимых устройств для фитнеса и медицинских целей до автономных и беспилотных автомобилей. Это обуславливает бурный рост числа подключенных M2M-устройств с текущих 10 миллиардов до 24–50 миллиардов

в самом ближайшем будущем, что подтверждается прогнозами таких компаний и ассоциаций, как Cisco, Ericsson и GSMA. В ближайшее десятилетие ожидается, что предметы повседневного обихода, включая мебель, пищевые контейнеры и даже бумажные документы, смогут получить возможность доступа в сеть Интернет. Для обозначения такой предстоящей трансформации Интернета было предложено несколько новых понятий, таких как Интернет Всего (компания Cisco), соединяющий воедино людей, процессы, информацию и вещи, а также Индустриальный Интернет (компания General Electric и др.), упрощающий работу приложений, имеющих жизненно важное значение и связанных с промышленной автоматизацией, управлением производственными процессами и т.д.

Ретроспектива развития технологий M2M. Исторически сложившаяся на рынке приложений IoT фрагментация, подкрепленная длительной практикой создания единичных нишевых решений, существенно усложняет управление системами M2M, замедляет развитие инновационных продуктов в этой области и затрудняет контроль над соответствующими накладными расходами. Как следствие, в настоящее время IoT представляет собой чрезмерно разветвленную и сложную неоднородную экосистему, которая включает в себя всевозможные радиотехнологии M2M, такие как ZigBee, WirelessHART, 6LowPAN, WiFi, Bluetooth, ISA100.11a, MiWi, BACnet, Z-Wave и многие другие. Таким образом, создание универсальной беспроводной инфраструктуры становится все более актуальной задачей, решение которой требует пересмотра традиционных подходов к построению, размещению, обслуживанию и использованию существующих и перспективных систем M2M.

Тогда как различные проводные и беспроводные средства связи сыграли в свое время важную роль для конструирования IoT в его нынешнем виде, существенная доля имеющихся решений не оправдала изначально возложенных на них ожиданий, прежде всего, с точки зрения их практической пригодности. Типовым примером может служить радиоборудование ZigBee, построенное на основе стандарта IEEE 802.15.4 с последующими техническими поправками, принятыми IEEE и IETF. Относительно низкая мощность передачи оборудования ZigBee существенно ограничивает эффективную зону покрытия устройств и в итоге обуславливает построение сложной многоскачковой топологии сети при ее развертывании на обширной территории. Организация такого покрытия требует высокой плотности размещения сетевых шлюзов и ретрансляторов, которые, в свою очередь, также нуждаются в электропитании. Это приводит к

быстрому истощению их батарей, которые необходимо заменять каждые один-два года, а также к низкой надежности с длительными периодами отказа и к высокой стоимости обслуживания такой системы.

Недавние попытки комитетов и групп по стандартизации IEEE 802.15.4e, IETF RoLL и 6TiSCH преодолеть отмеченные выше ограничения позволили некоторым образом улучшить работу сети в многоскачковом режиме за счет введения более жесткой синхронизации узлов при одновременном контроле расхода их энергии. Эти меры помогают продлить срок службы батарей, но не влияют на радиус зоны покрытия оконечных устройств. Таким образом, относительно невысокая степень надежности рассматриваемой системы становится препятствием к ее использованию в ситуациях, чувствительных к продолжительным интервалам отказа. Несмотря на то, что данная технология все еще применяется на практике, например, в системах контроля, построенных на основе стандартов WirelessHART и ISA 100.11a, некоторые крупные компании уже отказались от дальнейшего производства устройств ZigBee.

В результате беспроводные технологии для построения IoT находятся в настоящий момент на перепутье и насчитывают множество соперничающих решений M2M, претендующих на первенство в этой нише. Основную конкуренцию традиционным ZigBee-подобным подходам составляют энергоэффективные сети дальнего радиуса действия LPWAN, такие как LoRaWAN и SIGFOX, системы WiFi HaLow (IEEE 802.11ah) малой мощности, а также различные дополнения к сотовым технологиям 4G и 5G, предназначенные для поддержки M2M-приложений. Такие относительно более современные подходы становятся существенно привлекательнее для развертывания сетей IoT в отношении как доступности, так и надежности радиосоединения. Они позволяют снять ограничения при сборе данных во всевозможных приложениях M2M и, тем самым, становятся более пригодными для обеспечения перспективных IoT-сервисов.

Особенности межмашинного взаимодействия. Как следует из вышеизложенного, приложения M2M в значительной степени повлияли на ландшафт современных систем беспроводной передачи данных и сейчас быстро становятся одной из основных движущих сил, направленных на изменение сложившегося рынка услуг мобильной связи. Однако по сравнению с традиционными способами обмена информацией между людьми H2H организация соединений M2M обладает рядом уникальных особенностей, среди которых следует отметить очень большое число подключенных к сети устройств, относительно малый

объем данных при каждой отправке и сравнительно нечастые попытки передачи [231]. Таким образом, особые характеристики межмашинного трафика порождают целый ряд новых задач в области проектирования и развертывания сетей M2M [232], оказывая при этом влияние на параметры системы, связанные с ее стоимостью, сложностью построения и энергетической эффективностью, а также на соответствующие требования по качеству обслуживания.

Действительно, несмотря на потенциально большое число устройств, они в основном передают короткие фрагменты данных с достаточно длительными паузами [233]. Результаты практических измерений показывают, что M2M-устройства являются в большей степени производителями, нежели потребителями информации, что приводит к перекосу их требований по радиоресурсам сети доступа в сторону передачи по линии вверх. Такое поведение соответствует их типовому применению, зачастую приводящему к всплеску активности в рабочие часы, поскольку большинство участников межмашинного взаимодействия используется в коммерческой сфере. Важно также заметить, что в целях экономии батареи оборудование M2M проявляет активность только лишь в течение достаточно коротких интервалов времени, что приводит к нечастым и, как правило, устойчивым к задержке сессиям. К примеру, число счетчиков воды, электричества и газа может быть крайне велико (вплоть до 35 тысяч штук на соту), но интервалы между отправками данных от каждого из них бывают значительными и допускается, что информация прибудет с некоторым запаздыванием.

Все эти специфические особенности межмашинного взаимодействия требуют создания подходящих средств радиосвязи. При этом различаются режимы передачи устройством M2M своих данных непосредственно в сеть доступа или же на промежуточный агрегирующий узел с последующей пересылкой в сеть [234]. Кроме того, соседние M2M-устройства могут воспользоваться прямыми каналами D2D для отправки своих данных, минуя таким образом необходимость вовлечения инфраструктуры оператора. Для сокращения накладных расходов на сбор данных от устройств M2M изначально предполагалось использовать нелицензированный спектр при доставке их трафика (например, частотные диапазоны ISM, SRD, PLMR или даже телевизионные каналы). В результате возникло множество закрытых и нестандартных решений для организации локального межмашинного взаимодействия, т.н. капиллярных систем M2M. Однако подобные подходы накладывают жесткие ограничения на максимальную

излучаемую мощность и длительность рабочего цикла, а также вынуждают устройства прослушивать радиоканал перед началом каждой своей передачи.

Альтернативным и более современным решением по организации межмашинного взаимодействия становится использование технологий мобильной связи, т.н. сотовых систем M2M, которые могут быть доступны как в виде наружного покрытия, так и внутри помещений. Применение сотовых решений для подключения устройств M2M привлекательно по причине широкого охвата областей, встроенных механизмов управления и обеспечения безопасности, а также емкой транспортной сети. Однако при этом множество существующих систем M2M по-прежнему применяет сотовые технологии GSM, GPRS и cdma2000 предыдущих поколений, которые могут быть существенно улучшены по показателям своей универсальности, стоимости, сложности и энергетической эффективности. С учетом большого разнообразия приложений M2M и их требований (см. рисунок 4.1) комитетом 3GPP проводится непрерывная доработка спецификаций системы LTE для более эффективной поддержки различных M2M-сервисов.

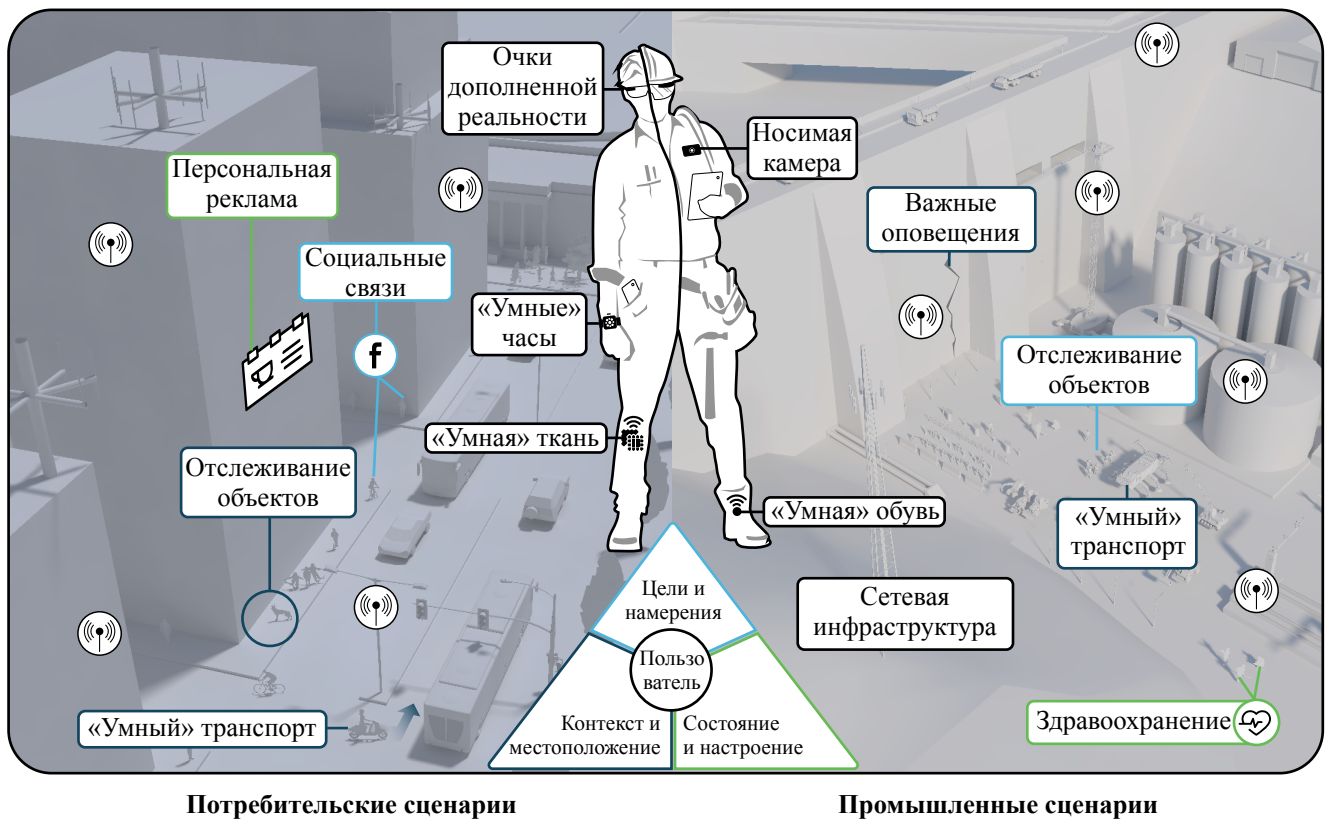


Рисунок 4.1 — Потребительские и промышленные применения M2M-устройств

По мере того как оборудование LTE становится все дешевле и проще в изготовлении, многие операторы услуг IoT рассматривают сотовые системы M2M в качестве наиболее перспективных, в основном благодаря более эффективному использованию ими радиоресурсов и высокой доступности сетевого покрытия. Таким образом, предпочтительный вариант развертывания сети представляет собой сопряжение капиллярной системы M2M с краем сотовой инфраструктуры и далее использование последней в качестве транспортной сети для передачи межмашинного трафика. При покрытии обширных территорий некоторые оконечные устройства могут быть подключены через капиллярный шлюз, тогда как другие соединяются с сотовой сетью напрямую, снижая тем самым сложность планирования системы. Дальнейшее изложение направлено на рассмотрение различных вариантов построения беспроводных соединений M2M в соответствии с современными стандартами радиосвязи.

4.2.2 Современные радиотехнологии для связи M2M

Поскольку Zigbee-подобные подходы заняли весьма ограниченную рыночную нишу и малопригодны для обеспечения работы перспективных M2M-приложений, исследовательский интерес в данной области сместился в сторону более эффективных решений [50]. Потребность в разработке новых технологий радиодоступа, которые позволили бы непосредственно учесть особенности межмашинного взаимодействия, стала очевидна уже в середине 2000-х годов. Низкая потребляемая мощность, доступная стоимость оборудования, широкая область охвата сети, а также возможность поддержки большого количества устройств с невысокой периодичностью отправки данных стали основными требованиями к современным системам M2M [235]. В этой связи создание подходящих технологических наработок (см. рисунок 4.2) было инициировано и продвигалось параллельно сразу в рамках нескольких комитетов по стандартизации, включая ETSI и 3GPP, а также группы IEEE 802.11 и 802.15. Текущие результаты работы некоторых из них обсуждаются далее.

Энергоэффективные сети дальнего радиуса действия. Сравнительно недавно возник новый класс беспроводных технологий для связи M2M, который принято называть энергоэффективной сетью дальнего радиуса дей-

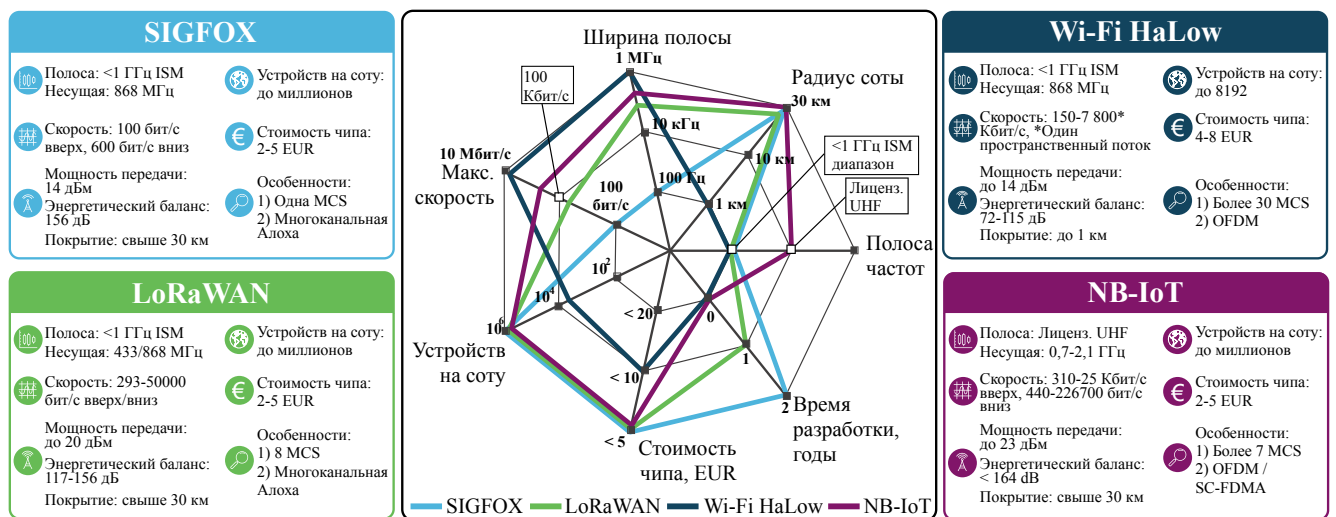


Рисунок 4.2 — Основные характеристики перспективных M2M-радиотехнологий

ствия, или LPWAN. Подобные решения реализованы на практике в нелицензированном спектре и поддерживают относительно невысокие скорости передачи данных небольшого объема [236], что, тем не менее, подходит для широкого класса M2M-приложений с низкой периодичностью отправки информации (см. рисунок 4.1). Многие решения LPWAN являются на сегодняшний день достаточно закрытыми и практически несовместимы между собой. Однако в последнее время наблюдаются попытки продвинуть некоторые компоненты LPWAN-технологий в комитет по стандартизации сотовых систем связи, например, через группу 3GPP GERAN. По аналогии с улучшениями, запланированными 3GPP в рамках эволюции сетей LTE, цель данных предложений состоит в расширении зоны покрытия существующих сетей M2M, снижении сложности построения окончательного оборудования и продлении времени жизни батарей.

На сегодняшний день наблюдается большое разнообразие всевозможных систем LPWAN, которое включает в себя стандартизованные технологии (IEEE 802.15.4g и 802.15.4k, ETSI LTN и т.д.), частично открытые протоколы, предложенные специальными рабочими группами и альянсами (к примеру, Weightless and LoRaWAN), а также закрытые нишевые решения (такие как SIGFOX, RPMA и WAVIoT). При таком множестве имеющихся вариантов подходы к разработке соответствующих технологических компонентов также значительно различаются. В качестве примера ниже рассматриваются два современных решения LPWAN, именуемые SIGFOX и LoRaWAN. Обе радиотехнологии ра-

ботают на частотах ниже 1 ГГц в нелицензированном спектральном диапазоне ISM (для промышленного, научного и медицинского применения) и используют метод доступа к каналу, построенный на принципах алгоритма Алоха с перестройкой по частоте, что приводит к ограничениям на длительность их рабочего цикла. Архитектурно обе системы имеют топологию, подобную сотовой.

Сеть SIGFOX оперирует с узкополосными сигналами ultra-narrowband 100 Гц, используя двоичную фазовую манипуляцию BPSK, а также 600 Гц, используя гауссовскую частотную манипуляцию GFSK, для передачи по линии вверх и вниз соответственно. Короткие пакеты данных отправляются в случайные моменты времени по произвольно выбранному каналу, что соответствует так называемому случайному множественному доступу с разделением по частоте и времени RFTDMA. Для обнаружения излученных радиосигналов оборудование БС имеет возможность прослушивать канал в достаточно широком диапазоне частот (от десятков до сотен кГц). Несмотря на то, что система SIGFOX накладывает жесткие ограничения на объем данных, передаваемых в каждом радиокадре (т.е. не более 12 байт), время отправки такого кадра может превышать 1 с. Из других узких мест данной технологии следует отметить наличие единственной модуляционно-кодовой схемы MCS, а также лимит на число пакетов данных, которые могут быть переданы на и от устройства M2M в день, равный 4 и 140 соответственно.

В отличие от технологии SIGFOX сеть LoRaWAN поддерживает множество MCS. При этом обязательные MCS базируются на собственной наработке LoRa, которая использует метод расширения спектра SS и основана на линейной частотной модуляции. Число соответствующих сигналов для отправки одного бита информации, а также полоса пропускания канала связи, которая влияет на длительность каждого сигнала, могут быть выбраны так, чтобы достичь требуемого соотношения между временем передачи и радиусом зоны покрытия. Кроме того, в некоторых географических регионах может применяться манипуляция GFSK со скоростью передачи данных в 50 кбит/с. Можно предположить, что рассмотренные выше технологии LPWAN будут полезны на относительно ранних стадиях развития экосистемы IoT, пока число подключенных к сети устройств еще достаточно мало. Однако при этом они могут сыграть важную роль для поддержки массовых приложений IoT, пока стандартизированные сотовые системы M2M окончательно не займут рынок и не получат достаточно широкое распространение.

Системы WiFi HaLow малой мощности. В недавние годы технология WiFi комитета IEEE 802.11 получила колоссальное развитие и фактически стала решением по умолчанию как для домашних, так и для корпоративных сетей связи. Однако системы на основе WiFi долгое время не применялись для приложений M2M в основном из-за своего достаточно высокого энергетического потребления. Сравнительно недавно данное обстоятельство изменилось, после того как поддерживающее стандартизацию IEEE 802.11 сообщество приняло решение наложить ограничения на длительность рабочего цикла сети, а также произвести модификацию радиооборудования WiFi в целях повышения энергетической эффективности системы. С другой стороны, поддержка мобильности и роуминга в рамках технологии WiFi остается достаточно скудной, а с точки зрения надежности связи отсутствуют как механизмы строгого гарантирования качества обслуживания, так и средства эффективного управления интерференцией на непредсказуемых нелицензированных частотах.

После соответствующего исследования стало понятно, что свойства распространения радиосигнала на частотах ниже 1 ГГц могут предоставить более благоприятные условия для установления соединений M2M по сравнению с традиционными подходами, использующими частотные диапазоны в районе 2,4 и 5 ГГц. Однако доступные в нелицензированном ISM-спектре на частотах ниже 1 ГГц радиоресурсы чрезвычайно невелики, что требует аккуратного и точного проектирования сети доступа. С учетом этого обстоятельства Комитетом по стандартизации IEEE 802.11 был запущен новый проект, который стартовал в 2010 году и получил обозначение IEEE 802.11ah. Целевая технология основана на модификации стандарта IEEE 802.11ac путем снижения тактовой частоты его работы в 10 раз и нацелена на предоставление более широкого охвата (до 1 км) для надежного обслуживания массовых приложений M2M с высокой спектральной и энергетической эффективностью.

На сегодняшний день в условиях городской застройки могут размещаться десятки тысяч устройств M2M, что приводит к необходимости поддерживать до 6000 узлов на каждой точке доступа IEEE 802.11ah. Достигая такой емкости, данная технология при этом не является обратно совместимой с другими протоколами линейки WiFi. В результате беспроводная сеть IEEE 802.11ah подходит для подключения продвинутых M2M-устройств, которые имеют более строгие требования по минимальной скорости передачи данных и задержке доступа. Технологически система IEEE 802.11ah представляет собой несовмести-

мое с другими WiFi-продуктами решение, функционирующее на частотах ниже 1 ГГц в нелицензированном спектральном диапазоне ISM и основанное на использовании мультиплексирования с ортогональным частотным разделением OFDM. Она поддерживает десятки различных настроек с учетом ширины полосы пропускания радиоканала, а также используемых модуляционно-кодовых схем MCS.

В настоящее время основная работа по созданию спецификаций технологии IEEE 802.11ah, также известной как WiFi HaLow малой мощности, фактически завершена, а тестовое оборудование, реализующее данный стандарт, уже начинает появляться на рынке. При этом одним из ключевых открытых вопросов становится будущий характер взаимоотношений между системой WiFi HaLow и другими продуктами линейки WiFi. В частности, представляет интерес, станут ли сети WiFi HaLow относительно независимыми или же будут интегрированы с традиционным оборудованием WiFi, например, в виде комбинированных устройств, как это было сделано в технологии Bluetooth Smart. К примеру, новые точки доступа WiFi могли бы поддерживать оба режима радиосвязи или пользовательские терминалы WiFi стали бы дополнительно обеспечивать функционал точки доступа HaLow для M2M-устройств.

Сотовые системы M2M. Применение сотовых технологий связи, разработанных комитетом 3GPP, становится все более привлекательным для организации массового доступа устройств M2M в сеть благодаря их обширной зоне обслуживания, сравнительно низкой стоимости развертывания, высокому уровню информационной безопасности, работе в контролируемом лицензированном спектре, а также относительной простоте и гибкости управления. Однако сети LTE изначально не были предназначены для эффективной поддержки M2M-приложений, поскольку исторически ориентировались на обеспечение услуг, сильно отличающихся от них как по характеристикам передаваемого трафика [237], так и по требованиям к качеству обслуживания [238]. Таким образом, комитету 3GPP пришлось значительно улучшить технологию LTE, чтобы позволить ей эффективно обслуживать всевозможные устройства M2M.

Учитывая тот факт, что число участников межмашинного взаимодействия продолжает стремительно возрастать, в состав спецификаций LTE потребовалось включить целый ряд специализированных механизмов для поддержки очень большого числа устройств [239]. На протяжении многих лет эффективное обеспечение межмашинного взаимодействия стало для различных рабочих

групп 3GPP предметом непрерывной деятельности. Уже с начала 2010-х годов были предприняты первые попытки по разработке новой технологии, широко известной под названием LTE-MTC, или LTE-M. Соответственно, в первую очередь был предложен метод управления перегрузкой, именуемый усовершенствованным запретом на доступ ЕАВ, который стал частью 3GPP Release 11 и позволил избежать ситуаций, когда сеть доступа отказывает из-за всплеска загрузки, связанного с одновременным подключением большого числа M2M-устройств.

Продолжая улучшать характеристики M2M-взаимодействия в отношении стоимости оборудования и его энергетического потребления, а также стремясь повысить фактическую дальность обслуживания и емкость сети, комитет 3GPP предпринял ряд решительных действий. Во-первых, были предложены более простые управляющие протоколы для подключения устройств M2M, которые позволили уменьшить их энергетические затраты, когда данные передаются нечасто и малыми порциями. В этой связи устройствам, отправляющим данные с длительными паузами, была также предоставлена возможность удлинять периоды ожидания в режиме непостоянного приема DRX, что привело к дальнейшему снижению энергетического потребления. Затем в 3GPP Release 12 был введен новый класс упрощенных терминалов Cat. 0, которые используют полосу пропускания в 20 МГц и получают предельную скорость передачи данных в 1 Мбит/с. Наконец, в состав 3GPP Release 13 был включен дополнительный класс Cat. M1, который имеет доступ к полосе пропускания в 1,4 МГц с ограниченной мощностью передачи, достигая при этом значительного прироста в максимальной дальности соединения.

Несмотря на существенное увеличение числа регионов, в которых уже развернуты сети LTE, множество областей по-прежнему использует системы предыдущих поколений, основанные на технологии GSM. В таком случае сотовые технологии M2M вынуждены использовать режим GPRS/EDGE для организации межмашинного взаимодействия. Поскольку ожидается, что сети GSM продолжат играть значительную роль в развитии экосистемы IoT ввиду своей высокой распространенности и низких издержек, комитет 3GPP предпринял попытку адаптировать систему GSM для нужд приложений IoT с целью предложить более эффективную альтернативу передаче M2M-трафика посредством GPRS. Возникшая в результате технология, известная как GSM с широким охватом EC-GSM, вводит поверх архитектуры GSM набор как управляющих каналов, так и каналов для передачи данных, а также включает в себя про-

двинутые механизмы передачи с повторениями и методы комбинирования сигналов. Как следствие, сети EC-GSM потенциально могут поддерживать до 50 тысяч устройств на соту, причем качество сигнала на частотах ниже 1 ГГц улучшается до 20 дБ.

Наконец, с 2015 года комитетом 3GPP началась работа по созданию новой сотовой системы M2M, которая стала известна как технология узкополосного IoT (NB-IoT). Соответствующий класс терминалов Cat. NB1 определяет устройства, совместимые с современными сотовыми системами и имеющие сложность порядка 10% той, которой обладают терминалы класса Cat. 1. Это стало возможно благодаря сужению полосы пропускания до 180 кГц, использованию схемы OFDMA с интервалом между поднесущими 15 кГц для передачи по линии вниз, а также схемы FDMA с интервалом между поднесущими 15 кГц и схемы SC-FDMA с интервалом между поднесущими 3,75 или 15 кГц для передачи по линии вверх. Технология NB-IoT может предложить три варианта развертывания сети M2M: с отдельным размещением, внутри полосы частот LTE и внутри защитной полосы частот LTE. Размер блока данных для отправки как по линии вверх, так и по линии вниз был сокращен, и при этом были упрощены многие управляющие протоколы. Стандартизация первой версии протокола NB-IoT завершилась в 2016 году, когда он был включен в состав 3GPP Release 13. После этого в рамках 3GPP Release 14 были введены еще два класса терминалов, Cat. M2 и Cat. NB2, с улучшениями для достижения дополнительной гибкости и производительности.

4.3 Анализ начального доступа в сотовых системах M2M

4.3.1 Характеристики работы канала случайного доступа в LTE

Организация соединения M2M-устройств с сетью. Для установления начального соединения с сетью LTE устройства M2M, такие как сенсоры, счетчики и измерители, могут использовать канал случайного доступа RACH [240]. Поскольку протокол функционирования данного канала исходно не учитывал особенности межмашинного взаимодействия, представляет инте-

рес исследование параметров его работы в случаях, когда сеть доступа подвержена всплескам загрузки, связанным с единовременным подключением большого числа M2M-устройств [241]. При этом конкуренция между множеством терминалов за возможность передачи в режиме перегрузки делает такой анализ нетривиальным [13]. Также важно сравнить между собой характеристики работы канала RACH в условиях перегрузки и в режиме нормальной работы, когда значительного всплеска загрузки не происходит [242].

Далее рассматривается типовой сценарий, соответствующий работе интеллектуальной энергосистемы, в котором большое число устройств M2M устанавливает соединение с сетью LTE в течение достаточно короткого интервала времени, например, для отправки тревожного сообщения [243] об изменении режима электропитания. Поскольку счетчики, сенсоры и измерители представляют собой устройства небольшого размера, питающиеся от батареи, учет их энергетического потребления крайне важен [244; 245]. В этой связи дальнейшее изложение направлено на изучение энергетической эффективности и других характеристик M2M-устройств при их массовом подключении к сети LTE в условиях перегрузки канала RACH. Данное исследование проводится на основании соответствующих технических документов комитета 3GPP, в которых заданы параметры рассматриваемого здесь сценария.

Процедура случайного доступа в системе LTE приведена на рисунке 4.3. Вначале терминал посылает на БС LTE сообщение Msg 1, называемое *преамбулой*, используя при этом физический канал случайного доступа PRACH. При этом конкретная преамбула для передачи выбирается случайным образом из множества предварительно зарезервированных для этой цели последовательностей [246]. Следует отметить, что число доступных для организации случайного доступа последовательностей может быть меньше, чем наибольшее возможное их число в 64 преамбулы, в зависимости от настроек сети доступа. Если два терминала или более выбрали идентичные преамбулы, а затем отправили их в один и тот же момент времени, на стороне БС может возникнуть *коллизия*. Преамбулы также могут не достичь БС при недостаточном значении излучаемой терминалом мощности.

Если преамбула была принята успешно, то БС LTE (называемая также eNodeB) подтверждает этот факт путем отправки в соответствующем интервале сообщения Msg 2, известного как ответ случайного доступа RAR. Для этого терминалу высылаются указатели на радиоресурсы физического коллективно-



Рисунок 4.3 — Порядок процедуры случайного доступа в системе LTE

го канала по линии вниз PDSCH, которые содержат его сообщение RAR [247]. Поскольку eNodeB должна установить, какой из терминалов передал какую преамбулу, инициируется так называемый процесс разрешения коллизии. По истечении некоторого времени обработки сообщения RAR участвующие терминалы передают сообщение Msg 3, которое представляет собой запрос на установление контроля за радиоресурсами RRC. Его отправка происходит в физическом коллективном канале по линии вверх PUSCH с использованием радиоресурсов, указанных в сообщении Msg 2. Процедура случайного доступа завершается после успешного приема от eNodeB сообщения Msg 4, которое называется настройкой соединения RRC.

Когда схожие сообщения Msg 3 отправляются более чем одним терминалом из-за предшествующей коллизии их сообщений Msg 1 (преамбул), на стороне eNodeB в лучшем случае может быть принято только одно из них, что приводит к отсылке соответствующего сообщения Msg 4 успешному терминалу. В противном случае, если терминал не получил ожидаемого им сообщения Msg 4, он запускает процедуру случайного доступа заново после некоторого интервала отсрочки, выбираемого случайным образом в рамках определенного окна, задаваемого значением индикатора отсрочки. Узкое место рассматриваемого протокола, особенно при наличии множества требующих одновременного доступа терминалов, связано с ростом вероятности возникновения коллизии (см. рисунок 4.4(a)). Однако доставка сообщения RAR внутри выделенного интервала ожидания также может потерпеть неудачу в силу ограниченных радио-

ресурсов канала PDCCH. Более того, возможно, что сообщения Msg 3 и Msg 4 с некоторой вероятностью искажаются при передаче.

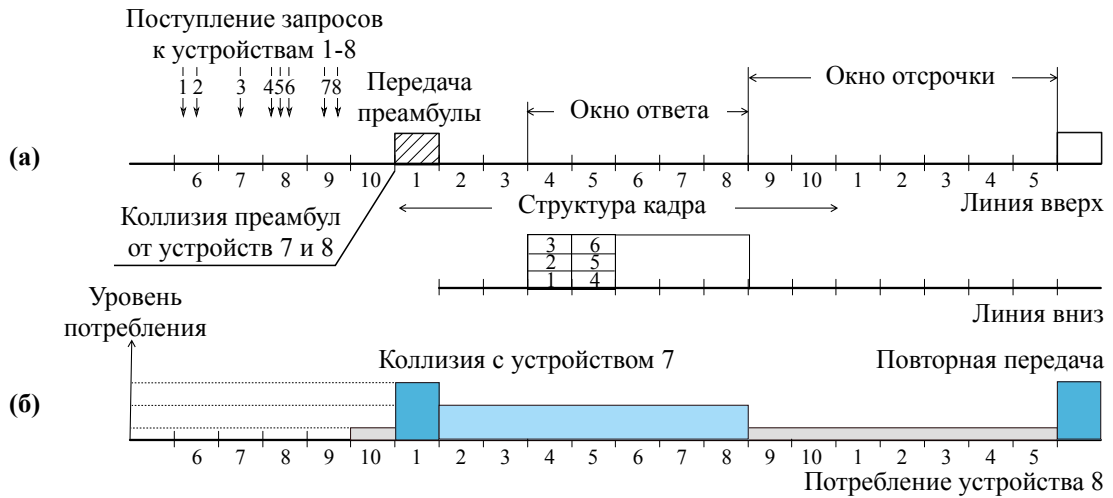


Рисунок 4.4 — Пример работы процедуры случайного доступа во времени (а) и пример энергопотребления для терминала номер 8 (б)

В результате целый ряд негативных событий может стать причиной неудачи при работе процедуры случайного доступа и тем самым вызвать дополнительную задержку начального входа в сеть. Очевидно, что вероятность коллизии возрастает с увеличением числа терминалов, участвующих в установлении соединения с БС, а также зависит от параметров их трафика. В условиях перегрузки RACH число конкурирующих за выход в этот канал устройств может достигать 30 тысяч штук на одну соту, что соответствует оценке [248], которая была использована в [249] и в итоге заимствована комитетом 3GPP в документе [250] в качестве рекомендованной максимальной плотности размещения терминалов. Такое огромное число конкурирующих M2M-устройств может привести к очень высокой вероятности коллизии и даже повлечь за собой исчерпание радиоресурсов сети доступа. В ответ на эту угрозу комитет 3GPP инициировал работу по оценке причин и последствий подобной перегрузки, анализ обстоятельств которой приведен ниже.

Моделирование перегрузок в сети доступа. В рассматриваемом типовом сценарии сота системы связи 3GPP LTE содержит M идентичных M2M-устройств. Каждое из них выбирает подкадр для начала процесса передачи по линии вверх согласно равномерному распределению (тип трафика 1) или бета-распределению (тип трафика 2) на интервале $[1; N]$. Передача преамбулы, занимающей в точности один подкадр, может осуществляться в каждом

подкадре с номером b , т.е. в слотах $1, b+1, \dots, b \cdot i + 1, i \in Z^+$. После активации устройство M2M считается занятым до момента завершения своей процедуры случайного доступа. В противном случае данное устройство называется неактивным. В тех подкадрах, где происходит обслуживание (т.е. когда запланирована возможность передачи PRACH), каждое занятое устройство выбирает одну из s преамбул случайным образом по равномерному закону и отправляет ее.

В соответствии с рекомендациями [250] далее подразумевается, что коллизия наступает, когда два или более устройства M2M выбирают идентичную преамбулу и передают ее в одном и том же слоте. При этом предполагается, что все вступившие в конфликт преамбулы будут потеряны на стороне БС по истечении некоторого времени обработки, что на практике соответствует худшей возможной ситуации. В противном случае отправка преамбулы считается успешной с вероятностью $1 - e^{-i}$ в силу постепенного повышения излучаемой терминалом мощности, где i – число попыток передачи [250]. Наибольшее допустимое число попыток передачи одной преамбулы составляет L_1 . Если ее отправка становится неуспешной из-за коллизии или же недостаточного уровня мощности, соответствующее M2M-устройство равномерным образом выбирает себе счетчик отсрочки из окна W . После K_0 подкадров ожидания начинается так называемый интервал ответа, имеющий размер K (см. рисунок 4.4).

Внутри данного интервала ответа eNodeB высылает терминалам сообщения RAR в подкадрах, выбранных равномерно из диапазона $[1; K]$. Если устройство M2M не получает сообщение RAR, то его текущая попытка отправки преамбулы объявляется неуспешной, и такой терминал инициирует процедуру отсрочки последующей передачи. В случае успешного получения M2M-устройством сообщения RAR данный терминал начинает формирование сообщения Msg 3 для отсылки в течение времени t_{pr} . Затем он отправляет сообщение Msg 3 и ожидает в течение времени $t_{tx} - 1$ приема сообщения Msg 4 (см. рисунок 4.3). В свою очередь, сообщения Msg 3 и Msg 4 доставляются успешно с вероятностями π_3 и π_4 соответственно. Наибольшее допустимое число попыток передачи сообщения Msg 3 составляет L_3 .

Анализ введенной выше модели системы связи с возможностью возникновения коллизий представляет собой непростую задачу, точное решение которой затруднено в силу «памяти» процесса. Это связано с наличием переменного времени отсрочки передачи, постоянных интервалов ожидания и, в особенности,

значительного числа преамбул в процедуре случайного доступа. Например, в классической системе случайного множественного доступа с единственной преамбулой [251] задержку входа в сеть можно оценить с помощью подхода, используемого ранее для анализа алгоритма Алоха [252]. Однако в изучаемой здесь системе этот популярный метод не дает точного приближения, поэтому далее применяется расширение техники из работы [253]. Для того чтобы абстрагироваться от наличия «памяти» и получить оценку величины $E[\tau^{(1)}]$ для рассматриваемой системы с коллизиями, вводится следующая эквивалентная модель.

- (1) Предполагается, что M2M-устройства активизируются согласно распределению Бернулли с вероятностью π . Тогда некоторый терминал может инициировать свое начальное соединение с сетью в каждом подкадре с соответствующей вероятностью $\pi = 1/N$, где N – число подкадров в исходной системе.
- (2) Далее опускается непосредственное рассмотрение интервала ожидания и окна отсрочки, которое заменяется предположением о том, что в каждом подкадре занятое устройство может активизироваться с некоторой вероятностью $\pi_0 = 1/(K_0 + K_1 + K + \bar{w})$. Данное соображение означает, что терминал активизируется единожды в течение интервала $(K_1 + K_0 + K + \bar{w})$, если его исходная попытка передачи оказалась неуспешной из-за коллизии или же недостаточного уровня мощности.
- (3) Вероятность успешного завершения процедуры начального доступа обозначается как μ , т.е. запрос на вход в сеть обслуживается в текущем подкадре с некоторой вероятностью μ . В противном случае терминал предпринимает попытку получения доступа к каналу в следующем доступном для него подкадре.
- (4) Для простоты наибольшее допустимое число попыток передачи одной преамбулы далее не рассматривается.

В рамках введенной выше упрощающей системы допущений можно получить аппроксимацию для средней задержки входа в сеть. Для этого вначале полезно рассмотреть ситуацию без возникновения коллизий, для которой вероятность получения обслуживания ($\tilde{\mu}$) может быть выведена из выражения $E[\tilde{\tau}^{(1)}] = E[\tau^{(1)}]$ как

$$\tilde{\mu} = \frac{1}{E[\tau^{(1)}]} = \frac{1}{c_1(K_1 + K_0 + K + \bar{w}) + \frac{b-K+1}{2} - \bar{w}}, \quad (4.1)$$

где $E[\tau^{(1)}]$ – средняя длительность интервала времени между моментом активизации терминала и моментом получения им сообщения RAR, а $E[\tilde{\tau}^{(1)}]$ – средняя длительность соответствующего интервала во введенной эквивалентной системе. Приведенное здесь выражение будет использовано в дальнейшем при описании загрузки системы для М2М-устройств, которые избежали коллизий.

Далее исследуется ситуация с возможностью возникновения коллизий, когда для некоторого подкадра предполагается, что устройство i активизировалось и уже выбрало свою преамбулу. Пусть система находится в состоянии j , которое представляет собой число занятых терминалов, включающих устройство i . В состоянии j поведение терминала i может быть описано простой марковской цепью с двумя состояниями, которые соответствуют тому, активизировано ли данное устройство [254]. Таким образом, Q_i может принимать значения 0 или 1. Матрицу переходных вероятностей для рассматриваемой марковской цепи легко записать в виде

$$\Pi = \begin{pmatrix} 1 - \pi & \pi \\ \mu_j & 1 - \mu_j \end{pmatrix}. \quad (4.2)$$

В результате стационарное распределение вероятностей $\omega = \{\omega_0, \omega_1\}$ выражается из соотношения $\Pi^T \omega = \omega$ и при этом $\omega_0 + \omega_1 = 1$. Таким образом, среднее значение параметра Q_i может быть найдено как

$$E[Q_i] = 1 \cdot \omega_1 = \frac{\pi}{\pi + \mu_j}, \quad (4.3)$$

где Q_i – величина, показывающая, активизировано ли устройство i , а μ_j – вероятность успешной передачи преамбулы.

Используя закон Литтла, возможно установить среднее время, проведенное системой в состоянии j ,

$$E[\tau_j^{(1)}] = \frac{E[Q_i]}{\pi} = \frac{1}{\pi + \mu_j}. \quad (4.4)$$

В состоянии j для $j - 1$ занятых терминалов вероятность получить доступ к каналу и выбрать ту же самую преамбулу, что и устройство i , составляет $\pi_0 \cdot 1/s$ (вероятность активизации устройства, умноженная на вероятность выбора идентичной преамбулы). Для неактивных $M - j$ терминалов соответствующая вероятность представляет собой $\pi \cdot 1/s$ (вероятность поступления нового запроса на доступ, умноженная на вероятность выбора идентичной преамбулы).

Следовательно, вероятность избежания коллизии π_j^* для устройства i в состоянии j может быть вычислена следующим образом:

$$\pi_j^* = (1 - \pi_0 s^{-1})^{j-1} (1 - \pi s^{-1})^{M-j}. \quad (4.5)$$

Далее учитывается механизм постепенного повышения излучаемой терминалом мощности. Тогда вероятность избежания коллизии при попытке передачи с номером n задается в виде

$$\begin{aligned} Pr\{\text{1-й успех}\} &= \left(1 - \frac{1}{e}\right) \pi_j^*, \\ Pr\{\text{2-й успех}\} &= \left(1 - \left(1 - \frac{1}{e}\right) \pi_j^*\right) \left(1 - \frac{1}{e^2}\right) \pi_j^*, \\ &\dots \\ Pr\{\text{n-й успех}\} &= \left(1 - \frac{1}{e^n}\right) \pi_j^* \prod_{i=1}^{n-1} \left(1 - \pi_j^* \left(1 - \frac{1}{e^i}\right)\right), \\ &\dots \end{aligned}$$

Здесь, как и ранее, не учитываются потерянные преамбулы, а производится усреднение по успешным попыткам передачи, когда искомые средние величины заменяют на условные математические ожидания. Таким образом, среднее число попыток передачи выражается в виде

$$\bar{n}_j = \pi_j^* \sum_{n=1}^{L_1} n \left(1 - \frac{1}{e^n}\right) \prod_{i=1}^{n-1} \left(1 - \pi_j^* \left(1 - \frac{1}{e^i}\right)\right). \quad (4.6)$$

Принимая во внимание механизм постепенного повышения мощности, теперь можно установить значение вероятности успешной передачи μ_j устройством i в следующем виде:

$$\mu_j = \left(\bar{n}_j \cdot (K_1 + K_0 + K + \bar{w}) + \frac{b - K + 1}{2} - \bar{w}\right)^{-1}, \quad (4.7)$$

где величина \bar{n}_j вычисляется при помощи выражения, приведенного выше.

Следовательно, среднее время обслуживания может быть получено как

$$E[\tau^{(1)}] = \sum_{j=1}^M \theta_j E[\tau_j^{(1)}] = \sum_{j=1}^M \theta_j \frac{1}{\pi + \mu_j}, \quad (4.8)$$

где $\{\theta_j\}_{j=1}^M$ – стационарное распределение вероятностей, а θ_j – стационарная вероятность пребывания в состоянии j .

Для нахождения введенного выше стационарного распределения вероятностей следует рассмотреть все переходы между состояниями системы и решить соответствующее матричное уравнение размерности M . Для упрощения вычислений можно опустить более сложные переходы между состояниями и усреднить θ_j с использованием биномиального распределения следующим образом:

$$\theta_j = \binom{M-1}{j-1} \rho^{j-1} (1-\rho)^{M-j}, \quad (4.9)$$

где ρ – загрузка рассматриваемой системы, а $\binom{M-1}{j-1} = \frac{(M-1)!}{(j-1)!(M-j)!}$.

При таком способе расчета не учитываются коллизии между другими терминалами и предполагается, что их испытывает только рассматриваемое устройство i . Тем самым можно найти загрузку системы $\rho = \pi/\tilde{\mu}$ при помощи выражения для вероятности получения обслуживания μ в случае без коллизий. Итоговое выражение, позволяющее оценить среднее время обслуживания, представляет собой

$$E[\tau^{(1)}] = \sum_{j=1}^M \frac{\binom{M-1}{j-1} \rho^{j-1} (1-\rho)^{M-j}}{\frac{1}{N} + (a_j(K_1 + K_0 + K + \bar{w}) + \frac{b-K+1}{2} - \bar{w})^{-1}}, \quad (4.10)$$

где π_j^* и a_j были определены выше.

Предложенный подход допускает целый ряд важных практических обобщений и расширений. В частности, можно изучить режим нормального функционирования сети доступа, когда перегрузок не возникает. При этом интервал времени между поступлениями новых запросов на доступ распределен по экспоненциальному закону с заданным параметром $1/\lambda$. Также стоит отметить, что представленная математическая модель может охарактеризовать работу методов управления перегрузкой, например таких, как предварительная отсрочка (pre-backoff), предложенная в [255].

4.3.2 Анализ производительности сотовых систем M2M

Поддержка большого числа устройств. Как было отмечено выше, интеллектуальная энергосистема представляет собой важное приложение M2M, включающее в себя большое число сенсоров, счетчиков и измерителей, которые

передают свои данные через сетевую инфраструктуру в автоматизированном режиме. Такое поведение является типовым для массовых M2M-сервисов и поэтому становится опорным сценарием, демонстрирующим ключевые особенности межмашинного взаимодействия (см. рисунок 4.5). Соответственно, входящие в состав интеллектуальной энергосистемы устройства M2M могут быть поделены на следующие классы в зависимости от приоритета отправляемой ими информации: высокоприоритетные (тревожные сообщения) и низкоприоритетные (данные измерений). Со стороны сети связи передача тревожных сообщений потенциально сопряжена с повышенной трудностью, поскольку они часто возникают практически синхронно и при этом имеют строгие ограничения на допустимую задержку доступа [48].

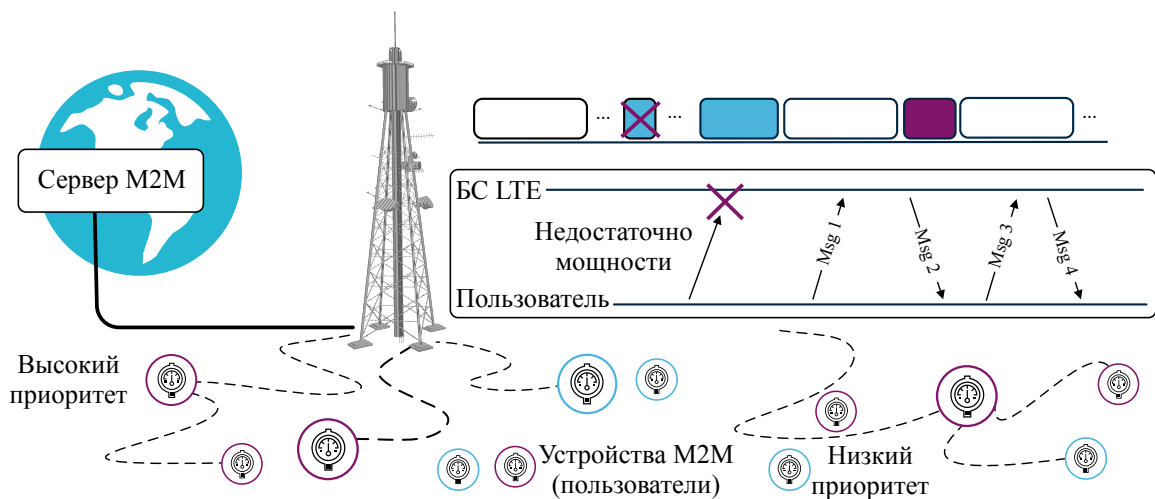


Рисунок 4.5 — Типовой сценарий M2M с большим числом устройств

На сегодняшний день система LTE предоставляет ряд выделенных каналов для передачи данных от M2M-устройств по линии вверх. В частности, канал PRACH используется терминалами для начального доступа, а также для того, чтобы запросить требуемые радиоресурсы, когда они еще не были предоставлены. Подключение к сети множества устройств M2M в течение короткого интервала времени может вызвать всплеск загрузки канала PRACH вследствие его ограниченной емкости. Важно напомнить, что работа системы PRACH организована в два этапа (см. рисунок 4.5). На первом из них происходит синхронизация отсчета времени по линии вверх (связанная с обменом сообщениями Msg 1 и Msg 2), когда при передаче преамбул в режиме случайного доступа используется постепенное повышение мощности для ее подстройки под состоя-

ние канала связи. На втором этапе применяется сообщение Msg 3 для отправки служебной информации по линии вверх на БС (называемую eNodeB или eNB), а также сообщение Msg 4 для последующего разрешения возможных коллизий.

Для понимания того, как большое число M2M-устройств влияет на характеристики процедуры случайного доступа в системе LTE, здесь и далее приводятся результаты моделирования работы канала PRACH на уровне протокола, которые были предварительно откалиброваны с тестовыми данными 3GPP [256]. Цель данного исследования – анализ функционирования системы PRACH в условиях перегрузки, когда большое число устройств M2M, имеющих различный приоритет, устанавливает свое соединение с сетью доступа (см. рисунок 4.6). В этой связи интересно наблюдать, что при добавлении порядка 40% высокоприоритетных терминалов, отправляющих тревожные сообщения, к исходному (типовому) множеству в 30000 низкоприоритетных терминалов, передающих данные своих измерений, резко ухудшается качество работы PRACH в отношении вероятности успешного доступа [257].

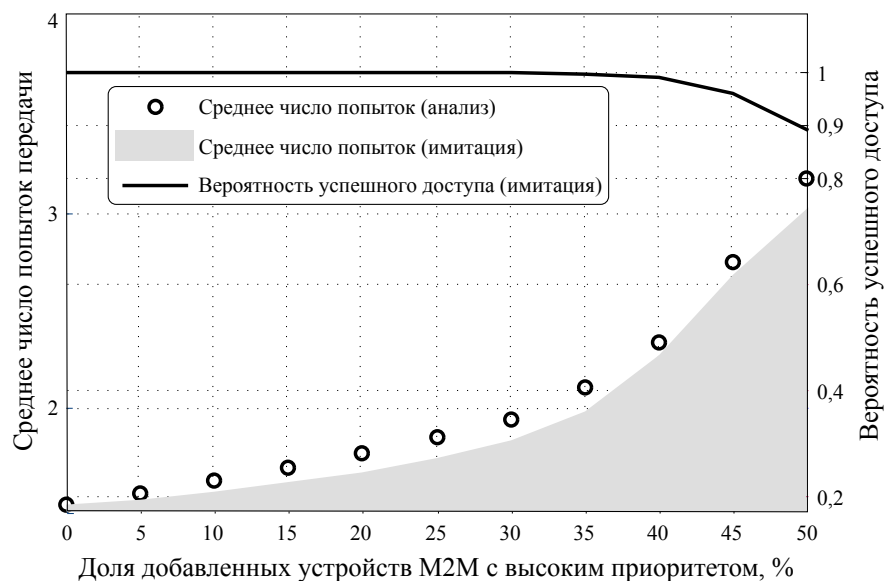


Рисунок 4.6 — Характеристики доступа M2M-устройств с различными приоритетами

Любопытно также заметить, что преамбулы PRACH, выбираемые M2M-устройствами случайным образом, можно рассматривать как неперекрывающиеся кодовые «каналы» (см. рисунок 4.7). В случае, когда два или более терминала выбирают идентичную преамбулу (канал), возникает ситуация, напоминающая типичную коллизию между пакетами данных. Подобная

интерпретация позволяет изучать процесс состязания между устройствами М2М, опираясь на известные результаты, полученные для многоканальных протоколов случайного доступа [258]. Как следствие, можно получить точную аппроксимацию в заданном диапазоне параметров работы системы PRACH (см. рисунок 4.6), таких как число доступных преамбул (M) и соперничающих терминалов (U), размер окна отсрочки и др. Однако подобные решения можно применять лишь в частных случаях, а вообще говоря, возникает необходимость проведения непосредственного численного анализа, который удастся только лишь для достаточно небольшого числа устройств/каналов из-за быстро растущей вычислительной сложности.

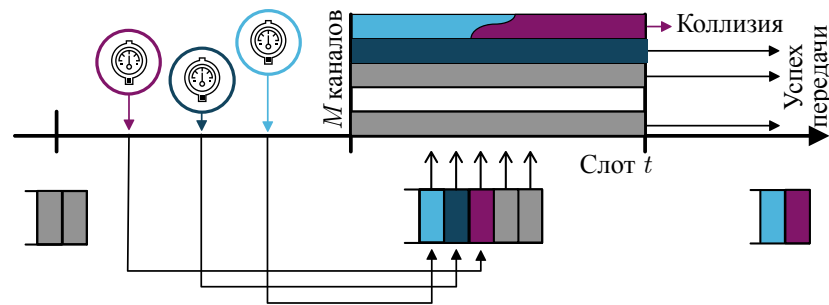


Рисунок 4.7 — Моделирование многоканального протокола случайного доступа

В качестве альтернативы для анализа подобных систем М2М можно применять методы диффузионной аппроксимации, которые позволяют охарактеризовать области стабильности соответствующей многоканальной сети случайного множественного доступа [259], а также использовать подходы с кооперацией [260] и рандомизацией [261]. В качестве целевого сценария полезно рассмотреть приложение М2М, связанное с промышленной автоматизацией (см. рисунок 4.8) и требующее определенных гарантий по задержке доступа и надежности доставки данных (например, при отправке приоритетных и/или критических сообщений). В этом контексте рисунок 4.9 демонстрирует задержки передачи, наблюдаемые в системе PRACH и полученные при помощи аналитического моделирования. Подобные результаты позволяют проводить оптимизацию процедуры случайного доступа путем выбора подходящего значения вероятности повторной передачи сообщения Msg 1 для заданного (наибольшего) числа М2М-устройств и множества используемых ими преамбул (каналов).

Более подробно на рисунке сравнивается оптимизированная задержка доступа с аналогичными данными, полученными для типовых значений индикато-

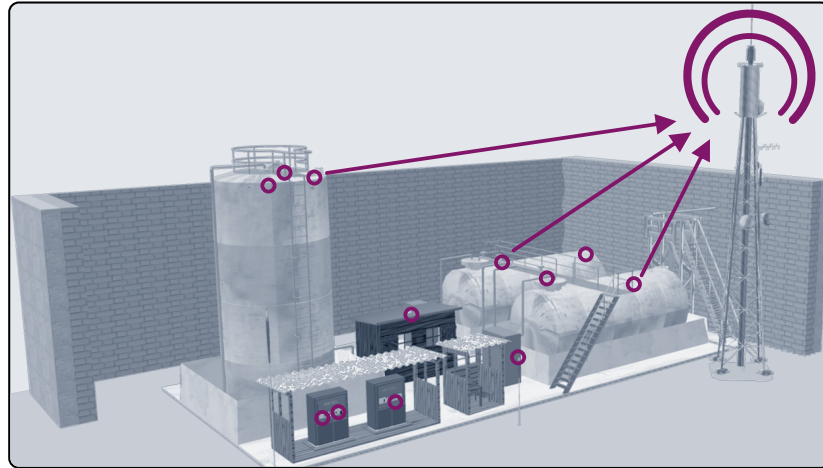


Рисунок 4.8 — Пример размещения M2M-системы для автоматизации производства

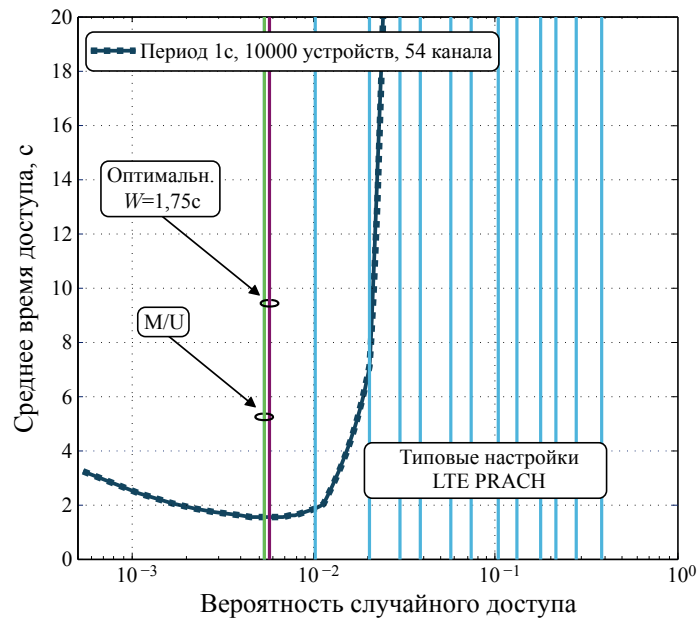


Рисунок 4.9 — Задержка доступа в сотовой системе с большим числом устройств M2M

ра отсрочки PRACH, заданных спецификациями 3GPP. Рассматриваемый подход позволяет БС LTE подстраивать параметры функционирования процедуры PRACH при наличии общесистемной информации обо всех подключенных устройствах M2M. Однако в случаях, когда подобная информация по какой-либо причине недоступна, конкурирующие терминалы могут использовать более простые (эвристические) методы контроля доступа в сеть, такие как выбор вероятности повторной передачи преамбулы в виде отношения M к U . Это зача-

стую позволяет достичь производительности, близкой к оптимальной. Предложенные здесь результаты позволят проектировать методы эффективного управления характеристиками работы сотовой системы M2M, например, задержкой доступа в сеть LTE, что оказывается крайне важно при поддержке приложений Индустриального Интернета с жесткими требованиями относительно доставки данных.

Энергетическая эффективность и передача данных малого объема. В тесной взаимосвязи с задержкой и вероятностью успешного доступа находится энергетическая эффективность соединений M2M, которую необходимо учитывать, т.к. M2M-устройства обычно имеют малые габариты и питаются от батареи. Приведенное здесь исследование рассматривает сценарий, когда некоторое приложение IoT вынуждает большое число терминалов выполнять определенное действие практически одновременно, например, передавать измерения интеллектуальных счетчиков. То же самое происходит, когда возникает непредвиденное, чрезвычайное или критическое событие, например, массовое отключение электричества или его восстановление, отказ сетевой инфраструктуры и т.д. В этих случаях множество устройств M2M подключаются к сети доступа в течение короткого интервала времени, используя в этих целях процедуру случайного доступа PRACH для установления синхронизации отсчета времени по линии вверх как при начальном доступе, так и при непосредственной передаче своих данных.

На рисунке 4.10 показаны характеристики начального входа в сеть для 30 тысяч устройств M2M относительно потребляемой ими мощности, вероятности коллизии и вероятности успешного доступа для различных значений индикатора отсрочки PRACH. Приведенные данные получены для типового процесса активизации терминалов согласно бета-распределению (тип трафика 2), как рекомендовано соответствующей документацией 3GPP (см. таблицу 6.1.1 в [256]). Используемая среда имитационного моделирования учитывает основные уровни мощности для каждого передающего терминала (неактивную, простоя, приема и передачи) на всех этапах работы процедуры PRACH (см. рисунок 4.11). Результаты показывают эффективность альтернативных методов управления перегрузкой (см. рисунок 4.10), поэтому можно сделать вывод о том, что комбинация увеличенного интервала отсрочки и предварительной отсрочки успешно справляется с перегрузкой, наступающей вследствие всплеска активности устройств M2M.

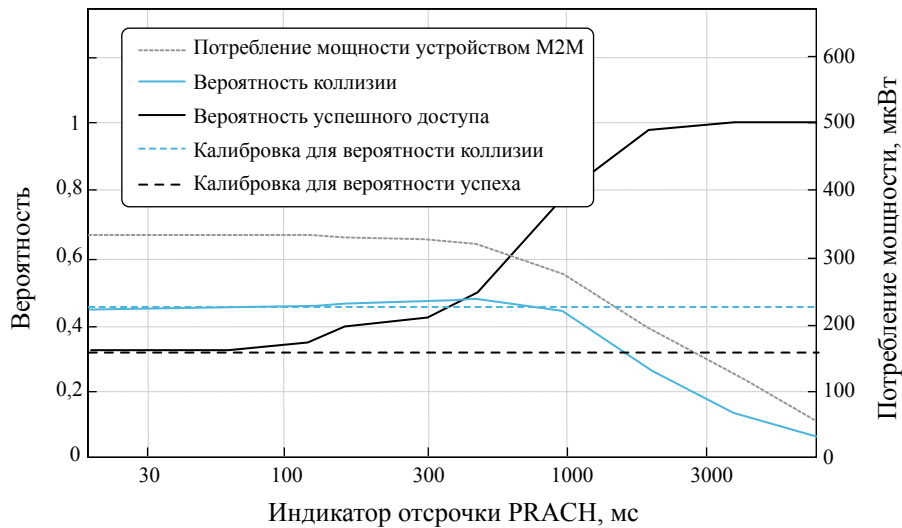


Рисунок 4.10 — Характеристики начального входа в сеть для сотовой системы M2M

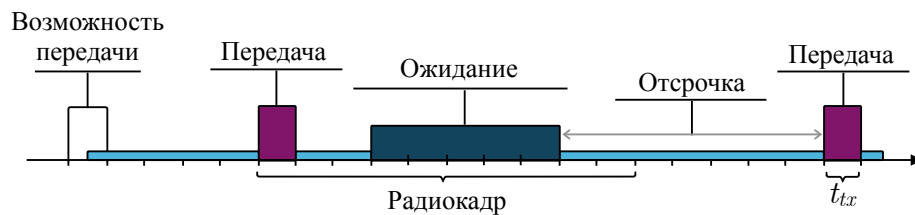


Рисунок 4.11 — Уровни потребляемой устройством M2M мощности

Далее в приведенном здесь анализе изучаются аспекты, связанные с энергетическим потреблением M2M-устройств. Важно отметить, что относительно короткая длительность цикла опроса (paging cycles) в системе LTE может быть крайне неэффективна для устройств M2M [262], особенно при значительных интервалах между поступлениями данных с учетом сравнительной устойчивости к задержке услуг M2M. Как следствие, увеличение длительности цикла опроса в режиме простоя позволяет таким терминалам находиться в режиме ожидания более продолжительное время и тем самым экономить ресурс своей батареи. Для определения потребляемой мощности требуется построить адекватную модель, которая могла бы учесть типовые параметры трафика M2M-устройств (см. рисунок 4.12). Приведенные в работе [263] данные показывают, что увеличение максимальной длительности цикла опроса и периода непостоянного приема DRX позволяет достичь существенного прироста энергетической эффективности терминалов (см. также рисунок 4.13).

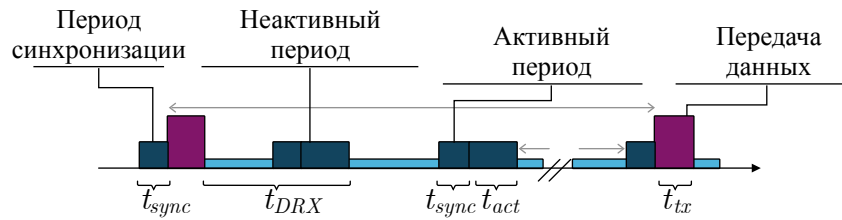


Рисунок 4.12 — Моделирование потребляемой устройством M2M мощности

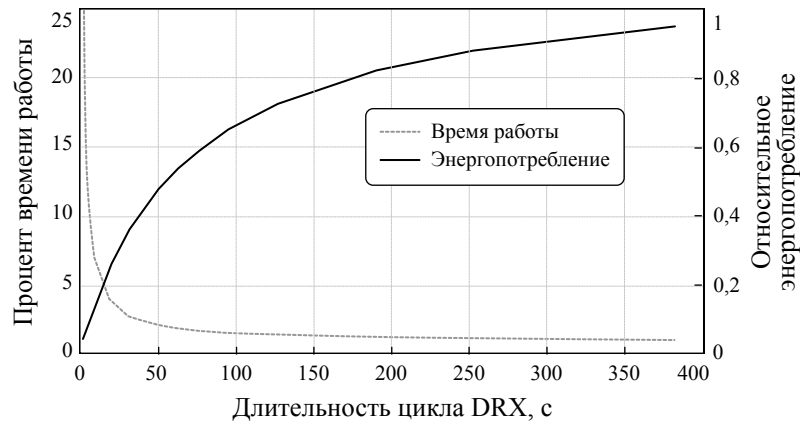


Рисунок 4.13 — Повышение энергетической эффективности M2M-устройств

В условиях, когда устройства M2M могут переносить некоторое повышение задержки, можно использовать прямые соединения D2D для дальнейшего снижения потребляемой ими мощности. К примеру, одно из M2M-устройств могло бы агрегировать и затем пересылать в сеть доступа данные от соседних терминалов, наблюдающих низкое качество инфраструктурного соединения. Подобные механизмы ретрансляции по каналам D2D позволяют уменьшать число повторных передач и тем самым снижать энергетическое потребление терминалов, особенно тех, которые размещаются на краю соты и передают данные малого объема. В противном случае, если введение дополнительной задержки недопустимо (как, например, в приложениях промышленной автоматизации), дополнительное снижение задержки доступа может быть достигнуто за счет сокращения процедуры PRACH, как показано на рисунке 4.11. Кроме того, представляет интерес разработка схем, позволяющих M2M-устройствам отправлять данные малого объема без значительных накладных расходов, характерных для технологии LTE.

Рисунок 4.14 воспроизводит текущую структуру кадра LTE (для полосы пропускания в 5 МГц) в виде прямоугольной сетки ресурсных блоков. Их группы составлены из различных каналов, предусмотренных системой, таких как канал PRACH с периодически выделяемыми радиоресурсами и физический канал управления по линии вверх PUCCH с непрерывно доступными радиоресурсами. Поскольку емкости обоих каналов, PUCCH и PRACH, могут оказаться недостаточными при обслуживании большого числа терминалов, то для передачи дополнительного трафика M2M возможно использовать долю радиоресурсов физического коллективного канала по линии вверх PUSCH, который применяется для обмена информацией между людьми H2H. Соответствующий протокол позволяет сократить число управляющих сообщений, необходимых для разрешения коллизий (см. описание в [264]). Его применение приводит к более эффективному использованию радиоресурсов, уменьшению задержки доступа, а также значительному снижению потребляемой устройствами M2M энергии (см. рисунок 4.15).

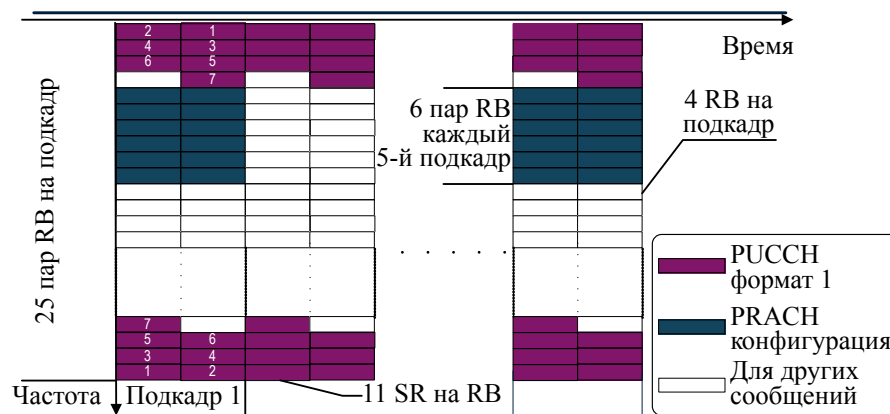


Рисунок 4.14 — Типовая структура кадра LTE для сотовой системы M2M

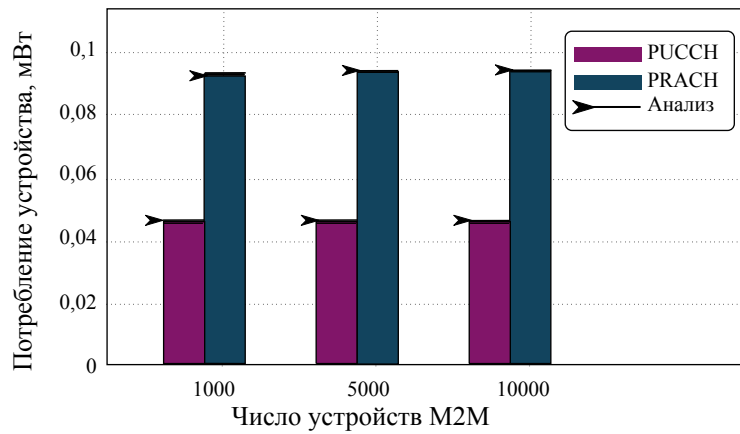


Рисунок 4.15 — Характеристики передачи данных малого объема

4.4 Предпосылки и последствия распространения носимых устройств

4.4.1 Возможности современного носимого оборудования

Носимые беспроводные устройства представляют собой пик развития современной миниатюрной электроники, позволяя встроить функционал устройства мобильной связи, например, в наручные часы. На сегодняшний день рынок носимых устройств стремительно набирает обороты по мере того, как на нем возникает все больше производителей оборудования, разработчиков приложений и создателей соответствующего контента [265]. Несмотря на то, что отдельные приложения носимой электроники появились уже более 20 лет назад (в основном, для специального применения в виде элементов военной формы, очков ночного видения, встроенных камер и т.д.), они долго оставались крупногабаритными, неудобными в использовании и недостаточно эстетичными [266]. В итоге широкое распространение получили лишь беспроводные наушники.

Однако недавний прогресс в технологиях мобильной связи и производстве аккумуляторных батарей, а также в области встраиваемой электроники, в использовании современных материалов и программного обеспечения сейчас позволяет создавать все более производительное, энергоэффективное, легкое и стильное носимое оборудование. В настоящее время беспроводные носимые устройства для фитнеса меняют сложившиеся практики занятия спортом. Ста-

ли популярными такие типы носимых устройств, как умные очки и часы, которые часто позиционируются на рынке в качестве аксессуаров к современным мобильным телефонам. Отдельного внимания заслуживают перспективные категории носимых устройств, например, кольца, датчики здоровья и умная ткань, которые смогут применяться в большом диапазоне будущих беспроводных приложений.

Вначале на рынке носимых устройств доминировали всего несколько крупных игроков, таких как компании Apple и Google, но по мере его развития появились различные производители, предлагающие многочисленные новые решения. Примерами являются умные часы, подключаемые к современным мобильным телефонам, продвинутое средство контроля за состоянием здоровья, а также различные вариации умных очков [267]. В результате сфера производства носимой электроники становится прибыльнее, являясь тем самым одной из самых быстро растущих областей бизнеса. Поскольку носимые устройства развиваются как более функциональные и стильные [268], основной акцент их дальнейшей эволюции смещается в сторону обеспечения удобства использования. Данное обстоятельство обуславливает необходимость создания соответствующих средств надежной беспроводной связи, обеспечивающей в том числе и высокие скорости передачи данных [269].

К примеру, умные очки предоставляют значительные новые возможности для работы и отдыха, позволяя получать доступ к информации исключительно при помощи зрения. Однако при этом поток данных должен поступать на устройство отображения непрерывно и с достаточно низкой задержкой. В производстве подобного оборудования, которое начала разрабатывать корпорация Google, на сегодняшний день насчитывается множество компаний, таких как Vuzix, GlassUp, Recon Instruments и Telepathy. В скором будущем ожидается, что диапазон носимых устройств пополнится так называемой умной одеждой, для производства которой уже имеются все необходимые компоненты. Ряд компаний планирует вход на данный рынок, в том числе OMSignal, Hexoskin и Athos. Для интеграции компонентов этих новых продуктов между собой станет ключевым эффективным беспроводное взаимодействие.

В заключение следует резюмировать, что в области разработки беспроводных носимых устройств недавно произошло несколько важных изменений. Они отчасти связаны с созданием новых пользовательских приложений, которые приведут к существенному увеличению плотности беспроводных соединений.

По мере того как возникают дополнительные типы носимых устройств, понадобятся средства обеспечения низкой задержки и высокой скорости передачи данных на этих соединениях [270]. Подобные средства станут решающим фактором для организации работы таких требовательных сервисов, как дополненная и виртуальная реальность [271]. Дальнейшее изложение направлено на обзор актуального состояния области современных радиотехнологий, пригодных для применения в носимом оборудовании, а также на изучение соответствующих характеристик производительности в сценариях с высокой плотностью беспроводных соединений.

4.4.2 Особенности применения носимых устройств в условиях их высокой плотности

Поскольку современные технологии позволяют создавать многофункциональные устройства небольшого размера, наблюдается их активная интеграция во все сферы жизни общества. Благодаря современным средствам сбора информации о внешнем мире, носимые устройства позволяют значительно повысить знания об окружающей среде [272]. Кроме того, они также усиливают способности, связанные с работой органов чувств пользователя, или же упрощают взаимодействие с техникой и устройствами, окружающими их носителя [273]. Однако по аналогии с беспроводными сенсорными сетями системы носимых устройств требуют создания соответствующих методов управления и оптимизации их работы. В этой связи одним из ключевых факторов является учет повышенной плотности размещения носимого оборудования. Если каждый пользователь имеет несколько носимых устройств, необходимо обеспечить работу сотен беспроводных соединений в массовых сценариях с большим количеством устройств.

На сегодняшний день в городских условиях, например, в переполненном общественном транспорте можно наблюдать чрезвычайно высокие плотности размещения абонентов (см. рисунок 4.16), использующих мобильные устройства (электронные книги и игры, проигрыватели музыки и многие другие). Ввиду ограниченной емкости современных технологий радиодоступа, большинство устройств функционирует без постоянного подключения к сети Интернет. Кро-

ме того, значительная часть соответствующей аудио- и визуальной информации доставляется посредством проводов без вовлечения радиоресурсов беспроводного канала связи. При данных ограничениях сотни пассажиров в непосредственной близости могут использовать всевозможные приложения, не создавая друг другу значительных сложностей. Эта ситуация резко меняется при переходе к беспроводному взаимодействию.



Рисунок 4.16 — Сценарий с высокой плотностью размещения носимого оборудования

Если предположить, что типичное число пассажиров, например, в городском автобусе (25–50 человек) одновременно используют радиотехнологию Bluetooth 2.0 для Hi-Fi стереонаушников, то возникнет ситуация перегрузки. Уже при числе 0,5 носимых устройств на пользователя в среднем, полоса пропускания канала связи может оказаться недостаточной для передачи соответствующего трафика. При повышении такой плотности до пяти устройств на человека, что отражает ситуацию, прогнозируемую в самом ближайшем будущем, система связи оказывается практически неработоспособной. Поэтому крайне своевременным становится рассмотрение массовых сценариев, в которых радиоресурсов системы доступа оказывается недостаточно для обеспечения текущей потребности в числе одновременно функционирующих беспроводных соединений.

Многие из задач, возникающих в системах носимых устройств с высокой плотностью, были решены ранее применительно к беспроводным сенсорным сетям. Однако плотность размещения сенсоров, как правило, существенно ниже, а зона покрытия – шире, чем при работе приложений, связанных с носимым оборудованием. Данная специфика накладывает существенные ограничения на

характеристики подходящих для организации их функционирования беспроводных технологий. Действительно, плотность беспроводных соединений не являлась решающим фактором при разработке предшествующих радиотехнологий в отличие от их емкости или радиуса зоны покрытия. Каждый беспроводной протокол представляет собой определенный компромисс в отношении своей простоты, эффективности и гибкости. Дальнейшее изложение направлено на обзор некоторых существующих решений, значительно различающихся в данных показателях, в плане их пригодности к использованию в массовых сценариях с высокой плотностью размещения носимых устройств.

В частности, рассматриваются типовые условия работы сети носимых устройств с высокой плотностью соединений, а также исследуются задачи, возникающие при разработке новых радиотехнологий, подходящих для таких массовых сценариев [274]. Для этой цели вначале проводится краткий обзор современных беспроводных решений с упором на более современные технологии в целях замены проводного соединения на беспроводное (т.н. «беспроводной провод»). В данном контексте рассматривается их параметр, характеризующий степень переиспользования пространства соседними устройствами, а также максимально достижимая скорость передачи данных в массовом сценарии с высокой плотностью размещения носимых устройств. В результате выявляются задачи, которые требуется решить при разработке новых технологий радиодоступа, учитывающих специфику носимого оборудования.

4.4.3 Альтернативные радиотехнологии для массового размещения носимого оборудования

Сравнение существующих беспроводных протоколов. В настоящее время технология WiFi представляет собой наиболее распространенное решение для беспроводной связи ближнего радиуса действия. По этой причине радиомодуль WiFi, основанный на семействе стандартов IEEE 802.11, размещается практически в каждом мобильном устройстве. Недавнее дополнение и расширение семейства технологиями IEEE 802.11ac и IEEE 802.11ah позволяет эффективно обслуживать сценарии, требующие высокой емкости сети и малого энергетического потребления устройств. На уровне доступа к среде передачи традици-

онная система WiFi обычно обеспечивает работу 5–10 абонентских терминалов, позволяя достичь времени использования канала связи вплоть до 90% [88]. Однако в случае поддержки приложений с высокой плотностью размещения устройств такой протокол наталкивается на серьезные ограничения.

Действительно, современные системы WLAN начинают испытывать сложности при поддержке 20 и более устройств, пытающихся получить доступ к беспроводному каналу одновременно, а эффективность использования радиоресурсов резко падает при наличии более 50 конкурирующих терминалов. Причиной такого поведения становится тот факт, что сеть доступа WiFi рассчитана на некоторое типовое число абонентских терминалов, которые соревнуются за доступ к среде передачи. Значительное превышение этого запланированного числа приводит к существенному снижению показателей производительности системы. Вообще говоря, для радиотехнологий ближнего радиуса действия характерно неявное допущение о наличии 10–20 интерферирующих между собой устройств, что легко видеть из соответствующего сравнения, приведенного в таблице 5.

Таблица 5 — Сравнение радиотехнологий ближнего радиуса действия

Технология	Скорость передачи, Мбит/с	Радиус действия, м	Число пользователей на кластер
Bluetooth	0,5–20	10–15	8
IEEE 802.11g/n	50–120	10–50	5–10
IEEE 802.11ac	до 400	10–30	10–15
ZigBee	0,5	10–100	5–10

К примеру, как следует из приведенного выше обзора, умные очки могут вскоре стать одним из самых распространенных носимых устройств, позволяя передавать видеоинформацию высокого разрешения без использования проводов. Такое применение требует скорости передачи данных по беспроводному каналу связи на уровне десятков Мбит/с. Как следствие, новые радиотехноло-

гии, которые придут на смену сегодняшним системам Bluetooth и WiFi, должны будут учесть эти требования, повысив свою емкость на несколько порядков до сотен Мбит/с. Далее рассматриваются исключительно такие, более современные протоколы, использующие соединения mmWave. Это, например, решения WiGig, построенные по стандарту IEEE 802.11ad, и наиболее новая разработка комитета IEEE, известная как 802.11ay, также функционирующая на частоте 60 ГГц.

Перспективные радиотехнологии ближнего радиуса действия.

Перспективные радиотехнологии ближнего радиуса действия используют, в основном, спектральный диапазон для промышленного, научного и медицинского применения с центральной частотой в 60 ГГц. В настоящее время существует несколько альтернативных протоколов, имеющих схожие принципы и пригодных для построения соответствующей системы связи [275]. Первый из них, выпущенный в 2008 году, известен под названием WirelessHD. Он реализует доступ к каналу связи с применением как методов случайного доступа, так и принципов, связанных с построением расписания передач. В каждой системе WirelessHD имеется один управляющий узел, который координирует ее работу подобно сотовой БС, но используя мультиплексирование во временной области. В отличие от сотовых технологий стандарт WirelessHD не предполагает дополнительных процедур для совместного планирования радиоресурсов между соседними сетями.

Когда множество сетей WirelessHD перекрываются в пространстве, времени или частоте, они не располагают средствами координирования своего общего расписания. В результате одновременно запланированные передачи с большой вероятностью окажутся неуспешны из-за неожиданно возникшей интерференции. Поскольку основным применением WirelessHD является обмен аудиовизуальной информацией (например, в системах домашнего кинотеатра), маловероятно, что данное ограничение будет преодолено последующими версиями стандарта, т.к. подобные системы редко находятся близко друг к другу. Следовательно, технологию WirelessHD нецелесообразно использовать для носимого оборудования, для которого характерно пересечение зон обслуживания соседних сетей. Ввиду наличия только четырех частотных каналов, число близко расположенных систем WirelessHD не может быть велико.

Альтернативой WirelessHD является технология WiGig, построенная на основе стандарта IEEE 802.11ad, который был завершен в 2009 году. Она пред-

лагает более гибкие возможности по доступу к среде передачи, позволяющие организовать отправку произвольных типов трафика, а также предоставляет поддержку устройств, имеющих более низкую стоимость и передающих данные сравнительно небольшого объема. Однако по своей архитектуре сеть WiGig очень схожа с системой WirelessHD: она также построена в рамках концепции центрального узла, который отвечает за построение расписания передач всех подключенных терминалов. Подобно WirelessHD, технология WiGig не предполагает ситуацию, когда соседствует более четырех сетей, и не предоставляет эффективных механизмов разрешения коллизий между управляющими узлами. Соответственно, стандарт WiGig не имеет средств надежного управления для множества перекрывающихся систем связи, но данное обстоятельство может измениться после выхода следующей версии данной технологии на основе стандарта IEEE 802.11ay.

Стандарт ECMA-387, вышедший в 2010 году, также может рассматриваться в качестве кандидата для организации сетей носимых устройств. По аналогии с предыдущими вариантами, он реализует управление системой при участии центрального узла. Однако кроме того, ECMA-387 располагает средствами для поддержки мобильности терминалов, предоставляя механизмы разрешения возможных коллизий, возникающих по ее причине (см. раздел 16.5.3.11 спецификации данной технологии). Отдельные подсети общей системы ECMA-387 реализуют два способа борьбы с коллизиями при наложении соответствующих зон обслуживания: гибкое переключение каналов и координирование. При обнаружении коллизии один из вовлеченных управляющих узлов предпринимает попытку переключения на другой частотный канал и в случае успеха переводит на него все подключенные к себе абонентские терминалы.

В противном случае, когда такое переключение не удастся, управляющие узлы ECMA-387 применяют механизм, позволяющий им синхронизировать способ отсчета времени так, чтобы скоординировать отправку служебных сообщений, называемых маяками. Это позволяет маякам всех перекрывающихся сетей передаваться один за другим. В результате управляющие узлы получают возможность отслеживать расписание передач в сетях друг друга и таким образом предотвращать коллизии. Однако дополнительные сообщения, предназначенные для подобных целей, увеличивают накладные расходы протокола, а также повышают сложность и стоимость соответствующего оборудования. Кроме того, данный процесс может потребовать значительного времени, в течение кото-

рого соседние системы ЕСМА-387 переходят в установившийся согласованный режим, а любое резкое изменение топологии сети может привести к необходимости повторного согласования.

Накладные расходы, связанные с координированием работы перекрывающихся сетей ЕСМА-387, составляют как минимум 21,3 мкс на суперкадр длительностью 16,384 мс (примерно 0,1%) для каждой из соседних систем. Кроме того, переиспользование терминалами радиоресурсов значительно затруднено. На практике накладные расходы на передачу дополнительных маяков, например, для 200 близко расположенных сетей составят до 26% от общего ресурса системы во времени, и это при отсутствии других вероятных сложностей. Можно заключить, что протокол ЕСМА-387 может функционировать удовлетворительно в системах носимых устройств с высокой плотностью, поддерживая вплоть до нескольких сотен потенциально перекрывающихся сетей. К сожалению, современные производители беспроводного оборудования не обеспечивают поддержку технологии ЕСМА-387 в своих устройствах.

По результатам проведенного выше обзора можно заключить, что существующие радиотехнологии ближнего радиуса действия, работающие на частоте 60 ГГц, не располагают эффективными средствами разрешения коллизий и управления радиоресурсами в масштабах сценариев, предполагающих массовое размещение носимого оборудования, таких как переполненный городской общественный транспорт. Основная причина этого состоит в неявном допущении о достаточно быстром затухании радиосигнала на крайне высоких частотах, принятом разработчиками сетей доступа. Долгое время считалось, что благодаря такому значительному затуханию число перекрывающихся сетей и подключенных к ним пользователей не будет слишком велико, и продвинутые механизмы разрешения коллизий между ними не понадобятся в соответствующих стандартах.

Однако данное предположение, очевидно, не выполняется для носимого оборудования, размещенного с высокой плотностью. По этой причине разработка эффективных средств координирования соседних сетей, передающих данные большого числа носимых устройств, представляет собой актуальную исследовательскую задачу [276], решение которой может быть востребовано в связи с выходом нового стандарта IEEE 802.11ay. При этом имеющиеся радиотехнологии ближнего радиуса действия, работающие в микроволновом диапазоне, не справляются с загрузкой в массовых сценариях с высокой плотностью размеще-

ния терминалов. Однако связь mmWave (например, на частоте 60 ГГц) потенциально имеет возможность предоставить требуемые скорости передачи данных порядка десятков Мбит/с на устройство при наличии в системах нового поколения адекватных средств обеспечения работы перекрывающихся в пространстве сетей.

4.5 Повышение энергетической эффективности доступа носимых устройств

4.5.1 Сети носимых устройств на основе mmWave-соединений

Как было показано в предыдущем разделе, технология WiGig предоставляет на сегодняшний день одно из наиболее зрелых решений mmWave, которое использует продвинутые методы формирования диаграммы направленности антенны для преодоления неблагоприятных условий распространения сигнала на частоте 60 ГГц. Соответствующий стандарт осуществляет переход от ненаправленных передачи и приема к узконаправленным соединениям. Развитие систем приема – передачи сигнала, а также антенных решеток обусловило возможность построения оборудования IEEE 802.11ad, способного преодолеть основные ограничения связи mmWave [123], такие как высокие потери при распространении радиоволны и чувствительность к блокированию радиосигнала небольшими объектами. В результате недорогая технология mmWave становится важной составной частью экосистемы 5G [216], и для ее применения в носимых устройствах существенный интерес представляет изучение характеристик, связанных с энергетической эффективностью при организации доступа такого оборудования в сеть [277].

Как известно, прогресс в технологии производства аккумуляторных батарей недостаточен по сравнению с растущими скоростями передачи данных и отстает на несколько порядков [278]. В свою очередь, энергетическая эффективность обычно измеряется в бит/Дж, а ее повышение пропорционально росту емкости системы связи становится актуальной задачей для сохранения приемлемого энергопотребления устройств. Дополнительным обстоятельством, приво-

дящим к необходимости увеличения энергетической эффективности, является стремление операторов мобильной связи обеспечить устойчивое экономическое развитие своих сетей [279]. Соответственно, большое число исследовательских инициатив (таких как GreenTouch) посвящено разработке так называемой «зеленой» связи, имеющей высокую энергетическую эффективность.

Современные носимые устройства (браслеты, часы, очки, камеры и т.д.) используют различные средства дистанционного зондирования, обработки информации, беспроводной связи, обеспечения электропитания и позиционирования в пространстве, демонстрируя при этом значительное разнообразие своих аппаратных характеристик, размеров, параметров энергопотребления и цены [280]. По отношению к традиционным абонентским терминалам носимое оборудование также следует тенденциям моды и стиля подобно тому, как карманные часы в свое время были вытеснены наручными [281]. Продвинутое носимые устройства могут включать в себя множество сенсоров и измерителей, что делает крайне широким диапазон их возможных применений [282]. Такая специфика приводит к необходимости разработки радиотехнологий, которые учитывали бы особенности носимого оборудования, в том числе массовое применение и высокую мобильность [283].

В этой связи сценарии с высокой плотностью размещения носимых устройств (например, переполненный городской общественный транспорт) становятся одними из самых сложных для существующих сетей доступа и требуют обстоятельного рассмотрения. Как следует из приведенного выше обзора, связь mmWave в виде протокола IEEE 802.11ad (WiGig) с учетом наработок в спецификациях IEEE 802.11ay сегодня подходит наилучшим образом для обслуживания массовых сценариев (см. рисунок 4.17), опережая аналогичные решения, представленные стандартами WirelessHD и ECMA-387, а также традиционными системами на базе Bluetooth или WiFi. Это заключение особенно справедливо для продвинутых носимых устройств, передающих потоковые мультимедиа-данные, например, очки дополненной реальности, которые активно завоевывают рынок.

Дальнейшее изложение посвящено математическому моделированию сети носимых устройств на основе соединений mmWave с целью изучения их энергетической эффективности в массовых приложениях, подверженных значительному влиянию интерференции. В частности, рассматривается сценарий, в котором множество персональных сетей соседних пользователей конкурируют

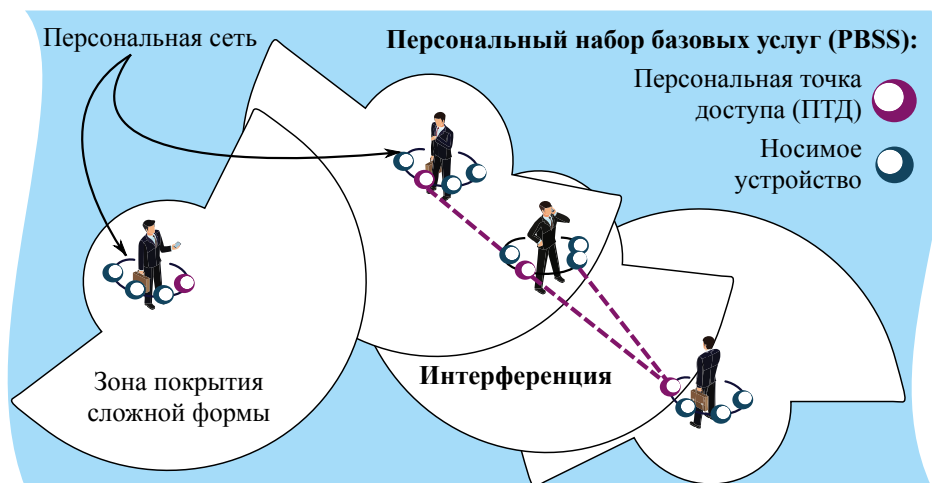


Рисунок 4.17 — Моделирование массового использования носимых устройств

за доступ к ограниченным ресурсам беспроводного канала связи. Проводится анализ вероятности успешного доступа к каналу, учитывающий реальную диаграмму направленности антенны, а также плотность размещения устройств и их геометрию. Кроме того, с учетом практической модели потребляемой мощности, выясняется качество методов планирования радиоресурсов и формирования диаграммы направленности относительно энергетической эффективности системы связи в различных условиях.

4.5.2 Моделирование mmWave-системы носимых устройств

Данный подраздел вводит в рассмотрение предлагаемую математическую модель системы носимых устройств на основе mmWave-соединений. Вначале описываются взаимодействующие узлы и особенности их размещения. Затем вводится специфика работы протоколов mmWave, характерная для обслуживания сценариев с высокой плотностью размещения носимых устройств [276]. Предлагаемый здесь аналитический подход (см. приложение Г) позволяет производить динамическое планирование радиоресурсов сети доступа (на основе стандартов WiGig или 5G NR) для достижения высоких показателей энергетической эффективности, что возможно на практике благодаря тесному взаимодействию сотовых и локальных сетей доступа [53]. Кроме того, проводимое

здесь исследование может быть полезно при интеграции носимых устройств с облачными системами [284].

Описание основных взаимодействий. Далее рассматривается набор персональных носимых сетей на плоскости, каждая из которых использует для передачи данных связь mmWave в нелицензированном спектре и именуется здесь персональным набором базовых услуг PBSS [285;286]. Каждая такая сеть включает в себя несколько идентичных беспроводных терминалов, являющихся носимыми устройствами, которые расположены в непосредственной близости друг к другу. Таким образом, один из терминалов (например, мобильный телефон пользователя) выполняет задачи центрального управляющего узла, координирующего работу остальных устройств. В частности, он отвечает за составление динамического расписания передач и оповещение о нем всех подключенных к нему терминалов. Данный управляющий узел называется персональной точкой доступа (ПТД).

Предполагается, что к одной ПТД подключено в точности M активных устройств, которые имеют данные, подготовленные к отправке по беспроводному каналу связи, используемому коллективно всеми PBSS. Ниже изучается исключительно передача информации по линии вверх, т.е. от носимых устройств и до обслуживающей их ПТД. Кроме того, не учитываются конкретные размеры и геометрические свойства ПТД и подключенных к ним устройств, а их местоположения обозначаются точками на плоскости. Однако при этом диаграмма направленности антенны каждого из терминалов моделируется в явном виде с помощью случайно выбранного азимута. Также предполагается, что как ПТД, так и все взаимодействующие с ней M носимых устройств расположены на теле их пользователя внутри окружности радиуса r_0 (см. рисунок 4.18).

Целью последующего моделирования является анализ характеристик работы mmWave-системы носимых устройств, включающей в себя множество интерферирующих между собой персональных сетей PBSS. Они распределены на плоскости так, что не все соседние терминалы имеют возможность принимать сообщения друг от друга ввиду значительных потерь при распространении радиоволны и блокирования их передач телом человека. Особый интерес представляет рассмотрение такой достаточно высокой плотности размещения ПТД, что в среднем внутри зоны покрытия каждого терминала функционирует несколько PBSS на каждом из доступных частотных каналов. Для учета пространственного компонента такой системы дополнительно предполагается,

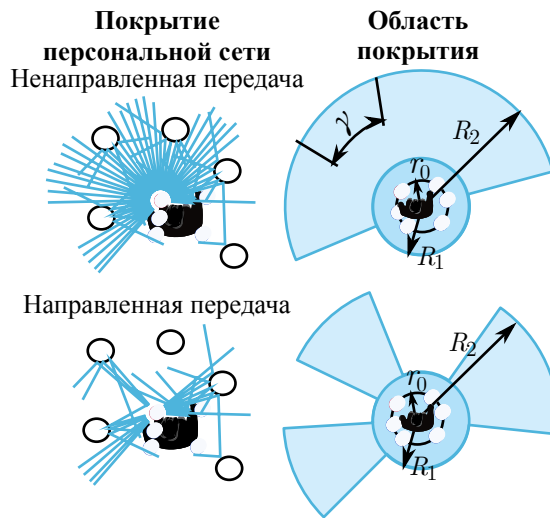


Рисунок 4.18 — Пример геометрического расположения системы PBSS

что местоположения ПТД на плоскости подчиняются пуассоновскому точечному процессу PPP интенсивности μ .

Учет направленности передачи. Для упрощения моделирования считается, что все взаимодействующие устройства используют антенны с идентичными параметрами и имеют одинаковые ограничения на максимальный уровень излучаемой мощности, что приводит к схожим характеристикам их зон покрытия. При передаче миллиметровых волн канал связи представляет собой среду, содержащую множество препятствий для их распространения, таких как тело человека – пользователя персональной носимой сети [287]. В этой связи важно отметить, что вне зависимости от диаграммы направленности антенны энергия может излучаться в пространство только наружу по отношению к телу человека. Количество такой энергии, испускаемой в каждом возможном направлении в зависимости от азимута диаграммы направленности антенны и порога ее чувствительности, определяет фактическую зону покрытия терминала.

Рисунок 4.18 демонстрирует типовую форму такой зоны покрытия для некоторого меченого устройства. Поскольку носимое оборудование может размещаться произвольно на теле своего владельца, для организации надежной связи внутри PBSS целесообразно использовать соединения, работающие на отраженных радиосигналах. Для целей последующего анализа предполагается, что такая отраженная энергия обеспечивает достаточное покрытие в непосредственной близости каждого терминала, ограниченное радиусом r_0 из-за плотного размещения и отражений. Распространение радиосигнала вблизи тела че-

ловека является самостоятельным исследовательским направлением, поэтому здесь предполагается, что мощность принятого сигнала представляет собой равномерно распределенную случайную величину, приводящую к средней скорости передачи данных R_{data} в рамках определенной сети PBSS.

В результате рассматриваемые диаграммы направленности антенн внутри PBSS имеют идентичную форму, которую можно описать при помощи функции $\rho(\Delta\alpha)$ относительного угла $\Delta\alpha$. Абсолютный угол направленности антенны α (азимут) по отношению к некоторой общей оси отсчета является случайной величиной, распределенной равномерно на интервале $[0; 2\pi]$, что соответствует случаю, когда носимые устройства могут быть размещены произвольно на теле человека. При этом зона покрытия определяется как область, в которой мощность принятого сигнала превышает заданный порог. Введенная таким образом модель позволяет терминалам связываться друг с другом вне зависимости от их фактического местоположения, тогда как возникающая при этом интерференция в сторону соседних сетей зависит от расстояния между устройствами и их ориентации.

Можно показать, что устройства принимают сообщения друг от друга (т.е. соединение существует) если и только если они расположены внутри области пересечения соответствующих зон покрытия (см. рисунок 4.19). Данное соображение предоставляет конструктивный подход для преобразования диаграммы направленности в критерии существования радиосоединения. К примеру, введенная здесь модель может быть расширена путем рассмотрения более сложных форм диаграммы направленности, которые соответствуют реальным антеннам и могут быть описаны полиномом Чебышева. Кроме того, в рамках данной модели вводится т.н. виртуальная диаграмма направленности. При этом для сформированной диаграммы возможно использовать достаточно простую форму зоны покрытия, которая затем пересчитывается в вероятность существования соединения в зависимости от расстояния.

Следует также отметить, что предложенная здесь модель распространения радиосигнала не рассматривает его отражения и замирания, вызванные внешними объектами (например, человеком, находящимся между двумя пользователями персональных носимых сетей). Однако оба эти эффекта могут быть неявно учтены либо виртуальной диаграммой направленности, либо вероятностью существования соединения. По существу, описанный здесь геометрический подход предоставляет гибкий и адаптивный способ моделирования ос-

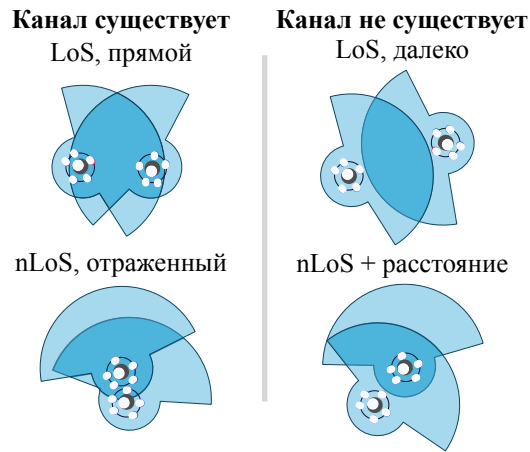


Рисунок 4.19 — Случаи наличия и отсутствия mmWave-соединения

новых практических свойств mmWave-системы носимых устройств. Дальнейшее изложение предполагает анализ зоны покрытия некоторой меченой ПТД $S_c = \int_0^{2\pi} \frac{\rho^2(\alpha)}{2} d\alpha$. Данное сужение связано с тем, что такая типовая ПТД является достаточно репрезентативной в силу стохастического характера построенной модели.

Практические примеры рассматриваемых диаграмм направленности. Для проведения последующего анализа полезно выписать общие выражения с учетом введенной для описания беспроводного покрытия функции $\rho(\Delta\alpha)$. Однако в качестве практического примера целесообразно рассмотреть отдельно более простой случай, когда радиус зоны покрытия $\rho(\Delta\alpha)$ представляет собой кусочно-непрерывную функцию (см. рисунок 4.18). Соответственно, она определяет зону покрытия с помощью двух кругов радиуса $R_1 < R_2$ в изотропной среде распространения, где сравнительно небольшой сектор внешнего круга $\gamma \cdot 2\pi$ (здесь γ выражает соответствующую долю) задает фактическую покрытую область

$$\rho(\Delta\alpha) = \begin{cases} R_1, & \text{если } \Delta\alpha \leq 2\pi(1 - \gamma), \\ R_2, & \text{если } \Delta\alpha > 2\pi(1 - \gamma). \end{cases} \quad (4.11)$$

Предлагаемая абстракция протокола mmWave-доступа. Далее предлагается краткий обзор основных механизмов работы протокола mmWave-доступа, которые представляют интерес в контексте анализа системы взаимодействующих носимых устройств. Основной целью данного описания является проведение моделирования технологий радиодоступа, поддерживающих сред-

ства формирования диаграммы направленности антенны, таких как, например, сеть WiGig (IEEE 802.11ad). Однако приведенные ниже соображения также справедливы для учета альтернативных технологий ближнего радиуса действия, использующих связь mmWave, включая WirelessHD и ECMA-387.

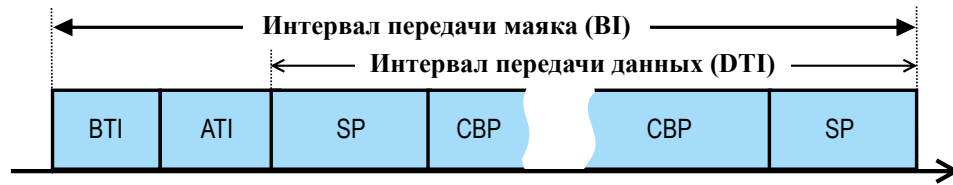


Рисунок 4.20 — Структура интервала передачи маяка в протоколе WiGig

Описание функционирования протокола. В протоколе доступа WiGig время работы системы поделено на интервалы маяка BI, в течение каждого из которых управляющий узел рассылает информацию о расписании передач [285]. Каждый BI, в свою очередь, включает в себя несколько служебных элементов (см. рисунок 4.20): интервал отправки маяка BTI и интервал отправки оповещения ATI, в течение которого согласовывается расписание передач внутри BI. Обмен данными может происходить в оставшейся части BI, называемой интервалом отправки данных DTI и включающей в себя период получения права на доступ CBP и период запланированного обслуживания SP. Здесь предполагается, что все устройства уже установили соединение с соответствующей ПТД, и их конкуренция за доступ внутри PBSS не рассматривается. Таким образом, ПТД может запланировать передачу своих устройств в любой части DTI (см. рисунок 4.21).

Для упрощения структуры рассматриваемого протокола интервалы BTI и ATI объединяются в один, который при проведении последующего анализа именуется просто маяком. Равные между собой интервалы BI называются далее суперкадрами, каждый из которых подразделяется на L слотов, а также считается, что передача как маяка, так и одного пакета данных от какого-либо терминала занимает в точности один слот (см. рисунок 4.21). Более того, для приближения такой модели к реальности также предполагается, что любое устройство может затребовать K слотов для передачи своих данных. Существенным условием в этой связи является синхронизация времени в системе на уровне слотов. Важно заметить, что введенные таким образом «аналитические» слоты не

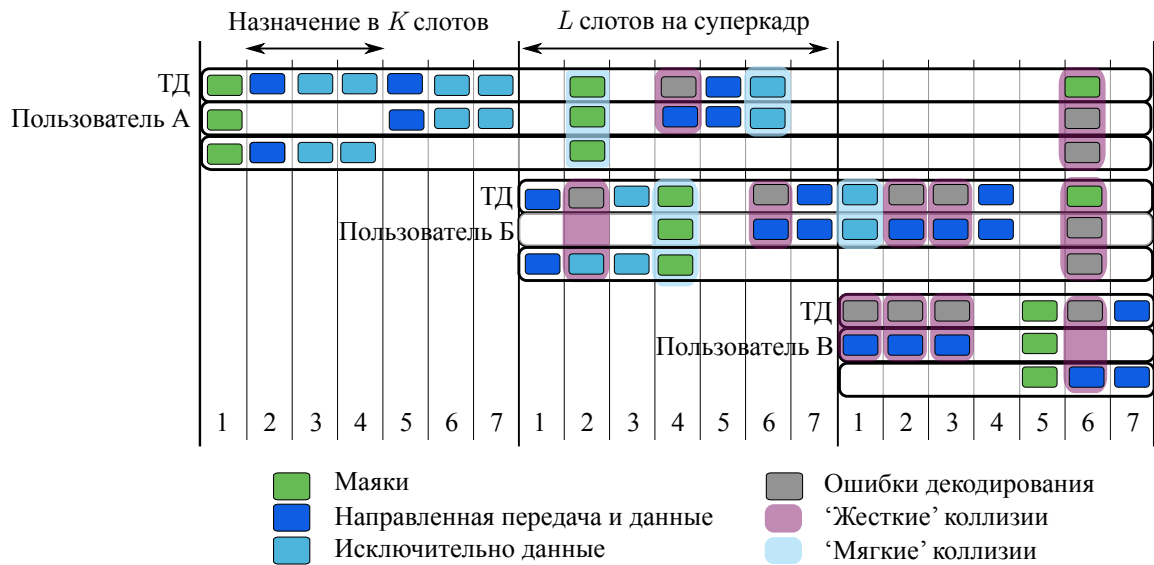


Рисунок 4.21 — Предлагаемая абстракция работы протокола WiGig

тождественны «физической» структуре времени передачи данных, задаваемой спецификациями протокола.

Поскольку передача маяка, по сути, представляет собой «рукопожатие» (т.к. сообщения отсылаются в обе стороны), каждое устройство, которое неспособно в нем участвовать, будет недоступно для последующего обмена данными. Вообще говоря, в начале каждого суперкадра ПТД выбирает слот для отправки своего маяка и затем планирует KM слотов для передачи данных в соответствии с некоторой дисциплиной планирования, задаваемой практической реализацией протокола mmWave-доступа. Как обсуждалось выше, в большинстве имеющихся современных решений координация между соседними сетями практически отсутствует. В результате ПТД и подключенные к ним устройства не осведомлены о текущем расписании других близлежащих PBSS и вынуждены конкурировать с ними за доступ к каналу. Соответственно, если данные или маяки запланированы к передаче двумя или более ПТД одновременно, в рассматриваемой системе может возникнуть коллизия одного из трех типов:

- коллизия маяк – маяк (предполагается, что оба маяка будут потеряны, а соответствующие устройства не смогут передавать в течение всего суперкадра; если таким образом пострадала ПТД, то все подключенные к ней устройства тоже будут затронуты);

- коллизия маяк – данные (данные считаются потерянными; передача маяка не затрагивается, поскольку в нем используется более надежная модуляционно-кодовая схема MCS);
- коллизия данные – данные (все вовлеченные данные будут считаться потерянными).

Стоит отметить, что тип возникающей коллизии зависит от фактического местоположения вовлеченных в нее устройств внутри области пересечения зон покрытия. При этом большинство передач приводит лишь к частичному блокированию терминалов соседних персональных сетей. Кроме того, для появления коллизии необходимо, чтобы интерференция воздействовала на приемную сторону некоторого соединения. Поскольку здесь предполагается, что информация о состоянии канала связи доступна мгновенно и безошибочно, то интерференция, действующая на передающую сторону, не рассматривается. Однако так как маяки подразумевают отправку сообщений в обе стороны, блокирование некоторой ПТД другим маяком приводит к потере содержащейся в нем информации для всех подключенных к ней устройств.

Формирование диаграммы направленности. Для учета специфики, связанной с формированием диаграммы направленности антенны, следует принять во внимание процедуру «прочесывания» секторов SLS, проводимую перед каждой отправкой данных. Ее общая длительность зависит от числа секторов N_s согласно соотношению $2((5 + 1)N_s + 3 + 5) \mu s$ [288]. В некотором секторе процедура «прочесывания» добавляет фиксированные накладные расходы для каждого из подключенных устройств, и в случае ее успеха соответствующий терминал может передавать свои данные без дополнительных ограничений. При этом функционирование протокола уточнения направления луча BRP не зависит от числа секторов и накладывает небольшие постоянные издержки на передачу данных. В результате помимо маяков и данных, в рамках рассматриваемой модели вводится дополнительная сущность, отвечающая за работу процедуры «прочесывания» и проходящая по всем секторам антенны (см. рисунок 4.21).

Кроме того, поскольку отправка данных происходит в направленном режиме, она может привести к коллизии данные – данные, если между интерферирующим устройством и соответствующей ПТД имеется соединение. Обмен данными также влияет на процедуру «прочесывания» секторов (см. НП на рисунке 4.22), так как она предполагает ненаправленную передачу. Однако здесь предполагается, что отправка маяков не затрагивается данными при коллизии

благодаря их более высокой надежности¹. Поскольку процедура «прочесывания» является ненаправленной, она также может повлиять на функционирование аналогичной процедуры в соседней сети, что повлечет за собой неудачу или же незапланированный отклик. Ввиду достаточно малой длительности такой ненаправленной передачи далее считается, что ни данные, ни маяки не должны при этом пострадать.

	М	Д		М	Д	НП	
М	1	1		1	1	1	
Д	0	1		0	1	1	
				НП	0	0	1

Рисунок 4.22 — Соответствие между компонентами модели

Также важно подчеркнуть, что введенная модель не предполагает непосредственного координирования между соседними (перекрывающимися) персональными носимыми сетями (хотя стандарты IEEE 802.11ay и ECMA-387 могут поддерживать подобные средства координирования). Причина данного упрощения заключается в том, что такие протоколы координирования не были реализованы на практике и не являются составной частью имеющегося на рынке оборудования. По этой причине отсутствуют сведения об их возможном функционировании, которое, кроме того, может вызвать дополнительную интерференцию, понизить скорость передачи данных, а также привести к росту энергопотребления. Применение узконаправленной передачи в принципе может смягчить последствия этих побочных эффектов, и поэтому проводимое ниже исследование общесистемной энергетической эффективности представляет в данном контексте существенный интерес.

¹На рисунке «1» обозначает коллизию наверняка, когда сущность по вертикали терпит неудачу при столкновении с сущностью по горизонтали; «0» обозначает успех сущности по вертикали. Сравниваются системы Тип-1 (слева) и Тип-2 (справа). Значения 0/1 рассматриваются здесь для примера и могут быть заменены на другие вероятности.

4.5.3 Основные численные результаты и их интерпретация

В данном подразделе исследуются общесистемные характеристики работы на основе предложенной выше модели. Для анализа плотного размещения носимых устройств, применяющих mmWave-соединения для обмена данными, используются типовые параметры, приведенные в таблице 6. В частности, изучаются зависимости, связанные с энергетической эффективностью и масштабируемостью системы.

Проведенное далее имитационное моделирование основывается на результатах, полученных при помощи имитатора распространения радиосигнала mmWave и представленных в главе 3. Была воссоздана работа беспроводного канала связи между множеством передающих и принимающих терминалов, которые устанавливают соединения на миллиметровых волнах с учетом отражения, рассеяния, дифракции и т.д. В рассмотренных сценариях распространение радиосигнала mmWave в наибольшей степени подвержено воздействию отражений от окружающих тел пользователей. Такие отражения могут быть достаточно точно представлены либо посредством введенной выше модели с двумя кругами (основанной на вероятности существования соединения $\xi(r)$), либо путем подстройки виртуальной диаграммы направленности антенны к виду, приведенному на рисунке Г.6 в приложении Г.

Дальнейшее исследование имеет своей целью оценку производительности системы в целом, а также изучение предпочтительных конфигураций диаграммы направленности антенны для повышения энергетической эффективности. Для этого моделируется зона покрытия, представленная двумя полукругами с радиусами $R_1 = 1$ м и $R_2 = 7$ м, а вероятность существования соединения подчиняется экспоненциальному распределению $\xi(r) = e^{-0,7 \max(r-0,4,0)}$. Число возможных секторов антенны ограничено значением 25. Выбранные уровни энергопотребления отражают успешный прием данных, отсутствие успеха при получении сообщений, испытавших коллизию, а также активное ожидание приема. При этом учитывается только энергопотребление радиомодуля терминала, а энергия, расходуемая другим функционалом носимого оборудования, не рассматривается.

Энергетическая эффективность изучаемой системы (Тип-2, см. рисунок 4.22) представлена на рисунке 4.23, где показаны ее оптимальные значения,

Таблица 6 — Параметры моделирования mmWave-системы носимых устройств

Обозначение	Описание	Значение
μ	Плотность размещения ПТД (на канал)	Перем.
M	Число носимых устройств на ПТД	4
K	Число слотов времени на устройство	5
γ	Угол раскрытия диаграммы направленности	45°
f	Рабочая частота	60 ГГц
ΔT	Длительность суперкадра	200 мс
ΔT_s	Длительность слота	0,3 мс
R_{data}	Скорость передачи данных	4,5 Гбит/с
Энергопотребление, Дж/слот (заимствовано из [289])		
e_{BTx}	Передача маяка	$2,43e-04$
e_{BRx}	Прием маяка	$1,83e-04$
$e_{B,idle}$	Ожидание маяка	$1,86e-04$
e_{DTx}	Передача данных	$2,43e-04$
e_{DRx}	Прием данных	$1,83e-04$
$e_{D,idle}$	Ожидание данных	$1,86e-04$
e_{idle}	Режим простоя	$0,81e-06$

а также направления сдвига при изменении плотности размещения терминалов. Аналитические результаты (сплошные линии) совпадают с данными имитационного моделирования (маркеры), что подтверждает корректность предложенного математического подхода. Интересно заметить, что энергетическая эффективность резко повышается, когда применяется направленная передача (45°), что связано со снижением числа коллизий. Также важно отметить, что эффективность направленной системы растет при сокращении длительности кадров. Однако при этом следует учитывать накладные расходы протокола, которые также зависят от размера кадра, что приводит к появлению его некоторой оптимальной длительности.

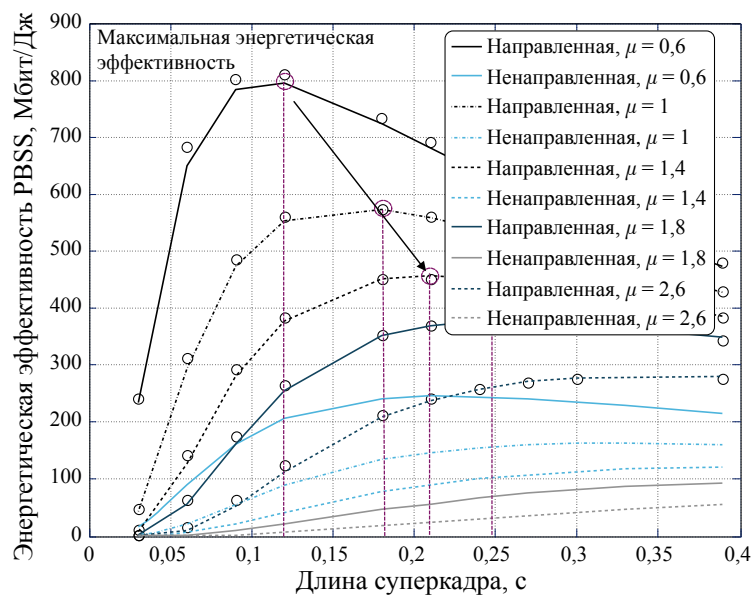


Рисунок 4.23 — Зависимость энергетической эффективности от длительности суперкадра

Далее, рисунок 4.24 демонстрирует выбор оптимального числа слотов при ограничении на минимальную требуемую скорость передачи данных (см. верхний правый угол). Здесь можно наблюдать, что для системы, состоящей из множества персональных носимых сетей, точка максимума энергетической эффективности не всегда находится внутри области значений, которые удовлетворяют ограничению на минимальную скорость передачи данных. Вообще говоря, чем выше требуется скорость передачи, тем менее вероятно получить точку оптимальной энергетической эффективности внутри рассматриваемой области. Следовательно, решение приходится искать либо на границе области удовле-

творительных скоростей передачи, либо в точке наивысшей энергетической эффективности, в зависимости от предпочтений производителя оборудования.

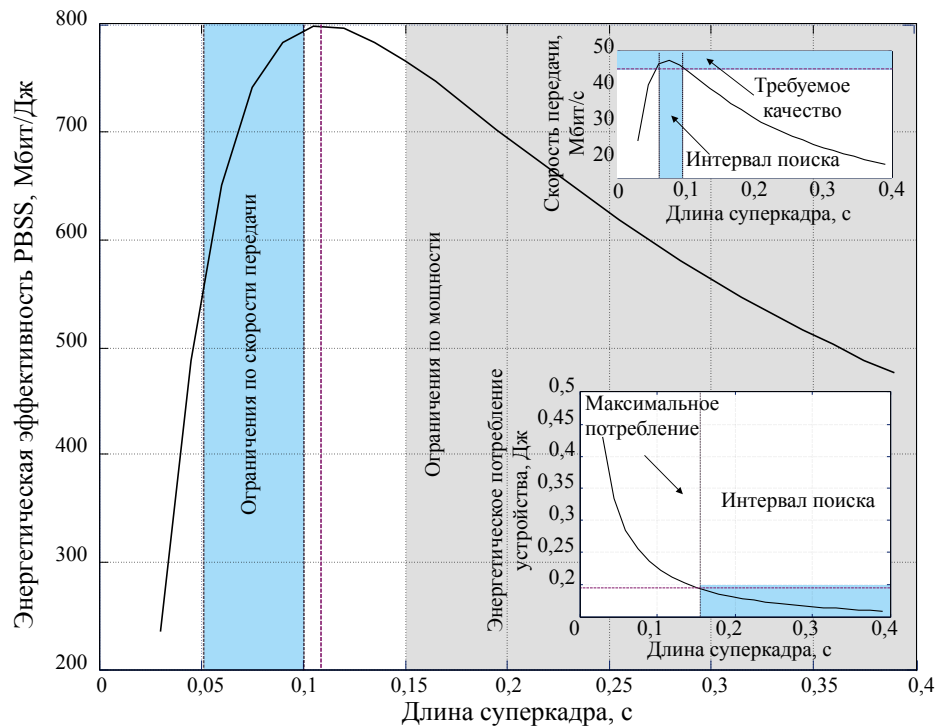


Рисунок 4.24 — Зависимость энергетической эффективности от числа слотов

Следующим важным решением, принимаемым системой, является выбор угла раскрытия диаграммы направленности антенны. Варьируя ее совместно с плотностью размещения терминалов (см. рисунок 4.25), возможно выявить вклад от использования направленной передачи. Здесь пунктирные линии соответствуют базовой системе с ненаправленными антеннами. Когда применяется направленная передача, она, как правило, повышает энергетическую эффективность системы. Однако в определенных условиях формирование узкой диаграммы направленности оказывается непродуктивно. Это видно при низких плотностях размещения терминалов и связано с неоправданным ростом накладных расходов, поскольку ненаправленная система функционирует удовлетворительно на достаточно коротких соединениях. По мере роста плотности применение направленной передачи становится все более критичным, так как оно позволяет снизить конкуренцию за доступ к каналу.

Рост плотности размещения терминалов влияет не только на скорость передачи, но и на энергетическую эффективность. Как следует из рисунка 4.26, последняя снижается для всех рассматриваемых систем. Видно, что суперкадры

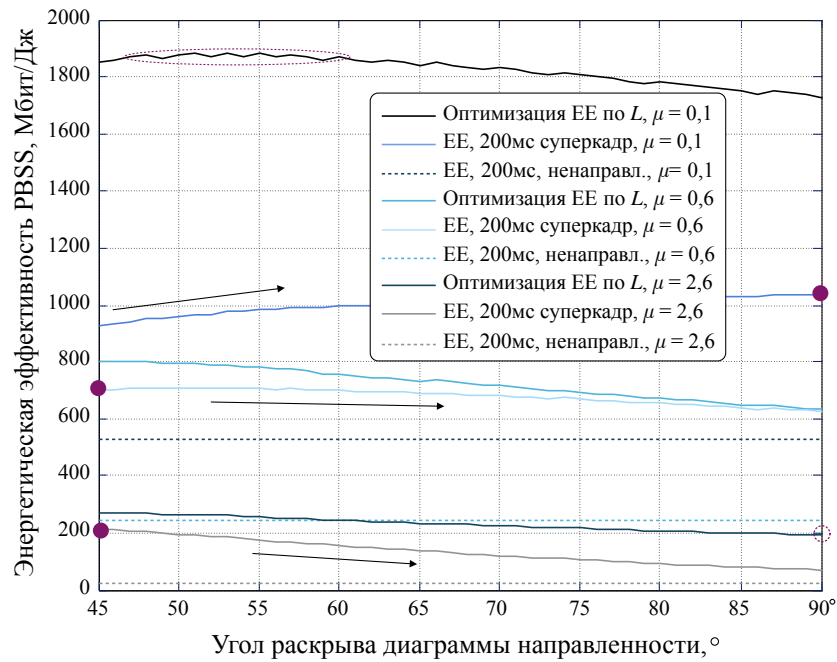


Рисунок 4.25 — Зависимость энергетической эффективности от угла раскрытия диаграммы направленности

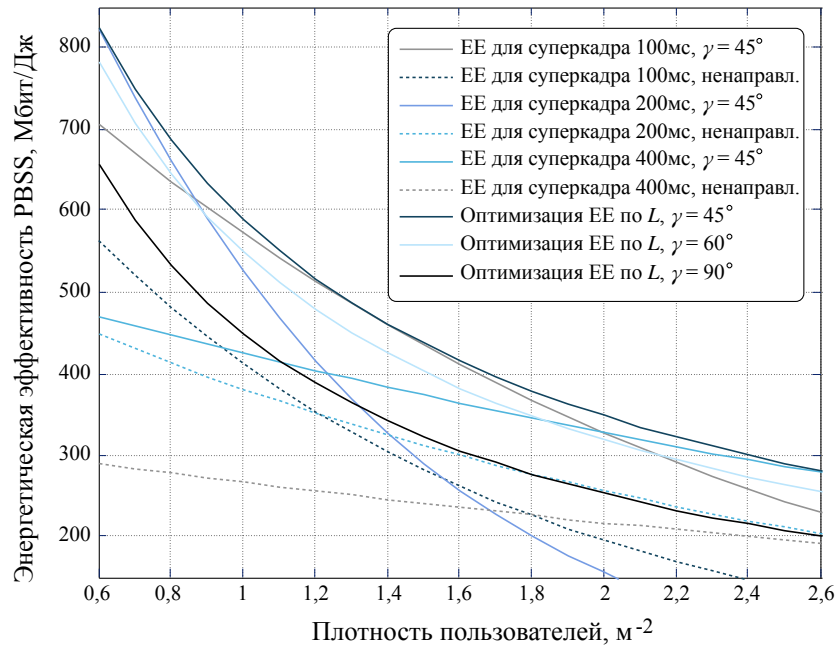


Рисунок 4.26 — Зависимость эффективности работы системы в целом от плотности размещения терминалов

большого размера позволяют достичь лучшей масштабируемости, а использование направленной передачи поддерживает достаточно высокую энергетическую эффективность при высоких плотностях размещения. При этом оптимизирован-

ная система оказывается существенно лучше своих аналогов вне зависимости от текущей плотности. Можно сделать вывод о том, что функционал по формированию диаграммы направленности значительно повышает эффективность работы системы в условиях ее высокой плотности, а также позволяет успешно справляться с ростом числа коллизий. С другой стороны, эффективное планирование радиоресурсов для передачи маяков и данных, т.е. выбор параметров L и K , представляет собой гораздо более сложную задачу, решение которой зависит от множества практических факторов.

4.6 Выводы

В результате предложенная модель сотовой сети с конкурентным обслуживанием большого числа пользователей, учитывающая особенности протокола множественного доступа, позволяет изучать показатели эффективности функционирования системы связи, включающие в себя среднее число попыток передачи, а также вероятность успешного установления соединения и задержку доступа в сеть. Кроме того, построенная модель сети радиодоступа для обслуживания множества плотно размещенных носимых устройств, учитывающая характер потребления ими энергии и особенности протокола конкурентного доступа, применима для оценки показателей эффективности функционирования системы связи и, в частности, энергетической эффективности передачи потокового трафика пользователей.

Глава 5. Применение гетерогенных сетей 5G и их перспективные приложения

5.1 Предварительные замечания

Механизмы сетевого содействия в гетерогенных сетях 5G позволяют эффективно интегрировать различные технологии радиодоступа, а также оперативно устанавливать прямое соединение между абонентскими терминалами. В частности, непосредственно управляя работой пользовательских устройств, поддержка со стороны сотовой системы 3GPP LTE упрощает контроль связи D2D в направлении ее инициализации и оптимизации прямых соединений, в том числе между «незнакомыми» терминалами. В данной главе рассматриваются практические функции управления D2D-взаимодействием в различных перспективных приложениях, а также приводятся соответствующие данные имитационного моделирования выгрузки трафика на соединения D2D и результаты работы, полученные в рамках модельной сети 3GPP LTE, управляющей системой WiFi-Direct.

Таким образом, применение связи D2D в сетях 5G позволяет эффективно разгружать сотовую систему доступа без дополнительных расходов на развертывание инфраструктуры мобильного оператора. При этом механизмы сетевого содействия помогают автоматизировать операции по обнаружению расположенных поблизости устройств/услуг, установлению и прекращению D2D-соединений, а также обеспечению безопасного взаимодействия. Поскольку радиотехнологии, пригодные для организации прямых соединений, получили наибольшее распространение в виде решений D2D в нелицензированном спектре, представляется особенно актуальным выяснение их практических ограничений с учетом реальных параметров пользователей и их контента. Использование при этом концепции сетевого кодирования позволяет существенно повысить доступность целевого контента на устройствах.

В данной главе рассматривается моделирование прямого взаимодействия абонентских терминалов с сетевым кодированием передаваемой ими информации. Для анализа такой системы предлагается новый подход, позволяющий охарактеризовать различные режимы распределения контента, а также сравнить

традиционные способы прямого обмена данными с методами, применяющими сетевое кодирование. В частности, демонстрируется, что даже простейшие формы сетевого кодирования позволяют значительно повысить степень доступности целевого контента на соседних абонентских терминалах и, тем самым, существенно увеличить долю выгруженного на соединения D2D трафика в рамках ряда реалистичных ограничений. Кроме того, даются некоторые практические рекомендации по их преодолению за счет кэширования и распределения данных с сетевым содействием.

В завершение описания перспективных приложений 5G данная глава рассматривает интеграцию сетей следующего поколения с системами удаленного хранения и обработки данных. В этой связи описываются новые пользовательские приложения и услуги 5G (например, использование средств дополненной и виртуальной реальности), а также предлагается их удобная классификация. Кроме того, прослеживаются тенденции дальнейшего развития архитектуры 5G, которые предполагают более тесное взаимодействие вычислительной и сетевой инфраструктур с учетом особенностей современного распределения контента. В продолжение анализа кэширования и распределения данных по D2D вводится понятие интегрированного узла 5G для вычисления и хранения с технологическим обоснованием его практической целесообразности.

5.2 Управление прямым соединением устройств в сотовых сетях

5.2.1 Контроль связи D2D через систему 3GPP LTE

Перспективные применения прямых соединений. Структура мобильного трафика на сегодняшний день такова, что значительная доля его потребления приходится на скачивание популярного контента, такого как видео-файлы или новые приложения [290]. Поэтому использование прямых соединений между близко расположенными пользователями становится предпочтительным для обмена такими локально востребованными данными, так как позволяет снизить нагрузку на сотовую сеть и сократить издержки мобильного оператора [291]. Использование более коротких и емких прямых соединений также вы-

годно и для самих пользователей [88]. Как следствие, применение связи D2D, которое началось как средство для разгрузки инфраструктурных соединений, становится сегодня одним из самых многообещающих направлений для создания новых мобильных приложений и сервисов в системах 5G (см. рисунок 5.1).

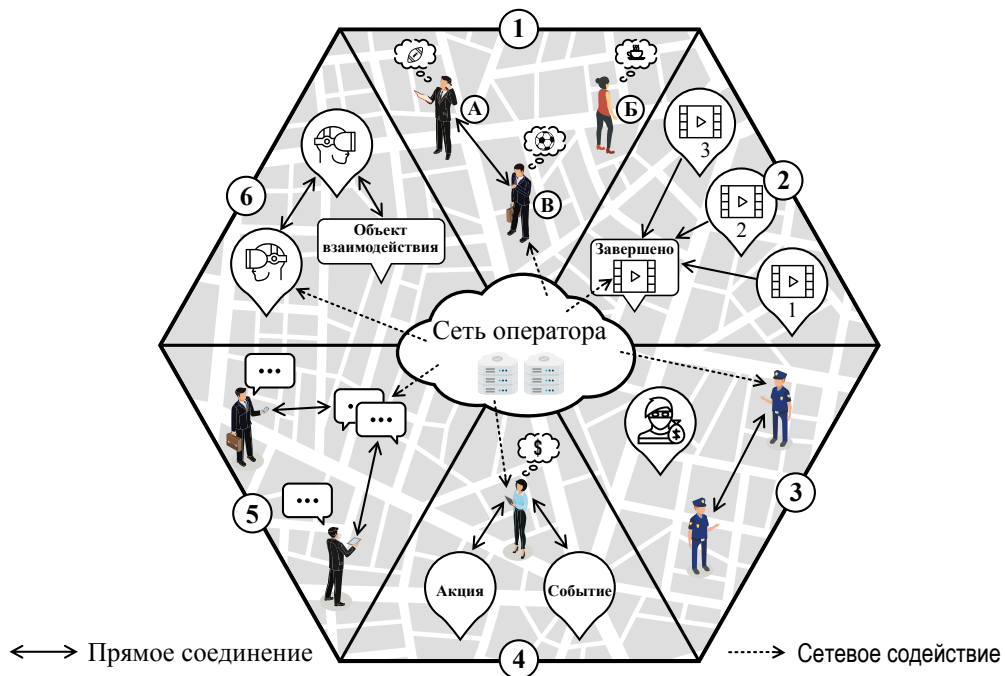


Рисунок 5.1 — Примеры использования прямых соединений D2D в сетях 5G

Обнаружение расположенных поблизости устройств/услуг. Такое применение связи D2D позволяет распространять информацию об идентификаторах абонентов, расположенных поблизости, для упрощения их последующего прямого взаимодействия. Данная группа функций связана, прежде всего, с повышением качества восприятия услуг пользователями в различных контекстах и предлагает широкие возможности по обнаружению подходящих партнеров для установления связи D2D. В частности, она осуществляет поддержку при восстановлении потерянных контактов и обнаружении «знакомых незнакомцев», у которых имеются схожие интересы, а также снабжает такие контакты необходимыми социальными рекомендациями.

Создание и распространение коллективного контента. Как следует из названия, соответствующие сценарии предполагают использование прямых соединений для загрузки контента и его обмена между абонентами, расположенными в непосредственной близости друг к другу. В данном случае речь идет, например, о всевозможной информации (фото, видео и т.д.), создаваемой

пользователями в рамках некоторого события или определенной деятельности. Это составляет общий контекст для всех взаимодействующих сторон, начиная с распространения мультимедиа-данных и вплоть до совместного выполнения корпоративных задач.

Профессиональные и промышленные сети связи. В данную категорию входят приложения, относящиеся к профессиональной деятельности или потребностям какого-либо местного бизнеса. Они могут нуждаться в оперативном установлении соединений и предъявлять повышенные требования к их безопасности. Поскольку связь D2D применяется, в том числе, и в промышленных целях, должны также выполняться требования регулирующих и законодательных органов. К примеру, при ликвидации последствий чрезвычайной ситуации, когда инфраструктура сотовой сети становится частично или полностью неработоспособной, соответствующие службы кризисного реагирования и послеаварийного восстановления могут продолжить свое взаимодействие по защищенным каналам прямой связи D2D.

Опережающее и предвидящее взаимодействие. Помимо прочего, система связи D2D предоставляет возможность опережающего оповещения своих пользователей о различных событиях, таких как специальные предложения и распродажи в торговых центрах, контекстная реклама и т.д. Более того, функционал прямого взаимодействия может представлять интересы самих абонентов, предвидя их намерения, и действовать от их имени (например, автоматизированный прием приглашений в группы, упреждающий обмен контентом и многое другое). Подобные продвинутые механизмы могут быть адаптированы под конкретные нужды и потребности владельцев устройств, предоставляющих возможность прямого взаимодействия.

Локальные социальные сети. Взаимодействие посредством прямых соединений дополнительно может служить основой для массового обмена информацией при организации досуга и отдыха. Действительно, если пользователи некоторой социальной сети достаточно локализованы в пространстве, то их контакты и коллективное общение может происходить поверх системы D2D. Данная категория применений дополнительно включает в себя всевозможные сервисы, использующие данные о локации, а также приложения, ответственные за распространение контента в ходе общественных мероприятий. Например, запись мгновенных повторов со спортивного соревнования может транслироваться по D2D всем заинтересованным пользователям, расположенным поблизости.

Проксимальные игры и другие не прямые взаимодействия. Наконец, системы D2D предоставляют дополнительные возможности для поддержки широкого круга сценариев, включающих в себя проксимальные игры, которые действуют как расположенных поблизости людей, так и всевозможные интерактивные объекты. Игровое взаимодействие становится все более привычным и массовым, так как позволяет его участникам избегать вступления в непосредственные социальные отношения или идентифицировать друг друга. Ожидается, что подобное общение в игровой форме станет существенно проще при использовании средств связи D2D, позволяя получить более богатый опыт от проксимальных игр и других не прямых взаимодействий.

Как обсуждалось в главе 3, технология связи D2D позволяет значительно расширить спектр возможностей для взаимодействия расположенных поблизости пользователей. Однако для ее практической реализации требуются достаточно простые и надежные механизмы управления, которые позволят встроить поддержку прямых соединений как неотъемлемую часть системы 5G [292]. Преимущества использования D2D-соединений включают в себя более высокие скорости передачи данных, сниженную задержку доступа и повышенную энергетическую эффективность, поэтому в сетях 5G такой режим взаимодействия является крайне желательным. Однако в прошлом для обеспечения работы прямых соединений были доступны лишь некоторые базовые средства поддержки, относящиеся к передаче устойчивого к задержке трафика, а общесистемный контроль за качеством соединения абонента отсутствовал.

Расширяя возможности имеющихся систем управления прямым взаимодействием, повсеместное сотовое покрытие можно использовать для обеспечения сетевого содействия при установлении и поддержании D2D-соединения, которое может существовать в двух вариантах [293]: (1) в качестве надстройки технологии 4G [294] в лицензированном (например, сотовом) спектральном диапазоне или (2) в качестве альтернативной (не сотовой) технологии в нелицензированном (например, ISM) спектральном диапазоне. Стоит напомнить, что первый вариант может быть реализован на практике в виде «подложки» сотовой системы связи, когда передача D2D осуществляется на тех же радиоресурсах, что и традиционное обслуживание пользователей, или в виде «обертки» сотовой сети, когда инфраструктурные и D2D-соединения используют отдельные радиоресурсы.

Оба варианта реализации D2D-соединений были в значительной степени исследованы в научно-технической литературе, а также рассматривались комитетом 3GPP, начиная с Release 12 и далее. При этом поддержка прямого взаимодействия спецификациями системы LTE изначально проводилась в контексте обнаружения расположенных поблизости устройств для служб общественной безопасности [295; 296], что существенно ограничивало практическое применение. В результате разработка сотовой технологии D2D, называемой также LTE-Direct, заняла существенное время, что создало предпосылки для появления альтернативных решений. Одним из них является возможность установления прямых соединений на нелицензированных частотах при помощи технологии IEEE 802.11 (WiFi), а сотовая система связи обеспечивает при этом функции сетевого содействия.

Большинство современных мобильных устройств уже имеет возможность одновременного использования систем LTE и WiFi. Если абонентские терминалы непрерывно ассоциированы с сотовой сетью, то они могут использовать это подключение в качестве управляющего при организации работы прямых D2D-соединений на основе технологии WiFi. Другими словами, сотовое соединение предоставляет возможность быстрого и эффективного обнаружения соседних устройств, их оперативного оповещения о результатах обнаружения, а также автоматизации процесса установления соединения. С выходом в свет более современных протоколов IEEE 802.11, таких как WiFi-Direct для взаимодействия в нелицензированном спектральном диапазоне без инфраструктуры ТД WiFi, а также IEEE 802.11ad/ay для передачи данных на миллиметровых волнах с повышенными скоростями, рассмотрение такого способа организации D2D-соединений становится все более привлекательным.

Функции управления D2D-взаимодействием. На сегодняшний день сотовые системы связи распространены настолько широко, что часто предполагается повсеместное сетевое покрытие. Наряду с этим, работа ключевого протокола IP неявно подразумевает, что взаимодействующие устройства имеют соответствующие адреса, привязанные к топологии сети. В отличие от такого традиционного подхода соединения D2D должны быть организованы принципиально иначе в силу своей непостоянной природы. Как следствие, обеспечение поддержки прямого взаимодействия приводит к пересмотру некоторых основополагающих принципов построения сетей связи. Таким образом, функционал системы D2D должен динамически перестраивать ее топологию для обеспече-

ния текущих требований обслуживаемых системой приложений, не создавая при этом трудностей пользователям.

В отличие от традиционных подходов, в которых сетевая архитектура предоставляется в качестве некоторого статического «сервиса» для обслуживаемых ей приложений, система D2D с сетевым содействием имеет возможность в явном виде учитывать фактическую социальную структуру определенной группы абонентов [297]. Как показано на рисунке 5.2, так называемая социальная плоскость воспроизводит характер общения пользователей в реальном мире в рамках общего физического пространства ①, которое как раз и мотивирует использование прямых соединений. Однако такие взаимоотношения между абонентами остаются скрытыми от традиционных телекоммуникационных сетей, так как последние не располагают подобными сведениями в силу ограничений современных технологий связи. С другой стороны, социальные сети имеют представление о фактических взаимоотношениях между своими пользователями, что может быть представлено в виде социального «отражения» соответствующих контактов ②.

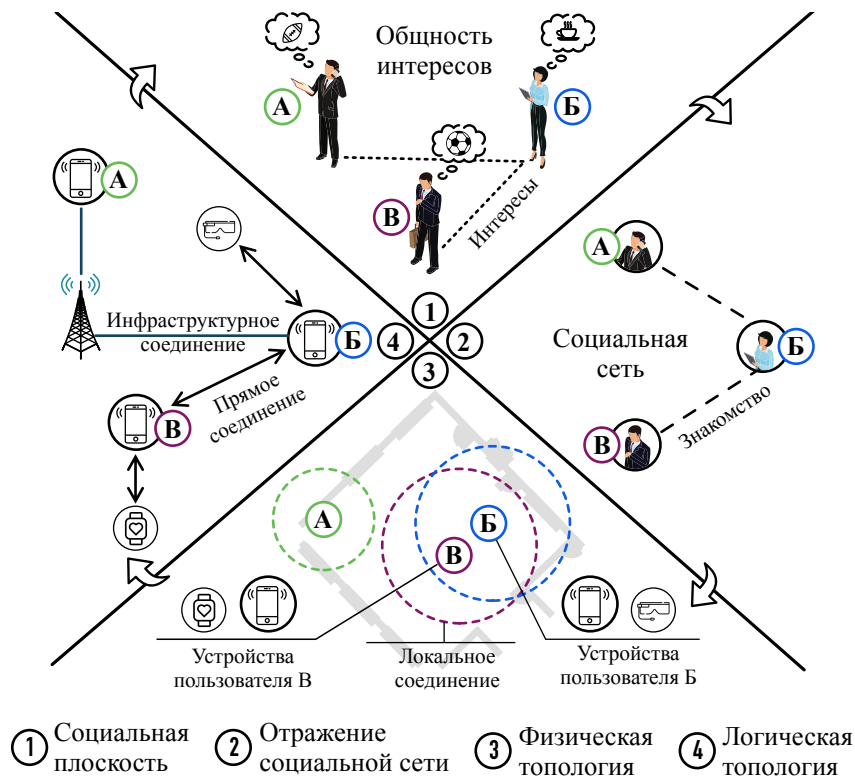


Рисунок 5.2 — Принципы организации прямого взаимодействия

Предлагаемый здесь подход к построению системы D2D позволяет ее абонентам выборочно предоставлять свое социальное «отражение» 3GPP-инфра-

структуре проксимальных услуг ProSe¹, разрешая ей «видеть» значительно дальше своих исходных возможностей. В рассматриваемой архитектуре сети каждое подключенное устройство представляет собой не только модуль SIM, но и характеризуется набором идентификаторов социальных сетей, позволяющих установить взаимосвязь между множеством взаимодействующих объектов, как это показано на уровне физической топологии ③. Наконец, используя сведения о некоторой физической топологии, обслуживаемые мобильные приложения могут взаимодействовать со своими целевыми социальными партнерами при помощи содействия со стороны сотовой функции ProSe ④.

В итоге обмен информацией происходит непосредственно по соединениям D2D, которые устанавливаются для этой цели по необходимости. Таким образом, логическая топология сети может быть построена по требованию так, чтобы возможности составляющих ее абонентских терминалов использовались бы в полной мере. В частности, предлагаемый подход позволяет отобразить социальные взаимодействия на логическую топологию сети. На этапе сетевой регистрации (см. рисунок 5.3) абонентский терминал авторизуется на соответствующем сервере D2D (например, с использованием протокола oAuth) и использует его доменное имя как часть своего профиля при взаимодействии с различными серверами приложений. При этом список используемых серверов приложений может храниться в системных настройках терминала. Если терминал переводится на другой сервер D2D, данная информация обновляется.

Публикация и поиск данных/услуг P2P. Когда пользователь намеревается получить доступ к определенным данным или услугам P2P, он авторизуется на стороннем сервере приложений, который предоставляет функционал отслеживания контента, и осуществляет поиск местоположения желаемых данных или услуг. Напротив, когда абонент имеет возможность предоставить определенный P2P-контент в общий доступ, он также авторизуется на стороннем сервере приложений и публикует соответствующую информацию. В качестве альтернативы пользователь также имеет возможность загрузки целевого контента в облачное хранилище, размещенное на сервере приложений (не показано на диаграмме). Когда абонент авторизуется на сервере приложений и осуществляет поиск определенных данных/услуг P2P, модуль отслеживания контента может составить

¹Функционал ProSe представляет собой способ реализации связи D2D в рамках технологии 3GPP LTE.

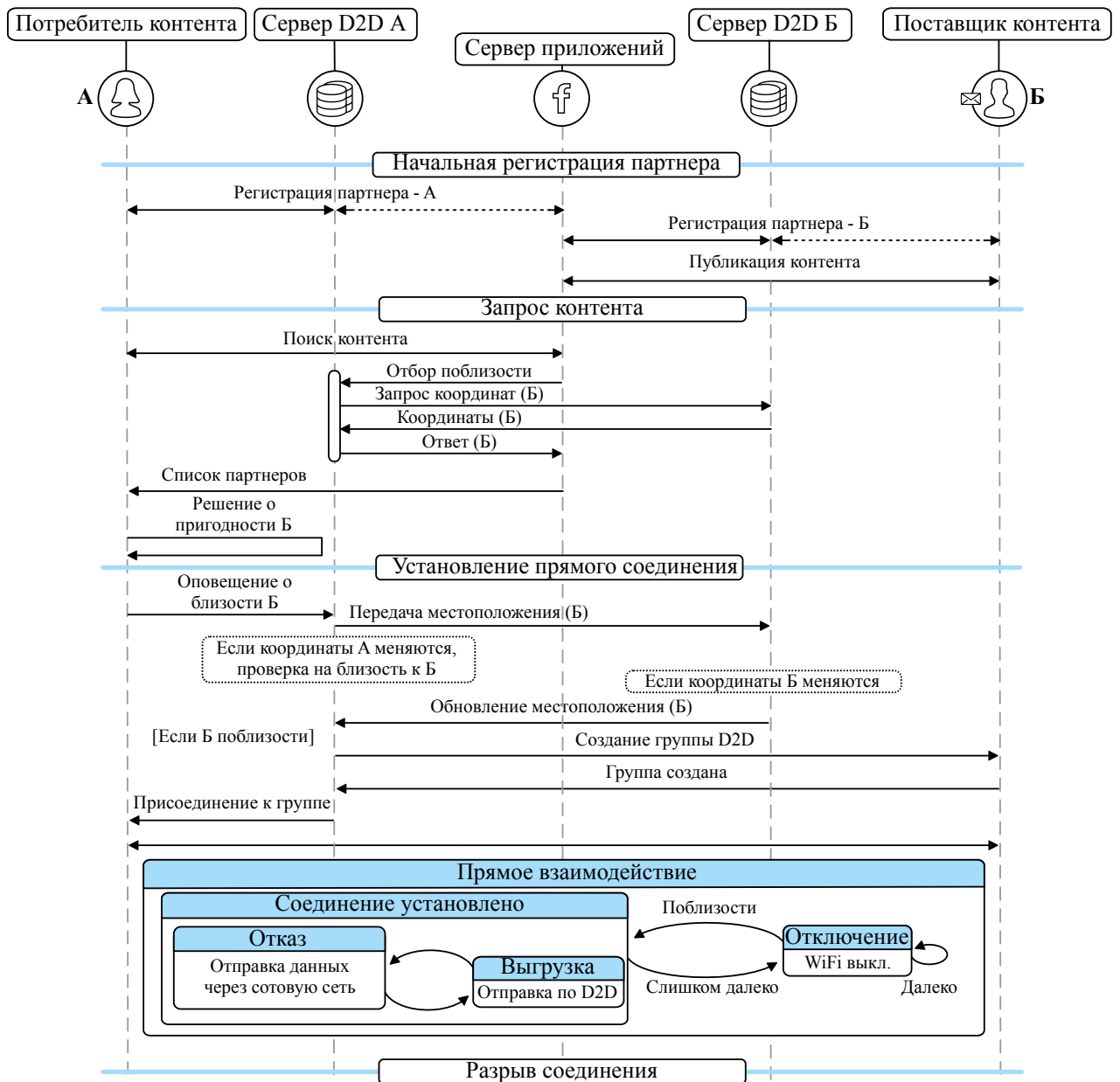


Рисунок 5.3 — Предлагаемый протокол для работы процедуры ProSe

список возможных местоположений контента на основании критериев, заданных пользователем.

Например, если пользователь предпочитает получать доступ к P2P-контенту посредством соединений D2D в нелицензированном спектре, то производится фильтрация списка местоположений контента с выявлением только тех вариантов доступа, которые предоставляют требуемую возможность и расположены поблизости. Для осуществления такой фильтрации сервер приложений должен вначале затребовать информацию от D2D-сервера, который обслуживает запрашивающего абонента, относительно предпочтительных на-

строек связи D2D и близости к различным поставщикам контента. Если такой потенциальный поставщик не управляется тем же самым D2D-сервером, то последний связывается с другим D2D-сервером, который обслуживает поставщика контента, и запрашивает соответствующую информацию. В сущности, взаимодействие D2D-серверов позволяет серверам приложений определять, находятся ли целевые абоненты достаточно близко друг к другу без необходимости получения их точного географического местоположения. Такая возможность важна для системы, поскольку пользователи зачастую предпочитают не раскрывать подобные сведения.

Установление и прекращение D2D-соединений. Когда пользователь переходит по ссылке от сервера приложений, открывающей доступ к запрашиваемому контенту, размещенному на стороне некоторого поставщика (D2D-партнера), соответствующий абонентский терминал обращается к своему D2D-серверу для оказания содействия при установлении соединения. При этом в зависимости от информации, доступной от служб определения местоположения пользователей, D2D-сервер принимает решение относительно предпочтительного режима работы сессии P2P (т.е. инфраструктурное или D2D-соединение). Пока P2P-сессия остается активной поверх соединения D2D, соответствующий D2D-сервер отслеживает текущее местоположение партнеров, а также свойства их D2D-соединения, и определяет, следует ли перевести данную сессию обратно на инфраструктурное соединение или нет.

Рассмотренный процесс на стороне сервера D2D может быть достаточно просто реализован с помощью конечного автомата, который включает в себя соответствующие механизмы гистерезисного управления при принятии решения относительно выгрузки некоторой сессии на D2D-соединение или же перевода ее на обслуживание сетевой инфраструктурой [172]. Эти механизмы предотвращают излишние попытки подключения и отключения радиоинтерфейсов из-за незначительных изменений состояния канала связи или местоположения терминалов. Процедура выгрузки сессии может быть реализована посредством техники «внедрения маршрута» или же более традиционными способами, предоставляемыми протоколом мобильного IP. При прекращении D2D-соединения абонентские терминалы и их серверы D2D информируются соответственно, причем детали этого процесса зависят от его реализации.

5.2.2 Общесистемное исследование работы прямых соединений

При функционировании системы D2D с сетевым содействием абонентские терминалы приобретают возможность определять, что они находятся в непосредственной близости друг к другу для установления прямого соединения и последующей выгрузки на него своей P2P-сессии, минуя сотовую инфраструктуру. Такой режим работы, построенный на основе описанного выше протокола, позволяет существенно повысить емкость сети, а также увеличить скорость передачи данных терминалом и его энергетическую эффективность. Дальнейшее изложение сосредоточено на выявлении количественного прироста производительности при выгрузке трафика из сети 3GPP LTE на прямые соединения, использующие технологию WFD (WiFi-Direct) в рамках ряда практических сценариев.

Для этих целей был разработан специализированный комплекс программ имитационного моделирования, построенный на основе актуальной методологии 3GPP LTE и текущих спецификаций IEEE 802.11. Он представляет собой гибкий программный инструмент, позволяющий моделировать различные способы расположения узлов сети, типы передаваемого трафика, особенности беспроводного канала связи, а также альтернативные протоколы радиодоступа. Данный комплекс предоставляет возможность изучения практических режимов взаимодействия между инфраструктурными и прямыми соединениями, а также исследования методов управления интегрированной системой LTE-D2D в случаях типового размещения взаимодействующих объектов. С его помощью далее демонстрируются характеристики работы рассматриваемой системы D2D с сетевым содействием.

Типовой сценарий применения связи D2D. Для иллюстрации возможных выигрышей от использования технологии WFD с содействием со стороны сети LTE далее воссоздается типовой сценарий, отражающий современные городские условия. В частности, согласно рекомендациям ITU/3GPP моделируется сотовое покрытие типа «городские малые соты» (UMi) и рассматривается достаточно высокая плотность размещения абонентов. Такой способ расположения узлов отражает ситуацию, в которой выгрузка трафика с инфраструктурных на D2D-соединения оказывается наиболее полезной, т.е. когда сотовая сеть испытывает затруднения при обслуживании многочисленных сессий, устанавли-

ливаемых пользователями. В рассматриваемой среде плотность расположения абонентских терминалов такова, что для каждого из них вероятность оказаться внутри зоны покрытия D2D-соединения какого-либо соседнего терминала достаточно высока.

В отсутствие явной информации о фактическом «спросе и предложении» при обмене P2P-контентом, здесь предполагается, что имеется некоторая доля x абонентских терминалов, запрашивающих некоторые данные/услуги P2P, которые находятся внутри зоны покрытия радиотехнологии D2D терминала-поставщика этого контента. Такой подход позволяет исследовать рассматриваемую систему без необходимости сужения до определенного способа обнаружения соседей или распространения контента и, тем самым, сравнивать соответствующие характеристики производительности с аналогичными параметрами работы сотовой сети LTE (без возможности установления связи D2D). Другими словами, считается, что все абоненты системы вовлечены в P2P-взаимодействие со своими партнерами, но только часть возможных партнеров находится в радиусе действия связи D2D.

Инфраструктура сети LTE состоит из 19 шестиугольных сот (см. рисунок 5.4), а расстояние между соседними БС составляет 200 м, так что радиус каждой соты приблизительно равен 110 м. Также используется подход с «оберткой» сотами для увеличения точности моделирования на границах рассматриваемой зоны покрытия. Система оперирует двумя полосами частот по 10 МГц, которые являются общими для всех сот, имеющих по три сектора каждая, и работает в режиме частотного дуплексирования (для передачи по линии вверх и вниз соответственно). Абонентские терминалы располагают радиомодулями LTE и WiFi, а также способны использовать технологии радиодоступа LTE и WFD одновременно [291]. Они равномерно распределены внутри зоны покрытия сотовой сети и ассоциированы с БС на основе своего значения SINR по линии вниз (около 20 терминалов на сектор). Пользователи остаются статичны в течение времени моделирования. Каналы связи моделируются с учетом необходимых свойств источника, получателя и среды распространения радиосигнала.

Каждая БС подключена к ядру сети и предоставляет сотовый доступ всем ассоциированным с ней абонентским терминалам. У каждого терминала имеется источник трафика, позволяющий генерировать несколько его типов. Для простоты в последующих примерах предполагается насыщенный трафик с па-

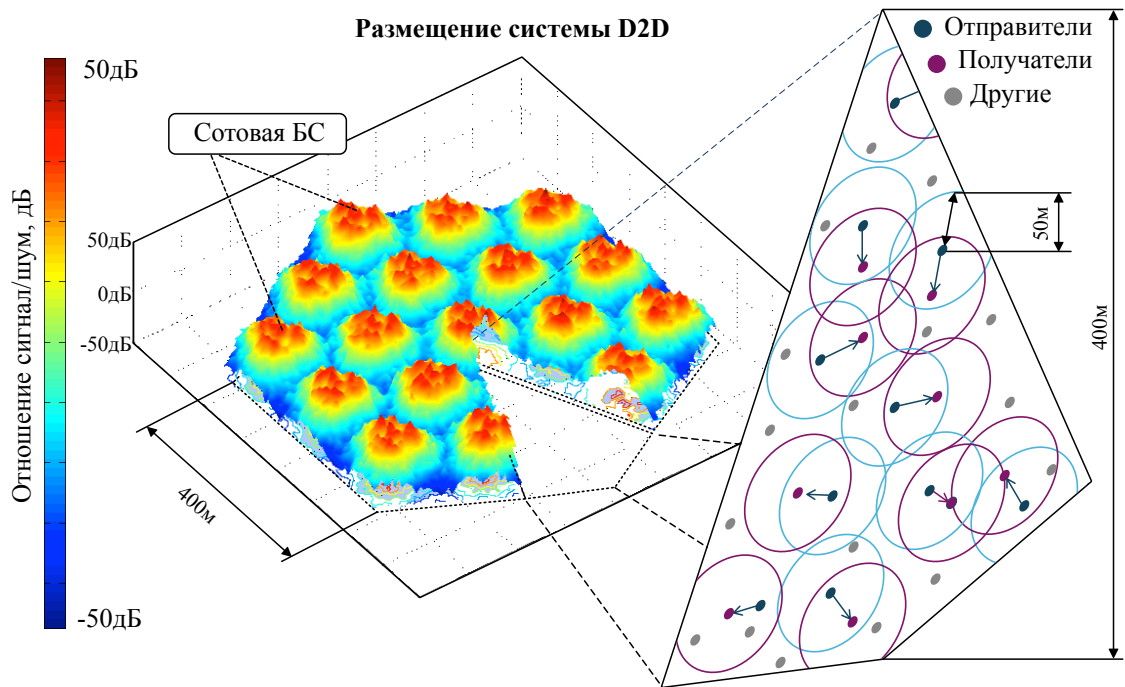


Рисунок 5.4 — Моделирование работы интегрированной системы LTE-D2D

кетами данных по 1500 байт. Дополнительные подробности относительно конфигурации рассматриваемой системы LTE приведены в таблице 7, а также в соответствующих документах стандартизации (например, 3GPP TR 36.814-900 и ITU-R M.2135-1). В целях верификации работы сотовой сети доступа был также реализован калибровочный сценарий из документа 3GPP TR 36.814-900, таблица A-2.1 и проведено проверочное моделирование. Полученные таким образом имитационные результаты хорошо согласуются с целевыми значениями спектральной эффективности на краю и в центре соты.

Когда абонентские терминалы используют технологию WFD, им приходится не только конкурировать за возможность передачи, но также сосуществовать с другими устройствами, использующими традиционную связь WLAN при помощи инфраструктурных ТД. Предполагается, что сторонние пользователи WLAN не ассоциированы с сотовой системой связи, поэтому их активность на нелицензированных частотах WiFi не может наблюдаться или управляться сетью LTE. Здесь и далее такие устройства называются неконтролируемыми. Неконтролируемые терминалы также передают насыщенный трафик с пакетами данных по 1500 байт, но эта информация всегда предназначена обслуживающей их ТД (т.е. они не используют технологию WFD). Для упрощения сценария моделирования передача данных по линии вниз в системе WiFi не рассматрива-

Таблица 7 — Параметры моделирования работы прямых соединений

Параметр	Значение/Источник
LTE	
Модель распространения	ITU-RM.2135-1, таблица A.2.2-1, A1-3
Модель замирания	ITU-RM.2135-1, раздел 1.3.1.1
Доступ к среде	Кольцевое планирование
Контроль мощности и скорости	Целевое SINR 15 дБ в замкнутом цикле
Частотные радиоресурсы	10 + 10 МГц, частотное дуплексирование
Структура кадра	2 из 20 специальных подкадров, кадр 10 мс
Приемопередатчик	ITU-R M.2135-1, таблица 8-4
Конфигурация антенны	1x2 (разнесенный прием на БС)
WiFi	
Модель распространения	Эмпирическая
Модель замирания	Только корреляция
Доступ к среде	CSMA/CA, порог доступа -76 дБм
Контроль мощности и скорости	Целевое SINR 25 дБ в открытом цикле
Частотные радиоресурсы	20 МГц с разделением времени
Структура кадра	Контрольные сообщения на 18 Мбит/с
Приемопередатчик	Фактор шума 7 дБ, порог шума -95 дБм
Конфигурация антенны	1x1 (одна антенна)

ется. Вместо этого изменяется число неконтролируемых терминалов так, чтобы получить различный уровень состязания за доступ к каналу связи WiFi.

Данное исследование предполагает, что все соединения WiFi (как традиционные с ТД, так и D2D) используют общие частотные радиоресурсы и, таким образом, должны уступать любой активной передаче, для которой мощность принятого сигнала превышает некоторый заданный порог (согласно правилам работы протокола IEEE 802.11). Также считается, что все ТД и подключенные к ним неконтролируемые терминалы используют ту же версию протокола доступа, что и WFD-терминалы, т.е. IEEE 802.11-2012. Для воссоздания реальных условий работы сети неконтролируемые терминалы сгруппированы вокруг обслуживающих их ТД. При этом сами ТД могут быть расположены произвольно внутри области моделирования, что соответствует WiFi-зонам в кафе, на вокзалах и т.д. Дополнительные подробности относительно конфигурации сети WiFi приведены в таблице 7. В целях калибровки результатов моделирования применялись полученные ранее данные из статей по исследованию работы системы WLAN [154].

Интерпретация полученных результатов. Результаты имитационного моделирования для общей емкости рассматриваемой сети представлены на рисунке 5.5. При этом переданные сессиями LTE и WFD данные суммируются в расчете на соту, связанную с запрашивающим абонентским терминалом. Видно, что выгрузка трафика из системы LTE на соединения WFD существенно повышает скорость передачи, увеличивая ее вплоть до четырех раз на уровне разгрузки в 30%. И чем больше данных выгружается, тем значительнее становится выигрыш. При наличии интерферирующих неконтролируемых терминалов (пять устройств на ТД в среднем) прирост скорости передачи снижается до уровня выигрыша в 2,5 раз при выгрузке в 30%. Полученные результаты также показывают, что соединения D2D имеют лучшую производительность, когда доля разгрузки невелика, а затем их работа ухудшается по мере увеличения числа выгружаемых на них сессий. Это происходит, в основном, из-за более сильной конкуренции между соединениями D2D, а также по причине роста общего уровня шума.

Рисунок 5.6 представляет результаты более детального сравнения значений скорости передачи для WFD- и LTE-соединений в рамках одного и того же сценария. Соответственно, качество соединений WFD сильно варьируется в зависимости от их длины, а также подвержено существенному влиянию со

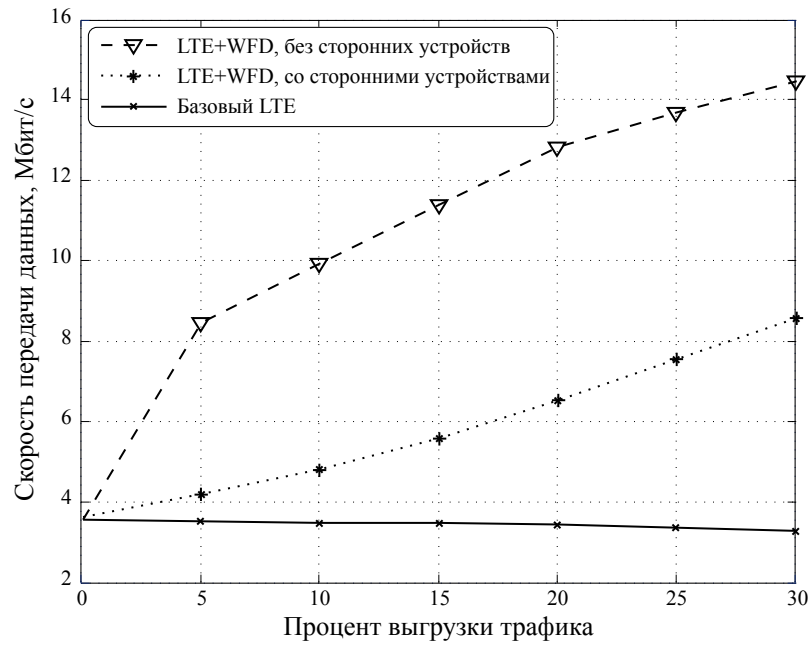


Рисунок 5.5 — Емкость сотовой + D2D сети в условиях насыщения

стороны соперничающих неконтролируемых терминалов. Однако на типовых дальностях работы WFD (т.е. до 50 м) такие соединения демонстрируют скорость передачи данных в несколько раз выше, чем через LTE. Учитывая это явное преимущество, можно сделать вывод о целесообразности использования радиотехнологии WFD при любой возможности, когда взаимодействующие терминалы находятся достаточно близко друг к другу, если только система не подвержена чрезмерному влиянию интерференции [298].

Поскольку энергетическая эффективность измеряется в битах на Джоуль, она зависит от типа используемой технологии радиодоступа. Здесь она вычисляется в предположении уровня мощности, связанного с потреблением электроцепи, в 100 мВт, а мощности приема – 200 мВт и 100 мВт плюс некоторая переменная составляющая при передаче. В этой связи рисунок 5.7 демонстрирует, что работа системы WiFi приводит к значительно более высокой энергетической эффективности по сравнению с LTE. Данное наблюдение справедливо благодаря более широкой полосе пропускания, доступной радиотехнологии WiFi. Кроме того, система LTE, как правило, назначает абонентским терминалам небольшие фрагменты радиоресурсов в разрозненных частотно-временных интервалах, что приводит к расходу энергии приемопередатчика в течение достаточно длительных периодов времени.

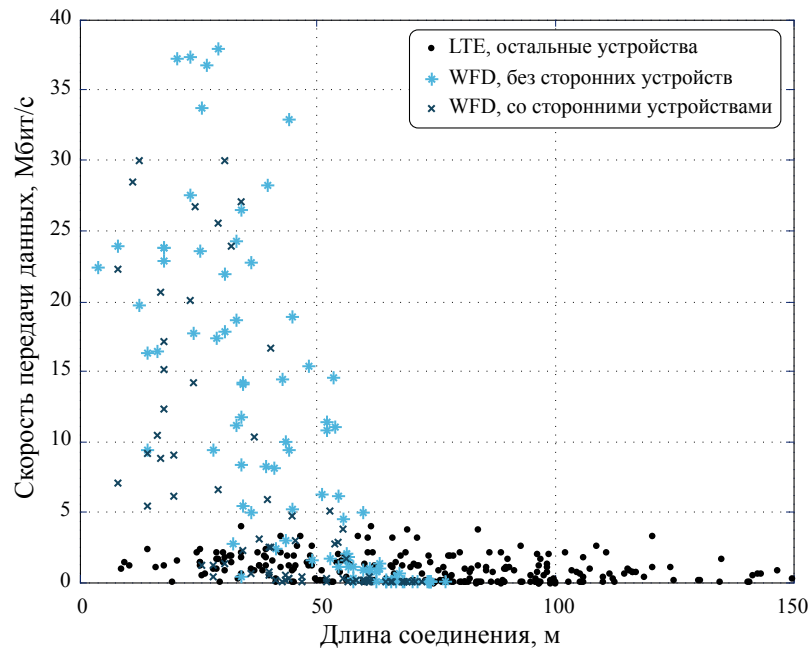


Рисунок 5.6 — Скорость передачи при выгрузке в 30% для насыщенной сотовой + D2D сети

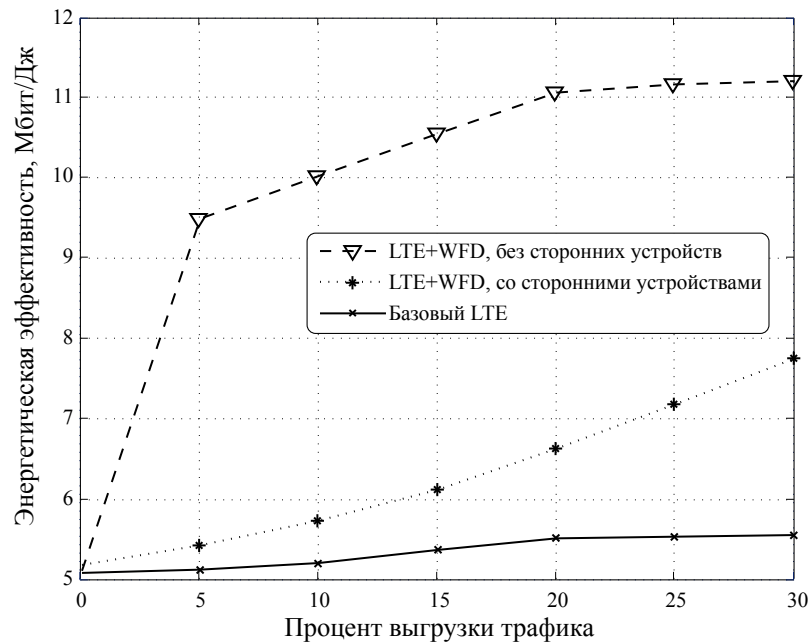


Рисунок 5.7 — Энергетическая эффективность соты для насыщенной сотовой + D2D сети

Напротив, протокол управления сетью WiFi активизирует приемопередатчик терминала исключительно в моменты времени, связанные с его доступом к среде. Когда пользователи WiFi вынуждены откладывать доступ к каналу

согласно информации, содержащейся в сообщениях RTS и CTS, они имеют возможность переходить в энергосберегающее состояние до начала приема или передачи. При этом в момент получения непосредственного доступа к каналу активный абонентский терминал использует всю его полосу пропускания. В результате, несмотря на значительное число активных абонентов в рамках рассматриваемой системы в целом, только некоторые из них фактически потребляют энергию, связанную с приемом или передачей данных, в то время как остальные могут находиться в энергосберегающем состоянии с низкой потребляемой мощностью.

Проведенное исследование подтверждает, что имеется существенный потенциал для повышения производительности как сети в целом, так и обслуживаемых ею абонентских терминалов при выгрузке сотового трафика на соединения WFD в городских условиях работы. Поскольку значительная доля ожидаемого роста мобильного трафика вызвана функционированием социальных приложений и услуг между абонентами, расположенными в непосредственной близости, применение подобных механизмов разгрузки сотовой инфраструктуры позволяет значительно увеличить фактическую емкость сети и качество пользовательского соединения. Для подтверждения этого важного факта на практике далее разрабатывается прототип реальной системы D2D с сетевым содействием, а также проводятся измерения характеристик ее работы в динамически меняющихся условиях применения.

5.2.3 Прототипирование системы выгрузки на соединения D2D

Организация связи WiFi-Direct при содействии LTE. Для реализации предложенного выше подхода к управлению прямыми соединениями и подтверждения целесообразности его использования функционал связи D2D встраивается в архитектуру системы 3GPP LTE. В частности, рассматривается выгрузка сотового трафика на управляемые соединения WFD и приводятся соответствующие характеристики работы системы. Для интеграции возможности установления прямых соединений использовалась экспериментальная сеть LTE [291], поддерживающая 3GPP Release 10 в следующем составе: (1) сеть радиодоступа (RAN), включающая в себя одну наружную и четыре внутренние

соты, работающие на частотах 700, 1800 и 2600 МГц; (2) ядра сети (ЕРС) и (3) мультимедиа-подсистемы IP (IMS), предоставляющие технологию передачи голоса через LTE (VoLTE).

Для реализации протокола прямого взаимодействия, описанного в главе 3, был разработан прототип системы связи D2D с сетевым содействием. Он включает в себя модуль отслеживания контента, позволяющий пользователям открывать некоторые данные/услуги для коллективного доступа, публикуя их в качестве ссылок вида `d2d://user@domain@prose_server/http:8080/path`. Такие ссылки используются сервером D2D для выяснения идентификаторов как абонентских терминалов, так и предоставляемых ими ресурсов (каталога с файлами, веб-камеры и т.д.). Модуль отслеживания контента представляет собой модифицированный веб-сервер со специализированным PHP-приложением, который конструирует ссылки для доступа к контенту и содержит базу данных SQL для хранения информации. Более подробное описание различных компонентов разработанного прототипа приведено в [291].

В частности, сервер D2D размещен в сети Интернет и осуществляет содействие в процессе установления D2D-соединения и выгрузки трафика. В качестве альтернативы можно было бы поместить D2D-сервер внутри ядра сети ЕРС, но это потребовало бы разрешения со стороны мобильного оператора. В разработанном прототипе сервер D2D реализован как веб-приложение на языке Python, позволяющее устанавливать связь посредством протокола HTTP как с абонентскими терминалами, так и с модулями отслеживания контента. При этом терминалы представляют собой смартфоны Android-4, снабженные функционалом сетевого содействия соединениям D2D, который предоставляет им возможность взаимодействия с D2D-сервером. Оно достигается путем изменения таблицы маршрутизации в ядре операционной системы терминала для сессий, устанавливаемых с D2D-партнерами так, что когда прямое соединение становится доступно, сессии перенаправляются на него автоматически.

В результате выгрузка на D2D становится «прозрачной» для терминала, т.е. все операции производятся без вмешательства пользователя. После авторизации абонента модулем отслеживания контента возможен поиск среди доступного контента, по итогам которого пользователю предоставляются все доступные местоположения выбранных данных/услуг (включая свойства поставщика контента, такие как возможность связи по D2D). После выбора абонентом определенного поставщика (местоположения) целевого контента его терминал

связывается с D2D-сервером и запрашивает IP-адрес поставщика для установления прямого соединения. Сервер D2D высылает в ответ текущий IP-адрес, а также информирует соответствующее приложение поставщика контента о возможности начала сессии. В случае запроса контента через соединение D2D, когда поставщик контента находится в непосредственной близости, D2D-сервер содействует терминалам-партнерам при установлении прямой сессии.

Рассматриваемая экспериментальная сеть LTE была снабжена всеми необходимыми модулями для поддержки прямой связи согласно описанию выше и в рамках текущих спецификаций 3GPP, именующих подобный функционал ProSe (см. рисунок 5.8). В этой связи была разработана собственная реализация архитектуры системы D2D с сетевым содействием, в которой роль D2D-сервера выполняет модуль e-SMLC. Он был размещен не внутри ядра сети EPC, а в подсистеме IMS как некоторая виртуальная сущность, по сути, являющаяся разновидностью IP-сервиса с возможностью доступа к функциям ядра сети. Далее приводится краткое описание основных шагов, связанных с реализацией описанной системы, а также основные выводы по итогам наблюдения ее практической работы.

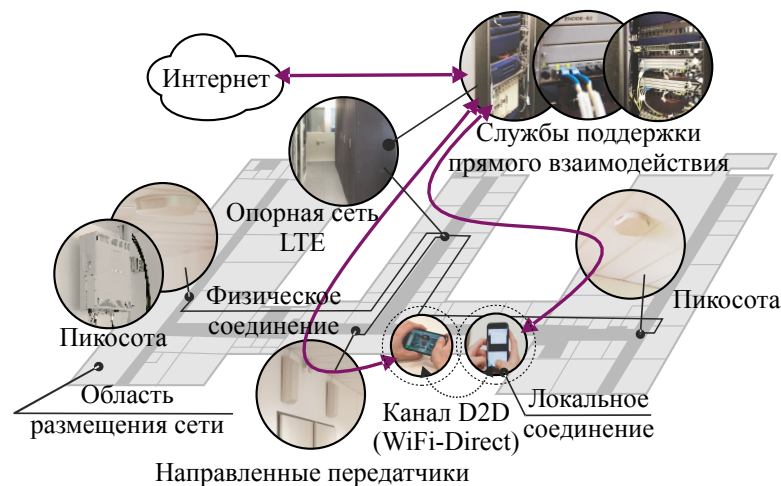


Рисунок 5.8 — Прототип системы связи WiFi-Direct с сетевым содействием

Влияние сети радиодоступа на D2D Использование возможностей RAN необходимо для обеспечения работы любых приложений и услуг мобильной связи, поскольку она отвечает за передачу данных и осуществляет управление. При этом прямые соединения оказываются не слишком чувствительными к свободной емкости сети доступа, так как управляющие сообщения являются достаточно короткими, а выгрузка данных на D2D-соединения как раз приво-

дит к снижению загрузки RAN. Однако передача управляющих сообщений D2D должна стать приоритетной по отношению к стороннему трафику RAN в целях снижения времени реагирования [299]. Кроме того, службы определения местоположения, располагающиеся внутри сети доступа, являются ключевыми при организации прямого взаимодействия. Их работа основана на взаимодействии БС сети с другими элементами ЕРС.

Преодоление ограничений связи по IP. Вообще говоря, прямое взаимодействие требует возможности передавать пакеты данных между целевыми устройствами без вмешательства промежуточных узлов на транспортном уровне и выше. Как правило, сотовые сети связи используют зарезервированные диапазоны IP-адресов для подключенных к ним абонентских терминалов, что теоретически не накладывает дополнительных ограничений, пока взаимодействующие D2D-партнеры находятся внутри сети одного мобильного оператора. Однако работа межсетевых экранов может помешать установлению непосредственного соединения между абонентскими терминалами, затрудняя тем самым функционирование системы D2D. На практике ограничения, связанные с присутствием межсетевых экранов, должны быть преодолены, что может потребовать изменений в правилах их работы.

Взаимодействие между D2D-серверами и терминалами. По своему построению система связи D2D с сетевым содействием изначально полагается на способность сотовой сети оперативно управлять абонентскими терминалами, вовлеченными в прямое взаимодействие. Такая способность должна быть реализована эффективно, с целью инициации D2D-соединений в кратчайшее время. Для этого, например, может использоваться протокол инициации сессий SIP, что, однако, потребует наличия активного радиопотока. В качестве альтернативы возможно более экономичное управление на уровне NAS, которое, в свою очередь, применяется для создания новых радиопотоков, но становится сложнее для реализации на практике.

Интеграция со службами определения местоположения. Для своей эффективной работы системе D2D требуется регулярное обновление сведений относительно фактического местоположения абонентских терминалов. В сети LTE подобная информация традиционно обрабатывается модулем e-SMLC, к которому может быть предоставлен безопасный доступ. Однако конкретный способ получения таких сведений может варьироваться от использования GPS-механизмов позиционирования до внедрения специализированных опорных сим-

волов непосредственно в структуру LTE-кадра и последующей триангуляции. Для целей данного прототипа [291] был использован дополнительный интерфейс между сервером D2D и модулем управления мобильностью ММЕ, позволяющий получать оперативную информацию о местоположении активных устройств. Вне зависимости от выбранного способа реализации текущие координаты абонентских терминалов могут быть отслежены во всех современных системах связи.

Управление соединением D2D. Наконец, в системе D2D требуется наличие эффективных средств контроля. В частности, передача управляющих сообщений может происходить при помощи выделенного сервиса, запущенного в качестве приложения на абонентском терминале, тогда как возможности для фактического исполнения полученных команд будут непосредственно интегрированы в драйверы ядра операционной системы. На практике для большинства имеющихся мобильных платформ (Android, большая часть систем на основе Linux и т.д.) это означает необходимость в использовании специализированной прошивки смартфона или же получении административных привилегий другими способами. Для других платформ, таких как iOS и Windows Mobile, использование подобных решений крайне затруднительно без соответствующей поддержки со стороны производителей оборудования.

Практическая демонстрация технологии D2D. Приведенное выше описание характеризует работоспособную систему D2D, управляемую сетью 3GPP LTE мобильного оператора. Соответствующая экспериментальная демонстрация пригодна для обеспечения следующих услуг: передача данных, связь VoIP поверх интегрированной сети LTE-WiFi и др. Также поддерживается функционал IMS. В ходе реализации были использованы актуальные спецификации 3GPP в рамках ограничений имеющегося практического оборудования [300]. Далее в дополнение к проведенной выше оценке производительности системы предлагается исследование эффективности ее работы по передаче управляющих сообщений в ходе обеспечения сетевого содействия. Такие результаты могут быть получены только при помощи реального прототипа, так как соответствующее имитационное моделирование затруднительно. Время установления прямого соединения выбрано в качестве целевого показателя эффективности.

Для достижения адекватной точности измерений разработанный протокол связи D2D был разделен на ряд этапов. Задержка передачи на сотовом

соединении замерялась между абонентским терминалом и сервером D2D при различных уровнях загрузки сети. Кроме того, измерялось время установления соединения D2D через WiFi-Direct между двумя смартфонами Sony Xperia ZL с модифицированной прошивкой Cyanogenmod 10.2. При этом функционирование связи D2D управлялось процедурой аутентификации PSK, использующей возможности протокола доступа WiFi. Замеры времени установления соединения D2D через WiFi-Direct (см. рисунок 5.9) проводились при помощи интерфейса `wpa_cli` путем наблюдения интервала между созданием новой сети и завершением конфигурации соединения, когда первый IP-пакет был успешно принят.

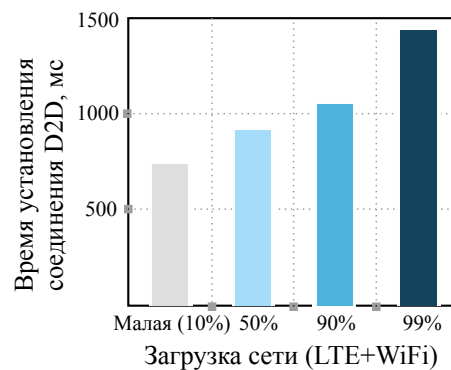


Рисунок 5.9 — Измеренное время установления прямого соединения

Полученные результаты свидетельствуют о том, что если общий уровень загрузки сети LTE не превышает 90%, то круговая задержка между абонентским терминалом и D2D-сервером сохраняется ниже 30 мс. Такая загрузка не приводит к заметному ухудшению работы управляющего протокола D2D по сравнению с временем установления соединения WiFi. При повышении загрузки задержка в системе LTE возрастает и становится соизмерима с аналогичным показателем для WiFi, обуславливая необходимость оптимизации управления сетью LTE. Это соображение особенно важно, поскольку выгрузка на D2D применяется как раз при высокой загрузке соты. При уровне загрузки в 99% очередь на передачу пакетов оказывается переполненной, и в результате круговая задержка в сети LTE становится порядка 0,5 с. Без надлежащего контроля качества обслуживания величина задержки в сети LTE может значительно повлиять на итоговую производительность технологии связи D2D.

На основании полученных выше характеристик работы системы D2D, а также исходя из накопленного в ходе проведения экспериментов опыта, можно

сформулировать следующие ограничения практической системы связи WiFi-Direct с сетевым содействием.

- Время установления соединения через WiFi-Direct можно существенно сократить, снижая задержки, возникающие при работе мобильной операционной системы. Такие задержки влияют, в основном, на момент начала выгрузки трафика и в большинстве случаев приводят к постоянным накладным расходам [301].
- Этап переговоров в процессе установления соединений D2D необходимо организовать так, чтобы он не стал «узким местом» в сети LTE, поскольку он предполагает вплоть до четырех циклов круговой задержки LTE. Ошибки при этом могут привести к быстро растущим накладным расходам и увеличению задержки.
- В загруженной сотовой сети взаимодействие с сервером D2D следует построить так, чтобы оно не занимало значительное время, поскольку обмен управляющими сообщениями происходит в том же канале, что и передача пакетов данных. При этом отсутствует дополнительный радиопоток, выделяемый программным интерфейсом приложения, отвечающего за работу сети D2D на стороне мобильной операционной системы.
- Если абоненты в значительной степени подвижны, то отслеживанию их текущего местоположения следует уделить особое внимание (особенно внутри зданий). Иначе могут возникнуть ложные уведомления о возможности использования прямой связи или же запоздалые попытки установления соединения из-за относительно коротких интервалов нахождения терминалов в непосредственной географической близости друг к другу.

В ходе изучения работы рассмотренного прототипа сети связи D2D наблюдалось значительное увеличение визуального качества передаваемых видеопотоков после их выгрузки на D2D-соединения даже при достаточно значительных расстояниях между взаимодействующими терминалами (до 20 м внутри помещений и до 50 м снаружи). Кроме того, задержки на инфраструктурном соединении LTE отмечались на уровне порядка 100 мс, что является недопустимым для ряда приложений (например, игровых), в которых приемлемая круговая задержка составляет не более 50 мс. При этом задержки на соединении D2D в типовой среде применения не превышали значения в 5 мс для большинства экспериментов.

5.3 Анализ практического использования прямых соединений для обмена данными

5.3.1 Прямое взаимодействие с сетевым кодированием

В связи с резким увеличением объемов передаваемого трафика мобильные операторы вынуждены производить активное уплотнение своих сотовых покрытий. Позволяя снизить уровень нагрузки на сеть, такое уплотнение в то же время требует повышения затрат на установку и обслуживание дополнительных БС. Использование прямой связи D2D между абонентскими терминалами, расположенными в непосредственной близости друг к другу, представляет собой в этом контексте привлекательную альтернативу для разгрузки сотовых соединений без возникновения дополнительных издержек на стороне оператора. Как было показано в главе 3, соединения D2D предлагают более высокую скорость передачи данных, более низкую задержку доступа и повышенную энергетическую эффективность. Их многочисленные приложения были рассмотрены выше.

В предыдущем разделе, а также в главе 3 было проведено аналитическое и имитационное моделирование процесса выгрузки трафика на технологию WiFi-Direct под управлением сети LTE, а также получены результаты практических измерений. При наличии сетевого содействия рассматриваемая система D2D может автоматизировать обнаружение целевых устройств/услуг и установление соединения, а также обеспечить безопасность связи D2D даже для тех пользователей, которые находятся вне социальных сфер друг друга. Таким образом, использование прямых соединений с сетевым содействием открывает новые возможности для взаимодействия абонентов, в том числе незнакомых друг другу ранее. Однако необходимо дополнительное исследование практической работы системы обмена данными D2D, учитывающее фактическое размещение целевого контента.

Поскольку мобильность абонентов может привести к непредсказуемым и краткосрочным контактам между их терминалами, остаются сомнения относительно того, что передача контента через соединения D2D будет эффективна на практике. Далее рассматривается применение подходов, связанных с сете-

вым кодированием, для увеличения степени доступности целевого контента на устройствах, взаимодействующих в режиме D2D. С момента ее введения [302] концепция сетевого кодирования сместила акцент в проектировании систем связи с принципа «сохранить и передать» в сторону принципа «обработать и передать». Соответственно, в отличие от классических подходов к кодированию источника или канала в сетевом кодировании любой задействованный узел может участвовать в кодировании, декодировании или перекодировании информации включая ее источник, получателя и промежуточные узлы.

В недавние годы было опубликовано значительное количество работ, посвященных сетевому кодированию, по таким различным дисциплинам, как математика, физика и биология [303]. Только в рамках теории информации спектр возможных применений сетевого кодирования огромен и включает в себя множество различных кодовых конструкций, от простейших методов на основе функции исключающего «ИЛИ» [304] до значительно более сложных фонтанных кодов [305] и случайного линейного сетевого кодирования RLNC [306]. Кроме того, были рассмотрены методы моделирования и виртуализации сетевого кодирования [307], а также изучено его применение в системах D2D [308], когда необходимо сделать упор на практических аспектах работы данной технологии.

Таким образом, на сегодняшний день были предприняты попытки применения сетевого кодирования для различных типов соединений [82; 309; 310]. Данные работы рассмотрели возможность коллективного обмена файлами в системах WLAN, эффективного назначения радиоресурсов с учетом свойств D2D-соединений, а также использования сетевого кодирования на физическом уровне для улучшения связи D2D. Однако имеющиеся подходы рассматривают исключительно ранние технологии D2D при отсутствии сетевого содействия. В этом разделе исследуется эффективность функционирования системы связи WiFi-Direct при содействии со стороны LTE с возможностью сетевого кодирования (см. рисунок 5.10), что позволяет изучать динамику процессов при обмене данными по D2D с упором на обеспечение доступности контента.

На рисунке 5.10 показан пример работы системы D2D, включающей сотовую БС и множество обслуживаемых ею абонентских терминалов. Они запрашивают, скачивают и сохраняют в своей внутренней памяти (т.н. кэше) фрагменты закодированных данных, соответствующие различным единицам контента. Такие фрагменты (представлены на рисунке в виде квадратов) могут быть либо загружены из сотовой сети, либо получены от других пользо-

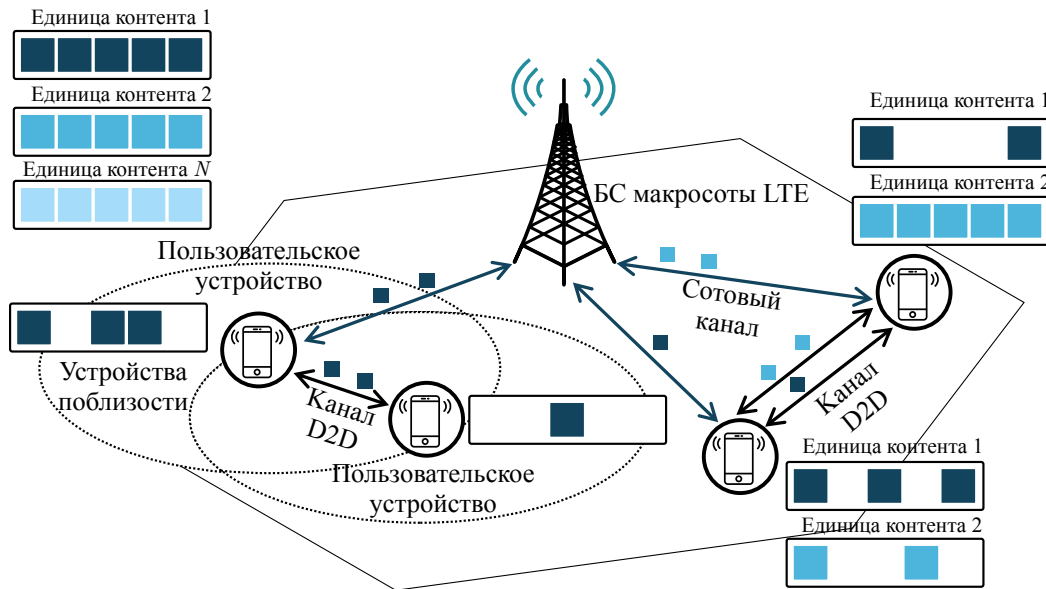


Рисунок 5.10 — Система выгрузки данных на D2D с сетевым кодированием

вателей поблизости при помощи прямых соединений. Далее демонстрируется, что использование уже простейших методов сетевого кодирования (например, RLNC [311]) позволяет существенно увеличить степень доступности целевого контента на соседних терминалах и, тем самым, повысить долю выгруженного на D2D-соединения трафика в рамках ряда практических ограничений.

Как показано в [158; 312], рассмотрение систем D2D с возможностью сетевого кодирования становится все более актуальным на практике в современной структуре мобильного трафика. Действительно, передача мобильного видео впервые превысила половину общего объема передаваемых по сети данных еще в 2012 году [313]. При этом растущая популярность услуги видео по запросу зачастую приводит к асинхронным попыткам получения объемных файлов [314; 315]. Данное обстоятельство не позволяет мобильным операторам использовать для этих целей синхронные механизмы, такие как широковещательная передача по LTE, и может привести к высокой загрузке. Все вышеизложенное создает предпосылки для использования абонентских терминалов в качестве промежуточных кэширующих узлов для часто востребованного контента, особенно в свете того факта, что значительную долю общего трафика составляет небольшое число популярных файлов, а доступные объемы дискового пространства возрастают быстрее, чем емкость беспроводных сетей доступа [158].

Вообще говоря, моделирование процесса обмена популярным контентом в кэширующих системах с возможностью взаимодействия по D2D является доста-

точно новым направлением [316]. В его рамках было рассмотрено применение сетевого кодирования для хранения популярных файлов [317], но в основном с более теоретической точки зрения, не привязанной к конкретным радиотехнологиям. Далее отмеченный пробел восполняется при помощи практического подхода к анализу кэширования и распределения данных по D2D с сетевым содействием, учитывающего особенности перспективных технологий радиодоступа. Последующее изложение ставит своей целью математическое моделирование системы D2D, в которой сетевое кодирование служит для обеспечения обмена данными между соседними терминалами, а также позволяет сделать их прямое взаимодействие более эффективным и надежным.

5.3.2 Моделирование прямого обмена контентом

Ниже описывается предлагаемая модель множественного доступа, которая пригодна для изучения динамического распределения контента поверх системы прямых соединений с сетевым содействием [318]. Дальнейшее изложение построено по принципу от общих замечаний к детальным допущениям относительно функционирования рассматриваемой системы с соответствующими комментариями относительно их правомерности.

Общие соображения. Здесь исследуется двухслойная гетерогенная сеть, в которой сотовая инфраструктура в лицензированном спектре (например, 3GPP LTE) интегрирована с системой D2D на нелицензированных частотах (например, WiFi-Direct). В частности, абонентские терминалы имеют возможность либо получать свой целевой контент путем его скачивания через сотовую БС, либо устанавливать соединения D2D с расположенными поблизости партнерами для загрузки доступного контента напрямую. На практике система управления процессом распределения контента, требуемая для обеспечения эффективной работы системы D2D с сетевым содействием, может быть реализована поставщиком целевых беспроводных услуг (например, с использованием функционала ProSe сотовой сети связи [88]).

Далее в качестве примера, но без потери общности рассматривается некоторая область моделирования, которая является изолированной сотой централизованной сети мобильного оператора. Эта сота с центром в точке размеще-

ния БС имеет радиус R , и никакие процессы, протекающие за ее пределами, не рассматриваются (см. рисунок 5.11). Такой сценарий является типовым для моделирования массовых мероприятий (т.е. концертов, спортивных соревнований, фестивалей и т.д.), где применение связи D2D будет наиболее полезно в силу повышенной плотности размещения пользователей [319] и относительно высокой степени переиспользования популярного контента. Последствиями сетевых взаимодействий вне данного участка можно пренебречь, поскольку сотовые системы, как правило, располагают средствами снижения межсотовой интерференции со стороны соседних сот, а влияние удаленных устройств на D2D-соединения быстро уменьшается с ростом расстояния.

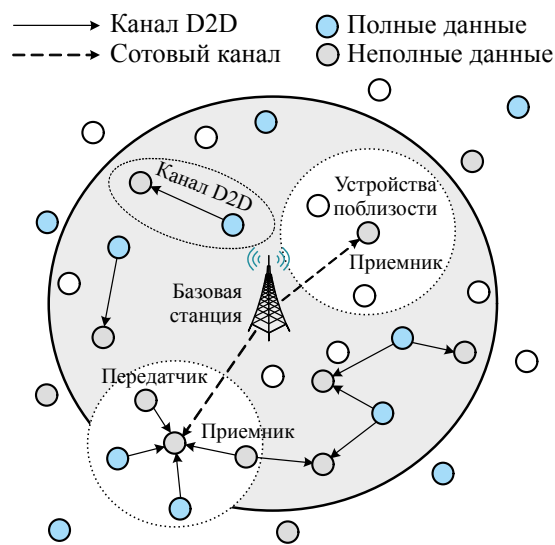


Рисунок 5.11 — Моделирование работы интегрированной двухслойной сети

Далее изучается отправка сотовых данных по линии вниз, тогда как соединения D2D предполагают обмен сообщениями в обе стороны. Кроме того, важно подчеркнуть, что диапазоны частот, используемые сотовой системой связи и сетью D2D, не перекрываются, так что соответствующие попытки передачи не интерферируют между собой. Однако при распределении информации между терминалами в слое D2D может возникать интерференция (поскольку технология WiFi-Direct использует режим случайного доступа к среде передачи), которая учитывается ниже. Относительно структуры целевого контента рассматриваются только сценарии, которые не требуют использования методов защиты от копирования. Подобный контент производится в изобилии в ходе массовых мероприятий и может быть ассоциирован с породившим его событи-

ем при помощи общего ключа безопасности, который раздается участникам, например, в виде штрихкода.

Допущения модели системы. Для анализа рассматриваемой двухслойной гетерогенной сети (см. приложение Д), объединяющей сотовые и D2D-соединения, далее исследуется процесс динамического создания контента, а также его распределения и доставки. Для описания функционирования такой интегрированной беспроводной сети строится соответствующая математическая модель, характеризующая распространение контента внутри замкнутой области моделирования, а также учитывающая динамику запросов на его получение пользователями системы связи. Для того чтобы строго сформулировать этапы жизненного цикла целевого контента, вводится набор основных допущений, составляющих модель системы.

Допущение (Жизненный цикл контента). Новые единицы целевого контента прибывают в систему в соответствии с некоторым входным процессом интенсивности λ_c и проводят в ней случайные интервалы времени (называемые временем жизни контента) со средней длительностью μ_c^{-1} вне зависимости от момента своего поступления. В течение всего времени жизни контент считается *востребованным* и покидает систему (перестает быть востребованным) немедленно по истечении своего времени жизни. Для наглядности предполагается, что все единицы контента имеют определенный одинаковый размер Z .

Важно заметить, что интервалы времени жизни контента и между моментами его поступления предполагаются независимыми и одинаково распределенными случайными величинами. Здесь и далее в качестве примера рассматриваются пуассоновский поток поступления и экспоненциально распределенное время жизни, но последующие рассуждения могут быть проведены аналогично и в случае альтернативных процессов. На рисунке 5.12 приведены примеры жизненного цикла нескольких единиц контента, которые прибывают в систему и покидают ее в соответствии с допущениями модели. Для учета структуры рассматриваемых единиц контента используется понятие *фрагмент*, который представляет собой часть единицы контента (например, файла данных). В этой связи принимается во внимание следующее допущение.

Допущение (Структура контента). Каждая единица контента C поделена на S фрагментов равного размера, так что она может быть восстановлена после получения всех этих фрагментов: $C = (C_1, \dots, C_S)$. Другими словами, поль-

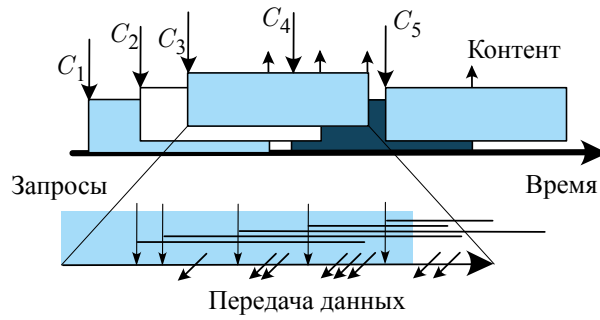


Рисунок 5.12 — Пример работы системы в течение жизненного цикла единицы контента C_4

зователь может считать некоторую единицу контента успешно полученной, если и только если были собраны все соответствующие ее фрагменты.

В течение случайного времени жизни τ некоторой единицы контента C любой абонент при помощи своего терминала может породить запрос на ее получение и тем самым стать *потребителем* контента. Предполагается, что один и тот же абонент может быть потребителем определенной единицы контента только однажды.

Допущение (Запрос на получение контента). Интервалы времени между последовательными запросами на получение контента от пользователей являются независимыми и одинаково распределенными случайными величинами и подчиняются экспоненциальному распределению с параметром λ . По истечении времени жизни τ новые запросы на доступ к соответствующей единице контента более не порождаются.

Абонентский терминал имеет возможность передать предварительно полученный им фрагмент некоторого контента своим D2D-соседям, находящимся внутри круга заданного радиуса R_0 . После получения всех необходимых фрагментов абонентский терминал восстанавливает исходную единицу контента и предполагает процесс ее получения успешно завершенным. Стоит заметить, что терминал может получить различные фрагменты одной и той же единицы контента от различных D2D-соседей. При этом только одно из соединений D2D может быть активным в каждый момент времени и использоваться для передачи данных.

Допущение (Выход из системы). Кроме того, предполагается, что каждый пользователь (как еще получающий, так и уже получивший контент) может покинуть рассматриваемую систему с интенсивностью θ . В любом случае такой

уходящий абонент уносит с собой все полученные им фрагменты контента и более не участвует в его распределении.

Естественная последовательность действий пользователя, заинтересованного в получении некоторой единицы контента C , состоит в сборе всех его фрагментов из доступных на текущий момент источников (либо D2D-соседей, либо сотовой БС) в течение некоторого случайного интервала времени со средним θ^{-1} . При этом прямое скачивание контента у соседних терминалов с помощью технологии D2D является предпочтительным благодаря более высокой скорости передачи данных, а также повышенной энергетической интенсивности [172;173]. Однако, таким образом, может потребоваться установление D2D-соединений с несколькими соседями подряд для получения всех требуемых фрагментов целевого контента.

Важно отметить, что предлагаемый здесь подход не различает абонентов внутри рассматриваемой области моделирования, а также абстрагируется от фактических местоположений их терминалов. Вместо этого предполагается, что географические координаты пользователя, породившего новый запрос на получение контента, представляют собой случайные величины, которые на практике могут быть уточнены при помощи некоторой подходящей модели мобильности абонента. Тогда для учета топологии системы связи (случайного характера размещения терминалов и динамической природы D2D-соединений) формулируется следующее важное допущение.

Допущение (Пространственное размещение терминалов). Абонентские терминалы, порождающие новые запросы на получение контента, размещены равномерно внутри рассматриваемой области моделирования, представляющей собой окружность заданного радиуса R . При этом предполагается, что во избежание необходимости учета граничных эффектов терминалы имеют возможность получать фрагменты контента у своих непосредственных D2D-соседей, находящихся вне данной области моделирования.

В отличие от имеющихся работ по тематике распределения контента, например, в проводных сетях [320], учет пространственного размещения терминалов является важным соображением в беспроводных системах связи, приводящим к различным скоростям передачи данных в зависимости от конкретного местоположения того или иного абонентского терминала, а также расстояния между ними.

Допущение (Скорость передачи данных). Скорость передачи данных $r_{i,j}$ на беспроводном соединении $i \rightarrow j$, а также соответствующая мощность передачи p_i связаны друг с другом через «обобщенную» формулу Шеннона

$$r_{i,j} = w \log \left(1 + p_i \frac{\gamma_{i,j}}{N_0 + I} \right), \quad (5.1)$$

где w – полоса пропускания канала (для сотового или D2D-соединения); N_0 – уровень шума в канале связи; I – уровень интерференции на приемной стороне и $\gamma_{i,j}$ – коэффициент передачи радиоканала между взаимодействующими узлами, зависящая от расстояния $d_{i,j}$ между ними:

$$\gamma_{i,j} = \frac{G}{d_{i,j}^{\kappa}}, \quad (5.2)$$

где $d_{i,j}$ – расстояние между взаимодействующими между собой узлами; κ – экспонента распространения, а G – константа распространения радиосигнала, которая может быть установлена согласно рекомендациям, приведенным в [107].

Следует заметить, что фактическая полоса пропускания канала связи w может быть подобрана таким образом, чтобы учесть соответствующие накладные расходы физического и канального уровней при работе некоторой технологии радиодоступа (т.е. 3GPP LTE или WiFi-Direct в рассматриваемой системе). При этом параметры κ и G также определяются типом радиотехнологии и используемой антенны. Стоит отметить, что влияние эффекта быстрого затухания сигнала не рассматривается здесь в явном виде, но может быть учтено неявно как эквивалентное увеличение уровня шума в канале. Эффект медленного затухания сигнала имеет в рассматриваемой модели лишь незначительный вклад и включен в величину G , поскольку местоположения абонентских терминалов уже являются случайными.

Для того чтобы ввести ограничение на скорость передачи данных, связанное с практическим набором модуляционно-кодовых схем, ее максимальное значение фиксируется как r_{lim} для $d < d_0$. Это фактически означает, что дальнейший рост SNR/SINR не приводит к неограниченному увеличению скорости передачи данных, что соответствует некоторому предельному расстоянию $d_0 = d(r_{\text{lim}})$, которое может быть легко найдено с использованием вышеприведенных выражений.

Различные режимы распределения контента. В рамках предлагаемого подхода к анализу рассматриваемой системы связи сравниваются различные режимы распределения беспроводного контента. Под влиянием идей

из [311], а также множества подобных результатов здесь вводятся в рассмотрение два альтернативных режима, называемых соответственно *базовым* и *кодированным*.

1. Базовый режим распределения (т.н. «BitTorrent», BT): абонентский терминал сохраняет в своей памяти полученные фрагменты контента C_1, \dots, C_S до тех пор, пока все они не будут собраны так, чтобы исходная единица контента могла бы быть восстановлена. Терминал U_i связывается со своими D2D-соседами внутри некоторого радиуса R_0 , получая от них все доступные фрагменты $C_j^{(i)}$, $j = \overline{1, S}$. Кроме того, каждый терминал имеет возможность загрузить отсутствующие у D2D-соседей фрагменты через сотовую БС.
2. Кодированный режим распределения (с сетевым кодированием, NC): вместо сохранения и последующего распределения отдельных фрагментов $\{C_i\}$ абонентский терминал отправляет их линейные комбинации, т.е. $\sum_{i=1}^S \alpha_i C_i$, где $\alpha_i, i = 1, \dots, S$ – независимые и равномерно распределенные случайные коэффициенты, выбранные в некотором конечном поле (обычно рассматривается поле Галуа $GF(2^S)$). При таком способе сетевого кодирования каждые S линейно независимых кодовых символов будут достаточны для восстановления исходной единицы контента.

Сравнение введенных выше режимов распределения контента полезно с учетом их различной чувствительности к наличию фрагментов контента у D2D-соседей. При этом ожидается, что базовый режим (BT) окажется более чувствительным, чем кодированный (NC). В частности, использование случайного линейного сетевого кодирования (RLNC) представляет собой более привлекательный подход к обмену фрагментами контента для пользователей, находящихся на удалении друг от друга. Таким образом, случайная линейная комбинация фрагментов контента хранится в памяти каждого абонентского терминала.

Кодирование RLNC изначально было предложено для повышения эффективности работы пиринговых и кооперативных приложений [321; 322]. Выбор RLNC в данном исследовании определяется гибкостью этого механизма и достаточно низкой задержкой при его использовании по сравнению с альтернативными блочными кодовыми конструкциями. Кроме того, RLNC больше подходит для нерегулярных топологий, характерных для сетей D2D, когда терминалы и взаимосвязи между ними демонстрируют высокую динамику в отличие от статических и детерминированных структур. В результате кодирование RLNC

становится основным кандидатом для применения в рассматриваемой системе D2D с сетевым содействием, которая предполагает распределение контента при помощи прямых соединений между абонентскими терминалами [323].

По сравнению с детерминированными методами сетевого кодирования, для организации работы системы с использованием кодирования RLNC не требуется понимание того, какие именно фрагменты были приняты, что несколько усложняет практическую реализацию. Поскольку кодирование, декодирование и перекодирование производится на стороне источника, получателя, а также на промежуточных узлах вычислительная сложность при работе RLNC возрастает. Однако с учетом практических измерений и демонстраций при помощи реальных прототипов в работах [322; 324; 325] можно уверенно предположить, что кодированный режим распределения контента может быть эффективно внедрен в существующие мобильные устройства.

Основываясь на вышеприведенных рассуждениях, интересно изучить работу рассматриваемой системы для прямого обмена контентом при помощи подхода, принципиально сходного с жидкостными моделями из [320]. Важной чертой предлагаемого подхода становится возможность анализировать динамику распределения фрагментов контента, а также моделировать процессы поступления и ухода единиц контента из системы. Кроме того, учитываются важные пространственные взаимодействия между множеством соседних D2D-терминалов, которые вынуждены конкурировать за возможность передачи. В качестве показателя производительности выбирается число терминалов, которые успешно завершили получение своего целевого контента до некоторого крайнего срока, а также доля трафика, который был передан по соединениям D2D (т.н. выигрыш от выгрузки).

5.3.3 Моделирование системы распределения контента

Изучение работы рассматриваемой системы распределения контента представляет собой любопытную задачу, решение которой требует моделирования необходимого функционала по передаче данных и формированию контента, а также реальных механизмов управления для установления и разрыва D2D-соединений. Как следствие, функционал используемого в предыдущих гла-

вах комплекса программ имитационного моделирования [135], который был откалиброван и верифицирован во многих работах по исследованию связи D2D [88; 172; 173], должен быть расширен дополнительными возможностями.

Особенности рассматриваемого сценария. Исследуемый далее сценарий воспроизводит систему прямого взаимодействия с сетевым кодированием в рамках допущений моделирования, приведенных выше, а также дополнительно рассматривает следующие практические особенности:

- При аналитическом моделировании абонентские терминалы порождают новые запросы на получение контента вне зависимости от общего числа пользователей в системе и предыдущей истории их взаимодействий. На практике новая сессия, связанная с получением целевого контента, устанавливается терминалом только в случае, если ему еще не удалось загрузить этот контент в свою память (кэш). Имитационное моделирование воспроизводит данный процесс, пока все терминалы не скачают требуемый ими контент.
- Реальная интерференционная картина между текущими D2D-сессиями, как правило, более сложна, чем непосредственное разделение радиоресурсов канала связи среди абонентских терминалов, которые могут интерферировать друг с другом. Имитационное моделирование учитывает фактическое взаимное влияние передач терминалов друг на друга, следуя правилам работы процедуры доступа к каналу в системе WiFi Direct, которая реализует принцип CSMA/CA на основе алгоритма двойной экспоненциальной отсрочки.
- Установление соединений D2D требует некоторого времени, и имитационное моделирование учитывает соответствующую задержку каждый раз, когда абонентский терминал принимает данные от своих D2D-соседей. При этом использование инфраструктурного сотового соединения также накладывает некоторые ограничения на время доставки информации, но они оказываются несущественными по сравнению с аналогичными показателями для прямых соединений.

Основные параметры, используемые в ходе имитационного моделирования, собраны в таблице 8.

Описываемый здесь сценарий имитационного моделирования воспроизводит условия, отражающие некоторое массовое мероприятие, проводимое на стадионе, в концертном зале или торговом центре, где большое число людей имеет

Таблица 8 — Параметры моделирования системы распределения контента

Велич.	Значение	Определение/Описание
K	10	Число фрагментов по 1 Мбайт в единице контента
N	100	Количество абонентов при проведении моделирования
r_{D2D}	40 Мбит/с	Целевая скорость передачи данных по D2D-соединению
N_0	-90 дБм	Оценка уровня шума в сотовой системе связи
P_{BS}	23 дБм	Максимальная мощность передачи сотового терминала
w	5 МГц	Полоса пропускания сотовой сети
r_{BS}	10 Мбит/с	Целевая емкость инфраструктурного соединения с БС
R	100 м	Радиус рассматриваемой области моделирования
R_0	30 м	Наибольшая длина D2D-соединения
δ	1 с	Время установления D2D-соединения
λ_c	$100^{-1}1/с$	Интенсивность поступления единиц контента в систему
μ_c	$100^{-1}1/с$	Интенсивность ухода единиц контента из системы
λ	1 1/с	Интенсивность появления пользователей в системе
θ	$300^{-1}1/с$	Интенсивность выхода пользователей из системы
v_r	0,5 м/с	Скорость перемещения пользователей

возможность перемещаться и обмениваться небольшими фрагментами мультимедиа-данных, произведенных в рамках текущего мероприятия. Используемые при этом параметры выбраны из расчета небольшого спортивного турнира, в течение которого пользователи заинтересованы в просмотре мгновенных повторов какого-либо события. Можно ожидать, что размер стадиона для проведения такого турнира составит приблизительно 200 м в длину, и внутри него мобильным оператором будет размещена БС малой соты для обслуживания своих абонентов.

Емкости сотовой сети может оказаться недостаточно для передачи потокового видео по запросу, поэтому представляет интерес поиск альтернативных способов ее повышения. Одним из них служит применение системы прямых соединений с сетевым содействием, которая может быть размещена в месте проведения мероприятия либо его владельцем, либо оператором сотовых услуг 3GPP LTE. Дальнейшее изложение направлено на изучение такой интегрированной сети LTE-D2D в плане выгрузки трафика с инфраструктурных на прямые соединения между пользователями, а также качества их обслуживания. Ниже сравниваются два режима распределения контента, базовый и кодированный (см. описание выше), а до этого даются некоторые замечания относительно работы сотовой сети в отдельности (без использования слоя D2D).

Можно подсчитать, что одна малая сота 3GPP LTE будет испытывать существенные затруднения при обслуживании ожидаемых объемов пользовательского трафика внутри рассматриваемой зоны моделирования. В частности, требуемая емкость сотовой сети для такого обслуживания составит 80 Мбит/с из расчета средней загрузки, что примерно соответствует $80/0,7 \approx 120$ Мбит/с, которые потребуются в условиях пиковой загрузки. Этот несложный расчет означает, что при типовой скорости передачи в 10 Мбит/с на одну соту понадобится около 12 сот для обслуживания всех абонентских терминалов, которые использовали бы исключительно инфраструктурные соединения. И это в случае достаточно оптимистичного допущения относительно линейного роста емкости сети при повышении плотности размещения ее БС.

В то время как такое сверхплотное покрытие возможно технически, управление им представляет собой существенную сложность из-за возникающей интерференции. Действительно, размещение 12 малых БС внутри области радиусом в 100 м приводит к образованию сот радиусом всего лишь 25 м, что значительно ниже типового размера сот в современных системах LTE [326]. Та-

кие небольшие соты будут функционировать наилучшим образом в случае пространственного переиспользования радиоресурсов $1/3$, которое и применяется для последующего сравнения. В качестве альтернативы такому сверхплотному размещению малых сот для разгрузки сотовых каналов связи используются прямые соединения D2D, которые оказываются короче и ниже, и поэтому снижается влияние интерференции на передачи соседних абонентских терминалов.

Базовый режим распределения контента ВТ. Для применения прямых соединений D2D в процессе распределения контента вначале требуется выбрать соответствующий протокол взаимодействия. Следуя описанию модели системы каждая единица контента (например, файл) разбивается на некоторое число фрагментов (по аналогии с популярной системой «BitTorrent»). При этом все фрагменты предоставляются в общее пользование получившими их терминалами, что повышает вероятность использования D2D-соединения внутри рассматриваемой области моделирования. Для осуществления прямых взаимодействий применяется технология связи ближнего радиуса действия (например, WiFi-Direct) с эффективным пространственным переиспользованием радиоресурсов. Имея типовую длину в $R_0 = 30$ м, D2D-соединения поверх WiFi-Direct предоставляют необходимую емкость, недостающую сотовой системе доступа. Результаты проведенного исследования представлены на рисунках 5.13 и 5.14.

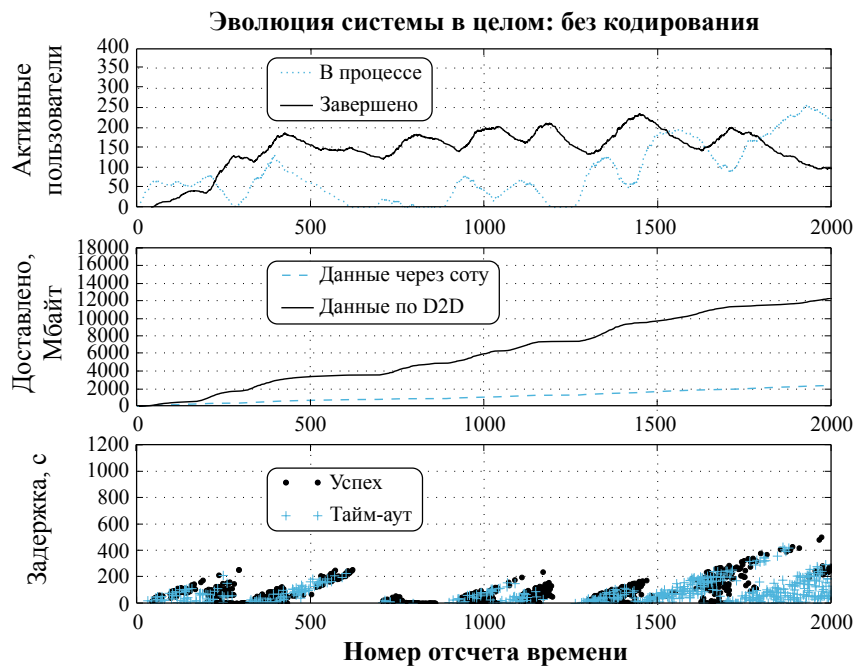


Рисунок 5.13 — Изучение режимов распределения контента: ВТ

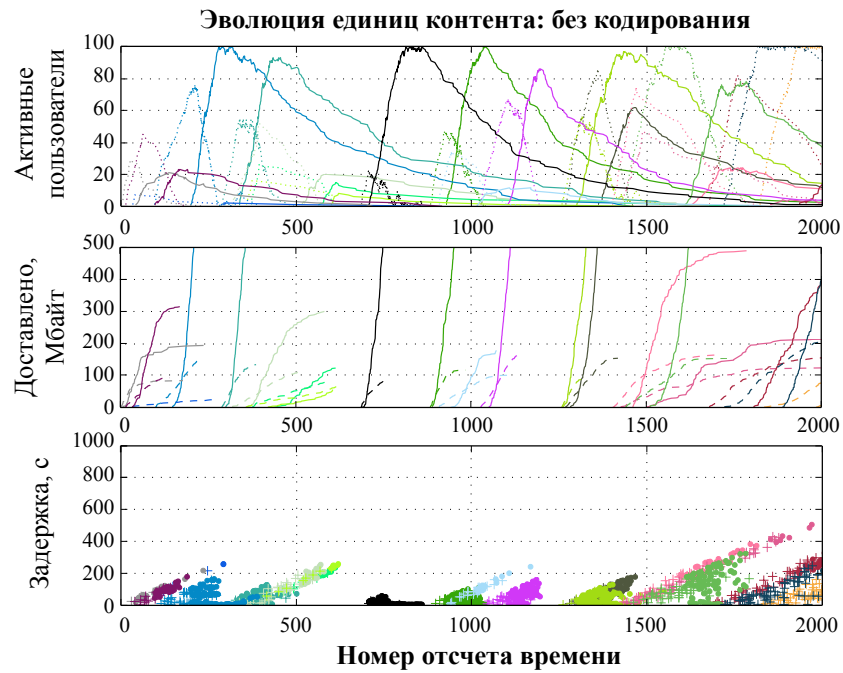


Рисунок 5.14 — Работа системы в процессе получения контента: ВТ

В частности, рисунки 5.13 и 5.15 демонстрируют реакцию системы на появление новых пользователей и запросов целевого контента от них. Видно, что использование режима распределения ВТ зачастую ведет к «разочарованию» абонентов, поскольку число терминалов, которые успешно завершили получение контента, растет достаточно медленно. С другой стороны, на рисунках 5.14 и 5.16 виден вклад отдельных единиц контента по мере того, как они поступают и уходят из системы. Легко видеть, что время жизни каждой единицы представляет собой один и тот же цикл: вначале контент загружается, в основном, через сотовые соединения, а затем – посредством взаимодействий D2D. Поскольку возможностей для установления прямого взаимодействия при использовании режима ВТ оказывается немного, выигрыш от выгрузки на слой D2D оказывается весьма ограниченным. Несмотря на то, что прямые соединения достаточно коротки и эффективно переиспользуют радиоресурсы в пространстве (а также предоставляют высокие скорости передачи), фактическая емкость системы D2D существенно ограничена.

Тогда как большая часть фрагментов действительно передается поверх D2D-соединений, общая емкость системы остается весьма ограниченной, а доля потерь в ней – достаточно велика. В итоге только лишь 34% всех пользовательских сессий обслуживаются вовремя, что связано со следующими двумя фак-

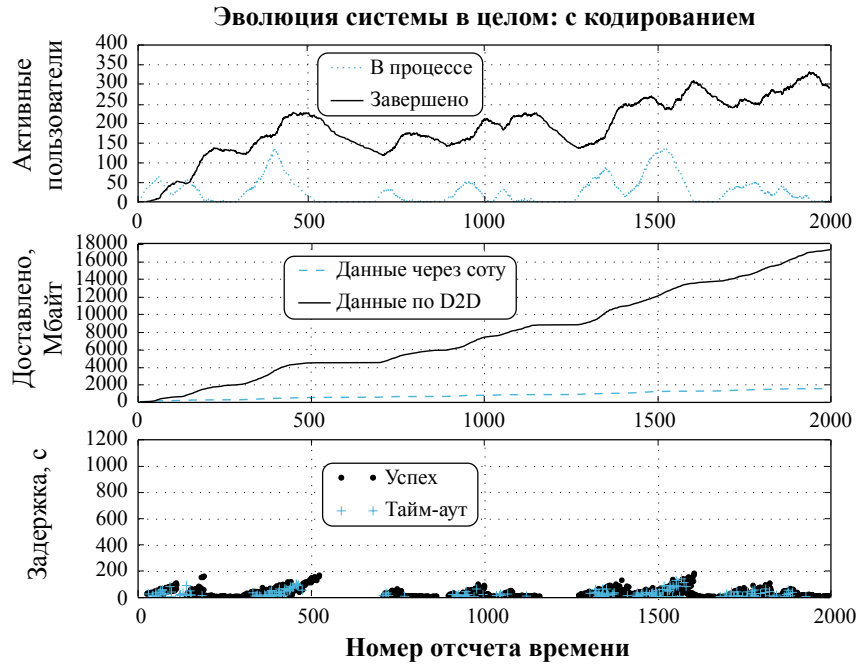


Рисунок 5.15 — Изучение режимов распределения контента: NC

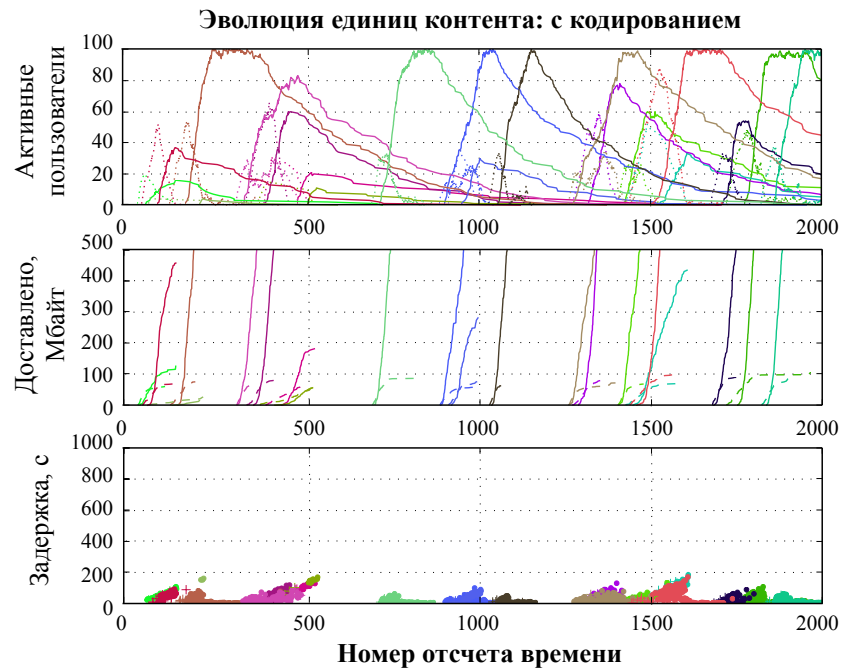


Рисунок 5.16 — Работа системы в процессе получения контента: NC

торами. Во-первых, практическая работа прямых соединений связана с рядом ограничений. В частности, скорость распределения контента через сеть D2D всегда ограничена той интенсивностью, с которой сотовая система «впрыскивает» новые фрагменты. К примеру, если для некоторой единицы контента уже

собрано 9/10 фрагментов и они доступны всем D2D-соседям, ни один из них не может успешно завершить получение контента. В результате каждый терминал будет вынужден запросить последний оставшийся фрагмент через сотовую сеть. Такой фрагмент может поступить с задержкой, не позволяющей получить контент вовремя, что и приводит к напрасной трате значительных радиоресурсов системы.

Кроме того, скорость распределения контента поверх соединений D2D при использовании режима ВТ в течение первых секунд работы остается пропорциональной скорости передачи данных в сотовой сети, поскольку начальные фрагменты доставляются посредством инфраструктурных соединений. Если при этом сотовая система связи начинает испытывать перегрузку, последняя приводит к образованию затора во всей интегрированной системе. Другими словами, последующие единицы контента, запрашиваемые пользователями, застанут сотовую сеть уже загруженной, что приведет к увеличению времени получения фрагментов через нее, и, в свою очередь, еще более усугубит ситуацию. В результате базовый режим разгрузки при помощи прямых соединений D2D (ВТ) предоставляет некоторую дополнительную емкость, но, вообще говоря, не реализует весь потенциал рассматриваемой системы.

Кодированный режим распределения контента NC. Для преодоления основных ограничений, возникающих при использовании базового режима распределения контента, далее ко всем его фрагментам применяется сетевое кодирование RLNC, так что их конкретный состав теряет значение при восстановлении исходной единицы контента. Такой подход позволяет абонентским терминалам обмениваться релевантными фрагментами контента с гораздо более высокой вероятностью, повышая таким образом эффективность работы D2D-соединений. Здесь в качестве иллюстрации выбирается простой метод сетевого кодирования, который, тем не менее, позволяет выявить соответствующие преимущества рассматриваемой системы без существенного увеличения ее сложности или же накладных расходов. Результаты для такого кодированного режима распределения контента (NC) представлены на рисунках 5.15 и 5.16.

Можно видеть, что более продвинутый режим распределения контента NC гораздо лучше справляется с пиками загрузки, возникающими при поступлении новых единиц контента в систему, а слой D2D демонстрирует при этом почти удвоенную скорость передачи. Как следствие, интегрированная сеть подвержена существенно меньшему влиянию перегрузки сотовых соеди-

нений, поскольку большую часть времени абонентские терминалы используют D2D-взаимодействие для получения своего целевого контента. В итоге такой режим работы приводит к снижению задержки передачи данных и повышению качества восприятия услуги: доля пользовательских сессий, которые завершили свое обслуживание вовремя, достигает 80%, что свидетельствует о привлекательности применения каналов D2D для выгрузки трафика. Оставшиеся потери можно компенсировать путем увеличения доступной емкости сотовой системы связи.

Рассматриваемый здесь пример подтверждает пригодность использования сетевого кодирования для повышения степени доступности контента при его распределении в беспроводных системах: когда начальные фрагменты очень ценятся в силу их дефицита, важно обмениваться ими при первой возможности с наибольшим числом доступных соседей, чего сложно достичь в базовой системе из-за потенциального наличия дублированных фрагментов. Уменьшение размера передаваемых фрагментов могло бы частично улучшить эту ситуацию, но тогда возросла бы доля времени на установление D2D-соединения, поскольку применение прямых взаимодействий для обмена небольшими фрагментами контента существенно менее эффективно. Напротив, желательно использование более крупных фрагментов для снижения накладных расходов при работе связи D2D.

Для сравнения изученных выше систем, использующих различные режимы распределения контента, их характеристики сведены в таблицу 9. Можно видеть, что применение подхода с сетевым кодированием поверх соединений D2D позволяет достичь производительности, схожей с аналогичным показателем в сверхплотной сети LTE из 12 сот². Однако в отличие от LTE использование сети D2D для выгрузки трафика не ведет к росту издержек мобильного оператора. Кроме того, в отличие от БС малых сот, режим распределения контента NC может быть развернут существенно быстрее, поскольку он не требует установки и обслуживания инфраструктурного оборудования, а также работы сложных механизмов координирования интерференции и управления радиоресурсами.

Некоторые практические соображения. Интересно отметить, что с увеличением производительности сотовой сети различие между системами под управлением режимов NC и VT становится менее значительным и фактически

²Эффективная полоса пропускания подобрана так, чтобы позволить единообразное сравнение с другими двумя системами.

Таблица 9 — Сравнение производительности систем распределения контента

Характеристика	12 сот	ВТ	НС
Время обслуживания	26,2 с	98,4 с	50,8 с
Вероятность успеха	73,1%	41,5%	67,4%
% трафика через D2D	0,0%	79,4%	88,7%
Полоса пропускания	3x10 МГц	10 + 20 МГц	10 + 20 МГц

исчезает, когда ее емкость становится достаточной для обслуживания всего поступающего пользовательского трафика. Таким образом, на практике режим ВТ может использоваться, пока соответствующие потери остаются допустимыми, а затем возможен переход к режиму НС, как только достигнуто ограничение емкости. Подобный переход может быть осуществлен и в обратную сторону, по мере того как система разгружается, в целях снижения накладных расходов на декодирование сетевого кода со стороны абонентских терминалов.

Особенности работы интегрированной системы связи в отношении задержки передачи и вероятности успешного обслуживания в зависимости от ее загрузки представлены на рисунках 5.17 и 5.18. Можно сделать вывод о том, что сравниваемые между собой режимы работы функционируют в схожем диапазоне интенсивностей входного потока, но демонстрируют различное поведение. В частности, использование режима ВТ приводит к снижению вероятности успеха и росту задержек, тогда как применение режима НС приближает характеристики производительности к аналогичным для сотовой сети во всем, за исключением времени доставки информации. На основе этих данных можно заключить, что выгрузка трафика на соединения D2D с сетевым кодированием представляет собой оправданную альтернативу сверхплотному сотовому покрытию в широком спектре входных параметров.

Наконец, важно заметить, что реально существующие технологии D2D, такие как система WiFi-Direct, на практике имеют доступ сразу к нескольким каналам по 20 МГц. В частности, при работе сети WiFi можно ожидать применение режима переиспользования частот 1/3 в диапазоне 2,4 ГГц, а также 23

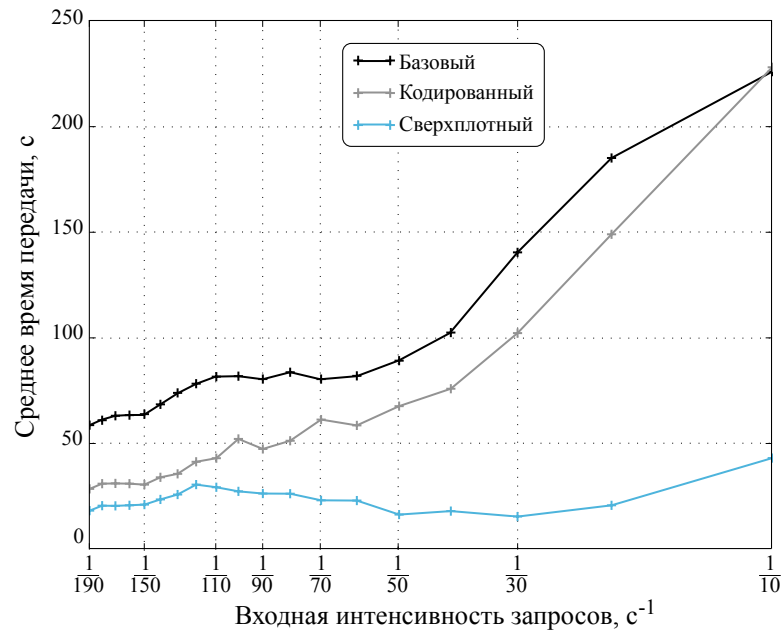


Рисунок 5.17 — Среднее время передачи для различных систем распределения контента

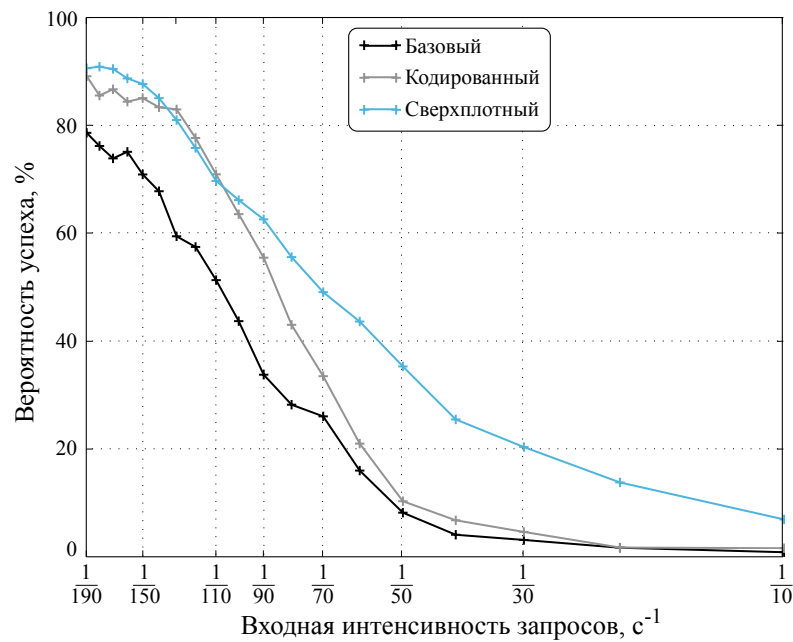


Рисунок 5.18 — Вероятность успеха для различных систем распределения контента

дополнительных канала в диапазоне 5 ГГц, что потенциально увеличивает уже значительную емкость слоя D2D (порядка 5 ГГц фактической полосы пропускания). Имеющимся на сегодняшний день сотовым технологиям радиодоступа

сложно сравниться по своей емкости с соответствующими возможностями сетей связи D2D в нелицензированном спектре. Это обстоятельство делает подобные решения естественным выбором при развертывании систем обмена объемными мультимедиа-данными в ходе массовых мероприятий (на концертах, фестивалях и спортивных соревнованиях).

5.4 Интеграция систем удаленного хранения и обработки данных

5.4.1 Дальнейшее развитие сетей следующего поколения

С учетом временных сроков, установленных союзом ИТУ в рамках инициативы ИМТ-2020, стандартизация сетей 5G в настоящее время приближается к своему завершению. В этой связи регулярные семинары RAN 5G, организуемые комитетом 3GPP, посвящены разработке технологий радиодоступа следующего поколения, что включает в себя дальнейшую эволюцию системы LTE и новую технологию NR (New Radio), которые призваны совместно удовлетворить многочисленным требованиям ИМТ-2020. Поскольку перспективная система NR представляет собой новую радиотехнологию без возможности обратной совместимости с уже имеющимися (например, LTE), становится понятно, что сети 5G объединят несколько различных технологий и систем. Данная особенность позволит мобильным операторам разворачивать более гибкие сетевые решения, пригодные для нужд определенных приложений и услуг мобильной связи.

Для этих целей системы 5G определяют новую концепцию сетевой «нарезки» (network slicing), которая предоставит выделенные слои для обслуживания отдельных сервисов в рамках единой сети [327], в отличие от имеющихся подходов с размещением отдельных сетей для различных целей. Такие относительно независимые сетевые слои позволят их приложениям гарантировать целевую скорость передачи данных и задержку, а также повысят эффективность использования спектральных радиоресурсов. С помощью динамически создаваемых и безопасных сетевых слоев гетерогенные системы 5G будут предоставлять требуемое качество обслуживания своим многочисленным приложениям и услугам [328]. После ответа на важные вопросы относительно применения данной

концепции, включая «каким образом осуществлять сетевую нарезку» и «с какой гранулярностью», операторы смогут использовать ее в своих системах, дополняя ею существующие технологии управления, такие как программно-определяемые сети SDN и виртуализация сетевых функций NFV.

В частности, технология SDN является достаточно известной (восходит к 1990-м годам) и позволяет динамически подключать и конфигурировать различные сетевые компоненты. Однако только сейчас благодаря достаточным вычислительным мощностям это решение начинает с успехом использоваться на практике. При этом технология NFV строится на основе широко распространенных серверов и коммутаторов, а также облачных систем хранения и обработки информации и позволяет осуществить переход к облачным сервисам при помощи виртуализации всевозможных функций сетевых узлов. С учетом дополнительных степеней свободы, предоставляемых системами SDN и NFV, будущие операторы сетей 5G получат возможность быстро разворачивать новые услуги, а также перемещать их в качестве виртуальных машин в ответ на динамические изменения потребности своих пользователей.

Опираясь на технологии SDN, NFV и сетевой «нарезки», возможно объединить передачу данных и их обработку в рамках быстро развивающейся экосистемы 5G, которую характеризуют как «вычислительные технологии для нужд средств связи». Данный подход строится на взаимодействии средств передачи и обработки информации и позволяет эффективно организовать удаленные вычисления по беспроводному каналу связи. Такая выгрузка функций по обработке данных в системах 5G разрешает им обслуживать устройства с ограниченными вычислительными возможностями и энергетическими ресурсами. Подобное оборудование теперь имеет возможность переводить свои вычислительно сложные задачи на (ближайшие) ресурсоемкие серверы. Таким образом, происходит фундаментальное изменение принципов сетевого взаимодействия, смещающее фокус с беспроводных соединений в сторону передаваемого контента и позволяющее сосредоточиться непосредственно на его распределении, хранении и получении в отличие от исторически сложившейся сетевой архитектуры, ориентированной исключительно на повышение емкости системы связи [290].

Однако возросшие потребности перспективных сетевых приложений в отношении доступной полосы пропускания (которые обсуждаются ниже) вынуждают мобильных операторов спешно наращивать емкость имеющихся сотовых покрытий. Для этих целей гетерогенные сети зарекомендовали себя в качестве

эффективного и надежного решения, опирающегося на макросоты для обеспечения бесшовного покрытия и осуществления функций управления сетью, а также на различные малые соты различных радиотехнологий и размеров в областях с высокой загрузкой [88]. При этом современные гетерогенные сети позволяют мобильным операторам размещать сотовую инфраструктуру ближе к абонентским терминалам, которые являются непосредственными производителями и потребителями контента, становясь тем самым эффективным средством уплотнения сети, особенно в городских условиях.

По мере того как сверхплотные покрытия малыми сотами позволяют повысить спектральную эффективность и емкость сети радиодоступа, они в то же время создают повышенную нагрузку на транзитные (backhaul) каналы [329]. Для большинства операторов мобильной связи размещение высокоскоростных транзитных каналов между множеством БС малых сот и опорной сетью становится экономически нецелесообразно. Для предотвращения заторов в условиях ограниченной емкости транзитных каналов, особенно в периоды пиковой загрузки, используется кэширование данных на БС, позволяющее предоставлять пользовательский контент локально, без использования транзитных соединений. В недавнее время множество исследований показало, что применение механизмов кэширования действительно является эффективным решением для снижения нагрузки на транзитные каналы. При этом часто востребованные данные размещаются на кэширующих БС в спокойные периоды, а затем скачиваются в часы пик расположенными поблизости пользователями.

Соответственно, на сегодняшний день имеется ряд альтернативных вариантов построения архитектуры кэширующей сети, которые различным образом интерпретируют информацию относительно степени востребованности контента и позволяют снять ограничения на емкость транзитных каналов за счет использования дополнительной памяти. В итоге системы 5G могут применять БС малых сот с высокой емкостью кэша для хранения наиболее популярного контента и низкоскоростными транзитными соединениями, которые служат в основном для обновления закэшированного контента со скоростью, пропорциональной изменению общей потребности в нем со стороны абонентов [158]. С учетом недавнего прогресса в производстве модулей памяти, такое «прозрачное» кэширование на стратегически важных участках сети позволяет повысить эффективность распределения контента, а также увеличить эффективность использования радиоресурсов системы, даже когда пользователи не запрашивают

некоторую единицу контента единовременно (т.е. использовать изменчивость сетевого трафика во времени).

5.4.2 Перспективные пользовательские приложения и услуги в 5G

Эволюция способов получения видеоконтента. По мере того как абоненты применяют все более многофункциональные смартфоны и планшеты, эти интеллектуальные устройства становятся неотъемлемой частью повседневного распорядка своих пользователей. На сегодняшний день использование смартфонов уже превышает по популярности использование компьютеров при доступе к сети Интернет. Действительно, типичной картиной прошлого века было чтение газеты пассажирами общественного транспорта, тогда как сегодня это занятие сменилось просмотром новостей онлайн при помощи всевозможных мобильных устройств. Вслед за переходом новостей в электронный вид изменяются бизнес-модели, связанные с распределением контента, позволяя ему публиковаться быстрее и чаще содержать мультимедиа-данные. Это, в свою очередь, создает *повторные запросы* на загрузку популярного контента, такого как срочные новости и лидеры видеопроката, что приводит к чрезвычайно избыточной передаче повторяющейся информации.

Чтение горячих новостей с мобильных терминалов представляет собой лишь один из множества примеров того, как люди демонстрируют схожесть интересов в получении контента по географическому принципу. Другим примером может стать массовое скачивание новой версии операционной системы iOS, которое порождает один из самых значительных пиков загрузки, известных сети Интернет. Помимо использования мобильных устройств для доступа к мультимедиа-контенту и коллективной загрузки популярных файлов, сегодняшний Интернет-трафик демонстрирует колоссальный рост доли видео, получаемого по запросу. Глобальные поставщики подобных услуг, такие как YouTube и Netflix, уже насчитывают многомиллионную аудиторию и фактически являются законодателями нового направления в потреблении видеoinформации. К примеру, Netflix, являясь мировым лидером на рынке видео по запросу, предоставляемого по подписке, насчитывал более 60 миллионов абонентов еще в середине 2015 года, и ожидается, что их число удвоится к концу 2020 года.

Однако в отличие от передачи потоковых мультимедиа-данных, при получении видео по запросу пользователи не загружают один и тот же контент одновременно. Это важное свойство, известное как *асинхронное переиспользование контента*, приводит к тому, что небольшое количество популярных единиц контента составляет подавляющее большинство фактически передаваемого трафика. Существуют практические свидетельства того, что типовые запросы со стороны пользователя концентрируются на сравнительно небольшом наборе файлов [330]. Более того, скачивание видео по запросу в высокой степени избыточно во времени и пространстве, что приводит к необходимости более глубокого исследования статистических свойств запрашиваемых данных, а также самой изменившейся структуры пользовательских запросов. Недавние исследования подтверждают, что современные сотовые технологии все еще не в состоянии предоставить бесшовную, эффективную и масштабируемую передачу видео по запросу, поскольку архитектура сети Интернет базируется на исторически сложившейся модели распределения контента, которая имеет ряд ограничений, обсуждаемых далее.

Имеющиеся ограничения при доставке контента. В существующих сотовых сетях мобильные терминалы, располагающиеся на краю соты, испытывают повышенное энергопотребление из-за высокой мощности передачи, а также имеют существенную задержку при доступе к целевому контенту через беспроводные соединения. При этом известно, что люди особенно чувствительны к величине задержки доступа и ее колебаниям. Наиболее крупные поставщики цифровой информации на современном рынке, такие как Google и Akamai, построившие многомиллиардный бизнес вокруг своих систем распределения контента, отмечают, что низкая задержка является ключевым ингредиентом успеха при удовлетворении потребности пользователей. Примеры из жизни подтверждают, что уже незначительный рост задержки доступа может значительно снизить качество восприятия от работы того или иного приложения или услуги в сети Интернет.

Для повышения качества восприятия пользователей можно применять кэширование часто востребованного контента *на краю сети* (т.е. на БС малых сот) [331], снижая тем самым разрыв между доступной и требуемой емкостью сети доступа [332]. При этом, отвечая на ключевые вопросы относительно того, что и где выгоднее кэшировать, важно учитывать непредсказуемую мобильность пользователей, которая становится важным усложняющим фактором,

особенно с учетом растущей популярности мобильных устройств при доступе к контенту. Помимо исследования характеристик такой мобильности, информация о местоположении абонентов становится немаловажной для их эффективного обслуживания в массовых сценариях: пассажиры с мобильными терминалами в общественном транспорте, скопления людей в торговых центрах, на стадионах, в аэропортах и т.д. В итоге важно учитывать множество практических особенностей доступа к контенту при организации эффективного процесса кэширования, таких как распределение популярности данных, местоположение, скорость и направление движения пользователей. Это позволит повысить качество обслуживания в сценариях, требующих обработки значительных объемов информации.

На сегодняшний день существует широкий спектр всевозможных приложений, которые требуют существенных вычислений, а также непрерывной обработки получаемой ими информации. Примерами являются системы навигации автомобилей, манипуляция данными в электронных играх, обработка видео в смартфонах, распознавание объектов подвижными роботами, синтез речи, обработка естественных языков и вычисления на носимых устройствах. Однако развитие подобных перспективных приложений и услуг 5G лимитировано вычислительной мощностью современного пользовательского оборудования, которая не может возрасти неограниченно в силу фундаментальных сдерживающих факторов, например, размеров и емкости аккумуляторной батареи. Как следствие, эти новые вычислительно-емкие сервисы неизбежно потребуют создания продвинутых механизмов для поддержки удаленных вычислений, которые должна предоставлять будущая архитектура 5G.

Типовые классы приложений и услуг в сетях 5G. Особенно важным сценарием в контексте развития сетей 5G становится использование *дополненной реальности*, позволяющей достичь поистине интерактивного восприятия услуг (см. рисунок 5.19). В отличие от виртуальной реальности, дополненная реальность направлена на модификацию объектов и явлений окружающего мира, нежели их воссоздание в искусственной среде. В этой связи предметы индивидуального окружения пользователя становятся как фоном, так и сущностью контекстного аннотирования при помощи компьютерно-сгенерированной информации, что требует сложных вычислений в реальном масштабе времени. В свете происходящей трансформации средств распределения цифрового контента, которая порождает мириады вычислительно-емких и чувствительных к

задержке приложений и услуг [333], далее предлагается различать два широких класса сервисов по способу порождения потока данных: пользователем или окружающей его сетевой инфраструктурой.

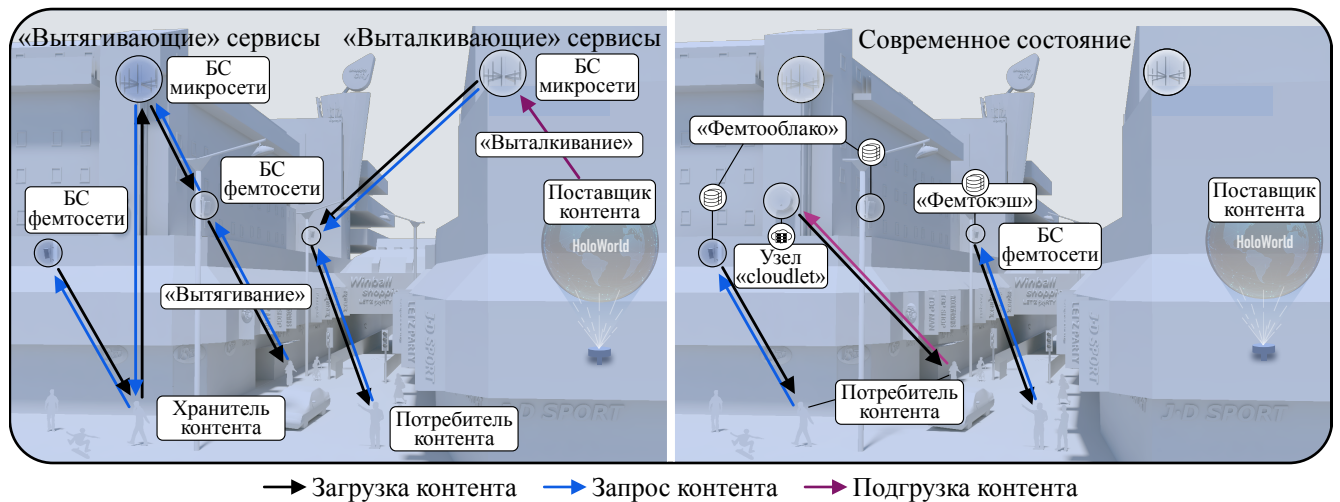


Рисунок 5.19 — Примеры использования перспективных сервисов 5G

«Вытягивающие» сервисы. Данная категория включает в себя сервисы, инициированные пользователем (см. рисунок 5.19), такие как обработка мультимедиа-данных для целей работы или отдыха. Соответствующие приложения варьируются от создания и редактирования любительского мультимедиа-контента в социальных сетях до обслуживания нужд абонентов «без рабочего места», позволяя профессионалам работать «со свободными руками» (например, в медицине, на производстве, в обслуживании и т.д.). Общим для всех подобных сценариев является «вытягивание» требуемого функционала из окружающей сетевой инфраструктуры некоторым пользователем или их группой по требованию (аналогично тому, что позволяет осуществлять решение PlanGrid для инженеров-строителей).

«Выталкивающие» сервисы. В данную категорию входят сервисы, инициируемые сетевой инфраструктурой (см. рисунок 5.19), что включает в себя услуги с учетом местоположения абонентов, «вирусную» рекламу, мониторинг опасной окружающей среды, контекстную обработку данных и мобильные приложения дополненной реальности. Соответствующие сценарии варьируются от навязывания пользователю контекстных новостей и информационных услуг до предоставления возможности распознавания объектов в реальном масштабе времени и визуализации цифровой информации (в таких областях, как игры,

развлечения и новости, коммунальные услуги, образование, профессиональная ориентация и т.д.). Такие сценарии, как правило, предполагают, что сетевая инфраструктура заранее «выталкивает» некоторый контент на своих абонентов, имея представление об их интересах или потребностях.

Далее предлагается обзор перспективных технологий для обеспечения описанных выше сценариев 5G.

5.4.3 Эволюция архитектуры систем связи 5G

Интеграция вычислительной и сетевой инфраструктур. В последние годы облачные вычисления стали широко распространены в качестве современной вычислительной инфраструктуры, которая позволяет эффективно производить удаленную обработку информации. Основываясь на виртуализации вычислительно-емкой обработки данных, облачные системы позволяют использовать множество операционных систем и приложений в рамках единой платформы, гарантируя при этом изоляцию и защиту как программ, так и их данных. Такой подход стал важным средством, позволяющим перевести вычислительно-емкие пользовательские приложения в более ресурсоемкое облачное пространство и тем самым добиться впечатляющего экономического успеха. На сегодняшний день поставщики облачных услуг предлагают своим пользователям функционал для эластичного использования системы по требованию, включая ресурсы самой инфраструктуры, облачных платформ и соответствующего программного обеспечения [334].

Применение облачных вычислений в мобильных сценариях привело к возникновению мобильных облачных вычислений, которые позволили преодолеть множество практических ограничений в отношении производительности, окружения и безопасности [335]. Однако основным сдерживающим фактором на пути современных систем мобильных облачных вычислений становится высокая *сквозная задержка* при передаче информации, включая компоненты, связанные с доступом, транспортировкой и обслуживанием данных. В этой связи текущая архитектура сетей доступа может наложить значительное запаздывание, а также ограничить пропускную способность соединения из-за правил использования спектральных радиоресурсов. В результате поставщики облачных услуг

заинтересованы в создании эффективных в плане затрат средств масштабирования полосы пропускания, доступной их пользователям.

Для улучшения характеристик работы своих сетей доступа, особенно при обслуживании портативных и носимых устройств, операторы мобильной связи принимают конкретные меры, рассмотренные выше. При этом бурное развитие технологий радиодоступа в эпоху 5G приводит к ситуации, когда соединения «последней мили» имеют более высокие параметры производительности, нежели соответствующая опорная инфраструктура сети Интернет. В итоге мобильные операторы могут испытывать резкую нехватку емкости транзитных каналов (тип размещения I на рисунке 5.20), и передача данных в сеть Интернет и обратно через их инфраструктуру может занять существенное время. Более эффективной, но затратной стратегией является наращивание емкости транзитных каналов наравне с соединениями сети доступа, что позволяет поддерживать более высокие скорости передачи данных и снизить загрузку (тип размещения II на рисунке 5.20).

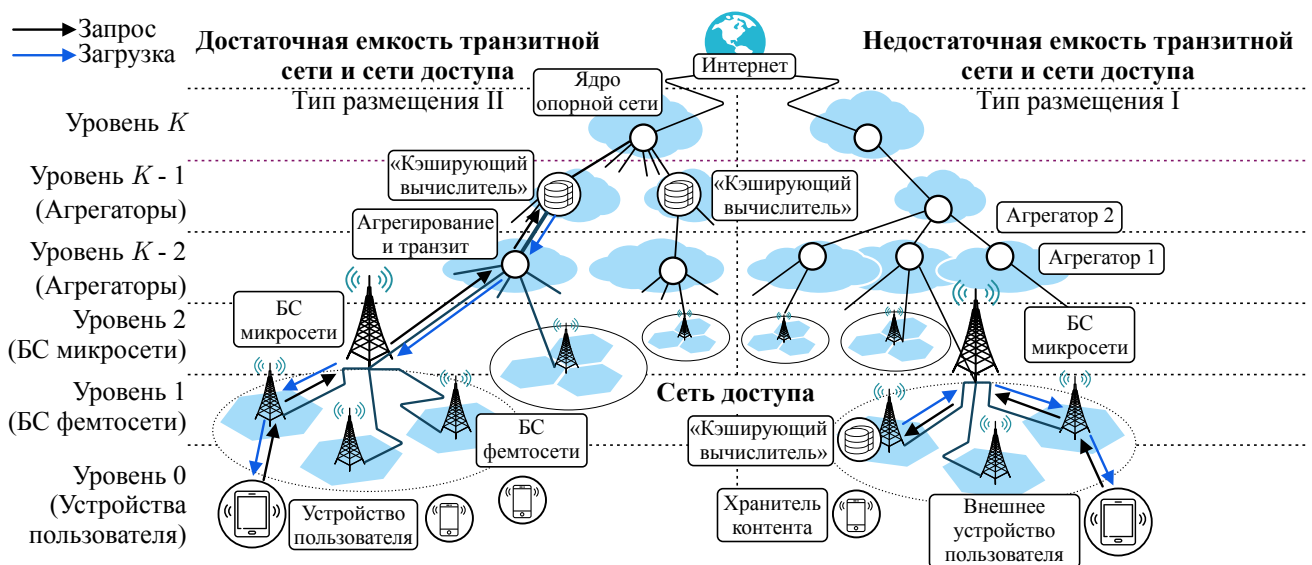


Рисунок 5.20 — Варианты размещения сетевой архитектуры и соответствующие уровни иерархии

Особенности современного распределения контента. Потребность в эффективном распределении контента является одним из важнейших факторов, определяющих развитие сетевой архитектуры [336]. Начиная с эпохи ранних пиринговых систем современный функционал сети Интернет сосредоточен на контенте как таковом, а не на его фактическом местоположении. Поддержи-

ваемые такими гигантами, как Akamai сети доставки контента CDN реализуют технологию «anycast», позволяющую выбрать наиболее подходящую (в смысле близости ее расположения) копию контента для обеспечения самоорганизующегося, адаптивного и устойчивого к сбоям распределения контента. Также была предложена концепция информационно-ориентированного взаимодействия в качестве общей инфраструктуры, предоставляющей кэширование контента внутри сети для достижения его масштабируемого, безопасного и экономичного распределения.

Суть описанных достижений заключается в отделении контента от его носителей и их местоположения не на уровне приложений, а уже на сетевом уровне. Такой подход разительно отличается от традиционной архитектуры «клиент – сервер», которая всегда перенаправляет свои вычислительные задачи на более ресурсоемкий сервер. Однако в случае, если узлы CDN размещены в опорной сети, может наблюдаться недостаток пропускной способности на краю сетевого покрытия, что приводит к снижению надежности и увеличению задержки на беспроводных соединениях. Такая ситуация на сегодняшний день типична, и пользователи, располагающиеся на краю сети, часто вынуждены выполнять всю необходимую обработку информации локально на своих терминалах, а затем использовать внешние носители памяти для последующего обмена результатами такой обработки. Как следствие, дальнейшее развитие сетевой архитектуры в направлении более эффективного кэширования и обработки данных является крайне актуальным.

В свете вышеизложенного интеграция возможностей для кэширования и обработки данных с гетерогенными сетями 5G нацелена на снижение времени доставки контента путем помещения его копий в различные участки сети и фактическое сокращение расстояния от места хранения ресурсов до места их использования [337]. Кроме того, кэширование внутри сети позволяет «сместить» максимальный поток трафика с часов пик на более удобное время, тем самым снижая колебания загрузки и уменьшая вероятность возникновения затора. Распределенное локальное кэширование обычно функционирует в две фазы: размещение контента на хранение и его доставка; оно также может учитывать дополнительную информацию о системе (например, число переходов и распределение популярности контента). Такое предиктивное (упреждающее) кэширование позволяет обеспечить более эффективное распределение контента с учетом его передачи, хранения и обработки.

Интегрированные узлы для вычисления и хранения в 5G. Поскольку современные сотовые сети основываются на базовой архитектуре «all-IP», возникают два альтернативных варианта для осуществления кэширования внутри сетей 5G [338]: (1) в ядре сети (EPC), содержащем обслуживающий шлюз (serving gateway), сетевой шлюз пакетных данных (packet data network gateway) и модуль управления мобильностью MME или (2) в сети радиодоступа (RAN), включающей в себя всевозможные БС, называемые eNodeB. В некоторых случаях может быть целесообразно объединять функционал по хранению данных со средствами для их обработки [339], особенно когда целевые сервисы требуют множества обращений к удаленному серверу в течение коротких интервалов времени или других сложных взаимодействий. Соответствующая архитектура, совмещающая функции кэширования и удаленных вычислений, предоставляет их в качестве виртуализированных ресурсов системы 5G, направленных на поддержку локальных пользовательских приложений.

В современной научно-исследовательской литературе были изучены всевозможные формы интеграции функционала для обеспечения передачи, хранения и обработки информации. После опубликования основополагающей работы [340] предлагались различные варианты того, как мобильное устройство может воспользоваться поддержкой удаленных серверов для выполнения своих задач. Для перемещения вычислительных ресурсов ближе к абонентским терминалам была предложена концепция «cloudlet», которая позволила мобильному оборудованию получать доступ к близко расположенным стационарным ресурсоемким серверам, подключенным к облачной системе при помощи высокоскоростных беспроводных соединений. При таком подходе серверы поддержки могут размещаться в общественных и коммерческих сетях, развернутых в аэропортах, на вокзалах, в кафе и на предприятиях, где наблюдаются массовые скопления людей. Таким образом, абонентские терминалы могут выгружать свои вычислительные задачи на ближайшие к ним серверы с низкой задержкой и на высокой скорости передачи, избегая взаимодействия с удаленной облачной системой.

Однако размещение узлов «cloudlet» влечет за собой дополнительные затраты на установку и обслуживание, а также не предоставляет средств поддержки мобильных пользователей. В качестве альтернативного подхода было предложено совмещение системы облачных вычислений и гетерогенной сети доступа, которое получило название фемто-облако. При этом БС малых сот фак-

тически берут на себя функции узлов «cloudlet», предоставляя, таким образом, эффективное распределение вычислительных ресурсов облака в непосредственной близости к мобильным пользователям. В дополнение рассматривалось кэширование библиотеки популярного контента на БС фемтосот (т.н. фемтокэширование [158]), а также в памяти самих абонентских терминалов, что приводит к дальнейшему снижению требований по емкости транзитных каналов в гетерогенных сетях [341]. В результате фемтокэширование и подобные ему подходы имеют возможность повысить масштабируемость сетей 5G, обеспечивая более высокие скорости передачи данных.

5.5 Выводы

В результате разработанный протокол управления системой прямых соединений для гетерогенной сети с интегрированными между собой технологиями радиодоступа, а также способ его реализации в рамках модельной сети с высокой плотностью позволяют снижать нагрузку на сотовую сеть путем выгрузки трафика в сети других технологий радиодоступа, повышая при этом скорость передачи пользователей. Кроме того, сформулированный алгоритм распределения трафика пользователей в гетерогенной системе прямых соединений с большим количеством устройств, основанный на применении методов сетевого кодирования и учитывающий структуру передаваемых сессий, применим для управления нагрузкой с целью снижения времени передачи пользовательских данных и повышения вероятности успешного доступа.

Заключение

Основные **научные результаты** диссертационной работы заключаются в следующем.

1. Разработана модель гетерогенной сети с набором интегрированных между собой технологий радиодоступа, применимая для оценки параметров качества обслуживания. Данная модель позволила определить, что вероятность блокировки сессии пользователя может быть снижена до 40% при наличии в рамках гетерогенной сети трех различных технологий радиодоступа.

2. Предложен алгоритм с разделением пользовательской сессии для совмещенной сети радиодоступа в условиях высокой плотности размещения узлов сетевой инфраструктуры, позволяющий повысить скорость отправки данных пользователями на величину 50–60% при входной интенсивности 10–20 сессий в секунду.

3. Разработана модель гетерогенной сети с возможностью установления прямых соединений между пользовательскими устройствами, применимая для оценки параметров качества обслуживания. Данная модель позволила определить, что среднее число передающих пользователей может быть увеличено на 20–25% при входной интенсивности 10–20 сессий в секунду.

4. Сформулирован алгоритм с одновременным подключением пользователя к нескольким узлам сетевой инфраструктуры при передаче на крайне высоких частотах, позволяющий понизить вероятность сброса пользовательской сессии от двух до десяти раз по сравнению с базовым алгоритмом без переподключений.

5. Предложена модель сотовой сети с конкурентным доступом большого числа одновременно подключающихся устройств, применимая для изучения показателей эффективности системы связи. Данная модель позволила определить, что вероятность коллизии может быть снижена до 50% при использовании интервала отсрочки длительностью от 1 до 3 секунд.

6. Разработана модель сети радиодоступа с конкурентным обслуживанием множества носимых устройств, применимая для оценки показателей эффективности системы радиодоступа на крайне высоких частотах. Данная модель позволила определить, что энергетическая эффективность персональной сети может быть повышена до трех раз в случае направленной передачи.

7. Предложен протокол управления системой прямых соединений, а также способ его реализации в гетерогенной сети с интегрированными между собой технологиями радиодоступа, позволяющие снизить нагрузку на сотовую сеть, повышая скорость передачи пользователей на 45% и более при выгрузке 20% трафика на прямые соединения.

8. Сформулирован алгоритм распределения трафика с сетевым кодированием, учитывающий структуру сессий, передаваемых в гетерогенной системе прямых соединений с большим количеством пользователей, и позволяющий повысить вероятность их успешного обслуживания до 20% по сравнению с передачей без использования сетевого кодирования.

Таким образом, в результате проведенных в диссертационной работе исследований решена **важная научно-техническая проблема** по созданию моделей и алгоритмов для обеспечения эффективного функционирования гетерогенных беспроводных сетей. В частности, предложены новые модели организации множественного доступа для пользователей гетерогенных сетей мобильной связи в условиях высокой плотности размещения устройств, а также созданы эффективные алгоритмы управления потоками трафика с целью снижения нагрузки на сотовую сеть передачи данных. Методология диссертации впервые объединяет методы теории массового обслуживания и стохастической геометрии, применяя их к современным гетерогенным сетям, и отличается от известных как своим комплексным подходом, так и учетом основных особенностей текущего этапа развития беспроводных систем.

Список литературы

1. Akyildiz, I. F. The evolution to 4G cellular systems: LTE-Advanced / I. F. Akyildiz, D. M. Gutierrez-Estevez, E. C. Reyes // *Physical communication*. — 2010. — Vol. 3, no. 4. — Pp. 217–244.
2. Ericsson. More than 50 billion connected devices [Электронный ресурс] // White paper. — 2011. — Режим доступа: https://www.akos-rs.si/files/Telekomunikacije/Digitalna_agenda/Internetni_protokol_Ipv6/More-than-50-billion-connected-devices.pdf (дата обращения: 05.11.2018).
3. Raychaudhuri, D. Frontiers of wireless and mobile communications / D. Raychaudhuri, N. B. Mandayam // *Proceedings of the IEEE*. — 2012. — Vol. 100, no. 4. — Pp. 824–840.
4. Tao, X. An overview of cooperative communications / X. Tao, X. Xu, Q. Cui // *IEEE Communications Magazine*. — 2012. — Vol. 50, no. 6. — Pp. 65–71.
5. Fodor, G. Design aspects of network assisted device-to-device communications / G. Fodor, E. Dahlman, G. Mildh, et al. // *IEEE Communications Magazine*. — 2012. — Vol. 50, no. 3. — Pp. 170–177.
6. Andreev, S. Network-assisted device-to-device connectivity: contemporary vision and open challenges / S. Andreev, D. Moltchanov, O. Galinina, et al. // *Proceedings of European Wireless 2015; 21th European Wireless Conference. / VDE*. — Budapest, Hungary. — 2015. — Pp. 1–8.
7. Tombaz, S. Energy-and cost-efficient ultra-high-capacity wireless access / S. Tombaz, A. Vastberg, J. Zander // *IEEE Wireless Communications*. — 2011. — Vol. 18, no. 5. — Pp. 18–24.
8. Andrews, J. Can cellular networks handle 1000x the data? [Электронный ресурс] / J. Andrews // Technical talk at University of Notre Dame. — 2011. — Режим доступа: http://users.ece.utexas.edu/~bevans/courses/realtime/lectures/Andrews_Cellular1000x_Nov2011.pdf (дата обращения: 05.11.2018).

9. Song, W. Multi-service load sharing for resource management in the cellular/WLAN integrated network / W. Song, W. Zhuang // IEEE Transactions on Wireless Communications. — 2009. — Vol. 8, no. 2. — Pp. 725–735.
10. Bellalta, B. Next generation IEEE 802.11 Wireless Local Area Networks: Current status, future directions and open challenges / B. Bellalta, L. Bononi, R. Bruno, et al. // Computer Communications. — 2016. — Vol. 75. — Pp. 1–25.
11. IEEE 802 Plenary Tutorial. — WPAN/WLAN/WWAN Multi-Radio Coexistence. — 2007, November.
12. Zetterman, T. Multi-radio coexistence and collaboration on an SDR platform / T. Zetterman, A. Piipponen, K. Raiskila, et al. // Analog Integrated Circuits and Signal Processing. — 2011. — Vol. 69, no. 2. — P. 329.
13. Sankaran, C. B. Data offloading techniques in 3GPP Rel-10 networks: A tutorial / C. B. Sankaran // IEEE Communications Magazine. — 2012. — Vol. 50, no. 6. — Pp.46–53.
14. Wang, L. Mathematical modeling for network selection in heterogeneous wireless networks — A tutorial / L. Wang, G. G. S. Kuo // IEEE Communications Surveys & Tutorials. — 2013. — Vol. 15, no. 1. — Pp. 271–292.
15. IEEE. Machine to machine (M2M) communication study report [Электронный ресурс] / H. Cho, J. Puthenkulam // IEEE 802.16 ppc-10/0002r6. — 2010. — Режим доступа: http://ieee802.org/16/ppc/docs/80216ppc-10_0002r7.doc (дата обращения: 05.11.2018).
16. IEEE. Informative text on Smart Grid applications for inclusion in IEEE 802.16p Systems Requirements Document (SRD) [Электронный ресурс] / N. Himayat, S. Talwar, K. Johnsson, et al. // IEEE C802. 16p-10/0007r1. — 2010. — Режим доступа: http://wirelessman.org/m2m/contrib/C80216p-10_0007r1.doc (дата обращения: 05.11.2018).
17. Cisco Systems Inc. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 [Электронный ресурс] // White paper. — 2017. — Режим доступа: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html> (дата обращения: 05.11.2018).

18. Lei, L. Operator controlled device-to-device communications in LTE-advanced networks / L. Lei, Z. Zhong, C. Lin, et al. // *IEEE Wireless Communications*. — 2012. — Vol. 19, no. 3. — Pp. 96–104.
19. Özgür, A. Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks / A. Özgür, O. Lévêque, D. N. C. Tse // *IEEE Transactions on information theory*. — 2007. — Vol. 53, no. 10. — Pp. 3549–3572.
20. Nosratinia, A. Grouping and partner selection in cooperative wireless networks / A. Nosratinia, T. E. Hunter // *IEEE Journal on Selected Areas in Communications*. — 2007. — Vol. 25, no. 2. — Pp. 369–378.
21. Ng, T. C. Joint optimization of relay strategies and resource allocations in cooperative cellular networks / T. C. Ng, W. Yu. // *IEEE Journal on Selected Areas in Communications*. — 2007. — Vol. 25, no. 2. — Pp. 328–339.
22. Wu, G. M2M: From mobile to embedded internet / G. Wu, S. Talwar, K. Johnson, et al. // *IEEE Communications Magazine*. — 2011. — Vol. 49, no. 4. — Pp. 36–43.
23. Zheng, K. Radio resource allocation in LTE-advanced cellular networks with M2M communications / K. Zheng, F. Hu, W. Wang, et al. // *IEEE Communications Magazine*. — 2012. — Vol. 50, no. 7. — Pp. 184–192.
24. Cheng, M. Overload control for machine-type-communications in LTE-advanced system / M. Cheng, G. Lin, H. Wei, et al. // *IEEE Communications Magazine*. — 2012. — Vol. 50, no. 6. — Pp. 38–45.
25. Pentikousis, K. In search of energy-efficient mobile networking / K. Pentikousis // *IEEE Communications Magazine*. — 2010. — Vol. 48, no. 1. — Pp. 95–103.
26. Li, G. Y. Energy-efficient wireless communications: tutorial, survey, and open issues / G. Y. Li, Z. Xu, C. Xiong, et al. // *IEEE Wireless Communications*. — 2011. — Vol. 18, no. 6. — Pp. 28–35.
27. Chen, Y. Fundamental tradeoffs on green wireless networks / Y. Chen, S. Zhang, S. Xu, et al. // *IEEE Communications Magazine*. — Jun. 2011. — Vol. 49, no. 6. — Pp. 30–37.

28. Hasan, Z. Green cellular networks: A survey, some research issues and challenges / Z. Hasan, H. Boostanimehr, V. K. Bhargava. // IEEE Communications Surveys & Tutorials. — 2011. — Vol. 13, no. 4. — Pp. 524–540.
29. Ponomarenko-Timofeev, A. Highly dynamic spectrum management within licensed shared access regulatory framework / A. Ponomarenko-Timofeev, A. Pyattaev, S. Andreev, et al. // IEEE Communications Magazine. — 2016. — Vol. 54, no. 3. — Pp. 100–109.
30. Mokrov, E. Modeling Transmit Power Reduction for a Typical Cell with Licensed Shared Access Capabilities / E. Mokrov, A. Ponomarenko-Timofeev, I. Gudkova, et al. // IEEE Transactions on Vehicular Technology. — 2018. — Vol. 67, no. 6. — Pp. 5505–5509.
31. Markova, E. Flexible Spectrum Management in a Smart City within Licensed Shared Access Framework / E. Markova, I. Gudkova, A. Ometov, et al. // IEEE Access. — 2017. — Vol. 5. — Pp. 22252–22261.
32. Talwar, S. Enabling technologies and architectures for 5G wireless / S. Talwar, D. Choudhury, K. Dimou, et al. // Proceedings of IEEE MTT-S International Microwave Symposium (IMS) / IEEE. — Tampa, FL, USA, 2014. — Pp. 1–4.
33. Andrews, J. G. What will 5G be? / J. G. Andrews, S. Buzzi, W. Choi, et al. // IEEE Journal on Selected Areas in Communications. — 2014. — Vol. 32, no. 6. — Pp. 1065–1082.
34. Karjalainen, J. Challenges and opportunities of mm-wave communication in 5G networks / J. Karjalainen, M. Nekovee, H. Benn, et al. // Proceedings of 9th International Conference on Cognitive Radio oriented Wireless networks and communications (CROWNCOM) / IEEE. — Oulu, Finland, 2014. — Pp. 372–376.
35. Rappaport, T. S. Millimeter wave mobile communications for 5G cellular: It will work! / T. S. Rappaport, S. Sun, R. Mayzus, et al. // IEEE Access. — 2013. — Vol. 1. — Pp. 335–349.
36. Boccardi, F. Five disruptive technology directions for 5G / F. Boccardi, R. W. Heath, A. Lozano, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 2. — Pp. 74–80.

37. Li, Q. C. Anchor-booster based heterogeneous networks with mmWave capable booster cells / Q. C. Li, H. Niu, G. Wu, et al. // Proceedings of Globecom Workshops (GC Wkshps) / IEEE. — Atlanta, GA, USA, 2013. — Pp. 93–98.
38. Nitsche, T. IEEE 802.11 ad: directional 60 GHz communication for multi-Gigabit-per-second Wi-Fi / T. Nitsche, C. Cordeiro, A. B. Flores, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 12. — Pp. 132–141.
39. METIS. Updated scenarios, requirements and KPIs for 5G mobile and wireless system with recommendations for future investigations [Электронный ресурс] / К. Kusume, M. Fallgren, O. Queseth, et. al // ICT-317669-METIS/D1.5. — 2015. — Режим доступа: <https://pdfs.semanticscholar.org/b5a5/01b7e58daafc923afe953e503e0dde27639e.pdf> (дата обращения: 05.11.2018).
40. 5G network technology: Putting Europe at the leading edge [Электронный ресурс] / R. Davies // European Parliamentary Research Service. — 2016. — Режим доступа: [http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/573892/EPRS_BRI\(2016\)573892_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/573892/EPRS_BRI(2016)573892_EN.pdf) (дата обращения: 05.11.2018).
41. NGMN Alliance. 5G White Paper — Executive Version [Электронный ресурс] / под. ред. J. Erfanian, B. Daly // White paper. — 2014. — Режим доступа: https://cdn3.scrvt.com/fokus/4e60fae4cbe2fea0/2fc1cf8cd1e9ad0c4c3f883ed9f181ad/141222_NGMN-Executive_Version_of_the_5G_White_Paper_v1_0.pdf (дата обращения: 05.11.2018).
42. ITU. IMT traffic estimates for the years 2020 to 2030 [Электронный ресурс] // Report ITU-R M.2370-0. — 2015. — Режим доступа: https://www.itu.int/dms_pub/itu-r/opb/rep/r-rep-m.2370-2015-pdf-e.pdf (дата обращения: 05.11.2018).
43. ITU. IMT Vision — Framework and overall objectives of the future development of IMT for 2020 and beyond [Электронный ресурс] // Recommendation ITU-R M.2083-0. — 2015. — Режим доступа: http://www.itu.int/dms_pubrec/itu-r/rec/m/r-rec-m.2083-0-201509-i!!pdf-e.pdf (дата обращения: 05.11.2018).

44. Gupta, A. A survey of 5G network: Architecture and emerging technologies / A. Gupta, R. K. Jha // *IEEE Access*. — 2015. — Vol. 3. — Pp. 1206–1232.
45. Гольдштейн, Б. С. Сети связи : Учебник для ВУЗов / Б. С. Гольдштейн, Н. А. Соколов, Г. Г. Яновский. — СПб. : БХВ-Петербург, 2014. — 400 с.
46. Da Silva, I. Tight integration of new 5G air interface and LTE to fulfill 5G requirements / I. Da Silva, G. Mildh, J. Rune, et al. // *Proceedings of IEEE 81st Vehicular Technology Conference (VTC Spring)* / IEEE. — Glasgow, UK, 2015. — Pp. 1–5.
47. Chen, S. The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication / S. Chen, J. Zhao. // *IEEE Communications Magazine*. — 2014. — Vol. 52, no. 5. — Pp. 36–43.
48. Condoluci, M. Enhanced radio access and data transmission procedures facilitating industry-compliant machine-type communications over LTE-based 5G networks / M. Condoluci, M. Dohler, G. Araniti, et al. // *IEEE Wireless Communications*. — 2016. — Vol. 23, no. 1. — Pp. 56–63.
49. Dahlman, E. 5G wireless access: requirements and realization / E. Dahlman, G. Mildh, S. Parkvall, et al. // *IEEE Communications Magazine*. — 2014. — Vol. 52, no. 12. — Pp. 42–47.
50. Andreev, S. Understanding the IoT connectivity landscape: a contemporary M2M radio technology roadmap / S. Andreev, O. Galinina, A. Pyattaev, et al. // *IEEE Communications Magazine*. — 2015. — Vol. 53, no. 9. — Pp. 32–40.
51. Palattella, M. R. Internet of things in the 5G era: Enablers, architecture, and business models / M. R. Palattella, M. Dohler, A. Grieco, et al. // *IEEE Journal on Selected Areas in Communications*. — 2016. — Vol. 34, no. 3. — Pp. 510–527.
52. Тихвинский, В. О. Сети 5G и IoT-инновационная среда для цифровой экономики России / В. О. Тихвинский // *Электросвязь*. — 2017. — № 8. — С. 18–24.

53. Andreev, S. Intelligent access network selection in converged multi-radio heterogeneous networks / S. Andreev, M. Gerasimenko, O. Galinina, et al. // IEEE Wireless Communications. — 2014. — Vol. 21, no. 6. — Pp. 86–96.
54. Lee, I. The Internet of Things (IoT): Applications, investments, and challenges for enterprises / I. Lee, K. Lee // Business Horizons. — 2015. — Vol. 58, no. 4. — Pp. 431–440.
55. Atzori, L. The internet of things: A survey / L. Atzori, A. Iera, G. Morabito // Computer Networks. — 2010. — Vol. 54, no. 15. — Pp. 2787–2805.
56. Li, S. The Internet of Things: a survey / S. Li, L. D. Xu, S. Zhao // Information Systems Frontiers. — 2015. — Vol. 17, no. 2. — Pp. 243–259.
57. Thierer, A. Projecting the Growth and Economic Impact of the Internet of Things [Электронный ресурс] / A. O’Sullivan, A. Thierer // Mercatus Center, George Mason University. — 2015. — Режим доступа: <https://www.mercatus.org/publication/projecting-growth-and-economic-impact-internet-things> (дата обращения: 05.11.2018).
58. World Economic Forum Industrial Internet of Things: Unleashing the Potential of Connected Products and Services [Электронный ресурс] // Industry Agenda. — 2015. — Режим доступа: http://www3.weforum.org/docs/WEFUSA_IndustrialInternet_Report2015.pdf (дата обращения: 05.11.2018).
59. Hermann, M. Design Principles for Industrie 4.0 Scenarios: A Literature Review [Электронный ресурс] / M. Hermann, T. Pentek, B. Otto // Working paper. — 2015. — no. 01. — Режим доступа: http://www.iim.mb.tu-dortmund.de/cms/de/forschung/Arbeitsberichte/Design-Principles-for-Industrie-4_0-Scenarios.pdf (дата обращения: 05.11.2018).
60. Varghese, A. Wireless requirements and challenges in Industry 4.0 / A. Varghese, D. Tandur // Proceedings of International Conference on Contemporary Computing and Informatics (IC3I) / IEEE. — Mysore, India, 2014. — Pp. 634–638.

61. Бородин, А. С. Сети связи пятого поколения как основа цифровой экономики / А. С. Бородин, А. Е. Кучерявый // Электросвязь. — 2017. — № 5. — С. 45–51.
62. Wang, S. Implementing smart factory of Industrie 4.0: an outlook / S. Wang, J. Wan, D. Li, et al. // International Journal of Distributed Sensor Networks. — 2016. — Vol. 12, no. 1. — Pp. 1–10.
63. Galinina, O. Leveraging heterogeneous device connectivity in a converged 5G-IoT ecosystem / O. Galinina, S. Andreev, M. Komarov, et al. // Computer Networks. — 2017. — Vol. 128. — Pp. 123–132.
64. 5G-PPP. 5G and the Factories of the Future [Электронный ресурс] / J.-S. Bedo, E. C. Strinati, S. Castellvi, et al. // White paper. — 2015. — Режим доступа: <https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White-Paper-on-Factories-of-the-Future-Vertical-Sector.pdf> (дата обращения: 05.11.2018).
65. Moskvitch, K. Tactile Internet: 5G and the Cloud on steroids / K. Moskvitch // Engineering & Technology. — 2015. — Vol. 10, no. 4. — Pp. 48–53.
66. Simsek, M. 5G-enabled tactile internet / M. Simsek, A. Aijaz, M. Dohler, et al. // IEEE Journal on Selected Areas in Communications. — 2016. — Vol. 34, no. 3. — Pp. 460–473.
67. Derhamy, H. A survey of commercial frameworks for the Internet of Things / H. Derhamy, J. Eliasson, J. Delsing, et al. // IEEE 20th Conference on Emerging Technologies and Factory Automation (ETFA) / IEEE Communications Society. — Luxembourg, Luxembourg, 2015. — Pp. 1–8.
68. Yilmaz, O. N. C. Analysis of ultra-reliable and low-latency 5G communication for a factory automation use case / O. N. C. Yilmaz, Y. E. Wang, N. A. Johansson, et al. // IEEE International Conference on Communication Workshop (ICCW) / IEEE. — London, UK, 2015. — Pp. 1190–1195.
69. Osseiran, A. Manufacturing reengineered: robots, 5G and the Industrial IoT [Электронный ресурс] / A. Osseiran, J. Sachs, M. Puleri, et al. // Ericsson Business Review. — 2015. — Режим доступа: <https://www.ericsson.com/assets/local/publications/ericsson-business-review/>

issue-4--2015/ebr-issue4-2015-industrial-iot.pdf (дата обращения: 05.11.2018).

70. Кучерявый, А. Е. Тактильный Интернет. Сети связи со сверхмалыми задержками / А. Е. Кучерявый, М. А. Маколкина, Р. В. Киричек // Электросвязь. — 2016. — № 1. — С. 44–46.
71. Никитин, А. В. Концепция построения мультисервисной сети оператора связи / А. В. Никитин, В. О. Пяттаев, И. Е. Никульский и др. // Вестник связи. — 2010. — № 5. — С. 47–49; № 7. — С. 41–45.
72. Ali, A. Next generation M2M cellular networks: challenges and practical considerations / A. Ali, W. Hamouda, M. Uysal // IEEE Communications Magazine. — 2015. — Vol. 53, no. 9. — Pp. 18–24.
73. Orsino, A. Exploiting D2D communications at the network edge for mission-critical IoT applications / A. Orsino, I. Farris, L. Militano, et al. // Proceedings of 23th European Wireless Conference / VDE. — Dresden, Germany, 2017. — Pp. 1–6.
74. Wu, J. A survey on high mobility wireless communications: Challenges, opportunities and solutions / J. Wu, P. Fan // IEEE Access. — 2016. — Vol. 4. — Pp. 450–476.
75. Grossglauser, M. Mobility increases the capacity of ad hoc wireless networks / M. Grossglauser, D. N. C. Tse // IEEE/ACM Transactions on Networking. — 2002. — Vol. 10, no. 4. — Pp. 477–486.
76. Андреев, С. Д. Оценка производительности простейшей системы абонентской кооперации / С. Д. Андреев, А. В. Винель, О. С. Галинина // Информационно-управляющие системы. — 2010. — Т. 49, № 6. — С. 33–41.
77. Пяттаев, А. В. Подходы к моделированию сетей связи с клиентской ретрансляцией / А. В. Пяттаев, Е. А. Кучерявый, С. Д. Андреев // Системы управления и информационные технологии. — 2013. — Т. 52, № 2. — С. 98–102.
78. Fodor, G. Providing quality of service in always best connected networks / G. Fodor, A. Eriksson, A. Tuoriniemi // IEEE Communications Magazine. — 2003. — Vol. 41, no. 7. — Pp. 154–163.

79. Orsino, A. Effects of heterogeneous mobility on D2D- and drone-assisted mission-critical MTC in 5G / A. Orsino, A. Ometov, G. Fodor, et al. // *IEEE Communications Magazine*. — 2017. — Vol. 55, no. 2. — Pp. 79–87.
80. Гайдамака, Ю. В. Модели и методы анализа и расчета показателей эффективности беспроводных гетерогенных сетей : монография / Ю. В. Гайдамака, Э. С. Сопин, И. А. Гудкова и др. — М. : ФИЦ ИУ РАН, 2018. — 71 с.
81. Shariatmadari, H. Machine-type communications: current status and future perspectives toward 5G systems / H. Shariatmadari, R. Ratasuk, S. Iraj, et al. // *IEEE Communications Magazine*. — 2015. — Vol. 53, no. 9. — Pp. 10–17.
82. Fodor, G. Device-to-device communications for national security and public safety / G. Fodor, S. Parkvall, S. Sorrentino, et al. // *IEEE Access*. — 2014. — Vol. 2. — Pp. 1510–1520.
83. Song, L. Game-theoretic resource allocation methods for device-to-device communication / L. Song, D. Niyato, Z. Han, et al. // *IEEE Wireless Communications*. — 2014. — Vol. 21, no. 3. — Pp. 136–144.
84. Yeh, S. Capacity and coverage enhancement in heterogeneous networks / S. Yeh, S. Talwar, G. Wu, et al. // *IEEE Wireless Communications*. — 2011. — Vol. 18, no. 3. — Pp. 32–38.
85. Marsch, P. Future mobile communication networks: Challenges in the design and operation / P. Marsch, B. Raaf, A. Szufarska, et al. // *IEEE Vehicular Technology Magazine*. — 2012. — Vol. 7, no. 1. — Pp. 16–23.
86. Galinina, O. Modeling Multi-Radio Coordination and Integration in Converged Heterogeneous Networks / O. Galinina, S. Andreev, A. Pyattaev, et al. // R. Vannithamby, S. Talwar (Eds.): *Towards 5G: Applications, Requirements and Candidate Technologies*. — John Wiley & Sons, Ltd. — 2016. — Pp. 99–128.
87. Andrews, J. G. An overview of load balancing in HetNets: Old myths and open problems / J. G. Andrews, S. Singh, Q. Ye, et al. // *IEEE Wireless Communications*. — 2014. — Vol. 21, no. 2. — Pp. 18–25.

88. Andreev, S. Cellular traffic offloading onto network-assisted device-to-device connections / S. Andreev, A. Pyattaev, K. Johnsson, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 4. — Pp. 20–31.
89. Baldemair, R. Evolving wireless communications: Addressing the challenges and expectations of the future / R. Baldemair, E. Dahlman, G. Fodor, et al. // IEEE Vehicular Technology Magazine. — 2013. — Vol. 8, no. 1. — Pp. 24–30.
90. Andrews, J. G. Seven ways that HetNets are a cellular paradigm shift / J. G. Andrews // IEEE Communications Magazine. — 2013. — Vol. 51, no. 3. — Pp. 136–144.
91. Bennis, M. When cellular meets WiFi in wireless small cell networks / M. Bennis, M. Simsek, A. Czylik, et al. // IEEE Communications Magazine. — 2013. — Vol. 51, no. 6. — Pp. 44–50.
92. 3GPP. 3GPP/WLAN RAN Interworking Study Item Report [Электронный ресурс] // 3GPP TR 37.842. — 2017. — Режим доступа: http://www.3gpp.org/ftp//Specs/archive/37_series/37.842/37842-d20.zip (дата обращения: 05.11.2018).
93. Urama, J. A Multi-Purpose Automated Vehicular Platform with Multi-Radio Connectivity Capabilities / J. Urama, M. Gerasimenko, M. Stusek, et al. // 2018 IEEE 87th Vehicular Technology Conference (VTC Spring) / IEEE. — Porto, Portugal, 2018. — Pp. 1–7.
94. Himayat, N. Design and performance analysis of multi-radio small cell networks / N. Himayat, S. Yeh, S. Talwar, et al. // A. Anpalagan, M. Bennis, R. Vannithamby (Eds.): Design and Deployment of Small Cell Networks. — Cambridge University Press. — 2015. — Pp. 31–57.
95. Gapeyenko, M. Flexible and Reliable UAV-Assisted Backhaul Operation in 5G mmWave Cellular Networks / M. Gapeyenko, V. Petrov, D. Moltchanov, et al. // IEEE Journal on Selected Areas in Communications. — 2018. — Vol. 36, no. 11. — Pp. 2486–2496.
96. Peng, M. Self-configuration and self-optimization in LTE-advanced heterogeneous networks / M. Peng, D. Liang, Y. Wei, et al. // IEEE Communications Magazine. — 2013. — Vol. 51, no. 5. — Pp. 36–45.

97. Dhillon, H. S. Modeling and analysis of K-tier downlink heterogeneous cellular networks / H. S. Dhillon, R. K. Ganti, F. Baccelli, et al. // *IEEE Journal on Selected Areas in Communications*. — 2012. — Vol. 30, no. 3. — Pp. 550–560.
98. Kim, H. Leveraging dynamic spare capacity in wireless systems to conserve mobile terminals' energy / H. Kim, G. De Veciana // *IEEE/ACM Transactions on Networking*. — 2010. — Vol. 18, no. 3. — Pp. 802–815.
99. Andrews, J. G. A primer on spatial modeling and analysis in wireless networks / J. G. Andrews, R. K. Ganti, M. Haenggi, et al. // *IEEE Communications Magazine*. — 2010. — Vol. 48, no. 11. — Pp. 156–163.
100. Kelly, F. *Stochastic Networks*. / F. Kelly, E. Yudovina. — Cambridge : Cambridge University Press, 2014. — 222 p.
101. Амбарцумян, Р. В. Введение в стохастическую геометрию / Р. В. Амбарцумян, Й. Мекке, Д. Штойян. — М. Наука, 1989. — 400 с.
102. ElSawy, H. Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey / H. ElSawy, E. Hossain, M. Haenggi. // *IEEE Communications Surveys & Tutorials*. — 2013. — Vol. 15, no. 3. — Pp. 996–1019.
103. Sankararaman, A. Spatial Birth-Death Wireless Networks / A. Sankararaman, F. Baccelli // *IEEE Transactions on Information Theory*. — 2017. — Vol. 63, no. 6. — Pp. 3964–3982.
104. Андреев, С. Д. Пространственно-временной подход к анализу гетерогенных систем связи / С. Д. Андреев, Е. А. Кучерявый, К. Е. Самуйлов // *Электросвязь*. — 2018. — № 9. — С. 20–26.
105. Galinina, O. Capturing spatial randomness of heterogeneous cellular/WLAN deployments with dynamic traffic / O. Galinina, S. Andreev, M. Gerasimenko, et al. // *IEEE Journal on Selected Areas in Communications*. — 2014. — Vol. 32, no. 6. — Pp. 1083–1099.
106. 3GPP. Study on WLAN/3GPP Radio Interworking [Электронный ресурс] // 3GPP TR 37.834. — 2013. — Режим доступа: http://www.3gpp.org/ftp/Specs/archive/37_series/37.834/37834-110.zip (дата обращения: 05.11.2018).

107. ITU. Guidelines for evaluation of radio interface technologies for IMT-Advanced [Электронный ресурс] // Report ITU-R M.2135-1. — 2009. — Режим доступа: <https://www.itu.int/pub/R-REP-M.2135-1-2009> (дата обращения: 05.11.2018).
108. Бородина, А. В. Оценивание эффективной пропускной способности узла в инфокоммуникационной тандемной сети / А. В. Бородина, Е. В. Морозов // Системы и средства информатики. — 2014. — Т. 24, № 2. — С. 37–54.
109. Galinina, O. Energy-efficient operation of a mobile user in a multi-tier cellular network / O. Galinina, A. Trushanin, V. Shumilov, et al. // Proceedings of International Conference on Analytical and Stochastic Modeling Techniques and Applications / Springer. — Ghent, Belgium, 2013. — Pp. 198–213.
110. Клейнрок, Л. Теория массового обслуживания / Л. Клейнрок. — М. : Машиностроение, 1979. — 432 с.
111. Бочаров, П. П. Теория массового обслуживания / П. П. Бочаров, А. В. Печинкин. — М. :Изд-во РУДН, 1995. — 529 с.
112. Xu, J. On the accuracy of the Wyner model in cellular networks / J. Xu, J. Zhang, J. G. Andrews // IEEE Transactions on Wireless Communications. — 2011. — Vol. 10, no. 9. — Pp. 3098–3109.
113. Степанов, С. Н. Планирование ресурса передачи при совместном обслуживании мультисервисного трафика реального времени и эластичного трафика данных / С. Н. Степанов, М. С. Степанов // Автоматика и телемеханика. — 2017. — Т. 78, № 11. — С. 79–93.
114. 3GPP. Coordinated multi-point operation for LTE physical layer aspects [Электронный ресурс] // 3GPP TR 36.819. — 2013, September. — Режим доступа: http://www.3gpp.org/ftp/Specs/archive/36_series/36.819/36819-b20.zip (дата обращения: 05.11.2018).
115. Ghosh, A. Heterogeneous cellular networks: From theory to practice / A. Ghosh, N. Mangalvedhe, R. Ratasuk, et al. // IEEE Communications Magazine. — 2012. — Vol. 50, no. 6. — Pp. 54–64.

116. Park, J. Asymptotic behavior of ultra-dense cellular networks and its economic impact / J. Park, S.-L. Kim, J. Zander // Proceedings of Global Communications Conference (GLOBECOM) / IEEE. — Austin, USA, 2014. — Pp. 4941–4946.
117. Singh, S. Offloading in heterogeneous networks: Modeling, analysis, and design insights / S. Singh, H. S. Dhillon, J. G. Andrews // IEEE Transactions on Wireless Communications. — 2013. — Vol. 12, no. 5. — Pp. 2484–2497.
118. Dhillon, H. S. Load-aware modeling and analysis of heterogeneous cellular networks / H. S. Dhillon, R. K. Ganti, J. G. Andrews // IEEE Transactions on Wireless Communications. — 2013. — Vol. 12, no. 4. — Pp. 1666–1677.
119. Himayat, N. Multi-radio heterogeneous networks: Architectures and performance / N. Himayat, S. Yeh, A. Y. Panah, et al. // Proceedings of International Conference on Computing, Networking and Communications (ICNC) / IEEE. — Honolulu, HI, USA, 2014. — Pp. 252–258.
120. Osseiran, A. Scenarios for 5G mobile and wireless communications: the vision of the METIS project / A. Osseiran, F. Boccardi, V. Braun, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 5. — Pp. 26–35.
121. Xu, J. Cooperative distributed optimization for the hyper-dense small cell deployment / J. Xu, J. Wang, Y. Zhu, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 5. — Pp. 61–67.
122. Li, Q. C. 5G network capacity: Key elements and technologies / Q. C. Li, H. Niu, A. T. Papathanassiou, et al. // IEEE Vehicular Technology Magazine. — 2014. — Vol. 9, no. 1. — Pp. 71–78.
123. Bangerter, B. Networks and devices for the 5G era / B. Bangerter, S. Talwar, R. Arefi, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 2. — Pp. 90–96.
124. Öhmann, D. Transient flow level models for interference-coupled cellular networks / D. Öhmann, A. Fehske, G. Fettweis // Proceedings of 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton) / IEEE. — Monticello, IL, USA, 2013. — Pp. 723–730.

125. Galinina, O. 5G multi-RAT LTE-WiFi ultra-dense small cells: Performance dynamics, architecture, and trends / O. Galinina, A. Pyattaev, S. Andreev, et al. // IEEE Journal on Selected Areas in Communications. — 2015. — Vol. 33, no. 6. — Pp. 1224–1240.
126. Astely, D. LTE release 12 and beyond / D. Astely, E. Dahlman, G. Fodor, et al. // IEEE Communications Magazine. — 2013. — Vol. 51, no. 7. — Pp. 154–160.
127. Nakamura, T. Trends in small cell enhancements in LTE advanced / T. Nakamura, S. Nagata, A. Benjebbour, et al. // IEEE Communications Magazine. — 2013. — Vol. 51, no. 2. — Pp. 98–105.
128. Kojima, Y. LTE-WiFi link aggregation at femtocell base station / Y. Kojima, J. Suga, T. Kawasaki, et al. // Proceedings of World Telecommunications Congress 2014 / VDE. — Berlin, Germany, 2014. — Pp. 1–6.
129. Damnjanovic, A. A survey on 3GPP heterogeneous networks / A. Damnjanovic, J. Montojo, Y. Wei, et al. // IEEE Wireless communications. — 2011. — Vol. 18, no. 3. — Pp. 10–21.
130. Кучерявый, Е. А. Концепция фантомных пользователей для исследования гетерогенных сетей / Е. А. Кучерявый, С. Д. Андреев // Электросвязь. — 2018. — № 4. — С. 41–47.
131. Lopez-Perez, D. On the expanded region of picocells in heterogeneous networks / D. Lopez-Perez, X. Chu, İ Guvenc // IEEE Journal of Selected Topics in Signal Processing. — 2012. — Vol. 6, no. 3. — Pp. 281–294.
132. Hossain, E. Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective / E. Hossain, M. Rasti, H. Tabassum, et al. // IEEE Wireless Communications. — 2014. — Vol. 21, no. 3. — Pp. 118–127.
133. Bennis, M. Self-organization in small cell networks: A reinforcement learning approach / M. Bennis, S. M. Perlaza, P. Blasco, et al. // IEEE Transactions on Wireless Communications. — 2013. — Vol. 12, no. 7. — Pp. 3202–3212.
134. Borst, S. Optimal resource allocation in HetNets / S. Borst, S. Hanly, P. Whiting // IEEE International Conference on Communications (ICC) / IEEE. — Budapest, Hungary, 2013. — Pp. 5437–5441.

135. W.I.N.T.E.R. Group WINTERsim tool [Электронный ресурс] // Режим доступа: <http://winter-group.net/downloads/> (дата обращения: 05.11.2018).
136. Gerasimenko, M. Prioritized centrally-controlled resource allocation in integrated multi-RAT HetNets / M. Gerasimenko, D. Moltchanov, R. Florea, et al. // IEEE 81st Vehicular Technology Conference (VTC Spring) / IEEE. — Glasgow, UK, 2015. — Pp. 1–7.
137. 3GPP. Further advancements for E-UTRA physical layer aspects [Электронный ресурс] // 3GPP TR 36.814. — 2010. — Режим доступа: http://www.3gpp.org/ftp//Specs/archive/36_series/36.814/36814-920.zip (дата обращения: 05.11.2018).
138. Navaratnarajah, S. Energy efficiency in heterogeneous wireless access networks / S. Navaratnarajah, A. Saeed, M. Dianati, et al. // IEEE Wireless Communications. — 2013. — Vol. 20, no. 5. — Pp. 37–43.
139. Doppler, K. Device-to-device communication as an underlay to LTE-advanced networks / K. Doppler, M. Rinne, C. Wijting, et al. // IEEE Communications Magazine. — 2009. — Vol. 47, no. 12. — Pp. 42–49.
140. Al-Kanj, L. Energy-Aware Cooperative Content Distribution over Wireless Networks: Design Alternatives and Implementation Aspects. / L. Al-Kanj, Z. Dawy, E. Yaacoub // IEEE Communications Surveys & Tutorials. — 2013. — Vol. 15, no. 4. — Pp. 1736–1760.
141. Пяттаев, А. В. Разработка и исследование системы прямых соединений D2D с сотовой поддержкой LTE / А. В. Пяттаев, С. Д. Андреев, Е. А. Кучерявый // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление. — 2018. — Т. 11, № 2. — С. 19–34.
142. Corson, M. S. Toward proximity-aware internetworking / M. S. Corson, R. Laroia, J. Li, et al. // IEEE Wireless Communications. — 2010. — Vol. 17, no. 6. — Pp. 26–33.
143. Belleschi, M. Performance analysis of a distributed resource allocation scheme for D2D communications / M. Belleschi, G. Fodor, A. Abrardo // Proceedings

- of GLOBECOM Workshops (GC Wkshps) / IEEE. — Houston, TX, USA, 2011. — Pp. 358–362.
144. Janis, P. Device-to-device communication underlaying cellular communications systems / P. Janis, C. Yu, K. Doppler, et al. // International Journal of Communications, Network and System Sciences. — 2009. — Vol. 2, no. 3. — P. 169–178.
 145. Yu, C. Resource sharing optimization for device-to-device communication underlaying cellular networks / C. Yu, K. Doppler, C. B. Ribeiro, et al. // IEEE Transactions on Wireless Communications. — 2011. — Vol. 10, no. 8. — Pp. 2752–2763.
 146. Kaufman, B. Spectrum sharing scheme between cellular users and ad-hoc device-to-device users / B. Kaufman, J. Lilleberg, B. Aazhang // IEEE Transactions on Wireless Communications. — 2013. — Vol. 12, no. 3. — Pp. 1038–1049.
 147. Yu, C. On the performance of device-to-device underlay communication with simple power control / C. Yu, O. Tirkkonen, K. Doppler, et al. // Proceedings of 69th Vehicular Technology Conference (VTC Spring) / IEEE. — Barcelona, Spain, 2009. — Pp. 1–5.
 148. Doppler, K. Advances in D2D communications: Energy efficient service and device discovery radio / K. Doppler, C. B. Ribeiro, J. Knecht // Proceedings of 2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE) / IEEE. — Chennai, India, 2011. — Pp. 1–6.
 149. Seppälä, J. Network controlled device-to-device (D2D) and cluster multicast concept for LTE and LTE-A networks / J. Seppälä, T. Koskela, T. Chen, et al. // Proceedings of Wireless Communications and Networking Conference (WCNC) / IEEE. — 2011. — Cancun, Quintana Roo, Mexico, 2011. — Pp. 986–991.
 150. Min, H. Reliability improvement using receive mode selection in the device-to-device uplink period underlaying cellular networks / H. Min, W. Seo, J. Lee, et al. // IEEE Transactions on Wireless Communications. — 2011. — Vol. 10, no. 2. — Pp. 413–418.

151. Chen, T. Time hopping for device-to-device communication in LTE cellular system / T. Chen, G. Charbit, S. Hakola // Proceedings of Wireless Communications and Networking Conference (WCNC) / IEEE. — Sydney, NSW, Australia, 2010. — Pp. 1–6.
152. Xu, S. Device-to-Device Communication Underlying Cellular Networks: Connection Establishment and Interference Avoidance. / S. Xu, H. Wang, T. Chen, et al. // KSII Transactions on Internet & Information Systems. — 2012. — Vol. 6, no. 1. — Pp. 203–228.
153. Jung, M. Joint mode selection and power allocation scheme for power-efficient device-to-device (D2D) communication / M. Jung, K. Hwang, S. Choi // Proceedings of 75th Vehicular technology conference (VTC Spring) / IEEE. — Yokohama, Japan, 2012. — Pp. 1–5.
154. Vigato, A. Joint discovery in synchronous wireless networks / A. Vigato, L. Vangelista, C. Measson, et al. // IEEE Transactions on Communications. — 2011. — Vol. 59, no. 8. — Pp. 2296–2305.
155. Mukherjee, A. Energy-efficient device-to-device MIMO underlay network with interference constraints / A. Mukherjee, A. Hottinen // Proceedings of International ITG Workshop on Smart Antennas (WSA) / IEEE. — Dresden, Germany, 2012. — Pp. 105–109.
156. Osseiran, A. Advances in Device-to-Device Communications and Network Coding for IMT-Advanced / A. Osseiran, K. Doppler, C. Ribeiro, et al. // Proceedings of the ICT-MobileSummit. — Santander, Spain, 2009. — Pp. 1–8.
157. Yu, C. Device-to-device underlay cellular network based on rate splitting / C. Yu, O. Tirkkonen // Wireless Communications and Networking Conference (WCNC) / IEEE. — Shanghai, China, 2012. — Pp. 262–266.
158. Golrezaei, N. Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution / N. Golrezaei, A. F. Molisch, A. G. Dimakis, et al. // IEEE Communications Magazine. — 2013. — Vol. 51, no. 4. — Pp. 142–149.
159. Qualcomm Technologies. LTE Direct: The Case for Device-to-Device Proximate Discovery [Электронный ресурс] // Qualcomm Research.

- 2013, February. — Режим доступа: https://www.qualcomm.com/system/files/document/files/Research_LTE-D_White_Paper.pdf (дата обращения: 05.11.2018).
160. Красилов, А. Н. Аналитическая модель взаимодействия механизмов случайного и детерминированного доступа к каналу в сетях Wi-Fi Mesh / А. Н. Красилов, А. И. Ляхов, Ю. И. Мороз // Автоматика и телемеханика. — 2013. — Т. 74, № 10. — С. 119–136.
161. Xing, B. An experimental study on Wi-Fi ad-hoc mode for mobile device-to-device video delivery / B. Xing, K. Seada, N. Venkatasubramanian // Proceedings of INFOCOM Workshops / IEEE. — Rio de Janeiro, Brazil, 2009. — Pp. 1–6.
162. WiFi Alliance. Wi-Fi Peer-to-Peer (P2P) Specifications, v1.7 [Электронный ресурс] // Режим доступа: <https://www.wi-fi.org/file/wi-fi-peer-to-peer-p2p-technical-specification-v17> (дата обращения: 05.11.2018).
163. Степанов, С. Н. Основы телетрафика мультисервисных сетей / С. Н. Степанов. — М. : Эко-Трендз, 2010. — 392 с.
164. Вишневецкий, В. М. Системы массового обслуживания с коррелированными входными потоками и их применение для моделирования телекоммуникационных сетей / В. М. Вишневецкий, А. Н. Дудин // Автоматика и телемеханика. — 2017. — Т. 78, № 8. — С. 3–59.
165. Андреев, С. Д. Управление потоками в гетерогенных мобильных сетях радиодоступа с соединениями устройство — устройство / С. Д. Андреев, К. Е. Самуйлов, А. М. Тюрликов // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2018. — Т. 26, № 4. — С. 357–370.
166. Huang, K. Spectrum sharing between cellular and mobile ad hoc networks: transmission-capacity trade-off / K. Huang, V. K. N. Lau, Y. Chen // IEEE Journal on Selected Areas in Communications. — 2009. — Vol. 27, no. 7. — Pp. 1256–1267.
167. Vaccelli, F. On the design of device-to-device autonomous discovery / F. Vaccelli, N. Khude, R. Laroia, et al. // Proceedings of Fourth International Con-

- ference on Communication Systems and Networks (COMSNETS) / IEEE. — Bangalore, India, 2012. — Pp. 1–9.
168. Erturk, M. C. Distributions of transmit power and SINR in device-to-device networks / M. C. Erturk, S. Mukherjee, H. Ishii, H. Arslan // IEEE Communications Letters. — 2013. — Vol. 17, no. 2. — Pp. 273–276.
169. Pyattaev, A. Network-assisted D2D over WiFi direct / A. Pyattaev, O. Galinina, K. Johnsson, et al. // S. Mumtaz, J. Rodriguez (Eds.): Smart Device to Smart Device Communication. — Springer. — 2014. — Pp. 165–218.
170. Andreev, S. Analyzing assisted offloading of cellular user sessions onto D2D links in unlicensed bands / S. Andreev, O. Galinina, A. Pyattaev, et al. // IEEE Journal on Selected Areas in Communications. — 2015. — Vol. 33, no. 1. — Pp. 67–80.
171. ITU. Guidelines for evaluation of radio interface technologies for IMT-Advanced [Электронный ресурс] // Recommendation ITU-R M.2135-1. — 2009. — Режим доступа: <https://www.itu.int/pub/R-REP-M.2135-1-2009> (дата обращения: 05.11.2018).
172. Pyattaev, A. Proximity-based data offloading via network assisted device-to-device communications / A. Pyattaev, K. Johnsson, S. Andreev, et al. // Proceedings of 77th Vehicular Technology Conference (VTC Spring) / IEEE. — Dresden, Germany, 2013. — Pp. 1–5.
173. Pyattaev, A. 3GPP LTE traffic offloading onto WiFi direct / A. Pyattaev, K. Johnsson, S. Andreev, et al. // Wireless Communications and Networking Conference Workshops (WCNCW) / IEEE. — Shanghai, China, 2013. — Pp. 135–140.
174. Gebert, J. Probabilities for opportunistic networking in different scenarios / J. Gebert, R. Fuchs // Future Network & Mobile Summit (FutureNetw) / IEEE. — Berlin, Germany, 2012. — Pp. 1–8.
175. Qualcomm Atheros. AR4100 System in Package 802.11n [Электронный ресурс] // General Availability. — 2012, April. — Режим доступа: https://www.mouser.com/ds/2/302/AR4100P_DATASHEET-1125841.pdf (дата обращения: 05.11.2018).

176. Jensen, A. R. LTE UE Power Consumption Model: For System Level Energy and Performance Optimization / A. R. Jensen, M. Lauridsen, P. Mogensen, et al. // Proceedings of the IEEE Vehicular Technology Conference (VTC Fall). — Quebec City, QC, Canada, 2012. — Pp. 1–5.
177. 3GPP. Channel model for frequency spectrum above 6 GHz [Электронный ресурс] // 3GPP TR 38.900 V2.0.0. — 2016. — Режим доступа: http://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_72/Docs/RP-161222.zip (дата обращения: 05.11.2018).
178. 3GPP. Study on scenarios and requirements for next generation access technologies [Электронный ресурс] // 3GPP TR 38.913 V0.3.0. — 2016. — Режим доступа: http://www.3gpp.org/ftp//Specs/archive/38_series/38.913/38913-f00.zip (дата обращения: 05.11.2018).
179. Ometov, A. Modeling Unreliable Operation of mmWave-Based Data Sessions in Mission-Critical PPDR Services / A. Ometov, E. Sopin, I. Gudkova, et al. // IEEE Access. — 2017. — Vol. 5. — Pp. 20536–20544.
180. Gapeyenko, M. Effects of Blockage in Deploying mmWave Drone Base Stations for 5G Networks and Beyond / M. Gapeyenko, I. Bor-Yaliniz, S. Andreev, et al. // 2018 IEEE International Conference on Communications Workshops (ICC Workshops) / IEEE. — Kansas City, MO, USA, 2018. — Pp. 1–6.
181. Kovalchukov, R. Improved Session Continuity in 5G NR with Joint Use of Multi-Connectivity and Guard Bandwidth / R. Kovalchukov, D. Moltchanov, V. Begishev, et al. // Proceedings of Global Communications Conference (GLOBECOM) / IEEE. — Abu Dhabi, UAE, 2018. — Pp. 1–7.
182. 3GPP. RAN 5G Workshop — The Start of Something [Электронный ресурс] // 2015. — Режим доступа: <http://www.3gpp.org/news-events/3gpp-news/1734-ran5g/> (дата обращения: 05.11.2018).
183. Nurmela, V. METIS D1.2: Initial channel models based on measurements / V. Nurmela, A. Karttunen, A. Roivainen, et al. // ICT-317669 METIS 2020. 2014. — Pp. 1–17.
184. COST IC1004. Scientific challenges towards 5G mobile communications [Электронный ресурс] // COST IC1004. — 2016, April. — Режим доступа:

- <http://www.ic1004.org/uploads/Documents/COST%20IC1004%20White%20Paper%20on%20Mobile%20Comms%20Challenges%20towards%205G%20-%202013.pdf> (дата обращения: 05.11.2018).
185. Haneda, K. 5G 3GPP-like channel models for outdoor urban microcellular and macrocellular environments / K. Haneda, J. Zhang, L. Tan, et al. // Proceedings of 83rd Vehicular Technology Conference (VTC Spring) / IEEE. — Nanjing, China, 2016. — Pp. 1–7.
 186. Muhammad, N. A. Analytical model for outdoor millimeter wave channels using geometry-based stochastic approach / N. A. Muhammad, P. Wang, Y. Li, et al. // IEEE Transactions on Vehicular Technology. — 2017. — Vol. 66, no. 2. — Pp. 912–926.
 187. Haneda, K. Frequency-agile pathloss models for urban street canyons / K. Haneda, N. Omaki, T. Imai, et al. // IEEE Transactions on Antennas and Propagation. — 2016. — Vol. 64, no. 5. — Pp. 1941–1951.
 188. Sun, S. Investigation of prediction accuracy, sensitivity, and parameter stability of large-scale propagation path loss models for 5G wireless communications / S. Sun, T. S. Rappaport, T. A. Thomas, et al. // IEEE Transactions on Vehicular Technology. — 2016. — Vol. 65, no. 5. — Pp. 2843–2860.
 189. Sun, S. Path loss, shadow fading, and line-of-sight probability models for 5G urban macro-cellular scenarios / S. Sun, T. A. Thomas, T. S. Rappaport, et al. // Proceedings of Globecom Workshops (GC Wkshps) / IEEE. — San Diego, CA, USA, 2015. — Pp. 1–7.
 190. Kovalchukov, R. Analyzing Effects of Directionality and Random Heights in Drone-Based mmWave Communication / R. Kovalchukov, D. Moltchanov, A. Samuylov, et al. // IEEE Transactions on Vehicular Technology. — 2018. — Vol. 67, no. 10. — Pp. 10064–10069.
 191. Petrov, V. On the use of integral geometry for interference modeling and analysis in wireless networks / V. Petrov, D. Moltchanov, P. Kustarev, et al. // IEEE Communications Letters. — 2016. — Vol. 20, no. 12. — Pp. 2530–2533.
 192. Venugopal, K. Device-to-device millimeter wave communications: Interference, coverage, rate, and finite topologies / K. Venugopal, M. C. Valenti,

- R. W. Heath // IEEE Transactions on Wireless Communications. — 2016. — Vol. 15, no. 9. — Pp. 6175–6188.
193. Petrov, V. Interference and SINR in millimeter wave and terahertz communication systems with blocking and directional antennas / V. Petrov, M. Komarov, D. Moltchanov, et al. // IEEE Transactions on Wireless Communications. — 2017. — Vol. 16, no. 3. — Pp. 1791–1808.
194. Mumtaz, S. MmWave Massive MIMO: A Paradigm for 5G / S. Mumtaz, J. Rodriguez, L. Dai (Eds.). — Academic Press, 2017. — 372 p.
195. Björnson, E. Massive MIMO: Ten myths and one critical question / E. Björnson, E. G. Larsson, T. L. Marzetta // IEEE Communications Magazine. — 2016. — Vol. 54, no. 2. — Pp. 114–123.
196. Gao, X. Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays / X. Gao, L. Dai, S. Han, et al. // IEEE Journal on Selected Areas in Communications. — 2016. — Vol. 34, no. 4. — Pp. 998–1009.
197. Liu, W. A tractable approach to uplink spectral efficiency of two-tier massive MIMO cellular HetNets / W. Liu, S. Jin, C. Wen, et al. // IEEE Communications Letters. — 2016. — Vol. 20, no. 2. — Pp. 348–351.
198. He, S. Resource Efficiency: A New Beamforming Design for Multicell Multiuser Systems / S. He, Y. Huang, Y. Lu, et al. // IEEE Transactions on Vehicular Technology. — 2016. — Vol. 65, no. 8. — Pp. 6063–6074.
199. Andrews, J. G. Modeling and analyzing millimeter wave cellular systems / J. G. Andrews, T. Bai, M. N. Kulkarni, et al. // IEEE Transactions on Communications. — 2017. — Vol. 65, no. 1. — Pp. 403–430.
200. Zhang, Z. Two-timeslot two-way full-duplex relaying for 5G wireless communication networks / Z. Zhang, Z. Ma, M. Xiao, et al. // IEEE Transactions on Communications. — 2016. — Vol. 64, no. 7. — Pp. 2873–2887.
201. Wang, P. Multi-gigabit millimeter wave wireless communications for 5G: From fixed access to cellular networks / P. Wang, Y. Li, L. Song, et al. // IEEE Communications Magazine. — 2015. — Vol. 53, no. 1. — Pp. 168–178.

202. Samuylov, A. Characterizing spatial correlation of blockage statistics in urban mmWave systems / A. Samuylov, M. Gapeyenko, D. Moltchanov, et al. // Proceedings of Globecom Workshops (GC Wkshps) / IEEE. — Washington, DC, USA, 2016. — Pp. 1–7.
203. Radwan, A. Low-Cost On-Demand C-RAN Based Mobile Small-Cells. / A. Radwan, K. M. S. Huq, S. Mumtaz, et al. // IEEE Access. — 2016. — Vol. 4. — Pp. 2331–2339.
204. Giordani, M. Multi-connectivity in 5G mmWave cellular networks / M. Giordani, M. Mezzavilla, S. Rangan, et al. // Proceedings of Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net) / IEEE. — Vilanova i la Geltru, Spain, 2016. — Pp. 1–7.
205. Lema, M. A. Flexible dual-connectivity spectrum aggregation for decoupled up-link and downlink access in 5G heterogeneous systems / M. A. Lema, E. Pardo, O. Galinina, et al. // IEEE Journal on Selected Areas in Communications. — 2016. — Vol. 34, no. 11. — Pp. 2851–2865.
206. Tesema, F. B. Mobility Modeling and Performance Evaluation of Multi-Connectivity in 5G Intra-Frequency Networks / F. B. Tesema, A. Awada, I. Viering, et al. // Proceedings of Globecom Workshops (GC Wkshps) / IEEE. — San Diego, CA, USA, 2015. — Pp. 1–6.
207. Ford, R. Achieving ultra-low latency in 5G millimeter wave cellular networks / R. Ford, M. Zhang, M. Mezzavilla, et al. // IEEE Communications Magazine. — 2017. — Vol. 55, no. 3. — Pp. 196–203.
208. Moltchanov, D. Upper bound on capacity of 5G mmWave cellular with multi-connectivity capabilities / D. Moltchanov, A. Ometov, S. Andreev, et al. // Electronics Letters. — 2018. — Vol. 54, no. 11. — Pp. 724–726.
209. Petrov, V. Dynamic multi-connectivity performance in ultra-dense urban mmWave deployments / V. Petrov, D. Solomitchii, A. Samuylov, et al. // IEEE Journal on Selected Areas in Communications. — 2017. — Vol. 35, no. 9. — Pp. 2038–2055.

210. Наумов, В. А. Анализ сетей ресурсных систем массового обслуживания / В. А. Наумов, К. Е. Самуйлов // Автоматика и телемеханика. — 2018. — Т. 79, № 5. — С. 59–68.
211. Gapeyenko, M. On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios / M. Gapeyenko, A. Samuylov, M. Gerasimenko, et al. // IEEE Transactions on Vehicular Technology. — 2017. — Vol. 66, no. 11. — Pp. 10124–10138.
212. Gapeyenko, M. Analysis of human-body blockage in urban millimeter-wave cellular communications / M. Gapeyenko, A. Samuylov, M. Gerasimenko, et al. // IEEE International Conference on Communications (ICC) / IEEE. — Kuala Lumpur, Malaysia, 2016. — Pp. 1–7.
213. Nain, P. Properties of random direction models / P. Nain, D. Towsley, B. Liu, et al. // Proceedings of the IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM). — Miami, FL, USA, 2005. — Vol. 3 — Pp. 1897–1907.
214. Groenevelt, R. Stochastic Models for Mobile Ad Hoc Networks : Ph. D. dissertation — Université de Nice — Sophia Antipolis, 2005 — 204 p.
215. Daley D. J. The busy period of the $M/GI/\infty$ queue / D. J. Daley // Queueing Systems. — 2001. — June. — Vol. 38, no. 2. — Pp. 195–204.
216. Niu, Y. A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges / Y. Niu, Y. Li, D. Jin, et al. // Wireless Networks. — 2015. — Vol. 21, no. 8. — Pp. 2657–2676.
217. Rappaport, T. S. Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design / T. S. Rappaport, G. R. MacCartney, M. K. Samimi, S. Sun // IEEE Transactions on Communications. — 2015. — Vol. 63, no. 9. — Pp. 3029–3056.
218. Solomitckii, D. Characterizing the impact of diffuse scattering in urban millimeter-wave deployments / D. Solomitckii, Q. C. Li, T. Balercia, et al. // IEEE Wireless Communications Letters. — 2016. — Vol. 5, no. 4. — Pp. 432–435.
219. Molisch, A. F. Wireless Communications / A. F. Molisch — 2 edition. — John Wiley & Sons, 2011. — 884 p.

220. Remcom Inc. — Wireless InSite Reference Manual, 2014. — February.
221. Durgin, G. An advanced 3D ray launching method for wireless propagation prediction / G. Durgin, N. Patwari, T. S. Rappaport // Proceedings of 47th Vehicular Technology Conference / IEEE. — Vol. 2. — Phoenix, AZ, USA, 1997. — Pp. 785–789.
222. Bessis, N. Development of Distributed Systems from Design to Application and Maintenance / N. Bessis. — IGI Global, 2012. — 367 p.
223. WINNER II. Test scenarios and calibration cases issue 2 [Электронный ресурс] / M. Abaii, G. Auer, Y. Cho, et al. // IST-4-027756 WINNER II. — Tech. Rep. D6.13.7 v1.00. — 2006, December. — Режим доступа: <https://pdfs.semanticscholar.org/6537/5d4c6412952b951c7f5fe2fd59d5a3832fac.pdf> (дата обращения: 05.11.2018).
224. Semkin, V. Characterization of radio links at 60 GHz using simple geometrical and highly accurate 3-D models / V. Semkin, D. Solomitckii, R. Naderpour, et al. // IEEE Transactions on Vehicular Technology. — 2017. — Vol. 66, no. 6. — Pp. 4647–4656.
225. Solomitckii, D. Toward Massive Ray-Based Simulations of mmWave Small Cells on Open Urban Maps / D. Solomitckii, M. Gapeyenko, S. S. Szyszkowicz, et al. // IEEE Antennas and Wireless Propagation Letters. — 2017. — Vol. 16. — Pp. 1435–1438.
226. New York State Department of Transportation. Annual Average Daily Traffic in New York City [Электронный ресурс] // 2015. — Режим доступа: <https://www.dot.ny.gov/tdv> (дата обращения: 05.11.2018).
227. 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects [Электронный ресурс] // 3GPP TR 36.814 version 9.0.0. — 2010, March. — Режим доступа: <http://www.qtc.jp/3GPP/Specs/36814-900.pdf> (дата обращения: 05.11.2018).
228. 3GPP. Study on New Radio Access Technology Physical Layer Aspects [Электронный ресурс] // 3GPP TR 38.802 version 0.3.1. — 2016. —

- Режим доступа: www.3gpp.org/ftp/Specs/archive/38_series/38.802/38802-e20.zip (дата обращения: 05.11.2018).
229. Moltchanov, D. Improving Session Continuity with Bandwidth Reservation in mmWave Communications / D. Moltchanov, A. Samuylov, V. Petrov, et al. // IEEE Wireless Communications Letters. — 2018. — no. 7. — P. 1–4.
230. Кучерявый, А. Е. Интернет Вещей / А. Е. Кучерявый // Электросвязь. — 2013. — № 1. — С. 21–24.
231. Van, D. P. Machine-to-machine communications over FiWi enhanced LTE networks: A power-saving framework and end-to-end performance / D. P. Van, B. P. Rimal, S. Andreev, et al. // Journal of Lightwave Technology. — 2016. — Vol. 34, no. 4. — Pp. 1062–1071.
232. Парамонов, А. И. Модели потоков трафика для сетей m2m / А. И. Парамонов // Электросвязь. — 2014. — № 4. — С. 11–16.
233. Галинина, О. С. Учет специфики доступа большого числа устройств при межмашинном взаимодействии в современных сотовых сетях / О. С. Галинина, С. Д. Андреев, А. М. Тюрликов // Информационно-управляющие системы. — 2018. — № 4. — С. 105–114.
234. Ометов, А. Я. Анализ производительности беспроводной системы агрегации данных с состязанием для современных сенсорных сетей / А. Я. Ометов, С. Д. Андреев, А. М. Тюрликов и др. // Информатика и ее применения. — 2016. — Т. 10, № 3. — С. 23–31.
235. Petrov, V. When IoT keeps people in the loop: A path towards a new global utility / V. Petrov, K. Mikhaylov, D. Moltchanov, et al. // IEEE Communications Magazine. — 2018. — no. 11. — Pp. 2–9.
236. ETSI. Low Throughput Networks (LTN); Use Cases for Low Throughput Networks [Электронный ресурс] // ETSI GS LTN 001. — 2014. — Режим доступа: https://www.etsi.org/deliver/etsi_gs/ltn/001_099/001/01.01.01_60/gs_ltn001v010101p.pdf (дата обращения: 05.11.2018).
237. Лившиц, Б. С. Теория телетрафика : Учебник для вузов / Б. С. Лившиц, А. П. Пшеничников, А. Д. Харкевич — М. : Связь, 1979. — 224 с.

238. Башарин, Г. П. Математическая теория телетрафика и ее приложения к анализу мультисервисных сетей связи следующих поколений / Г. П. Башарин, Ю. В. Гайдамака, К. Е. Самуилов // Автоматика и вычислительная техника. — 2013. — Т. 2. — С. 11–21.
239. Zheng, K. Challenges of massive access in highly dense LTE-advanced networks with machine-to-machine communications / K. Zheng, S. Ou, J. Alonso-Zarate, et al. // IEEE Wireless Communications. — 2014. — Vol. 21, no. 3. — Pp. 12–18.
240. Orsino, A. Improving Initial Access Reliability of 5G mmWave Cellular in Massive V2X Communications Scenarios / A. Orsino, O. Galinina, S. Andreev, et al. // 2018 IEEE International Conference on Communications (ICC) / IEEE. — Kansas City, MO, USA, 2018. — Pp. 1–7.
241. Gerasimenko, M. Energy and delay performance of machine-type communications (MTC) in long-term evolution-advanced (LTE-A) / M. Gerasimenko, O. Galinina, S. Andreev, et al. // C. Antón-Haro, M. Dohler (Eds.): Machine-to-machine (M2M) Communications. — Woodhead Publishing. — 2015. — Pp. 315–335.
242. Gerasimenko, M. Impact of machine-type communications on energy and delay performance of random access channel in LTE-advanced / M. Gerasimenko, V. Petrov, O. Galinina, et al. // Transactions on Emerging Telecommunications Technologies. — 2013. — Vol. 24, no. 4. — Pp. 366–377.
243. Винель, А. В. Анализ алгоритмов распространения тревожного сообщения с глобальным знанием в беспроводных сетях передачи данных с линейной топологией / А. В. Винель, А. Н. Дудин, С. Д. Андреев и др. // Информационно-управляющие системы. — 2010. — № 3. — С. 56–60.
244. Анисимов, А. В. Управление энергоэффективной доставкой информации на абонентскую станцию / А. В. Анисимов, С. Д. Андреев, А. М. Тюрликов // Электросвязь. — 2011. — № 4. — С. 56–61.
245. Andersson, K. Optimized access network selection in a combined WLAN/LTE environment / K. Andersson, C. Åhlund // Wireless Personal Communications. — 2011. — Vol. 61, no. 4. — Pp. 739–751.

246. ETSI. Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification [Электронный ресурс] // 3GPP TS 36.321. — 2012. — Режим доступа: https://www.etsi.org/deliver/etsi_ts/136300_136399/136321/10.05.00_60/ts_136321v100500p.pdf (дата обращения: 05.11.2018).
247. Johnson, C. Long Term Evolution IN BULLETS / C. Johnson. — 2nd edition. — CreateSpace Independent Publishing Platform, 2012. — 592 p.
248. Cheng, R. RACH collision probability for machine-type communications / R. Cheng, C. Wei, S. Tsao, et al. // IEEE 75th Vehicular Technology Conference (VTC Spring) / IEEE. — Yokohama, Japan, 2012. — Pp. 1–5.
249. Maeder, A. The challenge of M2M communications for the cellular radio access network / A. Maeder, P. Rost, D. Staehle // Proceedings of Würzburg Workshop IP, Joint ITG Euro-NF Workshop “Visions of Future Generation Networks” EuroView. — Würzburg, Germany, 2011. — Pp. 1–2.
250. 3GPP. Study on RAN Improvements for Machine-Type Communications [Электронный ресурс] // 3GPP TR 37.868. — 2011. — Режим доступа: http://www.3gpp.org/ftp/Specs/archive/37_series/37.868/37868-b00.zip (дата обращения: 05.11.2018).
251. Введенская, Н. Д. Случайный множественный доступ пакетов в канал с ошибками / Н. Д. Введенская, Б. С. Цыбаков // Проблемы передачи информации. — 1983. — Т. 19, № 2. — С. 52–68.
252. Kleinrock, L. Packet switching in a multiaccess broadcast channel: Performance evaluation / L. Kleinrock, S. Lam // IEEE Transactions on Communications. — 1975. — Vol. 23, no. 4. — Pp. 410–423.
253. Sidi, M. Two interfering queues in packet-radio networks / M. Sidi, A. Segall // IEEE Transactions on Communications. — 1983. — Vol. 31, no. 1. — Pp. 123–129.
254. Лившиц, Б. С. Системы массового обслуживания с конечным числом источников / Б. С. Лившиц, Я. В. Фидлин. — М. : Связь, 1968. — 168 с.

255. 3GPP. R2-113013. Access barring for delay tolerant access in LTE [Электронный ресурс] // 3GPP TSG-RAN WG2. — 2011. — Режим доступа: http://www.3gpp.org/ftp/tsg_ran/wg2_r12/tsgr2_74/Docs/R2-113013.zip (дата обращения: 05.11.2018).
256. 3GPP. Study on provision of low-cost MTC UEs based on LTE [Электронный ресурс] // 3GPP TR 36.888. — 2013. — Режим доступа: http://www.3gpp.org/ftp/Specs/archive/36_series/36.888/36888-c00.zip (дата обращения: 05.11.2018).
257. Dementev, O. Analyzing the overload of 3GPP LTE system by diverse classes of connected-mode MTC devices / O. Dementev, O. Galinina, M. Gerasimenko, et al. // IEEE World Forum on Internet of Things (WF-IoT) / IEEE. — Seoul, South Korea, 2014. — Pp. 309–3012.
258. Galinina, O. Multi-channel random access with replications / O. Galinina, A. Turlikov, S. Andreev, et al. // 2017 IEEE International Symposium on Information Theory (ISIT). — Aachen, Germany, 2017. — Pp. 2538–2542.
259. Galinina, O. Stabilizing multi-channel slotted aloha for machine-type communications / O. Galinina, A. Turlikov, S. Andreev, et al. // 2013 IEEE International Symposium on Information Theory Proceedings (ISIT) / IEEE. — Istanbul, Turkey, 2013. — Pp. 2119–2123.
260. Ribeiro, A. Opportunistic multipath for bandwidth-efficient cooperative multiple access / A. Ribeiro, X. Cai, G. B. Giannakis // IEEE Transactions on Wireless Communications. — 2006. — Vol. 5, no. 9. — Pp. 2321–2327.
261. Фосс, С. Г. Протоколы с двойной рандомизацией для системы множественного доступа с обратной связью «успех–неуспех» / С. Г. Фосс, Б. Хайек, А. М. Тюрликов // Проблемы передачи информации. — 2016. — Т. 52, № 2. — С. 61–71.
262. Вишнеvский, В. М. Математические методы исследования систем поллинга / В. М. Вишнеvский, О. В. Семенова // Автоматика и телемеханика. — 2006. — Т. 67, № 2. — С. 3–56.
263. Tirronen, T. Machine-to-machine communication with long-term evolution with reduced device energy consumption / T. Tirronen, A. Larmo, J. Sachs,

- et al. // Transactions on Emerging Telecommunications Technologies. — 2013. — Vol. 24, no. 4. — Pp. 413–426.
264. Andreev, S. Efficient small data access for machine-type communications in LTE / S. Andreev, A. Larmo, M. Gerasimenko, et al. // Proceedings of International Conference on Communications (ICC) / IEEE. — Budapest, Hungary, 2013. — Pp. 3569–3574.
265. Bhas, N. Smart wearable devices: Fitness healthcare entertainment and enterprise 2013–2018 / N. Bhas // Juniper Research, Basingstoke, UK, Technical Report — 2013.
266. Billinghamurst, M. Wearable devices: new ways to manage information / M. Billinghamurst, T. Starner // Computer. — 1999. — Vol. 32, no. 1. — Pp. 57–64.
267. ABI Research. Wearable Computing Technologies [Электронный ресурс] // 2013. — Режим доступа: <https://www.abiresearch.com/market-research/product/1015644-wearable-computing-technologies/> (дата обращения: 05.11.2018).
268. Bally, G. ShoeSense: a new perspective on hand gestures and wearable applications / G. Bally, J. Müller, M. Rohs, et al. // Proceedings of Conference on Human Factors in Computing Systems (CHI). — Austin, TX, USA, 2012. — Vol. 12. — Pp. 1–10.
269. Foerster, J. Realizing Gbps wireless personal area networks-guest editorial / J. Foerster, J. Lansford, J. Laskar, et al. // IEEE Journal on Selected Areas in Communications. — 2009. — Vol. 27, no. 8. — Pp. 1313–1317.
270. Fong, B. Factors causing uncertainties in outdoor wireless wearable communications / B. Fong, P. B. Rapajic, G. Y. Hong, et al. // IEEE Pervasive Computing. — 2003. — Vol. 2, no. 2. — Pp. 16–19.
271. Feiner, S. K. The importance of being mobile: some social consequences of wearable augmented reality systems / S. K. Feiner // Proceedings of 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99) / IEEE. — San Francisco, CA, USA, 1999. — Pp. 145–148.

272. Fletcher, R. R. iCalm: Wearable sensor and network architecture for wirelessly communicating and logging autonomic activity / R. R. Fletcher, K. Dobson, M. S. Goodwin, et al. // IEEE Transactions on Information Technology in Biomedicine. — 2010. — Vol. 14, no. 2. — Pp. 215–223.
273. Klingeberg., T. Mobile wearable device for long term monitoring of vital signs / T. Klingeberg, M. Schilling // Computer Methods and Programs in Biomedicine. — 2012. — Vol. 106, no. 2. — Pp. 89–96.
274. Pyattaev, A. Communication challenges in high-density deployments of wearable wireless devices / A. Pyattaev, K. Johnsson, S. Andreev, et al. // IEEE Wireless Communications. — 2015. — Vol. 22, no. 1. — Pp. 12–18.
275. Galinina, O. Analyzing Effects of Directional Deafness on mmWave Channel Access in Unlicensed Bands / O. Galinina, A. Pyattaev, K. Johnsson, et al. // IEEE Globecom Workshops (GC Wkshps) / IEEE. — Singapore, Singapore, 2017. — Pp. 1–7.
276. Galinina, O. Assessing system-level energy efficiency of mmwave-based wearable networks / O. Galinina, A. Pyattaev, K. Johnsson, et al. // IEEE Journal on Selected Areas in Communications. — 2016. — Vol. 34, no. 4. — Pp. 923–937.
277. Demestichas, P. 5G on the horizon: Key challenges for the radio-access network / P. Demestichas, A. Georgakopoulos, D. Karvounas, et al. // IEEE Vehicular Technology Magazine. — 2013. — Vol. 8, no. 3. — Pp. 47–53.
278. I, C. Toward green and soft: a 5G perspective / C. I, C. Rowell, S. Han, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 2. — Pp. 66–73.
279. Wu, Y. Green transmission technologies for balancing the energy efficiency and spectrum efficiency trade-off / Y. Wu, Y. Chen, J. Tang, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 11. — Pp. 112–120.
280. Li, Q. Edge cloud and underlay networks: Empowering 5G cell-less wireless architecture / Q. Li, H. Niu, A. Papathanassiou, et al. // Proceedings of 20th European Wireless Conference European Wireless / VDE. — Barcelona, Spain, 2014. — Pp. 676–681.

281. Baker, M. Wearable Computing from Jewels to Joules / M. Baker, J. Hong, M. Billinghurst // IEEE Pervasive Computing. — 2014. — Vol. 13, no. 4. — Pp. 20–22.
282. Han, Q. Mobile cloud sensing, big data, and 5G networks make an intelligent and smart world / Q. Han, S. Liang, H. Zhang // IEEE Network. — 2015. — Vol. 29, no. 2. — Pp. 40–45.
283. Agyapong, P. K. Design considerations for a 5G network architecture / P. K. Agyapong, M. Iwamura, D. Staehle, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 11. — Pp. 65–75.
284. Chen, M. AIWAC: Affective interaction through wearable computing and cloud technology / M. Chen, Y. Zhang, Y. Li, et al. // IEEE Wireless Communications. — 2015. — Vol. 22, no. 1. — Pp. 20–27.
285. Cordeiro C. IEEE 802.11 ad: Introduction and performance evaluation of the first multi-Gbps WiFi technology / C. Cordeiro, D. Akhmetov, M. Park // Proceedings of the ACM international workshop on mmWave communications: from circuits to networks / ACM. — Chicago, Illinois, USA, 2010. — Pp. 3–8.
286. Perahia, E. Gigabit wireless LANs: an overview of IEEE 802.11ac and 802.11ad / E. Perahia, M. X. Gong // ACM SIGMOBILE Mobile Computing and Communications Review. — 2011. — Vol. 15, no. 3. — Pp. 23–33.
287. Kovalchukov, R. Modeling Three-Dimensional Interference and SIR in Highly Directional mmWave Communications / R. Kovalchukov, A. Samuylov, D. Moltchanov, et al. // GLOBECOM 2017 IEEE Global Communications Conference / IEEE. — Singapore, Singapore, 2017. — Pp. 1–7.
288. IEEE. Agilent Technologies application note, Wireless LAN at 60 GHz — IEEE 802.11ad Explained [Электронный ресурс] // 2013, May. — Режим доступа: <https://www.cs.odu.edu/~cs752/papers/milli-008.pdf> (дата обращения: 05.11.2018).
289. Analog Devices Welcomes Hittite Microwave Corporation, v00.1112 [Электронный ресурс] // HMC6000LP711E, HMC6001LP711E 2015. — Режим доступа: www.findrf.com/common/download?id=7380b12f40a541ff8bce68ce68675dd1 (дата обращения: 05.11.2018).

290. Zhang, M. A survey of caching mechanisms in information-centric networking / M. Zhang, H. Luo, H. Zhang // IEEE Communications Surveys & Tutorials. — 2015. — Vol. 17, no. 3. — Pp. 1473–1499.
291. Pyattaev, A. 3GPP LTE-Assisted Wi-Fi-Direct: Trial Implementation of Live D2D Technology / A. Pyattaev, J. Hosen, K. Johnsson, et al. // ETRI Journal. — 2015. — Vol. 37, no. 5. — Pp. 877–887.
292. Andreev, S. A unifying perspective on proximity-based cellular-assisted mobile social networking / S. Andreev, J. Hosen, T. Olsson, et al. // IEEE Communications Magazine. — 2016. — Vol. 54, no. 4. — Pp. 108–116.
293. Asadi, A. A survey on device-to-device communication in cellular networks / A. Asadi, Q. Wang, V. Mancuso // IEEE Communications Surveys & Tutorials. — 2014. — Vol. 16, no. 4. — Pp. 1801–1819.
294. Рыжков, А. Е. Системы и сети радиодоступа 4G: LTE, WiMAX / А. Е. Рыжков, М. А. Сиверс, В. О. Воробьев и др. — СПб. : Линк, 2012 — 226 с.
295. Prasad, A. Energy-efficient D2D discovery for proximity services in 3GPP LTE-advanced networks: ProSe discovery mechanisms / A. Prasad, A. Kunz, G. Velev, et al. // IEEE Vehicular Technology Magazine. — 2014. — Vol. 9, no. 4. — Pp. 40–50.
296. Ометов, А. Я. Обеспечение информационной безопасности прямых соединений 5G при изменении скорости движения абонентов и наличии сотового содействия / А. Я. Ометов, С. Д. Андреев, А. Б. Левина и др. // Научно-технический вестник информационных технологий, механики и оптики. — 2017. — Т. 17, № 1. — С. 100–109.
297. Ometov, A. A novel security-centric framework for D2D connectivity based on spatial and social proximity / A. Ometov, A. Orsino, L. Militano, et al. // Computer Networks. — 2016. — Vol. 107. — Pp. 327–338.
298. Гайдамака, Ю. В. Анализ характеристик интерференции в модели взаимодействия устройств с учетом среды распространения сигнала / Ю. В. Гайдамака, С. Д. Андреев, Э. С. Сопин и др. // Информатика и ее применения. — 2016. — Т. 10, № 4. — С. 2–10.

299. Алиев, Т. И. Моделирование ядра мультисервисной сети с относительной приоритизацией неоднородного трафика / Т. И. Алиев, И. Е. Никульский, В. О. Пяттаев // Научно-технический вестник СПбГУ ИТМО. — 2009. — Т. 62, № 4. — С. 88–96.
300. Ometov, A. Implementing secure network-assisted D2D framework in live 3GPP LTE deployment / A. Ometov, P. Masek, J. Urama, et al. // IEEE International Conference on Communications Workshops (ICC) / IEEE. — Kuala Lumpur, Malaysia, 2016. — Pp. 749–754.
301. Liang, X. Security and privacy in mobile social networks: challenges and solutions / X. Liang, K. Zhang, X. Shen, X. Lin // IEEE Wireless Communications. — 2014. — Vol. 21, no. 1. — Pp. 33–41.
302. Ahlswede, R. Network information flow / R. Ahlswede, N. Cai, S.-Y. R. Li, et al. // IEEE Transactions on Information Theory. — 2000. — Vol. 46, no. 4. — Pp. 1204–1216.
303. Yeung, R. W. Network coding theory: An introduction / R. W. Yeung // Frontiers of Electrical and Electronic Engineering in China. — 2010. — Vol. 5, no. 3. — Pp. 363–390.
304. Seferoglu, H. Video-aware opportunistic network coding over wireless networks / H. Seferoglu, A. Markopoulou // IEEE Journal on Selected Areas in Communications. — 2009. — Vol. 27, no. 5. — Pp. 713–728.
305. Altman, E. Forward correction and fountain codes in delay-tolerant networks / E. Altman, F. De Pellegrini // IEEE INFOCOM. — Rio de Janeiro, Brazil, 2009. — Pp. 1899–1907.
306. Lima, L. Topology matters in network coding / L. Lima, D. Ferreira, J. Barros // Telecommunication Systems. — 2012. — Vol. 51, no. 4. — Pp. 247–257.
307. Network Coding / Khaldoun Al Agha (Ed.). — Wiley, 2012. — 304 p.
308. Pahlevani, P. Novel concepts for device-to-device communication using network coding / P. Pahlevani, M. Hundebøll, M. Pedersen, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 4. — Pp. 32–39.

309. Militano, L. Network coding and evolutionary theory for performance enhancement in wireless cooperative clusters / L. Militano, F. Fitzek, A. Iera, et al. // European Transactions on Telecommunications. — 2010. — Vol. 21, no. 8. — Pp. 725–737.
310. Крук, Е. А. Расчет задержки при использовании кодирования на транспортном уровне сети передачи данных / Е. А. Крук, Д. А. Маличенко // Известия высших учебных заведений. Приборостроение. — 2013. — Т. 56, № 8. — С. 45–51.
311. Altman, E. Coding for caches in the plane [Электронный ресурс] / E. Altman, K. Avrachenkov, J. Goseling // arXiv preprint arXiv:1309.0604. — 2013. — Режим доступа: <https://arxiv.org/pdf/1309.0604.pdf> (дата обращения: 05.11.2018).
312. Golrezaei, N. Base-station assisted device-to-device communications for high-throughput wireless video networks / N. Golrezaei, P. Mansourifard, A. F. Molisch, et al. // IEEE Transactions on Wireless Communications. — 2014. — Vol. 13, no. 7. — Pp. 3665–3676.
313. Cisco. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018 [Электронный ресурс] // White paper. — 2014, February. — Режим доступа: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html> (дата обращения: 05.11.2018).
314. Pedersen, M. V. Mobile clouds: The new content distribution platform / M. V. Pedersen, F. Fitzek // Proceedings of the IEEE. — 2012. — Vol. 100, no. Special Centennial Issue. — Pp. 1400–1403.
315. Ji, M. Wireless device-to-device caching networks: Basic principles and system performance / M. Ji, G. Caire, A. F. Molisch // IEEE Journal on Selected Areas in Communications. — 2016. — Vol. 34, no. 1. — Pp. 176–189.
316. Golrezaei, N. Wireless device-to-device communications with distributed caching / N. Golrezaei, A. G. Dimakis, A. F. Molisch // Proceedings of International Symposium on Information Theory Proceedings (ISIT) / IEEE. — Cambridge, MA, USA, 2012. — Pp. 2781–2785.

317. Shanmugam, K. Femtocaching: Wireless content delivery through distributed caching helpers / K. Shanmugam, N. Golrezaei, A. G. Dimakis, et al. // IEEE Transactions on Information Theory. — 2013. — Vol. 59, no. 12. — Pp. 8402–8413.
318. Pyattaev, A. Understanding practical limitations of network coding for assisted proximate communication / A. Pyattaev, O. Galinina, S. Andreev, et al. // IEEE Journal on Selected Areas in Communications. — 2015. — Vol. 33, no. 2. — Pp. 156–170.
319. Etezov, S. On distribution of SIR in dense D2D deployments / S. Etezov, Y. Gaidamaka, K. Samouylov, et al. // Proceedings of 22th European Wireless Conference / VDE. — Oulu, Finland, 2016. — Pp. 1–5.
320. Qiu, D. Modeling and performance analysis of BitTorrent-like peer-to-peer networks / D. Qiu, R. Srikant // ACM SIGCOMM computer communication review / ACM. — Vol. 34. — 2004. — Pp. 367–378.
321. Chou, P. A. Network coding for the internet and wireless networks / P. A. Chou, Y. Wu // IEEE Signal Processing Magazine. — 2007. — Vol. 24, no. 5. — Pp. 77–85.
322. Paramanathan, A. Lean and mean: network coding for commercial devices / A. Paramanathan, M. V. Pedersen, D. E. Lucani, et al. // IEEE Wireless Communications. — 2013. — Vol. 20, no. 5. — Pp. 54–61.
323. Heide, J. Green mobile clouds: Network coding and user cooperation for improved energy efficiency / J. Heide, F. Fitzek, M. V. Pedersen, et al. // 1st International Conference on Cloud Networking (CLOUDNET) / IEEE. — Paris, France, 2012. — Pp. 111–118.
324. Vingelmann, P. Synchronized multimedia streaming on the iphone platform with network coding / P. Vingelmann, F. Fitzek, M. V. Pedersen, et al. // IEEE Communications Magazine. — 2011. — Vol. 49, no. 6. Pp. 126–132.
325. Fitzek, F. Implementation of network coding for social mobile clouds / F. Fitzek, J. Heide, M. V. Pedersen, et al. // IEEE Signal Processing Magazine. — 2013. — Vol. 30, no. 1. — Pp. 159–164.

326. Кучерявый, А. Е. Сети связи общего пользования. Тенденции развития и методы расчета / А. Е. Кучерявый, А. И. Парамонов, Е. А. Кучерявый. — М. : ФГУП ЦНИИС, 2008. — 296 с.
327. Iversen, V. B. Teletraffic engineering and network planning — DTU Fotonik, 2015. — 398 p.
328. Florea, R. Networking Solutions for Integrated Heterogeneous Wireless Ecosystem / R. Florea, A. Ometov, A. Surak, et al. // J. Sen (Eds.): Cloud Computing. — InTechOpen. — 2017. — Pp. 103–128.
329. Zhou, Y. Optimized backhaul compression for uplink cloud radio access network / Y. Zhou, W. Yu // IEEE Journal on Selected Areas in Communications. — 2014. — Vol. 32, no. 6. — Pp. 1295–1307.
330. Jeon, S. Wireless multihop device-to-device caching networks / S. Jeon, S. Hong, M. Ji, et al. // IEEE Transactions on Information Theory. — 2017. — Vol. 63, no. 3. — Pp. 1662–1676.
331. Mäkitalo, N., Safe, secure executions at the network edge: Coordinating cloud, edge, and fog computing / N. Mäkitalo, A. Ometov, J. Kannisto, et al. // IEEE Software. — 2018. — Vol. 35, no. 1. — Pp. 30–37.
332. Baştuğ, E. Living on the edge: The role of proactive caching in 5G wireless networks / E. Baştuğ, M. Bennis, M. Debbah // IEEE Communications Magazine. — 2014. — Vol. 52, no. 8. — Pp. 82–89.
333. Mirahsan, M. HetHetNets: Heterogeneous traffic distribution in heterogeneous wireless cellular networks / M. Mirahsan, R. Schoenen, H. Yanikomeroğlu // IEEE Journal on Selected Areas in Communications. — 2015. — Vol. 33, no. 10. — Pp. 2252–2265.
334. Barbarossa, S. Communicating while computing: Distributed mobile cloud computing over 5G heterogeneous networks / S. Barbarossa, S. Sardellitti, P. Di Lorenzo // IEEE Signal Processing Magazine. — 2014. — Vol. 31, no. 6. — Pp. 45–55.
335. Dinh, H. T. A survey of mobile cloud computing: architecture, applications, and approaches / H. T. Dinh, C. Lee, D. Niyato, et al. // Wireless Communications and Mobile Computing. — 2013. — Vol. 13, no. 18. — Pp. 1587–1611.

336. Su, Z. Content distribution over content centric mobile social networks in 5G / Z. Su, Q. Xu // IEEE Communications Magazine. — 2015. — Vol. 53, no. 6. — Pp. 66–72.
337. Hachem, J. Content caching and delivery over heterogeneous wireless networks / J. Hachem, N. Karamchandani, S. Diggavi // Proceedings of Conference on Computer Communications (INFOCOM) / IEEE. — Kowloon, Hong Kong 2015. — Pp. 756–764.
338. Wang, X. Cache in the air: exploiting content caching and delivery techniques for 5G systems / X. Wang, M. Chen, T. Taleb, et al. // IEEE Communications Magazine. — 2014. — Vol. 52, no. 2. — Pp. 131–139.
339. Andreev, S. Exploring synergy between communications, caching, and computing in 5G-grade deployments / S. Andreev, O. Galinina, A. Pyattaev, et al. // IEEE Communications Magazine. — 2016. — Vol. 54, no. 8. — Pp. 60–69.
340. Balan, R. The case for cyber foraging / R. Balan, J. Flinn, M. Satyanarayanan, et al. // Proceedings of the 10th workshop on ACM SIGOPS European workshop / ACM. — Saint-Émilion, France, 2002. — Pp. 87–92.
341. Zhou, B. Stochastic content-centric multicast scheduling for cache-enabled heterogeneous cellular networks / B. Zhou, Y. Cui, M. Tao // IEEE Transactions on Wireless Communications. — 2016. — Vol. 15, no. 9. — Pp. 6284–6297.
342. Ometov, A. Characterizing the effect of packet losses in current WLAN Deployments / A. Ometov, S. Andreev, A. Turlikov, et al. // Proceedings of 13th International Conference on ITS Telecommunications (ITST) / IEEE. — Tampere, Finland, 2013. — Pp. 331–336.
343. 3GPP. Terminal conformance specification; Radio transmission and reception (TDD) [Электронный ресурс] // 3GPP TS 34.122. — 2014, March. — Режим доступа: http://www.3gpp.org/ftp/Specs/archive/34_series/34.122/34122-bd0.zip (дата обращения: 05.11.2018).
344. Цитович, И. И. Расчет стационарных вероятностей трехпоточковой модели управления доступом к ресурсам БШС с гистерезисами / И. И. Цитович, А. В. Чернушевич // Информационные процессы. — 2011. — Т. 11, № 2. — С. 262–276.

345. Foss, S. G. On a Voronoi aggregative process related to a bivariate Poisson process / S. G. Foss, S. A. Zuyev // *Advances in Applied Probability*. — 1996. — Vol. 28, no. 4. — Pp. 965–981.
346. Ye, Q. User association for load balancing in heterogeneous cellular networks / Q. Ye, B. Rong, Y. Chen, et al. // *IEEE Transactions on Wireless Communications*. — 2013. — Vol. 12, no. 6. — Pp. 2706–2716.
347. Singh, S. On association cells in random heterogeneous networks / S. Singh, F. Baccelli, J. G. Andrews // *IEEE Wireless Communications Letters*. — 2014. — Vol. 3, no. 1. — Pp. 70–73.
348. Kalashnikov, V. V. *Mathematical Methods in Queuing Theory* / V. V. Kalashnikov. — Springer Science & Business Media, 1993. — 379 p.
349. de Lima, C. H. M. Statistical analysis of self-organizing networks with biased cell association and interference avoidance / C. H. M de Lima, M. Bennis, L. Matti. // *IEEE Transactions on Vehicular Technology*. — 2013. — Vol. 62, no. 5. — Pp. 1950–1961.
350. Weiss, I. Limiting distributions in some occupancy problems / I. Weiss // *The Annals of Mathematical Statistics*. — 1958. — Vol. 29, no. 3. — Pp. 878–884.
351. Gadrich, T. The sequential occupancy problem through group throwing of indistinguishable balls / T. Gadrich, R. Ravid // *Methodology and Computing in Applied Probability*. — 2011. — Vol. 13, no. 2. — Pp. 433–448.
352. Sevastyanov, B. A. Limit Theorems in a Scheme for Allocation of Particles in Cells / B. A. Sevastyanov // *Theory of Probability & its Applications*. — 1966. — Vol. 11, no. 4. — Pp. 614–619.
353. Mikhailov, V. G. Estimate for the accuracy of the Poisson approximation for the number of empty cells in an equiprobable scheme for group allocation of particles, and applications / V. G. Mikhailov // *Proceedings of the Steklov Institute of Mathematics*. — 2013. — Vol. 282, no. 1. — Pp. 157–171.
354. Feller, W. *Introduction to Probability Theory and its Applications* / W. Feller. — John Wiley & Sons, Inc., New York, 1968. — Vol. I, 3rd edition — 528 p.

355. Ширяев, А. Н. Вероятность / А. Н. Ширяев. — М. : МЦНМО, 2011 — В 2-х книгах. 5-е издание. — 552 с., 416 с.
356. Galinina, O. A capacity bound for mmWave-based channel access in ultra-dense wearable deployments / O. Galinina, A. Turlikov, A. Pyattaev, et al. // Proceedings of 7th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT) / IEEE. — Brno, Czech Republic, 2015. — Pp. 298–304.
357. Venugopal, K. Interference in finite-sized highly dense millimeter wave networks / K. Venugopal, M. C. Valenti, R. W. Heath // Information Theory and Applications Workshop (ITA), 2015 / IEEE. — San Diego, CA, USA, 2015. — Pp. 175–180.
358. Haenggi, M. Mean interference in hard-core wireless networks / M. Haenggi // IEEE Communications Letters. — 2011. — Vol. 15, no. 8. — Pp. 792–794.
359. ElSawy, H. A modified hard core point process for analysis of random CSMA wireless networks in general fading environments / H. ElSawy, E. Hossain // IEEE Transactions on Communications. — 2013. — Vol. 61, no. 4. — Pp. 1520–1534.
360. Serrano, P. Per-frame energy consumption in 802.11 devices and its implication on modeling and design / P. Serrano, A. Garcia-Saavedra, G. Bianchi, et al. // IEEE/ACM Transactions on Networking (ToN). — 2015. — Vol. 23, no. 4. — Pp. 1243–1256.

Список рисунков

1.1	Особенности перспективных гетерогенных сетей 5G	32
1.2	Основные способы увеличения емкости систем 5G	41
1.3	Цикл разработки и развертывания систем ИМТ	43
1.4	Перспективные сценарии применения систем ИМТ к 2020 году и далее	46
1.5	Примеры 5G-приложений Индустриального Интернета для автоматизации производства	47
1.6	Ключевые особенности сценариев применения технологии 5G	55
2.1	Топология типовой гетерогенной сети	63
2.2	Архитектура гетерогенной сети связи	64
2.3	Модель гетерогенной сети (слои представлены в виде сечений)	71
2.4	Каскадный принцип работы системы со слоями WLAN, пико- и макро-	77
2.5	Среднее число сессий в макрослое с учетом затухания	81
2.6	Вероятность блокировки в макрослое с учетом затухания	81
2.7	Зависимость среднего числа сессий от плотности БС в пикослое	82
2.8	Зависимость среднего числа сессий от плотности ТД в слое WLAN	82
2.9	Вероятность блокировки в интегрированной гетерогенной сети с тремя слоями	83
2.10	Структура совмещенной гетерогенной сети и область моделирования	86
2.11	Управление интерференцией в сверхплотных сетях LTE	92
2.12	Зависимость коэффициента переиспользования радиоресурсов от фактической плотности сети	96
2.13	Пример активных соединений WiFi	97
2.14	Учет переменной плотности размещения	98
2.15	Среднее число абонентских терминалов на кв. м	100
2.16	Среднее время передачи данных терминалом	100
2.17	Средняя скорость передачи данных терминалом	102
2.18	Алгоритмы управления трафиком в интегрированной гетерогенной сети	103
2.19	Скорость передачи для различных алгоритмов управления трафиком	111
2.20	Сравнительный анализ алгоритмов управления трафиком	112

2.21	Индивидуальная скорость передачи абонентским терминалом	113
2.22	Распределение индивидуальной скорости передачи данных	114
3.1	Пример работы системы выгрузки на основе соединений D2D	124
3.2	Архитектура системы выгрузки на основе соединений D2D	124
3.3	Модель выгрузки сотового трафика на соединения D2D	128
3.4	Уровни потребляемой передающим терминалом мощности	130
3.5	Организация работы рассматриваемой системы связи	133
3.6	Емкость системы и энергопотребление абонентских терминалов	140
3.7	Вероятность блокировки (отказа) для сессии	141
3.8	Распределение числа соединений LTE	141
3.9	Распределение числа соединений D2D	142
3.10	Структура предлагаемого подхода к моделированию системы mmWave-связи	147
3.11	Динамическое блокирования канала прямой видимости	149
3.12	Рассматриваемая модель городского сценария (вид сверху)	156
3.13	Доля покрытых абонентских терминалов на частоте 28 ГГц	157
3.14	Пример использования связи mmWave в реалистичном сценарии	159
3.15	Детализированная трехмерная модель области моделирования	160
3.16	Рассматриваемые локации БС (слева); дополнительная функция распределения для значения принятой мощности (в центре); пример выбора MCS для БС номер 5 в условиях nLoS (справа)	161
3.17	Зависимость вероятности отказа в обслуживании от интенсивности входного потока	167
3.18	Зависимость вероятности разрыва сессии от интенсивности входного потока	168
3.19	Зависимость вероятности отказа в обслуживании от плотности препятствий	169
3.20	Зависимость вероятности разрыва сессии от плотности препятствий	170
3.21	Зависимость среднего числа активных сессий от плотности препятствий	170
4.1	Потребительские и промышленные применения M2M-устройств	179
4.2	Основные характеристики перспективных M2M-радиотехнологий	181
4.3	Порядок процедуры случайного доступа в системе LTE	188

4.4	Пример работы процедуры случайного доступа во времени (а) и пример энергопотребления для терминала номер 8 (б)	189
4.5	Типовой сценарий M2M с большим числом устройств	195
4.6	Характеристики доступа M2M-устройств с различными приоритетами	196
4.7	Моделирование многоканального протокола случайного доступа . . .	197
4.8	Пример размещения M2M-системы для автоматизации производства	198
4.9	Задержка доступа в сотовой системе с большим числом устройств M2M	198
4.10	Характеристики начального входа в сеть для сотовой системы M2M	200
4.11	Уровни потребляемой устройством M2M мощности	200
4.12	Моделирование потребляемой устройством M2M мощности	201
4.13	Повышение энергетической эффективности M2M-устройств	201
4.14	Типовая структура кадра LTE для сотовой системы M2M	202
4.15	Характеристики передачи данных малого объема	203
4.16	Сценарий с высокой плотностью размещения носимого оборудования	206
4.17	Моделирование массового использования носимых устройств	214
4.18	Пример геометрического расположения системы PBSS	216
4.19	Случаи наличия и отсутствия mmWave-соединения	218
4.20	Структура интервала передачи маяка в протоколе WiGig	219
4.21	Предлагаемая абстракция работы протокола WiGig	220
4.22	Соответствие между компонентами модели	222
4.23	Зависимость энергетической эффективности от длительности суперкадра	225
4.24	Зависимость энергетической эффективности от числа слотов	226
4.25	Зависимость энергетической эффективности от угла раскрытия диаграммы направленности	227
4.26	Зависимость эффективности работы системы в целом от плотности размещения терминалов	227
5.1	Примеры использования прямых соединений D2D в сетях 5G	231
5.2	Принципы организации прямого взаимодействия	235
5.3	Предлагаемый протокол для работы процедуры ProSe	237
5.4	Моделирование работы интегрированной системы LTE-D2D	241
5.5	Емкость сотовой + D2D сети в условиях насыщения	244

5.6	Скорость передачи при выгрузке в 30% для насыщенной сотовой + D2D сети	245
5.7	Энергетическая эффективность соты для насыщенной сотовой + D2D сети	245
5.8	Прототип системы связи WiFi-Direct с сетевым содействием	248
5.9	Измеренное время установления прямого соединения	251
5.10	Система выгрузки данных на D2D с сетевым кодированием	255
5.11	Моделирование работы интегрированной двухслойной сети	257
5.12	Пример работы системы в течение жизненного цикла единицы контента C_4	259
5.13	Изучение режимов распределения контента: ВТ	267
5.14	Работа системы в процессе получения контента: ВТ	268
5.15	Изучение режимов распределения контента: NC	269
5.16	Работа системы в процессе получения контента: NC	269
5.17	Среднее время передачи для различных систем распределения контента	273
5.18	Вероятность успеха для различных систем распределения контента	273
5.19	Примеры использования перспективных сервисов 5G	280
5.20	Варианты размещения сетевой архитектуры и соответствующие уровни иерархии	282
A.1	Моделирование работы слоя WLAN	336
A.2	Моделирование работы пикослоя	338
A.3	Моделирование работы макрослоя	340
A.4	Пространство состояний системы для макрослоя	342
A.5	Пространство состояний системы для слоя WLAN/пико-	343
B.1	Упрощение структуры марковской цепи для совмещенной гетерогенной сети доступа	363
B.2	Диаграмма Вороного и пример работы сети WiFi: область моделирования (слева) и меченая малая сота, содержащая зону покрытия СБС (справа)	366
B.3	Диаграмма Вороного и пример работы сети LTE: область моделирования (слева) и меченая малая сота, содержащая зону покрытия СБС с виртуальными пользователями (справа)	368

Б.4	Пример состояния двумерного марковского процесса для системы LTE	370
Б.5	Пример назначения радиоресурсов в совмещенной сети: зоны обслуживания	376
Б.6	Пример назначения радиоресурсов в совмещенной сети: плотность покрытия	377
В.1	Пространство состояний системы D2D	382
Г.1	Пример работы систем Тип-1.а (случайное размещение всех фрагментов данных), Тип-1.б (случайное размещения начала группы фрагментов данных) и Тип-1.в (группа фрагментов данных следует непосредственно за успешной отправкой маяка)	397
Г.2	Пропускная способность для систем Тип-1.а, Тип-1.б и Тип-1.в	398
Г.3	Иллюстрация геометрических свойств сети	403
Г.4	Пояснение для вероятности наличия соединения	403
Г.5	Вычисление доли $\int_r \psi(r) dF_r(r)$	404
Г.6	Примеры зоны покрытия и функции $\rho(\alpha)$ для различных диаграмм направленности	408
Г.7	Взаимосвязь между возможными событиями, связанными с энергопотреблением устройства	410
Д.1	Сравнение аналитических (А) и имитационных (И) результатов для режимов ВТ (а) и НС (б)	420
Д.2	Структура предложенного решения для двух фаз процесса: в течение времени жизни контента (фаза роста) и когда контент покидает систему (фаза сокращения)	421

Список таблиц

1	Параметры моделирования сверхплотной совмещенной сети	99
2	Параметры моделирования алгоритмов управления трафиком	110
3	Параметры моделирования системы прямых соединений	139
4	Параметры моделирования связи на крайне высоких частотах	150
5	Сравнение радиотехнологий ближнего радиуса действия	208
6	Параметры моделирования mmWave-системы носимых устройств	224
7	Параметры моделирования работы прямых соединений	242
8	Параметры моделирования системы распределения контента	265
9	Сравнение производительности систем распределения контента	272

Список приложений

- А. Анализ интегрированной гетерогенной сети доступа.
- Б. Анализ совмещенной гетерогенной сети доступа.
- В. Анализ интегрированных сотовой системы и сети D2D.
- Г. Анализ энергетической эффективности mmWave-системы носимых устройств.
- Д. Анализ применения сетевого кодирования в системе D2D с сетевым содействием.
- Е. Документы, подтверждающие внедрение основных результатов диссертационной работы.

Приложение А

Анализ интегрированной гетерогенной сети доступа

А.1 Особенности слоев гетерогенной сети доступа

В данном приложении изучаются особенности и отличительные черты каждого из трех слоев, рассматриваемых в рамках интегрированной гетерогенной сети доступа: WLAN, пико- и макро-.

А.1.1 Модель слоя WLAN

Ниже рассматривается модель слоя WLAN, включающего в себя случайное число ТД WiFi N_w , $A_j, j = \overline{1, N_w}$, размещенных в соответствии с допущениями модели системы, т.е. равномерно на плоскости (см. рисунок А.1).

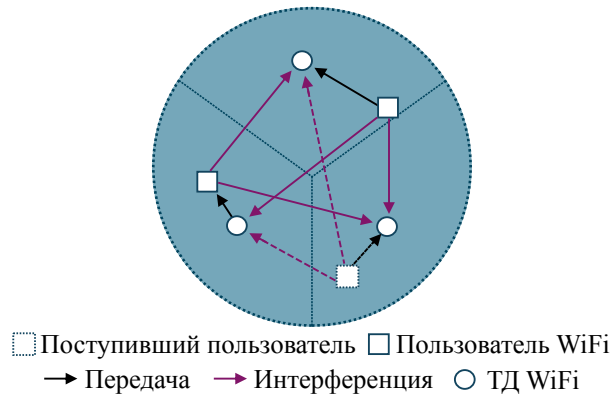


Рисунок А.1 — Моделирование работы слоя WLAN

Слой WLAN функционирует в соответствии с алгоритмом «Максимальная мощность», который устанавливает мощность передачи абонентских терминалов равной своему наибольшему возможному значению. Учитывая тот факт, что передача терминалов в системе WLAN подвержена существенному влиянию интерференции, далее в качестве отображения мощность – скорость пе-

передачи данных целесообразно использовать соответствующую вариацию формулы Шеннона. При этом принимается в расчет текущее число абонентских устройств, обслуживаемых некоторой ТД. Тогда мгновенная скорость передачи данных для сессии i определяется максимальной излучаемой мощностью p_{\max} как

$$r_i = r_i^{\max} = w \log(1 + \text{SINR}_i) = w \log\left(1 + \frac{\eta p_{\max} \gamma_{i,i}}{K N_0}\right),$$

где N_0 – известный уровень шума, а K – масштабирующий коэффициент, учитывающий вклад интерференции.

Таким образом, сеть доступа принимает на обслуживание новую сессию и ассоциирует ее с некоторой ТД, если выполняется следующее условие:

$$\sum_{\text{all AP sessions}} \left(\frac{r_0}{r_i^{\max}}\right) \leq \delta_w,$$

где δ_w – доступные радиоресурсы в рамках группы узлов, включающих данную ТД, но без учета накладных расходов, связанных с протоколом доступа канального уровня (отражающих вклад управляющих сообщений и издержки на конкуренцию терминалов за доступ к каналу связи, см. подробности в [342]), а r_{\max} – наибольшее достижимое значение скорости передачи данных при использовании максимальной излучаемой мощности.

Специфика способа управления интерференцией в сетях WLAN связана с работой функции ССА, которая определяет наличие передачи в канале связи. Таким образом, пороговый уровень шума можно условно принять равным значению, приводящему к срабатыванию функции ССА. В этой связи процедура контроля доступа системы WLAN проверяет, не превысит ли значение интерференции, которая будет произведена с учетом дополнительной мощности p_{\max} от вновь допущенного абонентского терминала, некоторого заданного порога по уровню шума N_0 на каждой из соседних ТД.

А.1.2 Модель пикослота

Далее рассматривается пикослой, включающий в себя случайное число малых БС N_p , P_j , $j = \overline{1, N_p}$, которые размещены равномерно на плоскости в соответствии с допущениями модели системы (см. рисунок А.2).

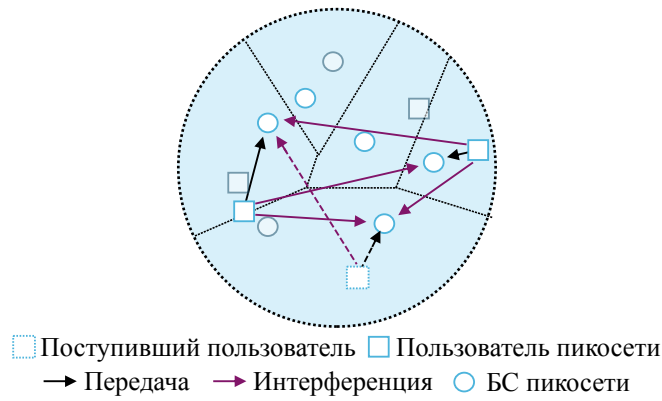


Рисунок А.2 — Моделирование работы пикослоя

В рамках описанной модели системы связи все пикосоты работают на одной и той же частоте, но если бы использовалось несколько различных частот, то их можно было бы рассматривать как множество отдельных пикослоев. По аналогии со слоем WLAN на соседних малых БС происходит учет уровня интерференции, производимой допущенными на обслуживание абонентскими терминалами.

Мощность p_i , излучаемая терминалом, и соответствующая скорость передачи данных r_i также связаны формулой Шеннона, учитывающей вклад интерференции:

$$r_i = w \log (1 + \text{SINR}_i) = w \log \left(1 + \frac{\eta p_i \gamma_{i,i}}{K N_0} \right).$$

В рамках пикослоя производится учет работы процедуры управления мощностью передачи в разомкнутом цикле (open-loop) как описано в [343], раздел 5.4.1.3. Соответственно, предполагается, что некоторый централизованный модуль управления оценивает параметры системы связи и сообщает абонентским терминалам подходящее значение излучаемой мощности p_i . Такая оценка основывается как на характеристиках, связанных с мощностью, так и на характеристиках, связанных с интерференцией, и производится следующим образом. В качестве предельно допустимого уровня интерференции принимается наивысшее ее значение, наблюдаемое на сильно загруженной малой БС. Являясь функцией числа соседних пикосот, такой уровень интерференции предполагает независимость от числа активных пользователей. Данное допущение упрощает учет интерференции на этапе контроля доступа, приводя к оценке ее уровня в худшем случае.

В отличие от слоя WLAN, пикослой использует алгоритм «Кольцевое планирование» для управления мощностью и планирования радиоресурсов. Как следствие, для приема на обслуживание новой сессии n_0 малая БС должна повысить мощность передачи других активных абонентских терминалов так, чтобы они «уложились» в меньшие интервалы времени после переупаковки радиоресурсов. Если такое повышение невозможно хотя бы для одной из уже допущенных сессий (включая эту новую сессию) и хотя бы на какой-то малой БС, т.е. $r_i^{\max} = w \log \left(1 + \frac{\gamma_i p_{\max}}{K N_0} \right) < n_0 r_0$, тогда рассматриваемой сессии будет отказано в доступе (в пикослой). В противном случае система перераспределит время своей работы для активных сессий, тогда как абонентским терминалам будет установлен более высокий уровень излучаемой мощности

$$p_i = \frac{1}{\eta \gamma_i} \left(e^{n_0 r_0 / w} - 1 \right) (K N_0),$$

где $p_i \leq p_{\max}$, а $K N_0$ – совокупный уровень шума и интерференции от соседних малых БС.

Таким образом, пикослой принимает на обслуживание только такие новые сессии, для передачи которых имеется достаточное количество радиоресурсов. Другими словами, каждая активная сессия на некоторой малой БС будет занимать в точности долю $1/n_0$ от общего времени работы системы, тогда как для всех текущих сессий должно выполняться следующее:

$$\sum_{\text{all sessions}} \left(\frac{r_0}{r_i^{\max}} \right) = \delta_p,$$

где δ_p – общие доступные радиоресурсы в рамках группы узлов, включающих данную малую БС.

А.1.3 Модель макрослоя

Наконец, ниже рассматривается макрослой, включающий в себя единственную макросоту, учет межсотовой интерференции для которой не производится (см. рисунок А.3). Это равносильно замене уровня интерференции, поступающей от соседних макросот, на эквивалентный прирост уровня шума в канале. Такая постановка приводит к рассмотрению сети связи, не подверженной

влиянию интерференции, поскольку интервалы передачи данных от абонентских терминалов по линии вверх разделены между собой согласно протоколу доступа к среде.

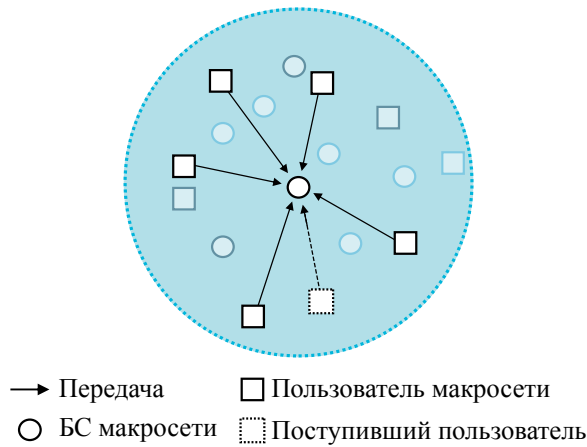


Рисунок А.3 — Моделирование работы макрослоя

Следовательно, в отличие от слоя WLAN и пикослоя, для макрослоя соответствующее допущение модели системы интерпретируется как следующее (упрощенное) отображение мощность – скорость передачи данных:

$$r_i = w \log (1 + \text{SNR}_i) = w \log \left(1 + \frac{\eta \gamma_i}{K N_0} p_i \right),$$

где SNR – отношение сигнал – шум; γ_i – коэффициент передачи радиоканала между абонентским терминалом и БС макросоты, а N_0 – уровень шума.

С учетом функционирования процедуры управления мощностью передачи OLPC, контроль доступа и управление мощностью в макрослое производятся по аналогии с тем, как они осуществляются в пикослое. При этом каждой допущенной к обслуживанию сессии предоставляется равная доля общего времени работы системы

$$\frac{r_0}{r_i} = \frac{\delta_m}{n}, r_i = r_0 n, \quad \forall i = \overline{1, n},$$

где δ_m – общие доступные радиоресурсы в рамках макросоты.

Однако ввиду отсутствия интерференции, процедура контроля доступа новых сессий в макрослой упрощается соответствующим образом. Как следствие, вновь поступившая сессия может быть принята на обслуживание в макрослой если и только если

$$r_i^{\max} = w \log \left(1 + \frac{\eta \gamma_i p_{\max}}{K N_0} \right) \geq n \frac{r_0}{\delta_m}.$$

Если выполняется приведенное выше условие, то общее время работы системы перераспределяется для n активных в текущий момент сессий, тогда как уровень мощности, излучаемой абонентскими терминалами, устанавливается в соответствии с $p_i = \frac{KN_0}{\eta\gamma_i} (e^{nr_0/(w\delta_m)} - 1) \geq p_{\max}$. В противном случае рассматриваемая сессия считается окончательно заблокированной и покидает систему, не получив обслуживания. Далее проводится анализ показателей эффективности гетерогенной сети доступа в рамках ее модели, введенной в главе 2. При этом три вышеописанных слоя изучаются по отдельности, а сам анализ построен по принципу разделения слоев согласно допущениям модели системы. Такой подход широко применяется для исследования сложных систем и связан с декомпозицией их на ряд более простых аналитических компонентов.

А.2 Подход к аналитическому моделированию

Ниже предлагается подход к аналитическому моделированию интегрированной гетерогенной сети доступа, включающей в себя слои WLAN, пико- и макро-, который построен на допущениях главы 2. Предполагается, что поступление новых сессий в три последовательно «упаковываемых» слоя системы связи происходит согласно пуассоновскому процессу с интенсивностями $\lambda_w = \lambda$, λ_p и λ_m соответственно. При этом макрослой рассматривается в моменты времени t , соответствующие поступлению новой сессии или окончанию обслуживания текущей сессии. Ввиду пуассоновского характера поступлений и экспоненциально распределенной длительности сессий (процесса обслуживания), поведение изучаемой системы может быть представлено как марковский процесс $S(t)$, будущее состояние которого определено набором активных сессий, обслуживаемых в текущий момент времени в некотором слое.

А.2.1 Анализ процессов в макрослое

Для макрослоя состояние процесса $S(t)$ задано характеристиками текущих сессий, обслуживаемых макросотой. Для удобства такие абстрактные ха-

рактические характеристики обозначаются ω и зависят от местоположения абонентского терминала. Таким образом, состояние системы может быть представлено как вектор $(\omega_1, \dots, \omega_n)$, в котором n – число активных сессий (см. рисунок А.4).

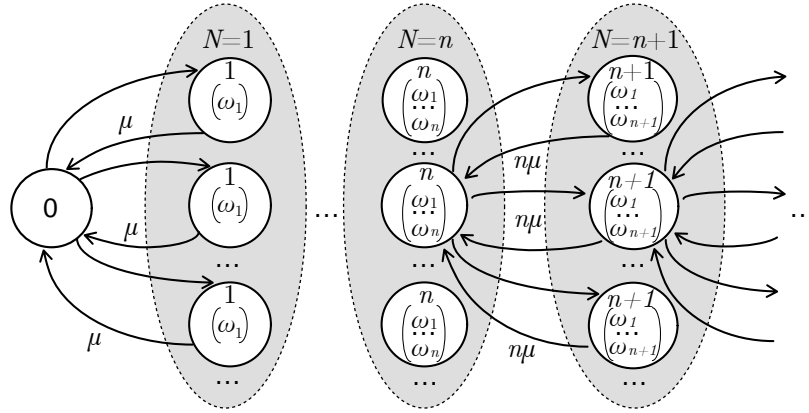


Рисунок А.4 — Пространство состояний системы для макрослоя

Пусть в макрослое имеется n активных сессий в состоянии s . Вероятность блокировки (отказа) в состоянии s для вновь поступившей сессии обозначается как $Q_{n+1|s}$. Тогда переходы из состояния $s = (\omega_1, \dots, \omega_n)$ в состояние $(\omega_1, \dots, \omega_n, \omega_{n+1})$ и обратно происходят с интенсивностями $\lambda_m (1 - Q_{n+1|s})$ и $(n + 1)\mu$ соответственно.

А.2.2 Анализ процессов в слое WLAN и пикослое

Поскольку как слой WLAN, так и пикослой подвержены воздействию интерференции, их моделирование производится схожим образом и описано ниже. Как следствие, соответствующие случайные процессы схожи в отношении своего пространства состояний. В частности, слой WLAN/пико- включает в себя некоторое число ТД/БС, которые размещены на плоскости с плотностями L_w и L_p . Таким образом, состояния марковского процесса $S(t)$ характеризуются набором сессий, обслуживаемых определенной ТД/БС. По аналогии с макрослоем вводится обозначение ω , тогда как состояние s слоя WLAN/пико- определяется в соответствии с

$$(\omega_1, \dots, \omega_{n_1}; \omega_{n_1+1}, \dots, \omega_{n_1+n_2}; \dots; \omega_{s_n+1}, \dots, \omega_{s_n+n_{N_w/f}}),$$

где $s_n = \sum_{i=1}^{N_w/p-1} n_i$, а n_1 , n_2 и $n_{N_w/p}$ – число абонентских терминалов, обслуживаемых первой, второй и последней ТД/БС соответственно. Случайная величина N_w/p представляет собой число ТД/БС в определенной области, распределенное по пуассоновскому закону. Пространство состояний рассматриваемой системы представлено на рисунке А.5.

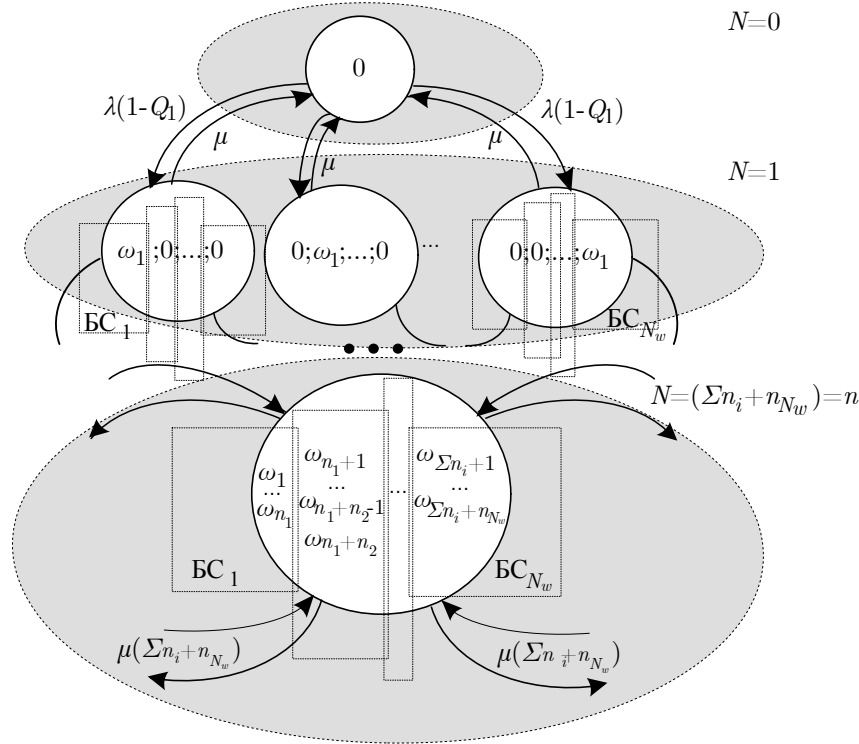


Рисунок А.5 — Пространство состояний системы для слоя WLAN/пико-

Представляет интерес состояние s , в котором слой WLAN/пико- обслуживает n активных сессий при помощи некоторого случайного числа N_w/N_p ТД/БС. Вероятность блокировки (отказа) вновь поступившей сессии обозначается $Q_{n+1|s}$. Тогда переходы из состояния s в состояние $n+1$, соответствующее текущему числу сессий, происходят с интенсивностью $L_w/p (1 - Q_{n+1|s})$. Обратные переходы происходят с интенсивностью $(n+1)\mu$, поскольку обслуживание сессии зависит не от конкретного состояния системы, а от числа одновременно активных сессий.

А.3 Стационарное состояние системы

В связи с потенциально неограниченным числом состояний рассматриваемой системы связи, нахождение стационарного распределения вероятностей для нее представляет собой сложную задачу [344]. Однако соответствующий марковский процесс может быть существенно упрощен за счет использования подхода *укрупнения состояний*.

Допущение (Укрупнение состояний).

1. Для макрослоя все состояния системы, содержащие n активных сессий, могут быть объединены в некоторое обобщенное состояние n вне зависимости от фактического местоположения абонентских терминалов.

2. Для слоя WLAN и пикослоя все состояния системы, содержащие n активных сессий, могут быть объединены в некоторое обобщенное состояние n вне зависимости от фактического местоположения абонентских терминалов и их ассоциации с определенной ТД/БС. Описываемый здесь подход проиллюстрирован на рисунке А.5.

3. Для сохранения свойства отсутствия «памяти» в рассматриваемой системе применяется упрощение, в соответствии с которым активные в состоянии n терминалы не сохраняют свое местоположение от состояния к состоянию, не изменяя при этом остальных своих свойств. В целях достижения аналитичности эти местоположения генерируются заново для каждого нового состояния n .

Следует отметить, что такой упрощенный подход тем не менее позволяет вести учет допущенных в прошлом сессий благодаря использованию вероятности Q_{n+1} блокировки (отказа) сессии, поступившей в состоянии n при том условии, что текущая n -я сессия удовлетворяет критериям процедуры контроля доступа. В результате укрупнения состояний изучение исходного процесса сводится к рассмотрению процессов гибели и размножения для каждого из трех слоев с интенсивностями $\lambda_{m/w/p}(1 - Q_{n+1})$ и $(n + 1)\mu$. Далее для простоты общая интенсивность поступления сессий в систему $\lambda_{m/w/p}$ обозначается как λ , а также формулируется следующее утверждение.

Утверждение. Стационарное распределение вероятностей $\{\pi_i\}_{i=0}^{\infty}$ для рассматриваемого процесса $S(t)$ с интенсивностями переходов в виде $\lambda(1 - Q_{n+1})$

и $(n + 1)\mu$ может быть аппроксимировано как

$$\pi_n = \pi_0 \frac{\lambda_{m/w/p}^n \prod_{i=1}^n (1 - Q_n)}{\mu^n n!}, \quad (\text{A.1})$$

где

$$\pi_0 = \left(\sum_{i=0}^{\infty} \frac{\lambda_{m/w/p}^i \prod_{i=1}^i (1 - Q_n)}{\mu^i i!} \right)^{-1},$$

а Q_{n+1} – вероятность блокировки (отказа) сессии при переходе из состояния n в состояние $n + 1$.

Доказательство. Приведенные выше выражения соответствуют стационарному распределению вероятностей в широко известной системе $M/M/c$, рассмотрению которой посвящено значительное количество соответствующей литературы.

Основываясь на стационарном распределении вероятностей и предполагая, что оно существует, можно оценить ряд общесистемных характеристик изучаемой сети связи в стационарном режиме. К примеру, можно вычислить среднее число активных сессий, а также вероятность того, что сессия будет окончательно заблокирована, следующим образом:

$$E[N] = \sum_{n=0}^{\infty} n\pi_n, \quad P_{block} = \sum_{n=0}^{\infty} Q_{n+1}\pi_n. \quad (\text{A.2})$$

Предложенный аналитический подход не рассматривает историю индивидуальных сессий, а исходные процессы с «памятью» заменяет на эквивалентные процессы без «памяти», в которых исследуется некоторый набор соответствующих случайных величин в каждый момент времени t . Если вычислены вероятности блокировки (отказа) сессии Q_{n+1} для каждого из трех слоев, то можно с легкостью получить стационарное распределение вероятностей. Таким образом, далее показано, как установить значения вероятностей Q_{n+1} . При этом для косвенного учета «памяти» системы, которая была опущена введенным выше упрощением, последующее изложение построено на рассмотрении соответствующих условных вероятностей.

А.4 Основные переходы в системе

В данном разделе производится вычисление вероятности Q_{n+1} и интенсивности переходов $\lambda_{m/w/p}(1 - Q_{n+1})$, требуемых для оценки стационарного распределения вероятностей. В свою очередь, это позволяет получить важные характеристики работы системы в стационарном режиме, включая вероятность окончательной блокировки сессии и среднее число активных сессий. Последнее может также использоваться для нахождения емкости системы связи в расчете на единицу площади при достаточно высоких значениях интенсивности входного потока сессий.

Последующее изложение построено следующим образом. Во-первых, предлагаемый подход иллюстрируется на примере более простого для понимания макрослоя. Во-вторых, рассматривается более сложный в силу необходимости учета интерференции слоев WLAN. В-третьих, изучается пикослой, который наследует ряд своих свойств как от макрослоя, так и от слоя WLAN.

А.4.1 Переходы в макрослой

Здесь рассматривается макрослой, для которого производятся вычисления с учетом работы алгоритма «Кольцевое планирование» для управления мощностью и планирования радиоресурсов. Соответственно, переходы из состояния n в состояние $n + 1$ определяются как

$$\begin{aligned} & \lambda_m (1 - Q_{n+1}) = \\ & = \lambda_m \left(\Pr \left\{ \frac{r_0}{r_i^{\max}} \leq \frac{\delta_m}{n+1}, \quad \forall i = 1, n+1 \mid \frac{r_0}{r_i^{\max}} \leq \frac{\delta_m}{n}, \quad \forall i = 1, n \right\} \right), \quad (\text{A.3}) \end{aligned}$$

где приведенная выше условная вероятность косвенно учитывает историю функционирования системы, тогда как сам упрощенный процесс, моделирующий ее работу, не имеет «памяти». Другими словами, оценивается вероятность того, что радиоресурсы системы могут быть распределены между $n + 1$ случайными сессиями при условии, если n других в стохастическом смысле сессий уже были допущены в систему в ее предшествующем состоянии.

Основываясь на предыдущем выражении, значение искомой вероятности вычисляется как

$$\begin{aligned} & \Pr \left\{ \frac{r_0}{r_i^{\max}} \leq \frac{\delta_m}{n+1}, \quad \forall i = 1, (n+1) \mid \frac{r_0}{r_i^{\max}} \leq \frac{\delta_m}{n}, \quad \forall i = 1, n \right\} = \\ & = \Pr \left\{ r_i^{\max} \geq \frac{r_0}{\delta_m} (n+1), \quad \forall i = 1, (n+1) \mid r_i^{\max} \geq \frac{r_0}{\delta_m} n, \quad \forall i = 1, n \right\}. \end{aligned}$$

Приведенный выше результат можно разделить на компоненты, которые соответствуют вновь поступившим и уже обслуживаемым сессиям:

$$\begin{aligned} & \Pr \left\{ r_{n+1}^{\max} \geq \frac{r_0(n+1)}{\delta_m} \right\} \prod_{i=1}^n \Pr \left\{ r_i^{\max} \geq \frac{r_0(n+1)}{\delta_m} \mid r_i^{\max} \geq \frac{r_0 n}{\delta_m} \right\} = \\ & = \Pr \left\{ r_{n+1}^{\max} \geq \frac{r_0(n+1)}{\delta_m} \right\} \left(\frac{\Pr \left\{ r_i^{\max} \geq \frac{r_0}{\delta_m} (n+1) \right\}}{\Pr \left\{ r_i^{\max} \geq \frac{r_0 n}{\delta_m} \right\}} \right)^n. \end{aligned} \quad (\text{A.4})$$

Вычисление вероятностей $\Pr \left\{ r_i^{\max} \geq \frac{r_0}{\delta_m} (n+1) \right\}$ и $\Pr \left\{ r_i^{\max} \geq \frac{r_0}{\delta_m} n \right\}$ основано на распределении $F_r(r)$ случайной величины r_i^{\max} и описано в заключительном разделе данного приложения. Зная эти вероятности, можно найти интенсивности переходов, а затем и стационарное распределение вероятностей, которое позволяет получить искомые общесистемные характеристики.

A.4.2 Переходы в слое WLAN

Далее рассматривается слой WLAN, для которого производятся вычисления с учетом работы алгоритма «Максимальная мощность» для управления мощностью и планирования радиоресурсов. Соответственно, слой WLAN принимает на обслуживание новую сессию, если выполняются следующие два условия: скорость передачи данных абонентским терминалом гарантируется не ниже своего минимального требуемого значения r_0 и уровень интерференции, производимой данным терминалом на соседние ТД, не превышает заданного порога N_0 . Тогда переходы из состояния n в состояние $n+1$ могут быть определены как

$$\lambda_w (1 - Q_{n+1}) = \lambda_w \Pr \left\{ A_j^{(n+1)}, j = \overline{1, n+1} \mid A_j^{(n)}, j = \overline{1, n} \right\}, \quad (\text{A.5})$$

где событие $A_j^{(n)}$ задано следующим образом:

$$A_j^{(n)} = \left\{ \frac{r_0}{r_j^{\max}} \leq \delta_w - \sigma_n \text{ и } \gamma_{j,k} p_{\max} \leq N_0, \quad \forall k \neq j \right\}.$$

Здесь δ_w – общие радиоресурсы, доступные на определенной ТД (без учета накладных расходов на передачу управляющих сообщений и разрешение конфликтов), а σ_n – доля радиоресурсов, выделенных другим сессиям, ассоциированным с той же самой ТД в текущем состоянии. Затем $r_0/(\delta_w - \sigma_n)$ обозначается как $\tilde{r}_{0,n}$, а вычисление σ_n приводится в заключительном разделе данного приложения по отдельности для слоя WLAN и для пикослоя на основании следующего допущения.

Допущение (Упрощение ассоциации с ТД). Для упрощения ассоциирования некоторой сессии с обслуживающей ее ТД в состоянии n предполагается, что в момент ее поступления в систему данная сессия наблюдает среднее (типовое) число активных соединений с ближайшей ТД. Это среднее значение зависит от количества активных сессий, т.е. от номера состояния n , а также от параметра \tilde{r}_0 .

Здесь полезно напомнить, что ТД размещены на плоскости согласно процессу PPP. Как следствие, число ТД k , не являющихся ближайшими к абонентскому терминалу, может быть достаточно велико. Для каждой сессии можно перенумеровать множество ТД в порядке увеличения расстояния между ними и соответствующим абонентским терминалом так, чтобы $k = 1$ соответствовало ближайшей ТД. При этом условие допуска сессии по скорости передачи должно выполняться для этой ближайшей ТД, тогда как условие по уровню интерференции должно выполняться для всех остальных ТД. Таким образом, интенсивность переходов может быть получена как

$$(1 - Q_{n+1}) = \frac{\Pr\{A_j^{(n+1)}, j=\overline{1, n+1}\}}{\Pr\{A_j^{(n)}, j=\overline{1, n}\}},$$

где $\{A_j^{(n)}, j = \overline{1, n}\}$ далее обозначается как $\{\mathbf{A}_n\}$.

Тогда вероятность $\Pr\{\mathbf{A}_n\}$ может быть выражена в следующем виде:

$$\Pr\{r_i^{\max} \geq \tilde{r}_{0,n} \text{ и } \gamma_{i,k} p_{\max} \leq N_0, \quad i = \overline{1, n}, k > 1\}.$$

Также предполагается, что значения $\gamma_{i,j}$ являются независимыми, что приводит к

$$\begin{aligned} \Pr\{\mathbf{A}_n\} &= \prod_{i=1}^n \Pr\{r_i^{\max} \geq r_{0,n}\} \prod_{i=1}^n \Pr\{\gamma_{i,k} p_{\max} \leq N_0, k > 1\} = \\ &= \prod_{i=1}^n \Pr\left\{\gamma_{i,i} \geq \frac{KN_0}{\eta p_{\max}} \left(e^{\frac{\tilde{r}_{0,n}}{w}} - 1\right)\right\} \prod_{i=1}^n \Pr\{\gamma_{i,k} p_{\max} \leq N_0, k > 1\}. \end{aligned}$$

Соответствующее значение $\gamma_{i,k}$ обозначается как γ_k , где γ_1 представляет собой коэффициент передачи радиоканала до ближайшей ТД, находящейся на расстоянии d_1 . Предполагается, что распределение величины γ_1 одинаково для всех сессий i , что выполняется для любого γ_k . Следовательно, можно установить

$$\Pr \{ \mathbf{A}_n \} = [\Pr \{ d_1 \leq d_{r,n} \}]^n \left[\Pr \left\{ \gamma_k \leq \frac{N_0}{p_{\max}}, k > 1 \right\} \right]^n,$$

где постоянная $d_{r,n}$ представляет собой $\left(\frac{p_{\max} \eta G}{K N_0} \right)^{\frac{1}{\kappa}} \left(e^{\frac{\tilde{r}_{0,n}}{w}} - 1 \right)^{-\frac{1}{\kappa}}$.

В результате значение искомой вероятности может быть получено через соотношение между значением коэффициента передачи радиоканала γ и расстоянием d , которое задано как

$$d = (G/\gamma)^{1/\kappa}, \quad (\text{A.6})$$

что приводит к следующему результату:

$$\Pr \left\{ \gamma_k \leq \frac{N_0}{p_{\max}}, k > 1 \right\} = \Pr \{ d_k \geq d_{thr}, k > 1 \} = \Pr \{ d_2 \geq d_{thr} \},$$

где $d_{thr} = \left[\frac{G p_{\max}}{N_0} \right]^{\frac{1}{\kappa}}$.

Тогда после подстановки получается

$$\Pr \{ \mathbf{A}_n \} = [F_{d_1}(d_{r,n})]^n [\Pr \{ d_2 < d_{thr} \}]^n.$$

Следует подчеркнуть, что условие $d_2 > d_{thr}$ выполняется тогда и только тогда, когда внутри окружности радиуса d_{thr} , описанной вокруг рассматриваемого абонентского терминала, находится в точности одна или же ни одной ТД. Следовательно, вероятность $\Pr \{ d_2 \geq d_{thr} \}$ можно получить, используя известные свойства пуассоновского процесса. В результате, основываясь на распределениях величин d_1 и d_2 , можно установить

$$\Pr \{ \mathbf{A}_n \} = \left[1 - e^{-\pi L_w d_r^2} \right]^n \left[L_w \pi d_{thr}^2 e^{-L_w \pi d_{thr}^2} + e^{-L_w \pi d_{thr}^2} \right]^n.$$

Наконец, можно вычислить интенсивность переходов как отношение $\Pr \{ \mathbf{A}_{n+1} \} / \Pr \{ \mathbf{A}_n \}$, т. е.

$$1 - Q_{n+1} = \frac{\left(1 - e^{-\pi L_w d_{r,n+1}^2} \right)^{n+1}}{\left(1 - e^{-\pi L_w d_{r,n}^2} \right)^n} \left(L_w \pi d_{thr}^2 e^{-L_w \pi d_{thr}^2} + e^{-L_w \pi d_{thr}^2} \right). \quad (\text{A.7})$$

Здесь значение $d_{r,n}$ зависит от $\tilde{r}_{0,n} = r_0/(\delta_w - \sigma_n)$, что, в свою очередь, является функцией числа активных сессий в слое WLAN с учетом занятых радиоресурсов σ_n . Способ нахождения σ_n , а также других необходимых параметров этого выражения представлен в заключительном разделе данного приложения. В итоге представляется возможным установить общесистемные характеристики, такие как среднее число активных сессий, а также вероятность того, что сессия будет окончательно заблокирована.

А.4.3 Переходы в пикослое

Текущий раздел завершается рассмотрением пикослоя и соответствующими вычислениями. По аналогии с макрослоем, также использующим алгоритм «Кольцевое планирование», переходы из состояния n в состояние $n + 1$ для пикослоя определяются как

$$\lambda_p (1 - Q_{n+1}) = \lambda_p \Pr \left\{ A_j^{(n+1)}, j = \overline{1, n+1} | A_j^{(n)}, j = \overline{1, n} \right\}. \quad (\text{A.8})$$

Событие $A_j^{(n)}$ представляет собой комбинацию двух факторов, т.е. выполнение условия допуска сессии по скорости передачи на ближайшей БС и выполнение условия по уровню интерференции на всех остальных БС. Таким образом, если малые БС пронумерованы в порядке возрастания расстояния от них до рассматриваемого абонентского терминала j , то

$$A_j^{(n)} = \left\{ r_j^{\max} \geq \frac{r_0 n_0}{\delta_p} \text{ и } \gamma_{j,k} p_j \leq N_0, \quad k > 1 \right\},$$

где n_0 – число терминалов, ассоциированных с ближайшей БС; δ_p – общие радиоресурсы, доступные на данной БС, а σ_n – доля радиоресурсов, выделенных другим сессиям, ассоциированным с той же самой БС в текущем состоянии.

Далее, используемая мощность передачи p_j может быть определена как

$$p_j = \frac{K N_0}{\eta \gamma_{j,j}} \left(e^{\frac{r_0 n_0}{w \delta_p}} - 1 \right). \quad (\text{A.9})$$

По аналогии со слоем WLAN малые БС размещены на плоскости согласно процессу PPP, а число БС k , не являющихся ближайшими к абонентскому терминалу, может быть достаточно велико для $k = \overline{1, \infty}$. Обозначая $\left\{ A_j^{(n)}, j = \overline{1, n} \right\}$

как $\{\mathbf{A}_n\}$, можно вычислить

$$\Pr \{\mathbf{A}_n\} = \Pr \left\{ r_i^{\max} \geq \frac{r_0}{\delta_p} n_0, \gamma_{i,k} p_i \leq N_0, i = \overline{1, n}, k > 1 \right\}.$$

Здесь предполагается, что значения $\gamma_{i,j}$ являются независимыми, что позволяет установить

$$\begin{aligned} \Pr \{\mathbf{A}_n\} &= \prod_{i=1}^n \Pr \left\{ r_i^{\max} \geq \frac{r_0}{\delta_p} n_0 \right\} \prod_{i=1}^n \Pr \{\gamma_{i,k} p_i \leq N_0, k > 1\} = \\ &= \prod_{i=1}^n \Pr \left\{ \gamma_{i,i} \geq \frac{KN_0}{\eta p_{\max}} \left(e^{\frac{r_0 n_0}{w \delta_p}} - 1 \right) \right\} \prod_{i=1}^n \Pr \{\gamma_{i,k} p_i \leq N_0, k > 1\}. \end{aligned}$$

Как и выше, соответствующее значение $\gamma_{i,k}$ обозначается γ_k , где γ_1 представляет собой коэффициент передачи радиоканала до ближайшей БС, находящейся на расстоянии d_1 . Следовательно, имеется

$$\Pr \{\mathbf{A}_n\} = [\Pr \{d_1 \leq d_{r,n}\}]^n \left[\Pr \left\{ \gamma_k \leq \frac{N_0}{p_i}, \quad \forall k > 1 \right\} \right]^n,$$

где $d_{r,n} = \left(\frac{p_{\max} \eta G}{KN_0} \right)^{\frac{1}{\kappa}} \left(e^{\frac{r_0 n_0}{w}} - 1 \right)^{-\frac{1}{\kappa}}$.

Наконец, вычисляется

$$\Pr \{\mathbf{A}_n\} = [F_{d_1}(d_{r,n})]^n \left[\Pr \left\{ \gamma_2 \leq \frac{\eta \gamma_1}{\left(e^{\frac{r_0 n_0}{w \delta_p}} - 1 \right)} \right\} \right]^n.$$

Способ нахождения компонентов, требуемых для вычисления приведенного выше выражения, представлен в заключительном разделе данного приложения. В итоге интенсивность переходов может быть найдена как

$$\lambda_p (1 - Q_{n+1}) = \lambda_p \frac{\Pr \{\mathbf{A}_{n+1}, \mathbf{A}_n\}}{\Pr \{\mathbf{A}_n\}} = \lambda_p \frac{\Pr \{\mathbf{A}_{n+1}\}}{\Pr \{\mathbf{A}_n\}}.$$

Используя этот результат, легко получить стационарное распределение вероятностей и затем – среднее число активных сессий, а также вероятность того, что сессия будет окончательно заблокирована.

A.5 Вспомогательные вычисления

В данном разделе приводятся некоторые вспомогательные распределения, необходимые для проведения вычислений, описанных в предыдущих разделах данного приложения.

А.5.1 Некоторые распределения для макрослоя

Проводя преобразования случайной величины r^{\max} , можно получить плотность вероятности для наивысшей мгновенной скорости передачи данных. В целях упрощения величина r_i^{\max} далее обозначается как r .

Вообще говоря, дальнейшие вычисления основываются на распределении скорости r , а также неявно используют распределение значений коэффициента передачи радиоканала γ . Если требуется описать различные точечные процессы, то данные распределения следует модифицировать с учетом соответствующей функции $f(x,y)$, что может представлять интересную техническую задачу, но не должно создать принципиальных трудностей.

Также важно отметить, что эффекты, связанные с затуханием радиосигнала, можно учесть в явном виде путем введения случайной величины h так, чтобы $\tilde{\gamma} = hGd^{-\kappa}$. Для учета медленного затухания h может быть описана логнормальным распределением со средним значением 1, тогда как для учета быстрого затухания принято использовать распределение Рэлея или Райса. В частности, новое распределение $\tilde{\gamma}$ может быть получено следующим образом. Поскольку процессы затухания сигнала и размещения абонентских терминалов на плоскости являются независимыми, то

$$f_{\tilde{\gamma}}(y) = \int_{\gamma_{\min}}^{\gamma_{\lim}} \frac{1}{x} f_{\gamma}(x) f_h\left(\frac{y}{x}\right) dx, \quad (\text{A.10})$$

где f_{γ} – базовое распределение значения коэффициента передачи радиоканала для заданного способа размещения терминалов на плоскости, которое описано ниже. Если требуется учесть более сложные условия, связанные с зависимостью данных процессов, целесообразно рассмотреть соответствующую совместную плотность вероятности.

В приведенном примере с равномерным размещением абонентских терминалов внутри окружности распределение расстояний между терминалом и обслуживающей его БС макросоты представляет собой $f_d(d) = 2d/R^2$, $0 \leq d \leq R$. Тогда, основываясь на соотношении между значением коэффициента передачи радиоканала γ и расстоянием d , можно выразить распределение величины γ в

виде

$$f_{\gamma}(\gamma) = \frac{1}{\kappa} \left[\frac{G}{\gamma} \right]^{\frac{1-\kappa}{\kappa}} \cdot \frac{2G^{\frac{1}{\kappa}}}{\gamma^{\frac{1}{\kappa}} R^2} = \frac{2}{\kappa R^2} \left[\frac{G}{\gamma} \right]^{\frac{2}{\kappa}-1}, \quad \gamma_R \leq \gamma \leq \gamma_{\text{lim}}, \quad (\text{A.11})$$

где γ_R и γ_{lim} соответствуют самому низкому уровню передаваемого сигнала на границе соты $d = R$ и самому высокому его уровню при $d = d_0$.

Для нахождения распределения случайной величины $r(d)$ используется следующее выражение для d :

$$d = \left[\frac{\eta G p_{\text{max}}}{K N_0} \right]^{1/\kappa} \left(e^{r/w} - 1 \right)^{-1/\kappa}, \quad (\text{A.12})$$

что позволяет вычислить распределение случайной величины r при отсутствии каких-либо ограничений.

$$F_r(r) = 1 - \frac{1}{R^2} \left[\frac{\eta G p_{\text{max}}}{K N_0} \right]^{2/\kappa} \left(e^{r/w} - 1 \right)^{-2/\kappa}, \quad r \geq r_R, \quad (\text{A.13})$$

где r_R – наибольшая возможная скорость передачи данных на границе R (т.е. нижняя граница для допустимых значений максимальной скорости передачи).

$$r_R = \min \left\{ r_{\text{lim}}, w \cdot \log \left(1 + \frac{\eta G p_{\text{max}}}{R^{\kappa} K N_0} \right) \right\} \geq r_i, \quad \forall i. \quad (\text{A.14})$$

Тогда, накладывая дополнительное ограничение r_{lim} , можно получить следующую функцию распределения:

$$F_r(r) = 1 - \frac{1}{R^2} \left[\frac{\eta G p}{N_0} \right]^{2/\kappa} \left(e^{r/w} - 1 \right)^{-2/\kappa}, \quad r_R \leq r < r_{\text{lim}}, \\ F_r(r_{\text{lim}}) = 1.$$

Таким образом, требуемые вероятности могут быть установлены как

$$\Pr \left\{ r \geq \frac{r_0 n}{\delta_m} \right\} = 1 - \Pr \left\{ r < \frac{r_0 n}{\delta_m} \right\} = 1 - F_r \left(\frac{r_0 n}{\delta_m} \right), \quad (\text{A.15})$$

а $\Pr \left\{ r \geq \frac{r_0}{\delta_m} (n + 1) \right\}$ может быть получена аналогично. Это позволяет вычислить стационарное распределение вероятностей.

A.5.2 Некоторые распределения для слоя WLAN

Далее выводятся вспомогательные вероятности и распределения для слоя WLAN. Вначале можно получить распределение расстояния d_1 от абонентского

терминала до ближайшей к нему ТД

$$F_{d_1}(d) = 1 - \Pr \{d_1 \geq d\} = 1 - \Pr \{N_w(d) = 0\}, \quad d \geq 0,$$

где $N_w(d)$ – случайное число ТД, размещенных внутри окружности радиуса d , описанной вокруг рассматриваемого терминала.

Аналогично можно найти распределение расстояния до второй ближайшей ТД

$$F_{d_2}(d) = 1 - \Pr \{d_2 \geq d_{thr}\} = 1 - \Pr \{0 \leq N_w(d_{thr}) \leq 1\},$$

где $\Pr \{0 \leq N_w(d_{thr}) \leq 1\}$ – вероятность того, что внутри окружности радиуса d , описанной вокруг рассматриваемого терминала, находится в точности одна или же ни одной ТД. С учетом того, что ТД распределены согласно пуассоновскому закону, для расстояний $d \geq 0$ можно вычислить

$$F_{d_1}(d) = 1 - e^{-\pi L_w d^2}, \quad f_{d_1}(d) = 2\pi L_w d e^{-\pi L_w d^2}. \quad (\text{A.16})$$

Тогда соответствующая вероятность для второго расстояния может быть установлена как

$$\begin{aligned} \Pr \{d_2 \geq d_{thr}\} &= \Pr \{N_w(d_{thr}) \leq 1\} = \\ &= L_w \pi d_{thr}^2 e^{-\pi d_{thr}^2} + e^{-L_w \pi d_{thr}^2}. \end{aligned} \quad (\text{A.17})$$

Для упрощения обозначений при рассмотрении наибольшей возможной скорости передачи данных предполагается, что распределение величины d связано с распределениями значений коэффициента передачи радиоканала γ и максимальной скорости передачи r . Следовательно, можно вывести распределение случайной величины

$$y = \frac{r_0}{r_i^{\max}} = \frac{r_0}{w} \left[\log \left(1 + p_{\max} \frac{\eta G}{K N_0} d^{-\kappa} \right) \right]^{-1}, \quad d \geq d_0. \quad (\text{A.18})$$

Выразив расстояние из предыдущего выражения, можно получить

$$d = \left(\frac{p_{\max} \eta G}{K N_0} \right)^{\frac{1}{\kappa}} \left(e^{\frac{r_0}{wy}} - 1 \right)^{-\frac{1}{\kappa}}, \quad y \geq y_0, \quad (\text{A.19})$$

где $y_0 = y(d_0)$, а производная от $d(y)$ принимает форму

$$d'_y = \frac{1}{\kappa} \left(\frac{p_{\max} \eta G}{K N_0} \right)^{\frac{1}{\kappa}} \left(e^{\frac{r_0}{wy}} - 1 \right)^{-\frac{1}{\kappa}-1} e^{\frac{r_0}{wy}} \frac{r_0}{wy^2}. \quad (\text{A.20})$$

Далее целесообразно вычислить следующую функцию распределения:

$$F_y(y) = 1 - e^{-\pi L_w \left(\frac{p_{\max} \eta G}{KN_0} \right)^{\frac{2}{\kappa}} \left(e^{\frac{r_0}{wy}} - 1 \right)^{-\frac{2}{\kappa}}}, \quad y \geq y_0. \quad (\text{A.21})$$

Тогда функция плотности вероятности

$$f_y(y) = 2\pi L_w d(y) d'(y) e^{-\pi L_w \left(\frac{p_{\max} \eta G}{KN_0} \right)^{\frac{2}{\kappa}} \left(e^{\frac{r_0}{wy}} - 1 \right)^{-\frac{2}{\kappa}}}. \quad (\text{A.22})$$

Следовательно, можно записать математическое ожидание случайной величины $E[y|y \leq \delta_w]$ в виде

$$\begin{aligned} E[y|y \leq \delta_w] &= \int_{y_0}^{\delta_w} y f_y(y|y \leq \delta_w) dy = \\ &= \frac{2\pi L_w}{C_3} \int_{y_0}^{\delta_w} y d(y) d'(y) e^{-\pi L_w \left(\frac{p_{\max} \eta G}{KN_0} \right)^{\frac{2}{\kappa}} \left(e^{\frac{r_0}{wy}} - 1 \right)^{-\frac{2}{\kappa}}} dy + \\ &+ \frac{2\pi L_w}{C_3} y_0 \int_0^{y_0} d(y) d'(y) e^{-\pi L_w \left(\frac{p_{\max} \eta G}{KN_0} \right)^{\frac{2}{\kappa}} \left(e^{\frac{r_0}{wy}} - 1 \right)^{-\frac{2}{\kappa}}} dy, \end{aligned} \quad (\text{A.23})$$

где $C_3 = \Pr\{y \leq \delta_w\} = F_y(\delta_w)$, а значение y_0 предполагается меньше, чем δ_w .

Ниже рассматривается состояние системы, в котором имеется n активных сессий. Плотность размещения абонентских терминалов при этом составляет $\lambda_N = n/(\pi R^2)$. Можно было бы продолжить заполнение плоскости терминалами до достижения данной плотности так, чтобы каждая ТД получила на обслуживание некоторое число сессий.

Утверждение. Среднее число обслуживаемых одной ТД сессий n_0 стремится к $\frac{n}{L_w(\pi R^2)} = \frac{n}{E[N_w]}$ для достаточно больших областей, где $E[N_w]$ – среднее число ТД, размещенных внутри окружности радиуса R .

Доказательство. Для достаточно большой области $S \gg \pi R^2$ можно оценить среднее число обслуживаемых одной ТД сессий как $N_n(S)/N_w(S)$, где N_n и N_w случайное число терминалов и ТД, размещенных в области S . Математическое ожидание значения $N_n(S)/N_w(S)$ может быть найдено по определению как

$$\begin{aligned} E \left[\frac{N_n(S)}{N_w(S)} \right] &= \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} \frac{k}{n} \Pr\{N_n(S) = k, N_w(S) = n\} = \\ &= \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} \frac{k}{n} \frac{(L_w S)^n}{n!} e^{-L_w S} \frac{(\lambda_N S)^k}{k!} e^{-\lambda_N S} = \\ &= \lambda_N S \sum_{n=1}^{\infty} \frac{1}{n e^{L_w S}} \frac{(L_w S)^n}{n!} \xrightarrow{S \rightarrow \infty} \frac{\lambda_N}{L_w} = \frac{n}{L_w(\pi R^2)}, \end{aligned} \quad (\text{A.24})$$

где $\lambda_N = \frac{n}{(\pi R^2)}$ – плотность размещения терминалов на единицу площади; L_w – плотность размещения ТД на единицу площади, а S – обозначение целевой замкнутой области.

Суммируя вышеизложенное, можно заключить, что среднее число обслуживаемых одной ТД терминалов в предельных случаях стремится к отношению между переменной плотностью размещения терминалов и постоянной плотностью размещения ТД.

Следует отметить, что данное утверждение аналогично более раннему результату, приведенному в работе [345]. Следовательно, вновь поступившая сессия наблюдает состояние системы, в котором каждая ТД уже обслуживает $n_0 = \frac{n}{L_w(\pi R^2)}$ сессий в среднем. В таком случае σ представляет собой усредненную долю радиоресурсов, занятых в состоянии n , и вычисляется как

$$\sigma = E \left[\frac{r_0}{r_i^{\max}} \mid \frac{r_0}{r_i^{\max}} \leq \delta_w \right] \frac{n}{E[N_w]} = E[y \mid y \leq \delta_w] \frac{n}{E[N_w]}, \quad (\text{A.25})$$

где математическое ожидание $E[y \mid y \leq \delta_w]$ может быть получено, например, численным интегрированием.

А.5.3 Некоторые распределения для пикослоя

Распределение расстояния от абонентского терминала до ближайшей к нему малой БС может быть получено по аналогии с соответствующим распределением для слоя WLAN:

$$F_{d_1}(d) = 1 - e^{-\pi L_p d^2}, \quad f_{d_1}(d) = 2\pi L_p d e^{-\pi L_p d^2}, \quad d \geq 0.$$

Тогда искомую вероятность можно выразить, выполняя шаги вычислительной процедуры из предыдущего подраздела

$$\begin{aligned} \Pr \left\{ \gamma_2 \leq \frac{N_0}{p_i} \right\} &= \Pr \left\{ \gamma_2 \leq \frac{\eta \gamma_1 N_0}{K N_0 \left(e^{\frac{r_0 n_0}{w \delta_p}} - 1 \right)} \right\} = \\ &= \Pr \{ \gamma_2 + y \leq 0 \} = \Pr \{ z \leq 0 \}, \end{aligned} \quad (\text{A.26})$$

с учетом следующей замены величин: $y = -\frac{\eta \gamma_1}{K \left(e^{\frac{r_0 n_0}{w \delta_p}} - 1 \right)} = -\frac{\eta G}{K \left(e^{\frac{r_0 n_0}{w \delta_p}} - 1 \right)} d_1^{-\kappa}$ и $z = \gamma_2 + y$.

Для нахождения распределения параметра z вначале вычисляется распределение случайной величины y . Для этого используется преобразование

$$d_1 = \left(\frac{\eta G}{K \left(e^{\frac{r_0 r_0}{w \delta_p}} - 1 \right)} \right)^{\frac{1}{\kappa}} (-y)^{-\frac{1}{\kappa}}, \quad (\text{A.27})$$

а также соответствующая первая производная

$$d'_1 = \frac{1}{\kappa} \left(\frac{\eta G}{K \left(e^{\frac{r_0 r_0}{w \delta_p}} - 1 \right)} \right)^{\frac{1}{\kappa}} (-y)^{-\frac{1}{\kappa}-1}. \quad (\text{A.28})$$

Применяя приведенные выше выражения, можно записать требуемую функцию распределения

$$F_y(y) = 1 - e^{-\pi L_p d(y)^2}, \quad y_{\min} < y < 0, \quad (\text{A.29})$$

где y_{\min} соответствует предельному значению скорости передачи данных. Таким образом, искомая функция принимает следующий вид:

$$f_y(y) = 2\pi L_p d(y) d'(y) e^{-\pi L_p d^2(y)}, \quad y_{\min} \leq y < 0. \quad (\text{A.30})$$

Основываясь на соотношении между значением коэффициента передачи радиоканала γ и расстоянием d , можно выразить распределение случайной величины $0 < \gamma_2 \leq \gamma_{\max}$, которое является одним из компонентов выражения для z :

$$\begin{aligned} F_{\gamma_2}(\gamma) &= \Pr \{ \gamma_2 < \gamma \} = \Pr \{ d_2 > d(\gamma) \} = \\ &= \Pr \{ N_p(d(\gamma)) \leq 1 \} = L_p \pi d^2(\gamma) e^{-L_p \pi d^2(\gamma)} + e^{-L_p \pi d^2(\gamma)}, \end{aligned} \quad (\text{A.31})$$

где $d(\gamma) = G^{\frac{1}{\kappa}} \gamma^{-\frac{1}{\kappa}}$; $d'(\gamma) = -\frac{1}{\kappa} G^{\frac{1}{\kappa}} \gamma^{-\frac{1}{\kappa}-1}$, а γ_{\max} соответствует расстоянию d_0 .

С учетом выражения $F_{d_2}(d) = 1 - \Pr \{ d_2 \geq d_2(\gamma) \}$ можно получить функцию плотности вероятности для расстояния d_2 до второй ближайшей малой БС

$$f_{d_2}(d) = 2L_p^2 \pi^2 d^3 e^{-L_p \pi d^2}, \quad 0 \leq d < \infty.$$

Далее для плотности вероятности соответствующего значения коэффициента передачи радиоканала γ_2 (при наличии ограничения величиной d_0) легко найти

$$\begin{aligned} f_{\gamma_2}(\gamma) &= \frac{2}{\kappa} L_p^2 \pi^2 G^{\frac{4}{\kappa}} \gamma^{-\frac{4}{\kappa}-1} e^{-L_p \pi d^2(\gamma)}, \quad 0 < \gamma < \gamma_{\max} \\ f_{\gamma_2}(\gamma_{\max}) &= 1 - F_{\gamma_2}(\gamma_{\max}). \end{aligned} \quad (\text{A.32})$$

Принимая в расчет пределы $y_{\min} \leq y < 0$ и $0 < \gamma \leq \gamma_{\max}$, можно получить выражение для $F_z(z)$:

$$\begin{aligned} F_z(z) &= \int_0^{\gamma_{\max}} f_\gamma(\gamma) \int_{y_{\min}}^{\min(0, z-\gamma)} f_y(y) dy d\gamma = \\ &= \int_0^{\gamma_{\max}} f_\gamma(\gamma) F_y(\min(0, z - \gamma)) d\gamma, \end{aligned} \quad (\text{A.33})$$

где γ_{\max} соответствует случаю d_0 .

Подставляя в предыдущее выражение значение $z = 0$, возможно найти неизвестную вероятность $\Pr \{z < 0\}$:

$$\begin{aligned} \Pr \{z < 0\} &= F_z(0) = \int_0^{\gamma_{\max}} f_\gamma(\gamma) F_y(-\gamma) d\gamma = \\ &= \frac{2}{\kappa} L_p^2 \pi^2 G^{\frac{4}{\kappa}} \int_0^{\gamma_{\max}} \gamma^{-\frac{4}{\kappa}-1} e^{-L_p \pi G^{\frac{2}{\kappa}} \gamma^{-\frac{2}{\kappa}}} \left(1 - e^{C_1 \gamma^{-\frac{2}{\kappa}}}\right) d\gamma, \end{aligned} \quad (\text{A.34})$$

где при этом

$$C_1 = -\pi L_p \left(\frac{\eta G}{K \left(e^{\frac{r_0 n_0}{w \delta_p}} - 1 \right)} \right)^{\frac{2}{\kappa}}.$$

Следует отметить, что вероятность выполнения условия допуска сессии по интерференции $\Pr \{\gamma_2 p_i \leq N_0\} = F_z(0)$ составляет

$$\begin{aligned} 1 - \frac{2}{\kappa} L_f^2 \pi^2 G^{\frac{4}{\kappa}} \int_0^{\gamma_{\max}} \gamma^{-\frac{4}{\kappa}-1} e^{-L_p \pi G^{\frac{2}{\kappa}} \gamma^{-\frac{2}{\kappa}}} e^{C_1 \gamma^{-\frac{2}{\kappa}}} d\gamma = \\ = 1 - \frac{2}{\kappa} L_p^2 \pi^2 G^{\frac{4}{\kappa}} \int_0^{\gamma_{\max}} \gamma^{-\frac{4}{\kappa}-1} e^{(-L_p \pi G^{\frac{2}{\kappa}} + C_1) \gamma^{-\frac{2}{\kappa}}} d\gamma. \end{aligned} \quad (\text{A.35})$$

Кроме того, полезно вычислить значение интеграла $\int_0^{\gamma_{\max}} \gamma^{-\frac{4}{\kappa}-1} e^{-C \gamma^{-\frac{2}{\kappa}}} d\gamma$ для $C > 0$ путем подстановки $u = \gamma^{-\frac{2}{\kappa}}$, $\gamma = u^{-\frac{\kappa}{2}}$, $d\gamma = -\frac{\kappa}{2} u^{-\frac{\kappa}{2}-1} du$:

$$\begin{aligned} - \int_0^{\gamma_{\max}} \frac{\kappa}{2} u^{-\frac{\kappa}{2}-1} u^{-\frac{\kappa}{2}(-\frac{4}{\kappa}-1)} e^{-Cu} d\gamma &= -\frac{\kappa}{2} \int_0^{\gamma_{\max}} u e^{-Cu} du = \\ &= -\frac{\kappa}{2} \left(-\frac{1}{C} u e^{-Cu} \Big|_0^{\gamma_{\max}} + \frac{1}{C} \int_0^{\gamma_{\max}} e^{-Cu} du \right) = \\ &= \frac{\kappa \gamma_{\max}}{2C} e^{-C \gamma_{\max}} + \frac{\kappa}{2C^2} e^{-Cu} \Big|_0^{\gamma_{\max}} = \frac{\kappa \gamma_{\max}}{2C} e^{-C \gamma_{\max}} + \frac{\kappa}{2C^2}. \end{aligned} \quad (\text{A.36})$$

Тогда можно установить

$$\Pr \{\gamma_2 p_i \leq N_0\} = 1 - L_p^2 \pi^2 G^{\frac{4}{\kappa}} \left(\frac{\gamma_{\max}}{C} e^{-C \gamma_{\max}} - \frac{1}{C^2} \right), \quad (\text{A.37})$$

где $C = L_p \pi G_{\kappa}^2 - C_1$.

Важно заметить, что приведенный выше результат справедлив только для множителя $b = \frac{\eta}{K(e^{\frac{r_0 n_0}{w}} - 1)} < 1$. В противном случае, если $b = 1$, тогда $\gamma_2 < \gamma_1$ по определению, а увеличение $b \geq 1$ ведет к $\Pr \{\gamma_2 p_i \leq N_0\} = 1$. Рассматривая состояние системы, в котором имеется n активных сессий, можно вычислить плотность размещения абонентских терминалов в пикослое $\lambda_N = n / (\pi R^2)$, как это было сделано для слоя WLAN.

Утверждение. Среднее число обслуживаемых одной малой БС сессий n_0 стремится к $\frac{n}{L_p(\pi R^2)} = \frac{n}{E[N_p]}$ для достаточно больших областей, где $E[N_p]$ – среднее число малых БС, размещенных внутри окружности радиуса R .

Доказательство проводится по аналогии с соответствующими рассуждениями для слоя WLAN.

С учетом $n_0 = \max\left(1, \frac{n}{E[N_p]}\right)$ можно получить интенсивности переходов для пикослоя.

Приложение Б

Анализ совмещенной гетерогенной сети доступа

Б.1 Подход к анализу и общие замечания

В этом приложении рассматривается одна типовая (меченая) малая сота, расположенная внутри заданной области моделирования. Предлагается аналитическая модель для учета динамики трафика обслуживаемых ей абонентов, основанная на применении теории массового обслуживания к ряду рабочих сценариев. Для наглядности последующее изложение вначале сосредоточивается на описании двух изучаемых технологий радиодоступа, LTE и WiFi, а также вводит в рассмотрение системные процессы, сформулированные в рамках математической модели, предложенной в главе 2. Далее проводится анализ сверхплотной совмещенной гетерогенной сети на примере простейшего способа размещения абонентских терминалов согласно PPP. Наконец, важное практическое расширение данного метода для смеси процессов PPP и кластерного приводится в заключительной части данного приложения.

Предлагаемый анализ совмещенной гетерогенной сети доступа имеет следующую общую структуру:

1. Предварительные сведения, описывающие динамику числа активных абонентских терминалов отдельно для каждой радиотехнологии:
 - формулирование предлагаемой математической абстракции и общего аналитического подхода на основе *укрупнения состояний* марковского процесса;
 - соответствующие вычисления отдельно для технологических особенностей WiFi и LTE (включая предлагаемую здесь концепцию виртуальных пользователей).

Последующий анализ учитывает различные подходы к управлению передачей в совмещенной гетерогенной сети.

2. Модель для базового алгоритма управления, основанного на предпочтительном использовании технологии WiFi.

3. Модель для предлагаемого алгоритма управления, основанного на одновременном использовании технологий LTE и WiFi.
4. Практическое расширение для смеси процессов PPP и кластерного, моделирующих размещение абонентских терминалов на плоскости.

Ниже приводится описание предлагаемого подхода к аналитическому моделированию совмещенной сети LTE-WiFi, основанного на системе допущений, введенной в главе 2. Он основан на вычислении стационарного распределения вероятностей для числа абонентских терминалов (активных сессий) на обслуживании, которое затем применяется для нахождения показателей эффективности системы связи в стационарном режиме, усредненных с учетом рассматриваемого способа размещения узлов.

Ассоциация абонентского терминала с ближайшей к нему СБС по расстоянию (в предположении однородной среды передачи) эквивалентна соединению с СБС, соответствующей максимальному уровню мощности принятого сигнала [346]. Геометрическое множество точек, соответствующих местоположению терминалов, ассоциированных с одной и той же СБС, составляет выпуклый многоугольник (называемый *ячейкой Вороного* [347]), который называется далее *зоной обслуживания* СБС. Иными словами, зона обслуживания задает множество точек плоскости, которые расположены ближе к рассматриваемой СБС, чем к любой другой. Ниже предполагается, что меченая СБС имеет круговую зону покрытия радиуса R , а фактическая зона обслуживания тогда представляет собой некоторый многоугольник, расположенный внутри зоны покрытия.

Б.2 Моделирование динамики системы связи

Для обеих рассматриваемых технологий радиодоступа вычисляется стационарное распределение вероятностей для текущего числа активных сессий в системе $N(t)$, а также для некоторых зависящих от него случайных величин. В исследуемой меченой малой соте динамику изменения числа передающих абонентских терминалов можно описать следующим марковским процессом. Данная типовая малая сота рассматривается во вложенных точках по оси времени, в которых новые сессии появляются в системе или же текущие сессии покидают ее. В качестве *состояния системы* принимается некоторый набор сессий (тер-

миналов) на обслуживании с учетом их индивидуальных параметров (скорость передачи данных, длина соединения с СБС и т.д.), заданных распределением местоположений соответствующих пользователей. Тем самым будущее состояние процесса полностью определено текущим набором расстояний между узлами сети.

Следовательно, состояние системы $\tilde{S}(t)$ в момент времени t может быть записано как

$$\tilde{S}(t) = (n; \xi_1, \dots, \xi_n),$$

где ξ_i – расстояние от меченой СБС до абонентского терминала i , а n – текущее число терминалов, ассоциированных с меченой СБС. Состояние системы, таким образом, включает в себя сведения относительно общего числа обслуживаемых терминалов (сессий), а также их непосредственного местоположения.

Исходный процесс $S(t)$ (см. верхнюю часть рисунка Б.1) может быть достаточно сложен для анализа. Поэтому здесь применяется укрупнение состояний [348], которое позволяет предложить адекватную аппроксимацию для «усредненного» поведения системы связи. В частности, состояния $\tilde{S}(t)$ объединяются в новое состояние $S(t) = N(t)$ (см. нижнюю часть рисунка Б.1), где $N(t)$ – общее число активных сессий в текущий момент времени. Это позволяет абстрагироваться от конкретного местоположения терминалов и их ассоциации с СБС. Это также означает, что состояние $N(t)$ включает в себя все возможные состояния исходного процесса $(N(t); \xi_1, \dots, \xi_n)$.

Упрощенная таким образом модель системы связи представляет собой процесс гибели и размножения с интенсивностями переходов a_i и b_i из состояния $N(t) = i$ в состояния $N(t) = i + 1$ и $N(t) = i - 1$ соответственно. При этом переходы между состояниями включают в себя в неявном виде информацию о распределении местоположений узлов и связаны с рассмотрением соответствующих условных вероятностей. Важно отметить, что в главе 2 производится проверка правомерности такого укрупнения состояний и подтверждается высокая точность соответствия между аналитическими и имитационными результатами. В итоге $N(t)$ представляет собой усредненный по некоторому способу размещения узлов сети процесс, который при этом не имеет «памяти» и, тем самым, может быть описан в терминах теории массового обслуживания.

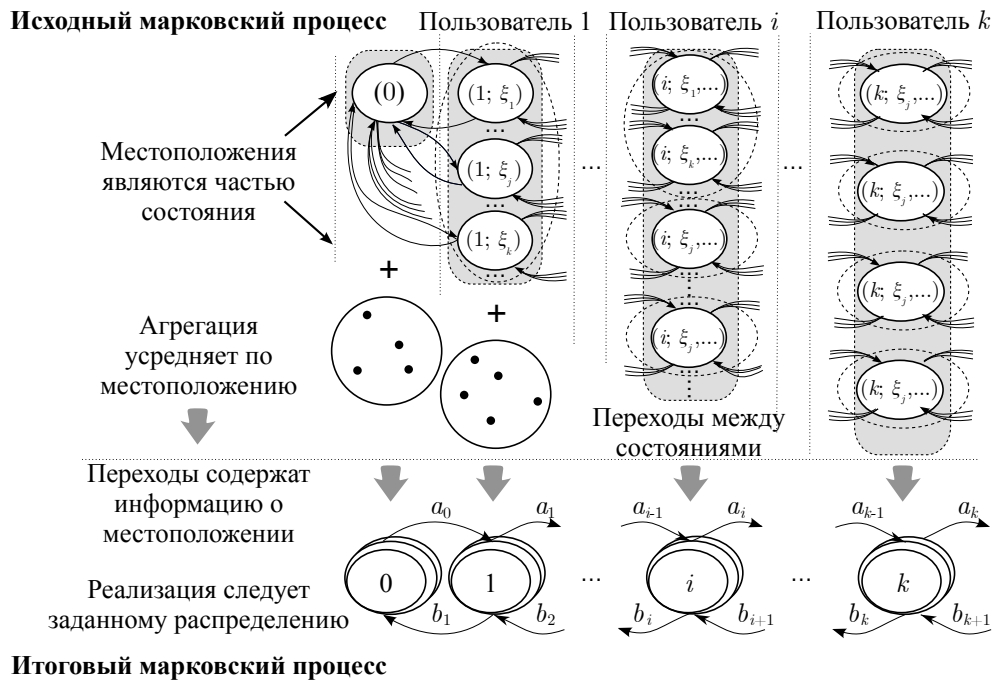


Рисунок Б.1 — Упрощение структуры марковской цепи для совмещенной гетерогенной сети доступа

Утверждение. Стационарное распределение вероятностей $\pi = \{\pi_n\}_{n=0}^{\infty}$ процесса $N(t)$ после укрупнения состояний может быть найдено как

$$\pi_n = \pi_0 \prod_{i=1}^n \frac{a_{i-1}}{b_i}, \quad (\text{Б.1})$$

где π_0 вычисляется из условий нормировки.

Приведенное утверждение позволяет получить общесистемные показатели эффективности в стационарном режиме – среднее число абонентских терминалов (сессий) на обслуживании и среднее время передачи данных. Однако данное выражение, в свою очередь, требует нахождения интенсивностей переходов a_i и b_i между соседними состояниями. При этом интенсивность переходов a_i для системы после укрупнения состояний определяется средним числом терминалов в единицу времени и на единицу площади зоны обслуживания, что может быть записано как

$$a_i = \lambda \frac{S_s}{D^2}, \quad (\text{Б.2})$$

где λ – интенсивность входного потока сессий в рамках рассматриваемой области моделирования D^2 , а S_s – площадь области, внутрь которой поступают новые сессии, обслуживаемые меченой СБС.

Определение интенсивности переходов из состояния i в состояние $i - 1$ существенно сложнее и связано с распределением фактической скорости передачи данных, т.е. с размером файла, подготовленного к отправке, и с мгновенной скоростью его передачи, заданной формулой Шеннона. Например, если i активных сессий получают одну и ту же скорость передачи данных $\frac{r}{n}$, то переход $(i) \rightarrow (i - 1)$ можно описать как $b_i = i \frac{1}{\theta i/r} = \frac{r}{\theta}$. При различных значениях мгновенной скорости передачи $r_i \neq const$ и интенсивность переходов b_i для системы после укрупнения состояний задается следующим утверждением.

Утверждение. Интенсивность переходов b_i процесса $N(t)$ после укрупнения состояний для равного разделения радиоресурсов между пользователями не зависит от состояния и может быть найдена как

$$b_i = \tilde{\delta}_{w/l} \left[\int_0^{\infty} x \left(\int_{r_R}^{r_{\lim}} \frac{r}{\theta} e^{-\frac{r}{\theta}x} f_r(r) dr - C \frac{r_{\lim}}{\theta} e^{-\frac{r_{\lim}x}{\theta}} \right) dx \right]^{-1}, \quad (\text{Б.3})$$

где $C = \int_{r_R}^{r_{\lim}} f_r(r) dr$, $\tilde{\delta}_{w/l} \leq \delta_{w/l} \leq 1$ – параметр, характеризующий долю доступных радиоресурсов, который вычисляется ниже; θ – средний размер передаваемого файла; $f_r(r)$, $r \in [r_R, r_{\lim}]$ – распределение мгновенной скорости передачи данных; r_R – минимально возможная мгновенная скорость передачи (для терминалов, расположенных на границе зоны покрытия, $w \log(1 + \frac{pG}{N_0 R^k})$) и r_{\lim} – предельно достижимая скорость передачи.

Доказательство приведено в заключительном разделе данного приложения.

Используя формулу Шеннона, распределение скоростей передачи $f_r(r)$ для обеих технологий радиодоступа можно получить в виде

$$f_r(r) = f_d(d(r)) |d'_r|,$$

где f_d – распределение расстояния между СБС и обслуживаемыми ей абонентскими терминалами, а $d(r)$ и соответствующая производная d'_r заданы как

$$d = \left(\frac{pG}{(e^{\frac{r}{w}} - 1) N_0} \right)^{\frac{1}{\kappa}}, \quad |d'_r| = \frac{1}{\kappa w} e^{\frac{r}{w}} \left(\frac{pG}{N_0} \right)^{\frac{1}{\kappa}} (e^{\frac{r}{w}} - 1)^{-\frac{1}{\kappa} - 1}.$$

Здесь p_{\max} обозначается как p для краткости, SINR можно заменить на SNR с учетом введенного в главе 2 процесса управления интерференцией. Стоит заметить, что предельно достижимая скорость передачи данных r_{\lim} связана с соответствующим расстоянием d_{\lim} и для абонентских терминалов на расстоянии

$d_{\text{lim}} = d(r_{\text{lim}})$ выполняется $w \log(1 + \text{SNR}) \geq r_{\text{lim}}$. Таким образом, распределение $f_d(d)$, $d \in [d_{\text{lim}}, R]$ описывает введенные ранее ограничения по расстоянию и скорости передачи.

Приведенные выше выражения задают стационарное распределение вероятностей и позволяют определить среднее число активных абонентских терминалов (сессий), их среднее время обслуживания и среднюю фактическую скорость передачи данных в расчете на сессию:

$$E_t[n] = \sum_{i=0}^{\infty} i\pi_i, \quad E_t[T] = \frac{E[n]D^2}{\lambda S_s}, \quad E_t[r] = E_s[r] \frac{\sum_{i=0}^{\infty} \frac{1}{i} \pi_i}{1 - \pi_0}, \quad (\text{Б.4})$$

где E_t/E_s описывает усреднение по временной и пространственной составляющим процесса соответственно, тогда как мгновенную скорость передачи данных, усредненную по способу размещения узлов сети, можно получить как

$$E_s[r] = \int_{r_R}^{r_{\text{lim}}} r f_d(d(r)) |d'_r(r)| dr + r_{\text{lim}} F_d(d_{\text{lim}}).$$

В свою очередь, для интенсивностей переходов a_i и b_i , не зависящих от состояния системы, можно преобразовать вышеприведенное выражение к виду

$$E_t[n] = \frac{a_i/b_i}{1 - a_i/b_i}, \quad E_t[r] = E_s[r] \frac{\log \frac{1}{1 - a_i/b_i}}{a_i/b_i}, \quad (\text{Б.5})$$

где $a_i/b_i < 1$ – загрузка системы в целом.

В последующих разделах приведены дополнительные способы вычисления величин a_i и b_i при рассмотрении зоны обслуживания некоторой меченой СБС, выбранной случайно.

Б.3 Основные вычисления для сети WiFi

Данный раздел посвящен изучению сети WiFi, особенности которой отражены на рисунке Б.2. Вначале рассматривается область моделирования в целом и определяется доля радиоресурсов, доступных в некоторой малой соте, а затем исследуется функционирование меченой СБС. Важно заметить, что поведение системы WiFi можно в точности описать в рамках предложенного выше подхода

к моделированию, если общий объем радиоресурсов, доступных в меченой малой соте, известен. Для учета работы функции DCF, регламентирующей доступ к каналу связи, предполагается, что каждое активное соединение WiFi «порождает» две перекрывающиеся окружности, заданные соответствующими зонами покрытия источника и получателя некоторой сессии (для передачи данных и подтверждений об их доставке соответственно).

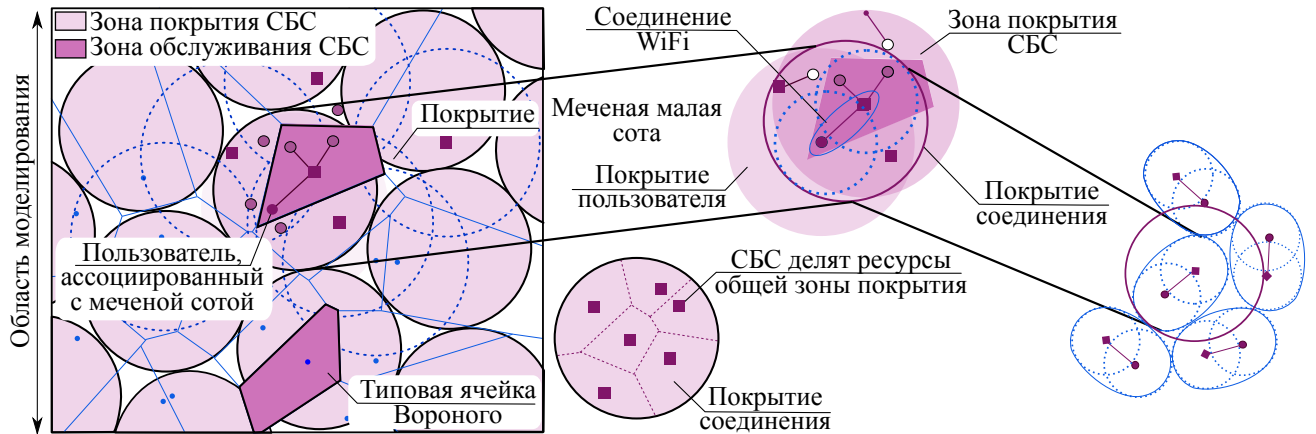


Рисунок Б.2 — Диаграмма Вороного и пример работы сети WiFi: область моделирования (слева) и меченая малая сота, содержащая зону покрытия СБС (справа)

В течение времени существования данного соединения WiFi другие абонентские терминалы или СБС не могут передавать свои данные в исключенной таким образом области. Геометрически это соответствует двум накладывающимся друг на друга кругам меньшего радиуса $R/2$, составляющим так называемую *зону покрытия соединения*. При этом зоны покрытия различных соединений в любой момент времени не должны перекрываться. Далее в качестве опорного параметра вычисляется среднее расстояние до ближайшей СБС, а затем зона покрытия соединения для простоты заменяется кругом площади \bar{S} .

В силу допущения, что все СБС размещены и/или контролируются одним и тем же мобильным оператором, соответствующие соединения WiFi могут быть координированы согласно некоторой общей процедуре планирования («выкалывания»), применяющейся после отработки протокола случайного множественного доступа [88]. Это означает, что зоны покрытия соединений WiFi в пределе составляют плотную упаковку области моделирования. Тогда число од-

новременно активных соединений можно получить на основе границы плотной упаковки с коэффициентом $\frac{\pi}{2\sqrt{3}}$.

Утверждение. Доля радиоресурсов, доступных на некоторой СБС, в сверхплотной сети WiFi может быть найдена как

$$\tilde{\delta}_w = \delta_w \frac{\pi}{2\sqrt{3}\mu\bar{S}}, \quad (\text{Б.6})$$

где δ_w – фактически доступные радиоресурсы за исключением накладных расходов в виде управляющих сообщений, а площадь зоны покрытия соединения S может быть вычислена как показано в заключительном разделе данного приложения, т.е. $\frac{1}{2}\pi R_2^2 - 2\bar{s}$.

Доказательство. Среднее число СБС, использующих одни и те же радиоресурсы одновременно, представляет собой произведение их общего числа в рассматриваемой области моделирования μD^2 на число зон покрытия соединений

$$\frac{\mu D^2}{\frac{\pi}{2\sqrt{3}}D^2/\bar{S}} = \frac{2\sqrt{3}\mu\bar{S}}{\pi}, \quad (\text{Б.7})$$

которые делят между собой общие радиоресурсы сети WiFi во времени. Доля радиоресурсов, доступных на одной СБС, является величиной, обратной данному выражению, умноженному на заданный параметр δ_w .

Распределение расстояния между СБС и равномерно размещенными внутри зоны обслуживания (заданной многоугольником) абонентскими терминалами выводится из известного распределения, которое следует дополнительно ограничить с учетом радиуса зоны покрытия R и расстояния d_{lim} , соответствующего предельной скорости передачи:

$$f_d(d) = \begin{cases} C_1 \cdot 2\pi\mu d e^{-\pi\mu d^2}, & d \in (d_{\text{lim}}, R], \\ C_1 \cdot (1 - e^{-\pi\mu d_{\text{lim}}^2}), & d = d_{\text{lim}}, \end{cases} \quad (\text{Б.8})$$

где $C_1 = \left(1 - e^{-\pi\mu R^2}\right)^{-1}$ – коэффициент нормировки. Основываясь на данном выражении, распределение скорости передачи данных можно получить с помощью

$$f_r(r) = f_d(d(r)) |d'_r|, \quad (\text{Б.9})$$

где компоненты $d(r)$ и $d'(r)$ были вычислены выше.

В итоге интенсивность переходов b_i выражается из распределения скорости передачи $f_r(r)$. При этом величина a_i с учетом параметра S_s может быть

записана как

$$a_i = \frac{\lambda}{\mu D^2}. \quad (\text{Б.10})$$

Б.4 Основные вычисления для сети LTE

Ниже исследуется сеть LTE согласно описанию на рисунке Б.3 и вычисляется стационарное распределение вероятностей для величины $N(t)$. Для этого определяются интенсивности переходов между состояниями рассматриваемой марковской цепи путем изучения меченой ячейки Вороного по аналогии с рассуждениями для системы WiFi. В отличие от последней здесь учитывается, что все абонентские терминалы, находящиеся внутри зоны покрытия данной СБС, но ассоциированные с другими (соседними) СБС, могут вызывать интерференцию. Также предполагается, что характеристики таких интерферирующих терминалов известны меченой СБС, которая исключает используемые ими радиоресурсы из своего процесса планирования.

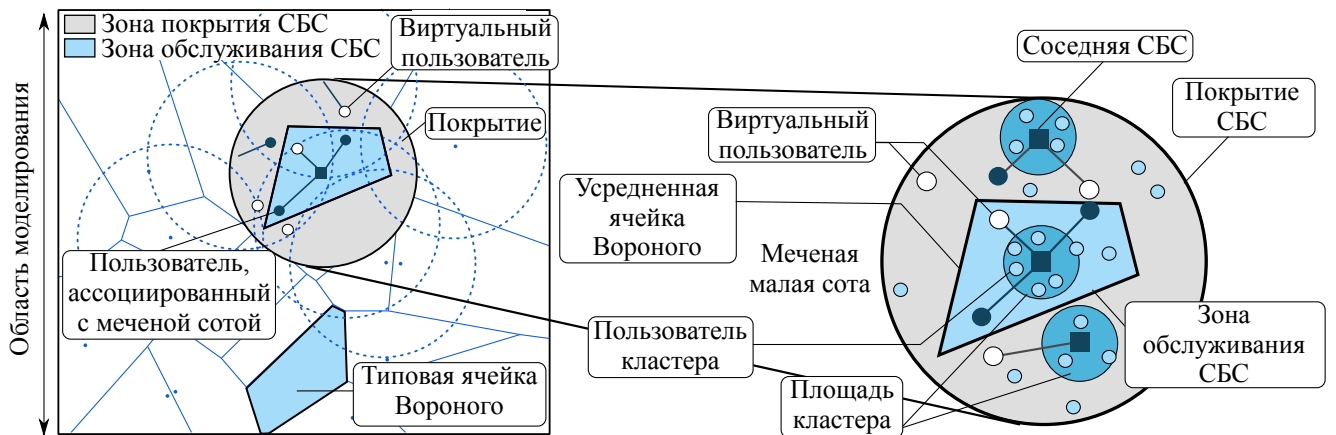


Рисунок Б.3 — Диаграмма Вороного и пример работы сети LTE: область моделирования (слева) и меченая малая сота, содержащая зону покрытия СБС с виртуальными пользователями (справа)

Учет каждого такого интерферирующего абонентского терминала, обслуживаемого соседней малой сотой, равносильно рассмотрению так называемого *виртуального* пользователя, принадлежащего меченой соте. Виртуальный терминал потребляет радиоресурсы СБС в соответствии со своим местоположением по отношению к ней, но при этом получает нулевую скорость передачи дан-

ных. Геометрически это эквивалентно «отражению» терминалов, находящихся внутри зоны покрытия, но вне зоны обслуживания, внутрь зоны обслуживания относительно ее границы. Далее, в отличие от сети WiFi в системе LTE нет необходимости разделять в явном виде общие радиоресурсы на множества, используемые отдельными СБС. Здесь радиоресурсы делятся между фактически обслуживаемыми данной СБС абонентскими терминалами и ее виртуальными терминалами. Следующее утверждение характеризует долю радиоресурсов, доступных на некоторой СБС.

Утверждение. Доля радиоресурсов, доступных на некоторой СБС в сверхплотной сети LTE, задана значением $\tilde{\delta}_l = \delta_l$. При этом радиоресурсы поделены между фактически обслуживаемыми и виртуальными терминалами данной СБС.

Далее вычисляется распределение расстояния от СБС до абонентских терминалов внутри зоны покрытия как в границах многоугольника, задающего зону обслуживания, так и вне их. При этом распределение расстояния от СБС до всех терминалов, для которых данная СБС является ближайшей, идентично аналогичному распределению в сети WiFi. Следует отметить, что с учетом предположения о способе управления излучаемой мощностью, поддерживающем заданный уровень SNR, предельно достижимая скорость передачи данных $r_{tar} \leq r_{lim}$ определяется целевым значением SNR η и остается постоянной. Кроме того, важно заметить, что скорость r_{tar} , так же как и расстояние d_{tar} , связана с заданным уровнем SNR η . Для простоты всем терминалам с таким целевым значением SNR ставится в соответствие расстояние d_{tar} . Как следствие, распределение расстояния до обеих групп абонентских терминалов (внутри зоны обслуживания и вне ее)

$$f_u(d) = \begin{cases} C_1 \cdot 2\pi\mu d e^{-\pi\mu d^2}, & d \in (d_{tar}, R], \\ C_1 \cdot (1 - e^{-\pi\mu d_{tar}^2}), & d = d_{tar}, \end{cases} \quad (\text{Б.11})$$

где $C_1 = (1 - e^{-\pi\mu R^2})^{-1}$.

Распределение скорости передачи данных $f_r(r)$ служит для нахождения интенсивности переходов b_i , тогда как интенсивность переходов a_i зависит от параметра S_s . С учетом виртуальных терминалов S_s определяется зоной покрытия πR^2 и тогда

$$a_i = \lambda \frac{\pi R^2}{D^2}. \quad (\text{Б.12})$$

В отличие от моделирования системы WiFi анализ системы LTE производится с помощью двумерного марковского процесса с пространством состояний $\{n_l, n_p, n_l = \overline{0, \infty}, n_p = \overline{0, \infty}\}$, где n_l и n_p – число фактически обслуживаемых и виртуальных терминалов соответственно. Данный процесс (см. рисунок Б.4) описывается соответствующей системой уравнений баланса. Для произвольного состояния (i, j) , где $0 < i < \infty$ и $0 < j < \infty$, можно записать

$$\begin{aligned} p_{i,j} & \left(aq_1 + a(1-q_1)(1-\nu) + \frac{i}{i+j}x + \frac{j}{i+j}x \right) \\ & = p_{i-1,j}aq_1 + p_{i,j-1}a(1-q_1)(1-\nu) + \\ & \quad + p_{i+1,j}\frac{i+1}{i+j+1} + p_{i,j+1}\frac{j+1}{i+j+1}, \end{aligned} \quad (\text{Б.13})$$

где ν – доля радиоресурсов, которые еще не были исключены из процесса планирования. Другими словами, рассматривается доля соседних абонентских терминалов, которые не становятся дополнительными виртуальными терминалами, занимая свободные радиоресурсы СБС. Скорость передачи данных как для обслуживаемых, так и для виртуальных терминалов пропорциональна доле $\frac{1}{n_l+n_p}$, поскольку они делят между собой общие радиоресурсы.

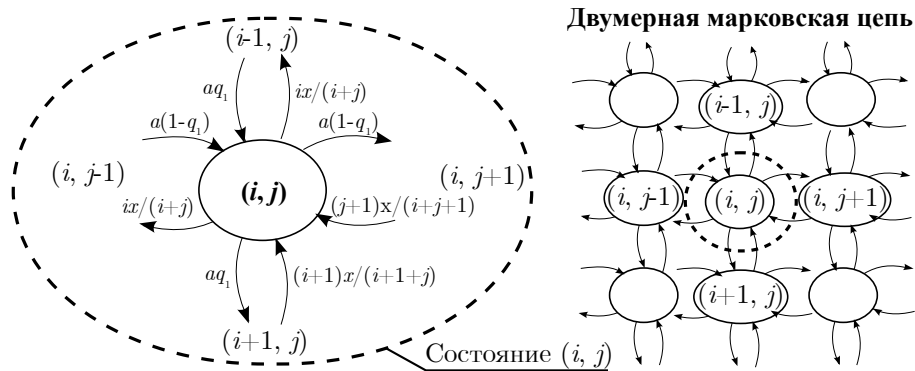


Рисунок Б.4 — Пример состояния двумерного марковского процесса для системы LTE

С учетом вышеизложенного искомое решение может быть получено в виде

$$\begin{aligned} p_{i,j} & = p_{0,0} \frac{a^i q_1^i}{x^i} \frac{\prod_{m=1}^{i+j} m}{\prod_{m=1}^i m \prod_{k=1}^j k} \frac{a^j (1-q_1)^j (1-\nu)^j}{x^j} \\ & = p_{0,0} \frac{(i+j)!}{i!j!} \frac{a^i q_1^i}{x^i} \frac{a^j (1-q_1)^j (1-\nu)^j}{x^j}, \end{aligned} \quad (\text{Б.14})$$

где $p_{0,0}$ вычисляется из условий нормировки $\sum p_{i,j} = 1$, а q_1 представляет собой вероятность оказаться внутри зоны покрытия, но вне зоны обслуживания.

Параметр a соответствует интенсивности a_i , тогда как x характеризует процесс обслуживания системой LTE i абонентских терминалов, использующих общее множество радиоресурсов. Последний параметр задается выражением, приведенным в заключительном разделе данного приложения, как $x = \frac{\tilde{\delta}_l}{E\left[\frac{s_j}{r_j}\right]}$.

Таким образом, среднее число обслуживаемых терминалов (сессий) составляет

$$\begin{aligned}
 \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} i p_{i,j} &= p_{0,0} \sum_{i=1}^{\infty} i \frac{a^i q_1^i}{x^i} \sum_{j=0}^{\infty} \frac{(i+j)! a^j (1-q_1)^j (1-\nu)^j}{i! j! x^j} = \\
 &= p_{0,0} \frac{a q_1}{x} \sum_{n=0}^{\infty} n \frac{(1-q_1)^n (1-\nu)^n a^n}{x^n} \sum_{i=0}^n \frac{n!}{i!(n-i)!} \frac{q_1^i}{(1-q_1)^i (1-\nu)^i} = \\
 &= p_{0,0} \frac{a q_1}{x} \sum_{n=0}^{\infty} n \frac{(1-q_1)^n (1-\nu)^n a^n}{x^n} \frac{((1-q_1)(1-\nu)+q_1)^n}{(1-q_1)^n (1-\nu)^n} = \\
 &= \frac{p_{0,0} \cdot a q_1}{x} \sum_{n=0}^{\infty} n \frac{a^n ((1-q_1)(1-\nu)+q_1)^n}{x^n} = \frac{p_{0,0} \cdot a q_1}{x(1-a \frac{1-\nu+q_1\nu}{x})}. \tag{B.15}
 \end{aligned}$$

Среднее число виртуальных терминалов, а также общее число терминалов в системе можно определить схожим образом. Однако в случае, когда работа сетей LTE и WiFi имеет явную зависимость (как, например, при рассмотрении одновременного использования обеих радиотехнологий), можно дополнительно упростить анализ системы LTE, заменяя рассмотрение процессов внутри зоны покрытия на эквивалентные процессы внутри зоны обслуживания. Таким образом, исследование различных алгоритмов управления потоками трафика можно проводить по аналогии, рассматривая только лишь зону обслуживания. **Утверждение.** Двумерный марковский процесс, моделирующий динамику фактически обслуживаемых и виртуальных терминалов, эквивалентен одномерному процессу с модифицированными интенсивностями переходов $a q_1$ и $x - a(1 - q_1)(1 - \nu)$.

Доказательство. Рассматривая одномерный процесс гибели и размножения с пространством состояний $\{n_l, n_l = \overline{0, \infty}\}$, можно описать динамику изменения числа абонентских терминалов, обслуживаемых меченой СБС в системе LTE. Интенсивность обслуживания таких терминалов пропорциональна площади усредненной зоны обслуживания и составляет $a q_1$, что характеризует интенсивность переходов из состояния $i - 1$ в состояние i . Тогда стационарное распределение вероятностей для рассматриваемого процесса может быть получено в виде

$$\pi_n = \pi_0 \prod_{i=1}^n \frac{a q_1}{b_i}, \tag{B.16}$$

где b_i – интенсивность переходов между состояниями i и $i - 1$; π_0 возможно вычислить из условий нормировки.

Для нахождения интенсивности переходов b_i следует преобразовать стационарное распределение вероятностей, основываясь на исходном двумерном процессе:

$$\begin{aligned}
 \pi_i &= \left(1 - \frac{a}{x}\right) \sum_{j=0}^{\infty} \frac{(i+j)!}{i!j!} \frac{a^i q_1^i}{x^i} \frac{a^j (1-q_1)^j (1-\nu)^j}{x^j} = \\
 &= \left(1 - \frac{a}{x}\right) \sum_{n=i}^{\infty} \left(\frac{q_1}{(1-q_1)(1-\nu)}\right)^i \frac{n!}{i!(n-i)!} \left(\frac{a^n (1-q_1)^n (1-\nu)^n}{x^n}\right) = \\
 &= \left(1 - \frac{a}{x}\right) \left(\frac{q_1}{(1-q_1)(1-\nu)}\right)^i \frac{\left(\frac{a(1-q_1)(1-\nu)}{x}\right)^i}{\left(1 - \frac{a(1-q_1)(1-\nu)}{x}\right)^{i+1}} = \\
 &= \frac{x-a}{x-a(1-q_1)(1-\nu)} \frac{(aq_1)^i}{(x-a(1-q_1)(1-\nu))^i}. \tag{B.17}
 \end{aligned}$$

Сравнивая вышеприведенные выражения, можно заключить, что для замены двумерного процесса на упрощенный одномерный значение b_i необходимо выбирать как $x - a(1 - q_1)(1 - \nu)$, что завершает доказательство.

В итоге можно сделать вывод, что относительно *средних значений* рассмотренный выше более сложный процесс эквивалентен процессу с пропорционально сниженными интенсивностями поступления и обслуживания. В частности, легко получить выражение для $E[n_i]$. Последующие разделы посвящены изучению динамики числа абонентских терминалов внутри типовой зоны обслуживания.

Б.5 Базовый алгоритм: предпочтительное использование технологии WiFi

В этом разделе полученные ранее результаты объединяются для изучения работы базового метода управления потоками трафика в совмещенной системе связи LTE-WiFi. В качестве примера типового подхода к выбору технологии радиодоступа рассматривается алгоритм предпочтительного использования одной из радиотехнологий. Без потери общности такой технологией здесь выбирается WiFi. Абонентский терминал использует ее с приоритетом, когда его текущее значение SNR превышает некоторое пороговое значение η_0 , а в противном случае передает свои данные по технологии LTE [53]. Данный способ управления отражает стремление мобильного оператора автоматизировать процесс выбора

технологии радиодоступа, которому на сегодняшний день следует большинство абонентов, а также сбалансировать нагрузку обеих сетей доступа, LTE и WiFi.

Порог SNR для переключения η_0 задает внутреннюю область радиуса $d_0 = \min \left[\left(\frac{p_w G}{\eta_0 N_w} \right)^{1/\kappa_w}, R_w \right]$ внутри зоны обслуживания, где абонентские терминалы используют исключительно технологию WiFi. Далее по тексту индекс « $.w$ » отмечает параметры, относящиеся к работе сети WiFi, а соответствующее значение максимального уровня излучаемой мощности для краткости записывается как p_w . Вероятность того, что абонентский терминал находится внутри зоны обслуживания WiFi, может быть при этом вычислена как

$$q = \Pr\{d < d_0\} = C_1(1 - e^{-\pi\mu d_0^2}), \quad C_1 = \frac{1}{1 - e^{-\pi\mu R^2}}. \quad (\text{Б.18})$$

Функционирование базового алгоритма управления в рамках одной меченой малой соты включает в себя состояния (n_w, n_l, n_p) , где n_w , n_l и n_p представляют собой число абонентских терминалов, использующих сети WiFi и LTE соответственно, а также число виртуальных терминалов, фактически обслуживаемых соседними малыми сотами LTE. В рассматриваемом случае процессы обслуживания в сетях WiFi и LTE являются независимыми, и тогда можно рассматривать динамику работы системы связи отдельно для WiFi и LTE подсистем.

С одной стороны, динамика функционирования системы WiFi описывается пространством состояний $\{n, n = \overline{0, \infty}\}$. Соответствующее распределение расстояния должно в этом случае учитывать предельную длину соединения d_0 . Однако R требуется заменить здесь на d_0 , тогда как $C_1 = (1 - e^{-\pi\mu d_0^2})$. С другой стороны, распределение расстояния $f_d(d)$ влияет на среднюю площадь зоны покрытия соединения как

$$\begin{aligned} \bar{S} &= 2\pi \left(\frac{d_0}{2}\right)^2 - \int_0^{\frac{d_0}{2}} s(x) f_d(x) dx \\ &= \frac{1}{2}\pi d_0^2 - C_1 \int_0^{d_0} s(x) \cdot 2\pi\mu x e^{-\pi\mu x^2} dx, \end{aligned} \quad (\text{Б.19})$$

где $s(x)$ вычисляется в заключительном разделе данного приложения. С учетом просеянного таким образом пуассоновского процесса поступлений, параметр a_i преобразуется к виду

$$a_i = q \frac{\lambda}{\mu D^2}. \quad (\text{Б.20})$$

По аналогии динамика функционирования системы LTE описывается на основе распределения расстояния, которое ограничено величиной d_0 , в виде

$$f_d(d) = \begin{cases} C_1 \cdot 2\pi\mu d e^{-\pi\mu d^2}, & d \in (d_m, R], \\ \max\left(C_1(e^{-\pi\mu d_0^2} - e^{-\pi\mu d_{tar}^2}), 0\right), & d = d_m, \end{cases} \quad (\text{Б.21})$$

где $d_m = \max(d_{tar}, d_0)$ и $C_1 = \left(e^{-\pi\mu d_0^2} - e^{-\pi\mu R^2}\right)^{-1}$.

Основываясь на приведенных выше выражениях, можно заменить исходный двумерный процесс на эквивалентный одномерный с интенсивностью переходов $a_i = q_1 a$, где a – общая интенсивность входного потока в систему LTE, составляющая

$$a = (1 - q)\lambda \frac{\pi R^2}{D^2}, \quad a_i = (1 - q)\lambda \frac{1/\mu}{D^2}. \quad (\text{Б.22})$$

Интенсивность переходов b_i должна быть выбрана так, чтобы $x - a(1 - q_1)(1 - \nu)$, и можно было бы подобным образом заменить исходный двумерный процесс на упрощенный одномерный. Важно отметить, что параметр q_1 требуется модифицировать с учетом распределения расстояния, а также выражения $E\left[\frac{s_j}{r_j}\right]$, зависящего от распределения скорости передачи данных. Последнее, в свою очередь, определяется через распределение расстояния, которое ограничено как $d > d_0$.

Б.6 Продвинутый алгоритм: одновременное использование технологий LTE и WiFi

В этом разделе приведен анализ предлагаемого метода управления потоками трафика в совмещенной гетерогенной сети доступа, который предполагает одновременное использование радиотехнологий LTE и WiFi. Такой подход к управлению обусловлен недавним развитием совмещенных LTE-WiFi модулей, а соответствующее обоснование его преимущества приведено в главе 2. Основываясь на приведенных ранее вычислениях, можно описать работу рассматриваемой системы связи и получить ее усредненные показатели эффективности. Для этого далее предполагается, что процессы, происходящие во всех соседних малых сотах при их сверхплотном размещении, являются статистически эквивалентными динамике некоторой меченой соты, выбранной произвольно. Соответственно, виртуальные терминалы покидают систему с той же интенсивностью,

что и фактически обслуженные терминалы. Эта интенсивность определяется агрегированием соответствующих интенсивностей для технологий WiFi и LTE.

Для математического описания системы следует рассмотреть двумерный процесс согласно пояснениям на рисунке Б.4, а также провести его замену на одномерный (по аналогии с подходом из предыдущего раздела) и переопределить значения параметров соответственно. Другими словами, процесс, имеющий пространство состояний $\{n, n_p, n = \overline{0, \infty}, n_p = \overline{0, \infty}\}$, где n – текущее число обслуживаемых абонентских терминалов и n_p – текущее число виртуальных терминалов, заменяется на процесс $\{n, n = \overline{0, \infty}\}$. Таким образом, интенсивности переходов с учетом соответствующей зоны обслуживания вычисляются согласно

$$a_i = \lambda \frac{1/\mu}{D^2}, \quad b_i = x - a(1 - q_1)(1 - \nu), \quad (\text{Б.23})$$

где x задается в виде

$$x = \frac{i}{E \left[\frac{s_j}{\frac{\tilde{\delta}_w r_w}{i} + \frac{\tilde{\delta}_l r_l}{i}} \right]} = \frac{1}{E \left[\frac{s_j}{\tilde{\delta}_w r_w + \tilde{\delta}_l r_l} \right]}. \quad (\text{Б.24})$$

Приведенное выше выражение можно переписать, основываясь на распределении расстояния

$$f_d(d) = \begin{cases} C_1 \cdot 2\pi\mu d e^{-\pi\mu d^2}, & d \in (d_{\text{lim}}, R], \\ \max(C_1 \cdot (1 - e^{-\pi\mu d_{\text{lim}}^2}), 0), & d = d_{\text{lim}}, \end{cases} \quad (\text{Б.25})$$

где $C_1 = (1 - e^{-\pi\mu R^2})^{-1}$.

Поскольку рассматриваемый алгоритм управления применяется на всех СБС сети доступа, далее можно следовать подходу с усреднением по аналогии с проведенным выше анализом базового алгоритма и таким образом пересчитать долю q_1 как

$$q_1 = \frac{S_{lte}}{S_{all}} = \frac{1/\mu - S_{wifi}}{\pi R_{lte}^2 - (\mu\pi R_{lte}^2) S_{wifi}} = \frac{1}{\mu\pi R_{lte}^2}. \quad (\text{Б.26})$$

Б.7 Важное практическое обобщение

Пример фактического назначения радиоресурсов для абонентских терминалов и СБС, размещенных равномерно, показан на рисунках Б.5 и Б.6. В качестве заключительного шага рассмотрим возможность учета смеси процессов

PPP и кластерного для моделирования поступления новых сессий (абонентских терминалов) в совмещенную сеть доступа. Следует напомнить, что терминалы в кластерах группируются внутри окружности r , что должно учитываться при вычислении распределения $f_d(d)$ с плотностью $f_c(d) = \frac{2d}{r^2}$ и коэффициентом $(1 - \varepsilon)$.

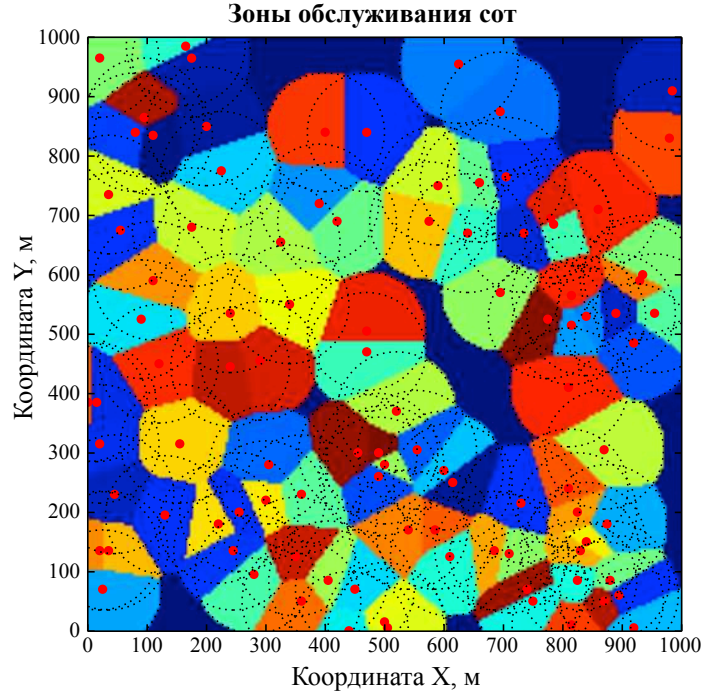


Рисунок Б.5 — Пример назначения радиоресурсов в совмещенной сети: зоны обслуживания

Утверждение. Для рассматриваемой смеси процесса PPP и кластерного итоговое распределение расстояния задается следующей кусочной функцией:

$$f_d(d) = \begin{cases} C_1 \varepsilon \cdot 2\pi\mu d e^{-\pi\mu d^2}, & r < d < R, \\ C_1 \varepsilon \cdot 2\pi\mu d e^{-\pi\mu d^2} + (1 - \varepsilon) \frac{2d}{r^2}, & d_{\text{lim}} < d \leq r, \\ C_1 \varepsilon \cdot (1 - e^{-\pi\mu d_{\text{lim}}^2}) + (1 - \varepsilon) \frac{d_{\text{lim}}^2}{r^2}, & d = d_{\text{lim}}, \end{cases} \quad (\text{Б.27})$$

где учитывается расстояние d_{lim} , соответствующее предельно достижимой скорости передачи данных r_{lim} , а величина $C_1 = (1 - e^{-\pi\mu R^2})^{-1}$ задает ограничение расстояния в соответствии с зоной покрытия R .

Кроме того, предлагаемый здесь подход позволяет вычислить вероятность блокировки при переходе из состояния $i - 1$ в состояние i как для базового, так и для предлагаемого алгоритма управления потоками трафика. При этом могут быть также учтены важные частные случаи, когда переходы зависят от двух

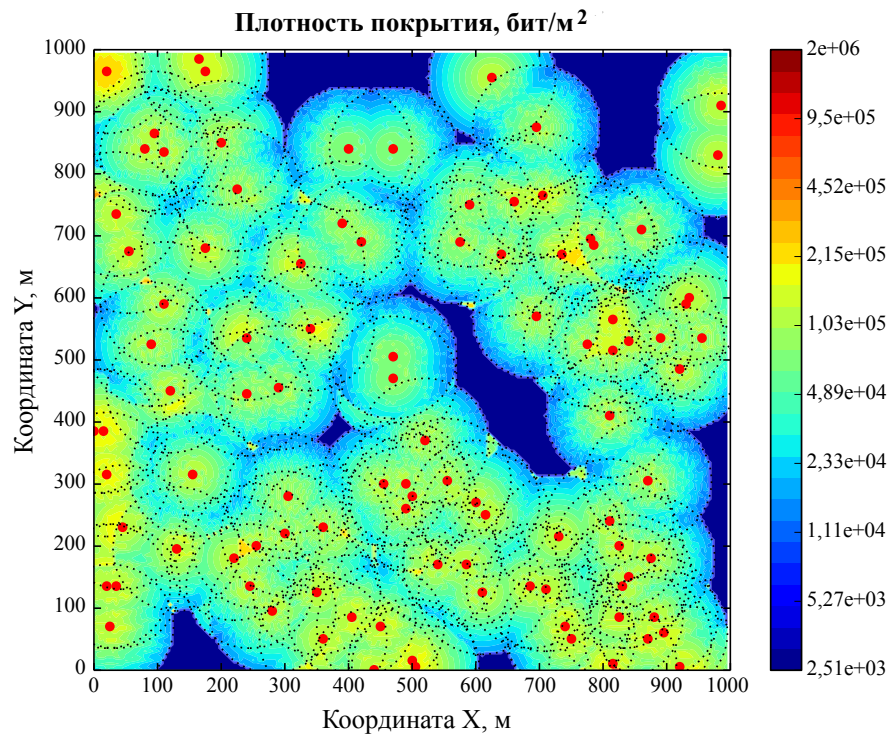


Рисунок Б.6 — Пример назначения радиоресурсов в совмещенной сети:
плотность покрытия

состояний, i и j . Подобные обобщения будут полезны, если мобильный оператор принимает решение отказать абоненту в обслуживании, когда его предполагаемая скорость передачи данных может оказаться ниже некоторого заданного порога, задающего минимальные требования к качеству пользовательского соединения [349].

Б.8 Вспомогательные вычисления

Б.8.1 Интенсивность переходов b_i для процесса $N(t)$ после укрупнения состояний

В силу предположения о равном разделении радиоресурсов, когда фактическая скорость передачи данных принимает значение $\tilde{r}_j = \frac{\delta}{i} r_j$, интенсивность

переходов b_i может быть получена как

$$b_i = \frac{i}{E \left[\frac{s_j}{r_j} \right]} = \frac{1}{E \left[\frac{s_j}{\delta r_j} \right]} = \frac{\delta}{E \left[\frac{s_j}{r_j} \right]}.$$

Далее вычисляется распределение случайной величины $x = s_j/r_j$, где r_j – предельно достижимая скорость передачи, а s_j – экспоненциально распределенная случайная величина, задающая размер передаваемого файла. Тогда распределение параметра x ($x > 0$) может быть найдено как

$$F_x(x) = \Pr \left\{ \frac{s_j}{r_j} < x \right\} = \Pr \{s_j < x r_j\}.$$

Таким образом, с учетом независимости величин r и s , а также границ $r \in [0, r_{\text{lim}}]$, $s \in [0, \infty]$ возможно записать вероятность $\Pr \{s < x r\}$ в виде

$$\begin{aligned} \Pr \{s < x r\} &= \int_0^\infty \frac{1}{\theta} e^{-\frac{1}{\theta} s} \left(\int_{\max(\frac{s}{x}, r_R)}^{r_{\text{lim}}} f_r(r) dr + F_r(r_{\text{lim}}) \right) ds = \\ &= \int_0^\infty \frac{1}{\theta} e^{-\frac{1}{\theta} s} \left(1 - \int_{r_R}^{\max(\frac{s}{x}, r_R)} f_r(r) dr \right) ds = \\ &= 1 - \int_0^{x r_{\text{lim}}} \frac{1}{\theta} e^{-\frac{1}{\theta} s} \int_{r_R}^{\max(\frac{s}{x}, r_R)} f_r(r) dr ds = \\ &= 1 - \int_0^{x r_{\text{lim}}} \frac{1}{\theta} e^{-\frac{1}{\theta} s} F_r(\max(\frac{s}{x}, r_R)) ds = \\ &= 1 - \int_{x r_R}^{x r_{\text{lim}}} \frac{1}{\theta} e^{-\frac{1}{\theta} s} F_r(\frac{s}{x}) ds = 1 - \int_{r_R}^{r_{\text{lim}}} \frac{x}{\theta} e^{-\frac{x}{\theta} r} F_r(r) dr = \\ &= 1 + e^{-\frac{x}{\theta} r_{\text{lim}}} \int_{r_R}^{r_{\text{lim}}} f_r(r) dr - \int_{r_R}^{r_{\text{lim}}} e^{-\frac{x}{\theta} r} f_r(r) dr, \end{aligned}$$

где $r_R = w \log \left(1 + \frac{p_{\text{max}}}{(N_0 + I)} \frac{G}{R^\kappa} \right)$ – фактическое минимальное значение предельно достижимой скорости передачи данных, соответствующее нахождению абонентского терминала на расстоянии R , а функция распределения для скорости передачи имеет такую структуру, что $F_r(r_R) = 0$ и $\max r = r_{\text{lim}}$.

Соответственно, основываясь на вероятности $\Pr \{s < x r\}$, плотность вероятности для случайной величины x можно задать как

$$f_x(x) = \frac{dF(x)}{dx} = \int_{r_R}^{r_{\text{lim}}} \frac{r}{\theta} e^{-\frac{r}{\theta} x} f_r(r) dr - C \frac{r_{\text{lim}}}{\theta} e^{-\frac{r_{\text{lim}} x}{\theta}}, \quad (\text{Б.28})$$

где $C = \int_{r_R}^{r_{\text{lim}}} f_r(r) dr$.

Из данного выражения следует математическое ожидание случайной величины s_j/r_j и, тем самым,

$$E \left[\frac{s_k}{r_k} \right] = \int_0^{r_{\text{lim}}} x \left(\int_{r_R}^{r_{\text{lim}}} \frac{r}{\theta} e^{-\frac{r}{\theta}x} f_r(r) dr - C \frac{r_{\text{lim}}}{\theta} e^{-\frac{r_{\text{lim}}x}{\theta}} \right) dx, \quad (\text{Б.29})$$

что полностью определяется распределением скорости передачи $f_r(r)$.

В итоге для предлагаемого алгоритма управления потоками трафика интенсивность переходов b_i с учетом фактической зоны обслуживания составляет

$$b_i = \frac{i}{E \left[\frac{s_j}{\frac{\bar{\delta}_w r_w}{i} + \frac{\bar{\delta}_l r_l}{i}} \right]} = \frac{1}{E \left[\frac{s_j}{\bar{\delta}_w r_w + \bar{\delta}_l r_l} \right]}. \quad (\text{Б.30})$$

Б.8.2 Средняя площадь пересечения накладывающихся зон покрытия

Рассматривая две накладывающиеся друг на друга окружности радиуса R_2 , соответствующие зонам покрытия согласно описанию выше, можно вычислить половинную площадь их перекрытия, которая задает объем радиоресурсов, исключенных из процесса планирования

$$s(x) = R_2^2 \arccos \frac{x}{2R_2} - \frac{1}{2} x r \sqrt{1 - \frac{x^2}{4R_2^2}}, \quad (\text{Б.31})$$

где x – расстояние между центрами зон покрытия. Интегрирование данного выражения по распределению расстояний $f_d(d)$ приводит к

$$\bar{s} = \frac{1}{2} \int_0^{2R_2} s(x) f_d(x) dx. \quad (\text{Б.32})$$

Таким образом, средняя площадь пересечения накладывающихся зон покрытия равна

$$2\pi R_2^2 - 2\bar{s} = 2\pi R_2^2 - \int_0^{2R_2} S(x) f_d(x) dx, \quad (\text{Б.33})$$

где R_2 – радиус зоны покрытия.

В результате данное приложение предлагает модель совмещенной гетерогенной сети LTE-WiFi, которая позволяет провести анализ ее уплотнения и учесть динамику ее загрузки. Она представляет собой гибкий инструмент для анализа достижимых пределов уплотнения гетерогенной сети.

Приложение В

Анализ интегрированных сотовой системы и сети D2D

В.1 Моделирование сети D2D

В данном приложении приводятся особенности анализа сети D2D, интегрированной с сотовой системой связи, и предлагается способ вычисления характеристик ее работы. Наиболее важные результаты оформляются в виде утверждений.

В.1.1 Вероятностная модель

Далее рассмотрим предлагаемый подход к анализу системы связи на примере слоя D2D и обсудим его применимость. Соответственно, система D2D рассматривается в некоторые моменты времени t , связанные с поступлением в нее новых сессий (абонентов) и их уходом по завершении обслуживания. Поведение такой модели описывается при помощи марковского процесса $S(t)$, будущее состояние которого определяется только лишь набором обслуживаемых сетью пользователей. С учетом того, что текущее состояние процесса зависит от набора активных сессий, он представлен в виде вектора с переменным числом элементов (см. пространство состояний процесса на рисунке В.1). Для удобства обозначим через $(\omega_1, \dots, \omega_n)$ вектор характеристик, связанных с набором обслуживаемых сессий размера n .

Пусть Q_{n+1} обозначает вероятность события, что вновь поступившей пользовательской сессии отказано в обслуживании системой D2D при условии, что в ней уже имеется n активных сессий. Тогда переходы из состояния $(n; \omega_1, \dots, \omega_n)$ в состояние $(n + 1; \omega_1, \dots, \omega_n, \omega_{n+1})$ и обратно происходят со следующими интенсивностями:

$$\lambda(1 - Q_{n+1}) \text{ и } (n + 1)\mu. \quad (\text{В.1})$$

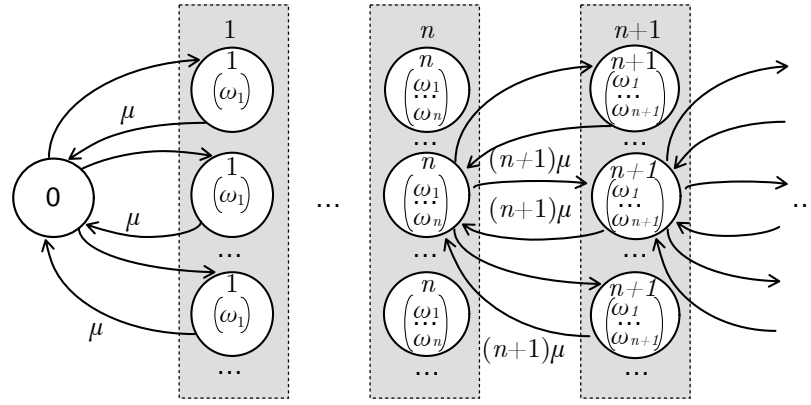


Рисунок В.1 — Пространство состояний системы D2D

В.1.2 Стационарное распределение вероятностей

Ввиду потенциально неограниченного числа состояний рассматриваемой системы связи, вычисление стационарного распределения вероятностей для нее может быть сложным. Однако соответствующий марковский процесс можно существенно упростить за счет использования подхода *укрупнения состояний*. Таким образом, состояния $\{(n; \omega_1, \dots, \omega_n)\}_{\omega \in \Omega}$ объединяются по n , где Ω представляет собой пространство всех возможных векторов $(\omega_1, \dots, \omega_n)$, $n \in N$. Кроме того, вводится следующее утверждение.

Утверждение. Стационарное распределение вероятностей $\{\pi_i\}_{i=0}^{\infty}$ для рассматриваемого процесса $S(t)$ может быть аппроксимировано как

$$\pi_n = \pi_0 \frac{\lambda^n \prod_{i=1}^n (1 - Q_n)}{\mu^n n!}, \quad (\text{В.2})$$

где

$$\pi_0 = \left(\sum_{i=0}^{\infty} \frac{\lambda^i \prod_{j=1}^i (1 - Q_j)}{\mu^i i!} \right)^{-1},$$

а Q_{n+1} – вероятность блокировки (отказа) сессии при переходе из состояния n в состояние $n + 1$.

Следует отметить, что здесь ключевым упрощением является отказ от учета конкретной истории развития процесса в отношении активных сессий, т.е. в каждом его состоянии рассматривается новый набор соответствующих случайных величин. Если известны вероятности блокировки Q_{n+1} , то можно легко получить искомое стационарное распределение вероятностей из данного

выражения. Таким образом, дальнейшее изложение нацелено на вычисление значений Q_{n+1} , а результат суммируется по соответствующему утверждению, которое использует распределения случайных величин $\gamma_{i,j}$ и $d_{i,j}$. Последние могут быть найдены после громоздких, но несложных преобразований, описание которых в этом приложении опускается.

Вначале рассматривается процедура контроля доступа в слой D2D. Пусть n сессий уже обслуживаются сетью связи. Тогда для всех $i = \overline{1, n}$ необходимо, чтобы было выполнено следующее условие, связанное с минимальной требуемой скоростью передачи данных:

$$r \leq w \log \left(1 + \frac{p_{\max} \gamma_{i,i}}{K N_0} \right) \Leftrightarrow p_{\max} \gamma_{i,i} \geq K N_0 (e^{\frac{r}{w}} - 1). \quad (\text{B.3})$$

Далее формулируется следующее основное утверждение.

Утверждение. Для процедуры контроля доступа вероятности блокировки (отказа) сессии Q_{n+1} можно аппроксимировать как

$$\begin{aligned} Q_{n+1} &= 1 - \Pr\{\text{принята} \mid \text{поступила}\} = \\ &= \left[F_{\gamma} \left(\frac{N_0}{p_{\max}} \right) \right]^{2n-1} \left[1 - F_{\gamma} \left(\frac{\theta_0}{p_{\max}} \right) \right], \end{aligned} \quad (\text{B.4})$$

где $\theta_0 = K N_0 (e^{\frac{r}{w}} - 1)$, а функция распределения для значения SNR на единицу мощности γ выражается как

$$\begin{aligned} F_{\gamma}(\gamma) &= 1 + \frac{G^4 \gamma^{-\frac{4}{k}}}{8R^4} - \frac{G^2 \gamma^{-\frac{2}{k}}}{R^2} \ln 2, \quad \text{если } \frac{G^k}{(2R^2)^{\frac{k}{2}}} \leq \gamma \leq \gamma_{\max}, \\ F_{\gamma}(\gamma) &= 1 - \frac{1}{R^2} \left(\frac{G^4 \gamma^{-\frac{4}{k}}}{8R^2} + G^2 \gamma^{-\frac{2}{k}} \ln \frac{4R^2 \gamma^{\frac{2}{k}}}{G^2} \right), \\ \text{если } \frac{G^k}{(2R)^k} &\leq \gamma \leq \frac{G^k}{(2R^2)^{\frac{k}{2}}}, \quad \gamma_{\max} = \frac{K N_0}{p_{\max}} (e^{\frac{r_{\max}}{w}} - 1). \end{aligned}$$

При этом среднее число сессий на обслуживании

$$E [N_{\text{сессий}}] = \sum_{n=0}^{\infty} n \pi_n, \quad (\text{B.5})$$

где $\{\pi_n\}$, $n \geq 0$ представляют собой стационарные вероятности.

В.1.3 Энергетическое потребление

Используя вышеприведенное утверждение, можно получить характеристики, связанные со средней энергетической эффективностью сессий. Математическое ожидание значения мощности, потребляемой абонентским терминалом, вычисляется как

$$E [p_{total}] = \sum_{n=1}^{\infty} p^{(n)} \pi_n, \quad (\text{B.6})$$

где $p^{(n)}$ – среднее значение потребляемой мощности в состоянии n , а π_n – стационарные вероятности, найденные ранее. Важно заметить, что потребляемая в состоянии $n = 0$ мощность принимается здесь нулевой, поскольку рассматриваются только активные сессии.

Таким образом, общесистемное значение потребляемой мощности в состоянии n можно легко выразить в следующем виде:

$$p^{(n)} = \left(\frac{p_{\max}}{\eta} + p_c \right) \sum_{i=1}^n \frac{r}{r_i^{\max}} + \sum_{i=1}^n \left(1 - \frac{r}{r_i^{\max}} \right) p_a. \quad (\text{B.7})$$

Тогда средняя потребляемая мощность в состоянии n

$$E[p^{(n)}] = \left(\frac{p_{\max}}{\eta} + p_c - p_a \right) \sum_{i=1}^n E \left[\frac{r}{r_i^{\max}} | n \text{ сессий} \right] + n p_a, \quad (\text{B.8})$$

где $E \left[\frac{r}{r_i^{\max}} | n \text{ сессий} \right]$ – математическое ожидание случайной величины $\frac{r}{r_i^{\max}}$ при условии, что n сессий уже было принято в систему. В итоге можно получить общее энергетическое потребление типовой сессии в слое D2D

$$E [\varepsilon] = \frac{\sum_{n=1}^{\infty} p^{(n)} \pi_n}{\lambda P_a}, \quad (\text{B.9})$$

где вероятность допуска сессии в слой D2D, обозначенная через P_a , определяется на основе формулы полной вероятности в виде

$$P_a = 1 - \sum_{n=0}^{\infty} \Pr \{ \text{заблокирована} \mid \text{поступила} \} \pi_n = 1 - \sum_{n=0}^{\infty} Q_{n+1} \pi_n. \quad (\text{B.10})$$

В.2 Моделирование сотовой системы связи

В данном разделе вычисляется стационарное распределение вероятностей для базовой сотовой системы, которая управляет работой слоя D2D, а также устанавливаются характеристики ее производительности. Вообще говоря, приведенные здесь рассуждения аналогичны анализу сети D2D, поэтому далее подчеркиваются только основные различия.

В.2.1 Вероятностная модель

Следует напомнить, что поступления новых сессий в сеть D2D, интегрированную с сотовой системой связи, подчиняются процессу PPP. Следовательно, поток точек, задающих местоположение пользователей сотовой сети доступа, представляет собой пуассоновский процесс, просеянный с вероятностью P_a и имеющий интенсивность $\lambda(1 - P_a)$. По аналогии с рассуждениями из предыдущего раздела, поведение рассматриваемой системы можно описать марковским процессом $S(t)$, заданным в моменты времени t поступления новых сессий и их ухода по завершении обслуживания. Интенсивности переходов между состояниями этого процесса используют параметры, соответствующие работе сотовой сети. Важно заметить, что для сотовой системы связи характеристики обслуживаемой ею пользовательской сессии полностью определяются местоположением передающего терминала.

В.2.2 Стационарное распределение вероятностей

Укрупняя состояния рассматриваемого здесь марковского процесса по аналогии с подходом для системы D2D и подставляя интенсивность входного потока для сотовой сети $\lambda(1 - P_a)$, можно получить стационарное распределение вероятностей $\{\pi_i\}_{i=0}^{\infty}$. Кроме того, следует учесть распределения случайных величин γ_i и d_i , которые можно найти, принимая во внимание равномерный

характер размещения узлов сети так, как это было сделано в слое D2D. Для вычисления стационарного распределения вероятностей необходимо установить вероятность блокировки (отказа) сессии Q_{n+1} для исследуемых алгоритмов передачи: MR (на максимальной скорости) и FU (с полным использованием).

В частности, для состояния n следует рассмотреть распределение случайной величины $z_n = \sum_{i=1}^n \frac{r}{r_i^{\max}}$. Поскольку сложно получить его непосредственно, далее не вычисляются свертки случайных величин $q_i = \frac{r}{r_i^{\max}}$, а значение z_n , $n > 2$ аппроксимируется следующим образом. Сумма z_n разбивается на два компонента $z_n = z_{n-1} + q_n$, где q_n – случайная величина, соответствующая новой сессии. Тогда значение z_{n-1} можно аппроксимировать случайной величиной, распределенной согласно усеченному нормальному закону на интервале $[0; \min(q_{\max}, 1)]$ со средним m_{n-1} и дисперсией σ_{n-1} . Здесь значение $q_{\max} = \frac{r}{w \log(1+p_{\max}(G/R)^k)}$ представляет собой максимум случайной величины q_i на краю соты. С учетом этого можно сформулировать следующее утверждение для алгоритма передачи MR и соответствующее утверждение для алгоритма передачи FU, которые позволяют получить аппроксимации для вероятности блокировки (отказа) сессии Q_{n+1} .

Утверждение. Для алгоритма передачи MR вероятности блокировки (отказа) сессии Q_1, Q_2, \dots, Q_{n+1} могут быть получены следующим образом:

а. При $n = 0$ следует непосредственно использовать функцию распределения случайной величины $q = \frac{r}{w \log(1+p_{\max}\gamma)}$ и установить

$$Q_1 = 1 - \frac{\left(Gp_{\max}^{\frac{1}{k}}\right)^2}{R^2} \left(e^{\frac{r}{wq_0}} - 1\right)^{-\frac{2}{k}}, \quad (\text{B.11})$$

где $q_0 = \min(q_{\max}, 1)$.

б. Для нахождения вероятности блокировки в случае, когда сессия поступает в состоянии 1, можно записать

$$Q_2 = 1 - \frac{\left(Gp_{\max}^{\frac{1}{k}}\right)^2}{R^2} \int_0^{q_0} \frac{e^{\frac{r}{wq_1}} \left(e^{\frac{r}{w \min(q_{\max}, 1 - q_1)}} - 1\right)^{\frac{2}{k}}}{q_1^2 \left(e^{\frac{r}{wq_1}} - 1\right)^{\frac{k+2}{k}}} dq_1. \quad (\text{B.12})$$

в. Вероятности Q_{n+1} , $n \geq 2$ могут быть аппроксимированы в виде

$$Q_{n+1} = 1 - \frac{\left(Gp_{\max}^{\frac{1}{k}}\right)^2}{\varphi_0} \int_0^{q'_0} \left(\int_0^{b'} \frac{e^{-\frac{(z_{n-1} - m_{n-1})^2}{2\sigma_{n-1}}} e^{\frac{r}{wq_n}}}{q_n^2 \left(e^{\frac{r}{wq_n}} - 1\right)^{\frac{k+2}{k}}} dq_n \right) dz_{n-1}, \quad (\text{B.13})$$

где $\varphi_0 = R^2 \sqrt{2\pi(n-1)}\sigma_{n-1} \left[\Phi\left(\frac{(n-1)q_{\max}-m_{n-1}}{\sigma_{n-1}}\right) - \Phi\left(\frac{-m_{n-1}}{\sigma_{n-1}}\right) \right]$, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{x^2}{2}} dx$, а остальные параметры (m_{n-1} , σ_{n-1} , q_{\max} , q'_0 , b') приведены ниже.

Далее рассматривается алгоритм передачи FU, причем полезно напомнить, что максимальная скорость передачи данных r_i^{\max} , $\forall i = \overline{1, n}$ определяется как

$$r_i^{\max} = \min [w \log(1 + \gamma_i p_{\max}), r_{\max}] \geq r_i,$$

где r_{\max} – максимальная достижимая скорость передачи данных, которая ограничена особенностями используемой технологии радиодоступа.

Утверждение. Для алгоритма передачи FU вероятность блокировки (отказа) сессии Q_{n+1} может быть получена из распределения случайной величины $r_i^{\max} = w \log(1 + \gamma_i p_{\max})$ следующим образом:

$$Q_{n+1} = 1 - \frac{G^2 p_{\max}^{\frac{2}{k}}}{R^2} \cdot \left[\frac{e^{\frac{rn}{w}} - 1}{\left(e^{\frac{rn+n}{w}} - 1\right)^2} \right]^{\frac{2}{k}}, \quad (\text{B.14})$$

где распределение случайной величины r_i^{\max} можно выразить из распределения γ_i .

Здесь следует принять во внимание условие $r_i \leq r_i^{\max}$, которое связано с ограничением на наибольшее число одновременно активных сессий, т.е. $n \leq \frac{1}{q_{\max}}$, где значение q_{\max} было определено выше. Также важно подчеркнуть, что для алгоритма передачи FU ранее было предложено точное выражение, а не аппроксимация, как в случае алгоритма передачи MR.

В.2.3 Энергетическое потребление

По аналогии с рассуждениями для сети D2D можно найти математическое ожидание значения мощности, потребляемой абонентским терминалом в сотовой системе связи. Для некоторого набора n активных сессий потребляемая мощность может быть найдена как

$$p^{(n)} = \sum_{i=1}^n \frac{r}{r_i} \left(\frac{p_i}{\eta} + p_c \right) + \sum_{i=1}^n \left(1 - \frac{r}{r_i} \right) p_a,$$

где p_i – уровень мощности, излучаемой терминалом i .

Среднее значение потребляемой мощности $p^{(n)}$ в состоянии n для алгоритма передачи MR составит

$$E[p^{(n)}] = \left(\frac{p_{\max}}{\eta} + p_c - p_a \right) E \left[\sum_{i=1}^n \frac{r}{r_i^{\max}} \mid \sum_{i=1}^n \frac{r}{r_i^{\max}} \leq 1 \right] + np_a, \quad (\text{B.15})$$

где компонент $E \left[\sum_{i=1}^n \frac{r}{r_i^{\max}} \mid \sum_{i=1}^n \frac{r}{r_i^{\max}} \leq 1 \right]$ вычисляется ниже.

Средняя потребляемая мощность в состоянии n для алгоритма передачи FU представляет собой

$$\begin{aligned} E[p^{(n)}] &= \frac{1}{\eta} E \left[p_i \mid \frac{r}{r_i} \leq \frac{1}{n} \right] + p_c + (n-1)p_a = \frac{1}{\eta} \left(e^{\frac{rn}{w}} - 1 \right) \cdot \\ &\cdot E \left[\frac{1}{\gamma_i} \mid \frac{1}{\gamma_i} \leq \max \left(\frac{R^k}{G^k}, \frac{p^{\max}}{e^{\frac{rn}{w}} - 1} \right) \right] + p_c + (n-1)p_a, \end{aligned} \quad (\text{B.16})$$

где выражение для $E[\cdot]$ также определено далее.

В итоге общее энергетическое потребление типовой сессии в сотовой системе связи может быть найдено с использованием соответствующего значения блокировки (отказа)

$$P_b = \sum_{n=0}^{\infty} Q_{n+1} \pi_n. \quad (\text{B.17})$$

В.3 Вспомогательные вычисления

В этом разделе собраны полезные разъяснения и подробности вычисления вспомогательных распределений, требуемых для применения сформулированных утверждений, а также приводятся выражения для расчетов энергетического потребления. Важно отметить, что все три рассматриваемые модели множественного доступа (сотовая сеть под управлением алгоритмов MR и FU, а также система D2D – см. главу 3) используют единый подход к вычислению вероятности блокировки (отказа) сессии и значения потребляемой мощности в некотором состоянии. Зная распределение местоположений абонентских терминалов внутри соты, можно получить распределение расстояний между передающим и принимающим пользователями (или же БС). Тогда, накладывая различные

в каждом случае ограничения, связанные с контролем доступа в сеть, устанавливают значения условных математических ожиданий соответствующих случайных величин при наличии n сессий на обслуживании. Для краткости ниже приводятся только итоговые выражения для обсуждаемых здесь результатов.

В.3.1 Слой D2D

Для проведения последующих преобразований можно выразить функцию распределения величины $z = d_{i,j}^2$ как

$$F_z(z) = -\frac{z^2}{8R^4} + \frac{z}{R^2} \ln 2, \quad \left(\frac{p_{\max} G^k}{K\theta \left(e^{\frac{r_{\max}}{w}} - 1 \right)} \right)^{2/k} \leq z \leq 2R^2,$$

$$F_z(z) = \frac{1}{R^2} \left(\frac{z^2}{8R^2} + z \ln \frac{4R^2}{z} \right), \quad 2R^2 \leq z \leq 4R^2.$$

Соответствующую функцию плотности вероятности можно получить непосредственно путем дифференцирования. Далее вводится следующее вспомогательное обозначение

$$y_{\min} = \frac{r}{w \log(1 + \text{SINR}_{\max})}, \quad z_1 = \frac{N_0}{p_{\max}}, \quad z_0 = \frac{\theta_0}{p_{\max}}.$$

Тогда, с использованием распределения случайной величины d_i^2 , вычисляется среднее значение потребляемой мощности в слое D2D

$$E \left[\frac{r}{r_i} | n \text{ сессий} \right] =$$

$$= [F_z(z_1)]^{1-n} \left[1 - F_z(z_0) \int_{y_{\min}}^1 y \psi(y) f_z(z(y)) dy \right]^{-1}.$$

Здесь выражение для функции $z(y)$ задается как

$$z = \left[\frac{p_{\max} G^k}{K\theta \left(e^{\frac{r}{wy}} - 1 \right)} \right]^{2/k} = \left[\frac{p_{\max} G^k}{K\theta} \right]^{2/k} \left(e^{\frac{r}{wy}} - 1 \right)^{-2/k},$$

а введенная в целях упрощения вспомогательная функция $\psi(y)$ представляет собой

$$\psi(y) = \left[\frac{p_{\max} G^k}{K\theta} \right]^{2/k} \frac{2}{k} \frac{r}{w} \left(e^{\frac{r}{wy}} - 1 \right)^{-\frac{2}{k}-1} e^{\frac{r}{wy}} \frac{1}{y^2}.$$

В.3.2 Сотовый слой: алгоритм MR

Ниже приводятся основные результаты и вспомогательные выражения, необходимые для применения первого утверждения, а также для нахождения значения мощности, потребляемой абонентским терминалом. В целях упрощения верхняя граница для возможных значений q_i , определяемая расстоянием R , обозначается как q_{\max} , а соответствующая нижняя граница, определяемая максимальным значением SNR, как q_{\min} . При этом $q_0 = \min(q_{\max}, 1)$, что дает

$$q_{\max} = \frac{r}{w \log \left(1 + \frac{G^k}{R^k} p_{\max} \right)}, \quad q_{\min} = \frac{r}{w \log (1 + \text{SNR}_{\max})}.$$

Функция распределения случайной величины q может быть вычислена следующим образом:

$$F_q(q) = \frac{\left(G p^{\frac{1}{k}} \right)^2}{R^2 \left(e^{\frac{r}{wq}} - 1 \right)^{\frac{2}{k}}}, \quad q \leq q_{\max}.$$

Тогда условная функция плотности вероятности принимает вид

$$F_q(q|q \leq 1) = \frac{\left(G p^{\frac{1}{k}} \right)^2}{R^2 \left(e^{\frac{r}{wq}} - 1 \right)^{\frac{2}{k}}} \cdot \frac{1}{F_q(1)} = \frac{\left(e^{\frac{r}{w}} - 1 \right)^{\frac{2}{k}}}{\left(e^{\frac{r}{wq}} - 1 \right)^{\frac{2}{k}}}, \quad q \leq 1,$$

где функция плотности вероятности $f_q(q|q \leq 1) = dF_q(q|q \leq 1)/dq$.

Для нахождения переходных вероятностей и энергетического потребления в некотором состоянии требуется выразить условное математическое ожидание $E[z_n | z_n \leq 1]$, $n \geq 1$ как показано далее. Зная распределение случайной величины q , можно установить ее первый момент как

$$\begin{aligned} E[q|q \leq 1] &= q_{\min} F_{q|q \leq 1}(q_{\min}) + \int_{q_{\min}}^1 q f_{q|q \leq 1}(q) dq = \\ &= q_{\min} \frac{\left(e^{\frac{r}{w}} - 1 \right)^{\frac{2}{k}}}{\left(e^{\frac{r}{wq}} - 1 \right)^{\frac{2}{k}}} + \frac{2r \left(e^{\frac{r}{w}} - 1 \right)^{\frac{2}{k}}}{kw} \int_{q_{\min}}^1 \frac{e^{\frac{r}{wq}}}{q \left(e^{\frac{r}{wq}} - 1 \right)^{\frac{k+2}{k}}} dq. \end{aligned}$$

В результате второй момент случайной величины q возможно получить в виде

$$\begin{aligned} E[q^2|q \leq 1] &= q_{\min}^2 F_{q|q \leq 1}(q_{\min}) + \int_{q_{\min}}^1 q^2 f_{q|q \leq 1}(q) dq = \\ &= q_{\min}^2 \frac{(e^{\frac{r}{w}} - 1)^{\frac{2}{k}}}{\left(e^{\frac{r}{wq}} - 1\right)^{\frac{2}{k}}} + \frac{2r (e^{\frac{r}{w}} - 1)^{\frac{2}{k}}}{kw} \int_{q_{\min}}^1 \frac{e^{\frac{r}{wq}}}{\left(e^{\frac{r}{wq}} - 1\right)^{\frac{k+2}{k}}} dq. \end{aligned}$$

Соответствующее значение дисперсии рассматриваемой случайной величины q вычисляется согласно

$$\sigma_q^2 = E[q^2|q \leq 1] - (E[q|q \leq 1])^2.$$

Выражения для получения условного математического ожидания $E[z_2|q \leq 1]$ и условной вероятности блокировки $Q_2 = Pr\{z_2 > 1|z_1 = q \leq 1\}$ представляют собой

$$\begin{aligned} E[q_1 + q_2|q_1 + q_2 \leq 1] &= \frac{\varphi(q_{\min})}{a} \int_0^{q_{\min}} \frac{e^{\frac{r}{wq_1}}}{q_1^2 \left(e^{\frac{r}{wq_1}} - 1\right)^{\frac{k+2}{k}}} dq_1 + \\ &+ \frac{1}{a} \int_{q_{\min}}^{q_0} \varphi(q_1) \frac{e^{\frac{r}{wq_1}}}{q_1^2 \left(e^{\frac{r}{wq_1}} - 1\right)^{\frac{k+2}{k}}} dq_1, \end{aligned}$$

где $q_0 = \min(q_{\max}, 1)$ и $a = Pr\{q_1 + q_2 \leq 1\}$, а также

$$\begin{aligned} \varphi(q) &= (q + q_{\min})(1 - C_0) + \\ &+ \int_{q_{\min}}^b (q + q_2) \frac{e^{\frac{r}{wq_2}}}{q_2^2 \left(e^{\frac{r}{wq_2}} - 1\right)^{\frac{k+2}{k}}} dq_2, \\ a &= \int_0^{q_0} \left(\int_0^b \frac{e^{\frac{r}{wq_2}}}{q_2^2 \left(e^{\frac{r}{wq_2}} - 1\right)^{\frac{k+2}{k}}} dq_2 \right) \frac{e^{\frac{r}{wq_1}}}{q_1^2 \left(e^{\frac{r}{wq_1}} - 1\right)^{\frac{k+2}{k}}} dq_1. \end{aligned}$$

При этом $b = \min(q_{\max}, 1 - q_1, 1 - q_{\min})$.

Для завершения предлагаемой здесь аппроксимации искомого распределения в случае $n > 2$ можно записать

$$E[z_{n-1} + q_n | z_{n-1} + q_n \leq 1] = \frac{\varphi_{n-1}(q_{\min})}{a'} \int_0^{(n-1)q_{\min}} e^{-\frac{(z_{n-1}-m_{n-1})^2}{2\sigma_{n-1}}} dz_{n-1} +$$

$$+ \frac{1}{a'} \int_{(n-1)q_{\min}}^{q'_0} \varphi_{n-1}(z_{n-1}) e^{-\frac{(z_{n-1}-m_{n-1})^2}{2\sigma_{n-1}}} dz_{n-1},$$

где $a' = \Pr\{z_{n-1} + q_n \leq 1\}$, $q'_0 = \min((n-1)q_{\max}, 1)$, а также

$$\varphi_{n-1}(q) = (q + q_{\min})(1 - C_0) +$$

$$+ \int_{q_{\min}}^b (q + q_n) \frac{e^{\frac{r}{wqn}}}{q_n^2 \left(e^{\frac{r}{wqn}} - 1\right)^{\frac{k+2}{k}}} dq_n,$$

$$a = \int_0^{q'_0} \left(\int_0^{b'} \frac{e^{\frac{r}{wqn}}}{q_n^2 \left(e^{\frac{r}{wqn}} - 1\right)^{\frac{k+2}{k}}} dq_n \right) e^{-\frac{(z_{n-1}-m_{n-1})^2}{2\sigma_{n-1}}} dz_{n-1}.$$

При этом $b' = \min(q_{\max}, 1 - z_{n-1}, 1 - (n-1)q_{\min})$.

По аналогии можно вычислить вторые моменты $E[z_2^2 | z_2 \leq 1]$, $E[z_n^2 | z_n \leq 1]$ и дисперсии σ_2^2 , σ_n^2 . Таким образом, параметры распределений m_n и σ_n^2 могут быть найдены с использованием преобразований, описанных выше.

В.3.3 Сотовый слой: алгоритм FU

Для получения среднего значения потребляемой мощности в рассматриваемом случае приводится выражение

$$E \left[\frac{1}{\gamma} \middle| \frac{1}{\gamma} \leq \frac{p_{\max}}{(e^{\frac{nr}{w}} - 1)} \right] =$$

$$= \frac{k}{k+2} \frac{p_{\max}}{e^{\frac{nr}{w}} - 1} \left(\frac{e^{\frac{nr}{w}} - 1}{e^{r_{\max}/w} - 1} \right)^{\frac{2}{k}+1} + \frac{2}{k+2} \frac{p_{\max}}{e^{\frac{nr}{w}} - 1}. \quad (\text{B.18})$$

В результате данное приложение содержит все вспомогательные величины и преобразования, требуемые для нахождения характеристик работы исследуемой системы связи. Вычисление некоторых приведенных здесь распределений становится при этом несложным техническим упражнением, которое выносится за рамки настоящего изложения.

Приложение Г

Анализ энергетической эффективности mmWave-системы носимых устройств

Г.1 Моделирование сети доступа в целом

В данном приложении приводятся основные результаты моделирования работы сети носимых устройств на основе mmWave-соединений. Дальнейшее изложение построено следующим образом. В начале, для оценки вероятности успешной доставки маяка, а также вероятности успешной передачи данных (когда удастся избежать коллизий типа «маяк – данные» и «данные – данные») учитываются особенности протокола доступа к каналу без рассмотрения геометрии размещения терминалов и направленности передачи. Далее предлагается аппроксимация для работы системы в целом, отражающая специфику формирования диаграммы направленности. После этого в явном виде вводятся в рассмотрение пространственные соображения. Наконец, учитываются все основные допущения, принятые главе 4, включая типовую форму диаграммы направленности, и после пространственного усреднения формулируются основные выводы. В завершение кратко обсуждаются возможные обобщения проведенного моделирования для более реалистичных допущений и сценариев.

Г.1.1 Протокол доступа к каналу

Здесь приводятся основные характеристики, необходимые для последующего анализа энергетической эффективности системы путем абстрагирования ее пространственной составляющей и рассмотрения сравнительно небольшой области моделирования, в которой все взаимодействующие терминалы могут устанавливать соединение друг с другом. Далее некоторое фиксированное число ПТД обозначается через N и проводится изучение такой системы взаимодействующих между собой PBSS. Принятые допущения позволяют получить

оценку снизу для значения общесистемной энергетической эффективности. Для этого предлагается полностью распределенный алгоритм, при котором ПТД не координируют передачу своих маяков и данных между собой, что может рассматриваться как худший случай практического функционирования системы.

Рассматривая введенный таким образом распределенный алгоритм и предполагая вначале $K = 1$, можно предложить два способа моделирования системы связи:

- *Практическая абстракция.* Все PBSS в системе синхронизированы только на уровне слотов. Начало суперкадра для каждой PBSS представляет собой случайную величину, которая генерируется однократно равномерным образом. В течение суперкадра ПТД периодически отправляет маяки, содержащие информацию о расписании передач. Располагая такой информацией, подключенные к ПТД устройства передают свои данные в отведенные им интервалы времени.
- *Теоретическая абстракция.* Все PBSS в системе синхронизированы как на уровне слотов, так и на уровне суперкадров. Однако время начала отправки маяка назначается случайно, так что в начале каждого суперкадра каждая ПТД равномерным образом выбирает слот для отсылки маяка, а также M слотов для передачи данных. Следует заметить, что в такой теоретической модели маяки могут следовать за фрагментами данных, но внутри одной PBSS возникновение коллизий по-прежнему исключено.

Хотя первая из введенных выше моделей и оказывается ближе к практике, дальнейшее исследование сосредоточивается на второй модели в силу ее аналитичности. При этом обе введенные абстракции функционируют эквивалентно в отношении своих усредненных характеристик производительности. Таким образом, далее изучается один суперкадр рассматриваемого протокола доступа к каналу, в котором слоты для передачи маяков и фрагментов данных выбираются ПТД равномерным образом из интервала $[1; L]$. Подобные постановки многократно анализировались в литературе и известны как задачи с последовательным размещением [350–352], в которых множество из M неразличимых между собой шаров последовательно занимает L ящиков.

Несмотря на то, что задачи с последовательным размещением, обсуждаемые, например, в [353], вообще говоря, аналогичны текущей постановке, основное ее отличие состоит в том, что требуется различать передачу маяков и

фрагментов данных. В частности, неуспешная доставка маяка приводит к невозможности дальнейшей отправки данных. На основании принципов работы описанного выше распределенного алгоритма доступа к каналу, далее предлагаются достаточно точные оценки его параметров, таких как вероятность успешной передачи маяка и фрагмента данных, а также среднее значение пропускной способности. В этой связи формулируется следующее важное утверждение.

Утверждение. Для описанной выше системы случайного доступа, включающей в себя передачу маяков и фрагментов данных, можно получить основные условные ее характеристики работы. Вероятность успешно принятого маяка и вероятность успешно принятого фрагмента данных при условии успешной отправки маяка могут быть вычислены при помощи выражений

$$P_{B|N} = e^{-\frac{N}{L}}, \quad P_{D|N,B} = e^{-\frac{N}{L}} e^{-M\frac{N}{L}e^{-\frac{N}{L}}}. \quad (\Gamma.1)$$

Пропускная способность сети PBSS определяется как среднее число слотов, содержащих успешную передачу данных, в расчете на один суперкадр

$$T = \frac{MP_{B|N}P_{D|N,B}}{L} = \frac{M}{L} e^{-2\frac{N}{L}} e^{-M\frac{N}{L}e^{-\frac{N}{L}}}. \quad (\Gamma.2)$$

Доказательство. Далее рассматривается базовая система, где все терминалы могут принимать передачи друг друга, и исследуется ее работа в рамках одного суперкадра. При этом некоторое фиксированное число ПТД предпринимает попытку передачи своих маяков и данных в соответствии с теоретической моделью алгоритма доступа, описанной выше (суперкадры системы синхронизированы, слоты выбираются случайно, данные могут быть отправлены как до, так и после маяка). Ниже предлагается краткое обсуждение способа получения асимптотических выражений для вероятностей успешной доставки маяка и фрагмента данных при заданном числе интерферирующих устройств с применением методов [354; 355]. С учетом того, что значение параметра L может быть достаточно велико, предложенная асимптотическая оценка используется в качестве аппроксимации для искомых вероятностей.

Если число потенциально интерферирующих маяков N , тогда в соответствии с принятыми ранее допущениями относительно функционирования протокола доступа к каналу вероятность успешной передачи маяка составляет $P_{b|N} = (1 - \frac{1}{L})^N$. Пусть L, N стремится к бесконечности так, чтобы $N/L = p = const$ (что соответствует числу устройств на слот в рамках суперкадра). Если значение p слишком велико, то вся система оказывается неработоспособной из-за

частых коллизий между маяками, в то время как при $p \rightarrow 0$ задержка доступа существенно повышается. Следовательно, представляет интерес изучение всего диапазона $0 < p < 1$.

Таким образом, вероятность успешной доставки маяка может быть введена как вероятность того, что интерферирующие маяки отсутствуют, т.е. $P_{B|N} = e^{-L/N}$. Успех при передаче маяков приводит к возникновению потока данных с интенсивностью $Me^{L/N}$. Тогда вероятность успешной отправки данных может быть определена как вероятность события, что слот данных свободен и от маяков, и от других фрагментов данных. Значит, $P_{D|N} = e^{-L/N}e^{-Me^{L/N}}$, что приводит к искомой, достаточно простой аппроксимации. Более тщательное доказательство данного факта можно провести на основе постановки и решения соответствующей комбинаторной задачи, а необходимые для этого вычисления можно найти в [356].

Здесь и далее модель сети доступа, в которой имеется два вида передач, маяки (М) и данные (Д), получает название Тип-1. Поскольку в спецификациях технологии WiGig отсутствуют сведения о том, как именно передачи маяков и фрагментов данных располагаются внутри суперкадра друг относительно друга, целесообразно рассмотреть три возможных альтернативы: (1) слоты для отправки данных выбираются произвольно; (2) фрагменты данных сгруппированы, но слот для отправки первого из них выбирается случайным образом; (3) группа фрагментов данных следует непосредственно после успешной отсылки маяка. Во всех этих случаях предполагается, что передача маяка происходит случайным образом на протяжении L слотов суперкадра.

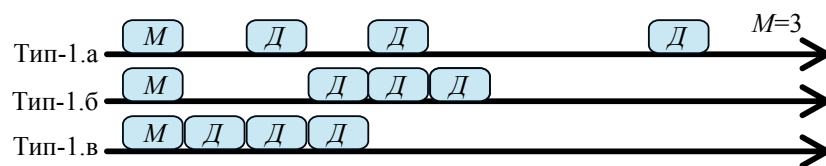


Рисунок Г.1 — Пример работы систем Тип-1.а (случайное размещение всех фрагментов данных), Тип-1.б (случайное размещения начала группы фрагментов данных) и Тип-1.в (группа фрагментов данных следует непосредственно за успешной отправкой маяка)

Соответственно, рисунок Г.1 демонстрирует различие между этими тремя вариантами, которые именуются Тип-1.а, Тип-1.б и Тип-1.в, а рисунок Г.2

приводит результаты сравнения пропускной способности для всех рассматриваемых альтернатив. Важно отметить, что в отношении пропускной способности и вероятности успешной доставки системы доступа Тип-1.а и Тип-1.б оказываются эквивалентны, тогда как Тип-1.в отличается от них ввиду корреляции между размещением маяка и фрагментов данных. Для анализа системы Тип-1.в представляется возможным сформулировать следующее утверждение, которое позволяет построить полезную аппроксимацию.

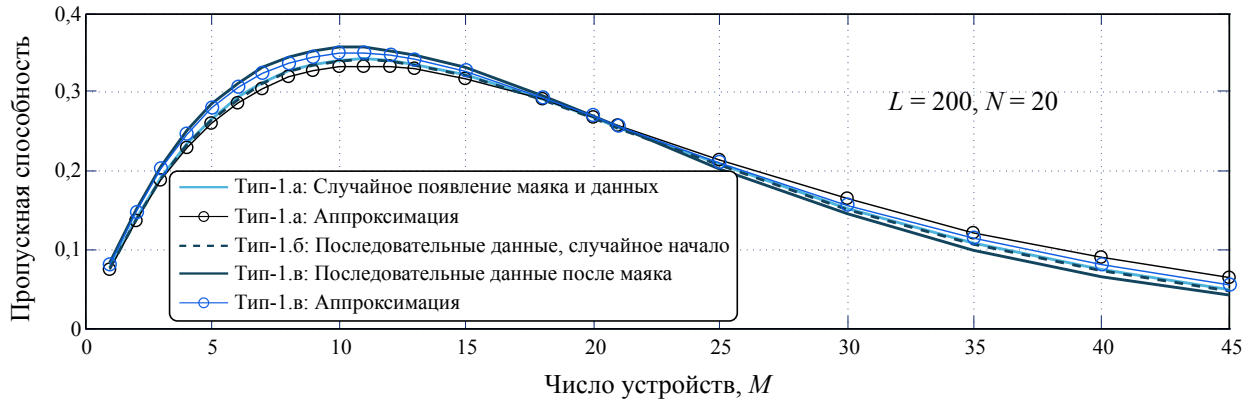


Рисунок Г.2 — Пропускная способность для систем Тип-1.а, Тип-1.б и Тип-1.в

Утверждение. В системе случайного доступа Тип-1.в соответствующие вероятности успеха с учетом требуемого общего количества передаваемых данных KM могут быть получены при помощи выражений

$$P_{B|N} = e^{-\frac{N}{L}}, \quad P_{D|N,B} = e^{-\frac{N}{L}} \left(1 - \frac{N}{L} e^{-\frac{N}{L}}\right)^{KM-1}. \quad (\text{Г.3})$$

Пропускная способность сети PBSS как среднее число успешных слотов для отправки данных на суперкадр может быть выражена как

$$T = \frac{MP_{B|N}P_{D|N,B}}{L} = \frac{M}{L} e^{-2\frac{N}{L}} \left(1 - \frac{N}{L} e^{-\frac{N}{L}}\right)^{KM-1}. \quad (\text{Г.4})$$

Доказательство основывается на рассуждениях, аналогичных приведенным при доказательстве предыдущего утверждения, которые опускаются здесь для краткости.

Следует заметить, что в актуальной документации, относящейся к технологии WiGig, не предлагается рекомендаций по выбору способа доступа к каналу связи. Таким образом, в последующем обсуждении для этих целей предполагается применение системы Тип-1.в, поскольку ее практическая реализация оказывается проще, а пропускная способность — немного выше в ненасыщенном

режиме. Далее для системы Тип-2 предполагается аналогичный подход при организации доступа, основанный на передаче группы фрагментов данных. Соответственно, за успешной отправкой маяка следует группа из M временных интервалов по K слотов каждый.

Естественно считать, что подстройка диаграммы направленности антенны, связанная с работой процедуры «прочесывания» секторов SLS, выполняется для каждого подключенного устройства единожды в первом слоте соответствующей группы из K слотов (т.н. слот НП). Если при этом происходит коллизия с другим маяком или фрагментом данных, рассматриваемая процедура проводится снова в следующем слоте при их наличии. В случае ее успеха данные передаются в выбранном секторе антенны в течение оставшейся части отведенного интервала. Эта процедура повторяется для других устройств во всех суперкадрах. Для изучения характеристик энергопотребления проводится оценка числа слотов Д и НП соответственно. При этом передача маяка имеет одинаковую вероятность успеха во всех рассмотренных выше системах.

Утверждение. В системе случайного доступа Тип-2 вероятность успешно избежать коллизии типа М, Д или НП с учетом требуемого общего количества передаваемых данных KM выражается как

$$q = e^{-\frac{N}{L}} \left(1 - \frac{N}{L} e^{-\frac{N}{L}}\right)^{KM-1}. \quad (\Gamma.5)$$

Число успешных слотов Д и НП на устройство может быть найдено в виде

$$\begin{aligned} E[n'_D] &= \left(qK - 1 + (1 - q)^K\right) e^{-\frac{N}{L}}, \\ E[n'_{BF}] &= q \left(1 - (1 - q)^K\right) e^{-\frac{N}{L}}. \end{aligned} \quad (\Gamma.6)$$

При наличии накладных расходов на подстройку диаграммы направленности ξ , когда доля $1 - \xi$ слотов НП может быть использована для обычной передачи фрагментов данных, пропускная способность сети PBSS (в слотах) составляет

$$T = \frac{1}{L} (E[n'_D] + (1 - \xi)E[n'_{BF}]). \quad (\Gamma.7)$$

Доказательство. Вероятность успешно избежать коллизии типа М, Д или НП следует непосредственно из вышеприведенного утверждения. Число успешных слотов НП можно получить при помощи среднего номера успешной попытки передачи НП среди K слотов, который следует из усеченного геометрического распределения с параметром q как

$$E[n_{attempts}] = \frac{1}{q} \left(1 - qK(1 - q)^K - (1 - q)^K\right). \quad (\Gamma.8)$$

После получения основных параметров системы можно продолжить исследование энергетической эффективности рассматриваемой сети доступа, анализируя ее пространственную составляющую.

Г.1.2 Геометрия сети связи

При анализе базовой модели с полной связностью системы учитываются ее геометрические компоненты, связанные с распределением на плоскости устройств, имеющих ограниченную зону покрытия. Вначале представляет интерес рассмотрение терминалов с ненаправленными антеннами, где функция $\rho(\Delta\alpha) = R$ определяет радиус соответствующей зоны покрытия. Далее изучается зона покрытия некоторой меченой ПТД, которая представляет собой круг радиуса R . Следовательно, соединение между двумя терминалами существует, если и только если они находятся в области пересечения двух кругов, задающих их зоны покрытия. Это можно записать как $\Pr\{\rho(\Delta\alpha) \geq d\} = \psi(d) = 1$ для любых двух устройств на пересечении зон покрытия. В силу пуассоновского характера распределения ПТД на плоскости, общее число (возможно интерферирующих между собой) терминалов N_{all} внутри данной области имеет распределение

$$P_N = \Pr\{N_{all} = N\} = \frac{(S_c\mu)^N}{N!} e^{-S_c\mu}, \quad (\text{Г.9})$$

где $S_c = \pi R^2$ – площадь зоны покрытия. Поскольку все устройства, расположенные внутри зоны покрытия меченой ПТД, интерферируют с меченым терминалом, распределение общего числа интерферирующих маяков (т.е. различных ПТД) N_b также следует соотношению, приведенному выше.

Далее, с учетом набора функций $f_i(N)$ случайной величины N , заданного предыдущими утверждениями, следующий очевидный шаг – получение соответствующих математических ожиданий. Полезно напомнить, что если существуют моменты некоторой случайной величины и рассматриваемая функция бесконечно дифференцируема, тогда ее можно представить в виде разложения в ряд Тейлора. Первый элемент такого разложения становится аппроксимацией $E[f(N)] \approx f(E[N])$, которая может быть уточнена при добавлении второго элемента $E[f(N)] \approx f(E[N]) + \frac{1}{2}f''(E[N])\text{Var}(N)$. Поскольку для последующего

упрощенного анализа оказывается пригоден уже первый компонент, то среднее число терминалов $N = \pi R^2 \mu$ подставляется в выражения из вышеприведенных утверждений. В результате можно сформулировать несколько полезных следствий.

Следствие. В системе случайного доступа Тип-1.в аппроксимация для числа успешных маяков и фрагментов данных на кадр одной сети PBSS может быть получена как

$$E[n_B] = e^{-p}, \quad E[n_D] = M e^{-2p} (1 - p e^{-p})^{KM-1}, \quad (\text{Г.10})$$

где $p = \frac{E[N_b]}{L}$.

Следствие. В системе случайного доступа Тип-2 аппроксимация для числа успешных фрагментов данных и попыток формирования диаграммы направленности на кадр одной сети PBSS может быть получена как

$$\begin{aligned} E[n_D] &= M \left(qK - 1 + (1 - q)^K \right) e^{-p}, \\ E[n_{BF}] &= Mq \left(1 - (1 - q)^K \right) e^{-p}, \end{aligned} \quad (\text{Г.11})$$

где $p = \frac{E[N_b]}{L}$, а $q = e^{-p} (1 - p e^{-p})^{KM-1}$.

Следует отметить, что использование любой из приведенных выше аппроксимаций может привести к отклонению результирующей оценки от показателей работы реальной системы. Однако в главе 4 при помощи данных имитационного моделирования показано, что предложенный здесь математический подход достаточно точен. С учетом этого далее используется аппроксимация, несмотря на то, что она не является точным решением. Кроме того, также используется вышеприведенное выражение.

Г.1.3 Направленность передачи

В дальнейшем изложении учитываются особенности передачи устройств при использовании mmWave-каналов. Важно напомнить, что соединение между меченой ПТД и любым другим терминалом внутри зоны покрытия существует, если и только если расстояние между ними составляет $d < \rho(\Delta\alpha_s)$ (в силу случайности азимута), где $0 \leq \Delta\alpha_s \leq 2\pi$ – относительный азимут устройства. Поскольку рассматривается некоторый меченый терминал, зона его покрытия

имеет площадь $S_c = \int_0^{2\pi} \frac{\rho^2(\alpha)}{2} d\alpha$, а внутри нее могут возникнуть интерферирующие устройства.

Распределение расстояний до меченой ПТД. Можно заметить, что наличие коллизий между маяками (когда терминал «слышит» другую ПТД) зависит от расстояния между ПТД и подключенными к ним устройствами, тогда как коллизии между фрагментами данных (когда ПТД «слышит» другую ПТД или ее терминалы) определяются расстоянием между ПТД или между ПТД и устройствами. Для упрощения этих взаимоотношений и ухода от явного моделирования граничных эффектов введенная модель дополняется предположением, что со стороны других сетей PBSS как ПТД, так и подключенные к ней терминалы расположены в одной и той же географической точке. В принципе данное допущение можно заменить на другое, согласно которому устройства размещаются случайным образом внутри определенной области вокруг ПТД, как обсуждается ниже.

Благодаря усреднению можно также исключить из рассмотрения случаи, когда терминалы находятся в рассматриваемой зоне обслуживания без соответствующих им ПТД или наоборот, когда не все устройства некоторой ПТД находятся внутри этой зоны обслуживания. Другими словами, для каждой ПТД и любой зоны обслуживания предполагается, что все ее M терминалов также размещаются внутри нее. Данное допущение касается исключительно взаимного расположения ПТД и ее устройств. Случайная направленность передачи терминалов может, тем не менее, привести к ситуации, когда меченый терминал «слышит» только часть устройств некоторой удаленной сети PBSS.

Далее характеризуется распределение расстояний между меченой ПТД и всеми остальными ПТД внутри зоны покрытия. С учетом равномерного размещения ПТД внутри этой зоны плотность вероятности представляет собой $f(x, y) = \frac{1}{S_c}$, $x, y \in S_c$. Тогда совместную функцию распределения можно выразить как

$$\begin{aligned} F(r, \alpha) &= \Pr\{z < r, \beta < \alpha\} = \int_{z < r, \beta < \alpha} f(x, y) z d\beta dz = \\ &= \frac{1}{S_c} \int_0^{\alpha} \int_0^{\min(\rho(\beta), r)} z dz d\beta = \frac{1}{S_c} \int_0^{\alpha} \frac{\min^2(\rho(\beta), r)}{2} d\beta, \end{aligned} \quad (\text{Г.12})$$

где $S_c = \pi(1 - \gamma) R_1^2 + \gamma\pi R_2^2$ – площадь зоны покрытия. Следовательно, маргинальная функция распределения, характеризующая расстояния между меченой

ПТД и другими ПТД внутри зоны покрытия, может быть записана как

$$F_r(r) = \frac{1}{S_c} \int_0^{2\pi} \frac{\min^2(\rho(\beta), r)}{2} d\beta. \quad (\text{Г.13})$$

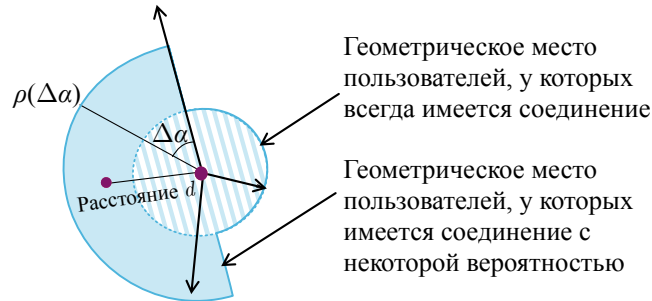


Рисунок Г.3 — Иллюстрация геометрических свойств сети

Для рассматриваемого частного случая можно записать

$$F_r(r) = \begin{cases} \frac{r^2}{(1-\gamma)R_1^2 + \gamma R_2^2}, & \text{если } r \leq R_1, \\ \frac{(1-\gamma)R_1^2 + \gamma r^2}{(1-\gamma)R_1^2 + \gamma R_2^2}, & \text{если } R_1 < r \leq R_2, \end{cases} \quad (\text{Г.14})$$

а функция плотности вероятности тогда представляет собой

$$f_r(r) = \begin{cases} \frac{2r}{(1-\gamma)R_1^2 + \gamma R_2^2}, & \text{если } r \leq R_1, \\ \frac{2\gamma r}{(1-\gamma)R_1^2 + \gamma R_2^2}, & \text{если } R_1 < r \leq R_2. \end{cases} \quad (\text{Г.15})$$

При этом среднее число возможных интерферирующих устройств составляет $E[N_{all}] = \mu S_c$.

Вероятность наличия соединения как функция расстояния d

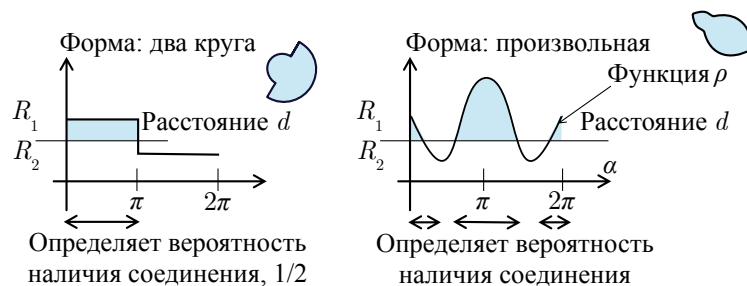


Рисунок Г.4 — Пояснение для вероятности наличия соединения

Вероятность наличия соединения и доля активных соединений.

Вероятность того, что соединение между двумя ПТД существует, может быть

получена из функции $\rho(\Delta\alpha_s)$ и представляет собой выражение $\Pr\{\rho(\alpha) \geq d\} = \Psi(d)$ для равномерно распределенной случайной величины $\alpha \in [0; 2\pi]$. Основываясь на вероятности наличия активного соединения $\psi(r)$ и усредняя по распределению расстояний, можно установить долю активных соединений как $\varepsilon = \int_r \psi(r) dF_r(r)$. В результате среднее число маяков, которое «слышит» меченая ПТД, составляет

$$E[N_b] = \varepsilon E[N_{all}] = \mu S_c \int_r \psi(r) dF_r(r), \quad (\text{Г.16})$$

что задает среднее число интерферирующих маяков.

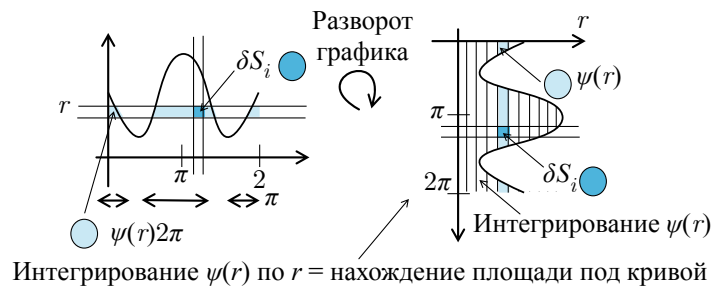


Рисунок Г.5 — Вычисление доли $\int_r \psi(r) dF_r(r)$

Поскольку функция $\psi(r)$ выражается через некоторую функцию, задающую форму зоны покрытия $\rho(\delta\alpha)$, то выражение $\int_r \psi(r) dF_r(r)$ можно вычислить непосредственно даже при отсутствии точного выражения для $\psi(r)$. Таким образом, $\psi(r)$ представляет собой долю азимутов $\delta\alpha$, которые располагаются внутри зоны покрытия радиуса свыше r . Тогда, обозначая площадь рассматриваемой зоны через S , можно записать следующее:

$$\begin{aligned} \varepsilon &= \int_r \psi(r) dF_r(r) = \int_S dF_r(r) = \frac{1}{2\pi} \int_0^{2\pi} \left(\int_0^{\rho(\alpha)} dF_r(r) \right) d\alpha = \\ &= \frac{1}{2\pi} \int_0^{2\pi} F_r(\rho(\alpha)) d\alpha. \end{aligned} \quad (\text{Г.17})$$

Далее, используя данное выражение, можно получить итоговое выражение для среднего числа интерферирующих терминалов

$$E[N_b] = \varepsilon E[N_{all}] = \frac{\mu}{4\pi} \int_0^{2\pi} \left(\int_0^{2\pi} \min^2(\rho(\beta), \rho(\alpha)) d\beta \right) d\alpha. \quad (\text{Г.18})$$

Альтернативный подход к моделированию особенностей радиоканала и отражений от тела человека заключается в непосредственном анализе убывающего значения вероятности наличия соединения с учетом выбранной виртуальной диаграммы направленности $\varphi(r) = \text{Pr}\{\text{соединение существует на расстоянии } r \mid \text{точка попадает внутрь виртуальной диаграммы}\}$. Тогда общая вероятность существования соединения на расстоянии r составляет $\varphi(r)\psi(r)$ и, тем самым, $\varepsilon = \int_r \varphi(r)\psi(r)dF_r(r)$. Следует отметить, что каждая ПТД из их случайного числа N_{all} имеет в точности M устройств, не все из которых интерферируют. Поскольку расстояние r_0 предполагается здесь пренебрежимо малым, каждый из M терминалов располагается на том же расстоянии от меченой ПТД, что и его ПТД.

Фактическое число интерферирующих устройств представляет собой биномиально распределенную случайную величину с параметром $\psi(r)$ за счет случайного азимута. Тогда соответствующее математическое ожидание, если ПТД расположена на расстоянии r , составляет $M\varphi(r)\psi(r)$. Важно подчеркнуть, что рассматриваемая модель уже является аппроксимацией реальной системы, а вероятность успеха также была получена приближенно в целях снижения вычислительной сложности, поэтому можно продолжить данный подход и предложить аппроксимацию, основанную на числе устройств, усредненном по площади зоны обслуживания. В этой связи можно сформулировать следующее утверждение.

Утверждение. Число устройств, которые может «слышать» некоторая меченая ПТД, усредненное по площади зоны обслуживания, составляет $E[N_d] = ME[N_b]$.

Доказательство основано на рассуждениях, аналогичных тем, которые использовались при вычислении $E[N_b]$, и для краткости здесь опускается.

Исключая ситуации, связанные с наличием устройств без своих ПТД или ПТД с меньшим числом устройств, далее предполагается, что число терминалов описывается соотношением $E[N_d]/E[N_b] = M$, так что все интерферирующие устройства оказываются распределенными между интерферирующими ПТД. В итоге, учитывая результаты, полученные при вычислении $E[N_b]$, и подставляя их в сформулированные следствия, можно получить аппроксимацию для основных параметров работы изучаемой системы и далее использовать их при оценке общесистемных показателей энергетической эффективности. Для рассматриваемого здесь частного случая, описываемого функцией $\rho(\Delta\alpha_s)$, когда $\varphi(r) = 1$,

соответствующее выражение для вероятности наличия соединения $\psi(r)$ может быть представлено как

$$\psi(r) = \begin{cases} 1, & \text{если } r \leq R_1, \\ \gamma, & \text{если } R_1 < r \leq R_2. \end{cases} \quad (\text{Г.19})$$

Тогда среднее число интерферирующих маяков в данном частном случае составит

$$E[N_b] = \mu S_c \int_0^{R_1} dF_r(r) + S_c \gamma \mu \int_{R_1}^{R_2} dF_r(r). \quad (\text{Г.20})$$

Используя данное выражение, можно получить

$$E[N_b] = \mu S_c \left[(1 - \gamma) \frac{R_1^2}{(1 - \gamma)R_1^2 + \gamma R_2^2} + \gamma \right]. \quad (\text{Г.21})$$

В качестве альтернативного подхода можно записать

$$E[N_b] = \mu S_c \frac{1}{2\pi} \left[\int_0^{2\pi(1-\gamma)} \frac{R_1^2}{(1-\gamma)R_1^2 + \gamma R_2^2} d\alpha + \int_{2\pi(1-\gamma)}^{2\pi} \frac{(1-\gamma)R_1^2 + \gamma R_2^2}{(1-\gamma)R_1^2 + \gamma R_2^2} d\alpha \right] = \mu S_c \left[\frac{(1-\gamma)R_1^2}{(1-\gamma)R_1^2 + \gamma R_2^2} + \gamma \right], \quad (\text{Г.22})$$

что приводит к искомому выражению.

Г.1.4 Расширения предложенной модели

В заключение обсуждаются некоторые полезные практические расширения разработанной модели. Во-первых, возможно рассмотреть более сложные модели канала связи, которые более полно характеризовали бы соответствующие эффекты распространения радиосигнала. Таким образом, предложенный подход к моделированию mmWave-системы носимых устройств можно обобщить для множества перекрывающихся зон покрытия, каждая из которых имеет свой уровень мощности принятого сигнала, что соответствует затуханию радиосигнала с увеличением расстояния. Кроме того, подобная форма представления позволяет учесть эффект так называемого захвата по мощности, когда из нескольких радиосигналов, имеющих сильно различающиеся уровни мощности, будет успешно принят тот, мощность которого максимальна.

Предложенный выше подход также позволяет учесть тот факт, что устройства располагаются на случайных расстояниях от своих ПТД. Предполагая, что такие расстояния подчиняются некоторой функции распределения $g_r(r), r \geq r_0$ (т.е. когда достаточно малые расстояния существенно не влияют на результирующую скорость передачи данных). Данное соображение, в свою очередь, ведет к изменению распределения расстояний между ПТД и интерферирующими терминалами, которое определяет интенсивность коллизий «данные – данные» на стороне меченой ПТД. При этом интенсивность как коллизий «маяк – данные», так и коллизий «маяк – маяк» на стороне устройства некоторой меченой сети PBSS остаются неизменными. Тогда дополнительное распределение расстояний $f_r(r)$ можно найти путем рассмотрения случайных треугольников.

Наконец, важным расширением является возможность использования статистических данных, полученных при помощи имитатора распространения радиосигнала, который в явном виде моделирует фактические местоположения ПТД и подключенных к ним устройств, углы раскрытия диаграммы направленности, а также физические явления, связанные с распространением mmWave-сигналов (включая отражение, рассеяние и дифракцию). Это позволяет более точно характеризовать эффекты, связанные с блокированием радиосигнала телом человека. В результате можно получить вид функций $\rho(\Delta\alpha)$ и $\xi(r)$, а также наблюдать их слабую зависимость от плотности размещения ПТД μ , тем самым непосредственно моделируя блокирование mmWave-соединения всевозможными препятствиями, расположенными между взаимодействующими узлами. Такое расширение может компенсировать отсутствие детальной модели препятствий, использующей методы стохастической геометрии [357]. Два дополнительных обобщения предложенного подхода описаны ниже.

Реалистичные диаграммы направленности. В среде распространения радиоволн на практике более целесообразно использовать антенны с диаграммой направленности, отличной от круговой. Поскольку с mmWave-сигналом могут взаимодействовать различные объекты, итоговая зона покрытия принимает, как правило, крайне сложную форму. При этом предложенный выше аналитический подход является достаточно общим и позволяет принять во внимание множество вариаций формы такой зоны покрытия. В частности, возможно использование ряда полиномиальных функций для задания требуемой формы зоны покрытия, которые могут как отражать результаты практических

измерений (например, при помощи подходящей аппроксимации), так и соответствовать модели некоторой сложной антенной решетки.

Далее предлагается рассмотреть линейную фазированную антенную решетку, для которой способ распространения радиосигнала может быть задан, например, в виде полиномиальной функции $\rho(x) = f_n(x) = \sum_{i=0}^{n/2} a_i x^i > 0$ для $x \in [-\pi, \pi]$. Такой полином определяет форму соответствующей зоны покрытия, а величина $\frac{n}{2}$ характеризует число лепестков антенны. Таким образом, среднее число интерферирующих между собой ПТД определяется в виде

$$E[N_b] = \frac{\mu}{4\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \min^2 \left[\sum_{i=0}^{n/2} a_i \alpha^i, \sum_{i=0}^{n/2} a_i \beta^i \right] d\beta d\alpha, \quad (\text{Г.23})$$

которое может быть решено численно, а в ряде случаев даже выписано в замкнутом виде. Некоторые примеры зоны покрытия, описываемой полиномиальной функцией, а также соответствующие выражения для радиуса $\rho(\alpha)$ приведены на рисунке Г.6.

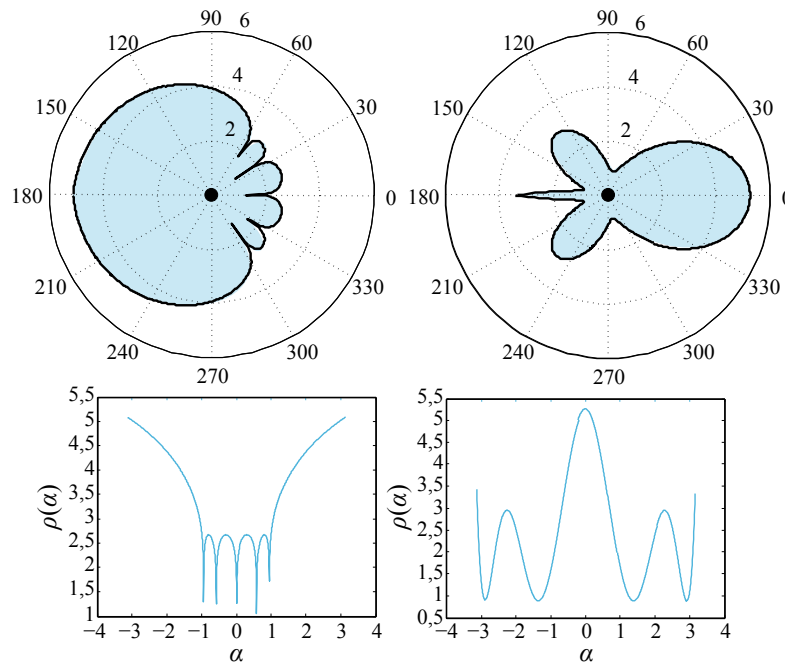


Рисунок Г.6 — Примеры зоны покрытия и функции $\rho(\alpha)$ для различных диаграмм направленности

Точечные процессы, отличные от пуассоновского. Следует также отметить, что исходное предположение о расположении узлов сети согласно пуассоновскому точечному процессу RPP на плоскости предполагает отсутствие

взаимосвязей в их размещении, что может быть уточнено для конкретного сценария использования носимых устройств, имеющего определенную зависимость. Помимо RPP для моделирования местоположения ПТД можно в качестве альтернативы использовать процессы твердого ядра [358; 359], которые не позволяют двум произвольно выбранным точкам находиться на расстоянии ближе, чем некоторое минимальное значение δ . Возможный способ применения данной альтернативы состоит в ограничении распределения расстояний $f_r(r)$ и последующем масштабировании плотности размещения ПТД пропорционально некоторой функции величины δ .

Подобный подход потребует дополнительного изучения вопроса, насколько близко окажется результирующая аппроксимация к аналогичным практическим значениям (здесь можно принять во внимание работу [358], в которой проведено схожее исследование интерференции для процессов Матерна). Однако здесь важно подчеркнуть, что поскольку в условиях плотного размещения две ПТД могут быть расположены сколь угодно близко друг к другу, использование процесса RPP представляется достаточно удовлетворительным.

Г.2 Анализ энергетической эффективности сети

В данном разделе на основании описанной ранее модели проводится оценка общесистемного значения энергопотребления и вычисляется соответствующий показатель энергетической эффективности сети, которая измеряется числом бит переданной информации в расчете на Джоуль затраченной энергии.

Для получения величины энергопотребления mmWave-сеть носимых устройств далее описывается в терминах так называемого энергетического отклика, представляющего собой величину энергопотребления системы в целом в ответ на то или иное атомарное событие в канале связи. При рассмотрении некоторой меченой ПТД и подключенных к ней устройств записывается значение энергетического отклика для всех событий, связанных с функционированием как ПТД, так и отдельно взятого устройства. В этой связи и по аналогии с [360] вычисляется значение мощности, потребляемой в расчете на один кадр с учетом всех необходимых расходов на обработку сообщений. Кроме того,

принимается во внимание специфика оборудования, работающего в диапазоне 60 ГГц, на основе характеристик усилителя мощности из [289].

В отношении отдельной ПТД целесообразно различать между собой два взаимоисключающих события: (1) успешная отправка маяка, связанная с потреблением энергии E_1 на его передачу e_{BTx} и прием e_{BRx} на всех устройствах данной сети PBSS, а также на всех соседних ПТД, предпринимающих попытку декодирования передаваемого маяка и (2) коллизия маяка, связанная с потреблением энергии E_2 на его передачу e_{BTx} и длительность простоя $e_{B,idle}$ (предполагая, что принимающее оборудование остается активно в течение всего времени неуспешного декодирования) для всех подключенных устройств сети, ожидающих доставку маяка.

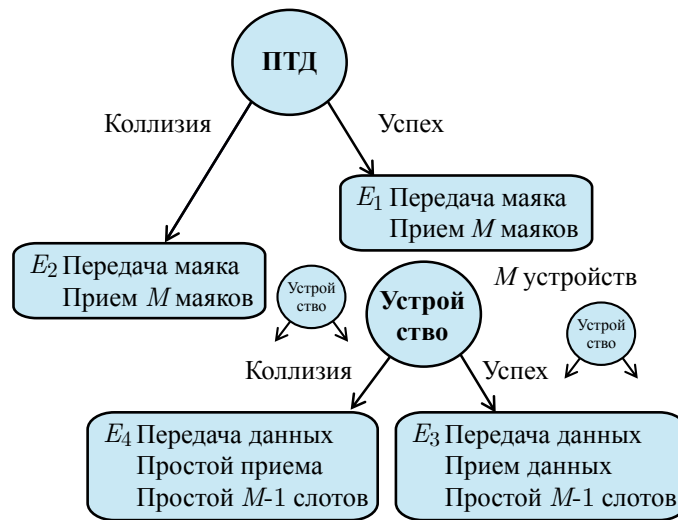


Рисунок Г.7 — Взаимосвязь между возможными событиями, связанными с энергопотреблением устройства

С другой стороны, похожая пара событий возникает при рассмотрении процесса относительно передаваемых сообщений. В частности, некоторое устройство вызывает в системе: (1) потребление энергии E_3 , которое представляет собой сумму компонентов, затрачиваемых на передачу e_{DTx} и на прием e_{DRx} данных на стороне ПТД; и (2) потребление энергии E_4 , которое также связано с передачей данных e_{DTx} , но включает в себя дополнительный компонент энергии $e_{D,idle}$ в течение времени простоя, когда ПТД прослушивает канал. Приведенное выше описание, представленное на рисунке Г.7, транслируется в

следующий набор выражений:

$$\begin{aligned}
E_1 &= e_{BTx} + Me_{BRx}, \\
E_2 &= e_{BTx} + Me_{B,idle}, \\
E_3 &= e_{DTx} + e_{DRx} + (M-1)e_{idle}, \\
E_4 &= e_{DTx} + e_{D,idle} + (M-1)e_{idle}.
\end{aligned} \tag{Г.24}$$

Если в канале связи отсутствует передача как маяков, так и фрагментов данных, то ПТД или подключенное к ней устройство потребляет e_{idle} , что в рамках всей сети PBSS приводит к вкладу $E_0 = (M+1)(L-1-KM)e_{idle}$. Основываясь на первом следствии, общее энергопотребление на PBSS в суперкадре можно определить как

$$\begin{aligned}
E_I &= E_0 + (1 - E[n_B])E_2 + E[n_B]E_1 + E[n_D]E_3 + \\
&+ (E[n_B]KM - E[n_D])E_4 + (1 - E[n_B])KME_{idle},
\end{aligned} \tag{Г.25}$$

а если обозначить число неуспешных слотов на формирование диаграммы направленности через $E[fBF]$, то на основании второго следствия возможно записать

$$\begin{aligned}
E_{II} &= E_0 + (1 - E[n_B])E_2 + E[n_B]E_1 + E[n_D]E_3 + \\
&+ (E[n_B]KM - E[n_D] - E[fBF] - E[n_{BF}])E_4 + \\
&+ E[n_{BF}]E_3 + E[fBF]E_4 + (1 - E[n_B])KME_{idle},
\end{aligned} \tag{Г.26}$$

где $E[n_B] = e^{-p}$ для обеих систем, тогда как $E[fBF] = E[n_B] \frac{M}{q} (1 - (1 - q)^K)$, а q было определено выше.

Наконец, с учетом мгновенной скорости передачи данных устройством R_{data} (Мбит/с), можно выразить энергетическую эффективность на PBSS для систем Тип-1 и Тип-2, которая определяется как количество бит, успешно переданных в расчете на Джоуль затраченной энергии

$$\eta_I = \frac{Me^{-2\frac{N}{L}} \left(1 - \frac{N}{L} e^{-\frac{N}{L}}\right)^{KM-1} R_{data} \Delta T_s}{E_I}, \tag{Г.27}$$

а также

$$\eta_{II} = \frac{(E[n_D] + (1 - \xi)E[n_{BF}])R_{data}\Delta T_s}{E_{II}}. \tag{Г.28}$$

Здесь $E[n_D]$ и $E[n_{BF}]$ приведены во втором следствии; ΔT_s – длительность слота, а $R_{data}\Delta T_s$ – объем данных, передаваемых за один слот.

Следует отметить, что полученные здесь выражения позволяют сформулировать множество оптимизационных задач на основе входящих в них параметров. Наиболее естественной из них становится находение такой длины суперкадра, которая являлась бы оптимальной в плане энергетической эффективности при определенных значениях плотности размещения узлов, некотором виде

диаграммы направленности и других заданных параметрах. Поскольку выбор длины суперкадра в значительной степени влияет на интенсивность коллизий между сообщениями соседних персональных сетей, такая постановка задачи может быть привлекательной при проведении оптимизации системы связи на практике.

Приложение Д

Анализ применения сетевого кодирования в системе D2D с сетевым содействием

Д.1 Моделирование системы распределения контента

В данном приложении при помощи методов жидкостной аппроксимации исследуется процесс распределения контента, описанный в главе 5. Для сохранения аналитичности отделяются друг от друга два процесса: жизненный цикл целевого контента и его доставка получателям. В частности, рассматривается динамика запросов на загрузку популярного контента. Важно отметить, что в изучаемой системе может не существовать стационарного состояния, поскольку время жизни единицы контента может быть недостаточно длительным для его достижения. Таким образом, целью дальнейшего изложения становится анализ динамики обозначенных процессов.

Для учета ключевых особенностей рассматриваемой системы в переходном режиме исходные дискретные процессы, связанные с числом пользователей, получающих некоторый целевой контент, заменяются их непрерывным аналогом. Для такого непрерывного процесса представляют интерес следующие отдельные фазы: (1) интервал времени, когда единица контента остается востребованной ($t \in [0; T]$), и (2) интервал времени, когда она перестает быть востребованной и покидает систему ($t \in [T; \infty)$). Первая фаза получает здесь название *фазы роста*, тогда как вторая именуется *фаза сокращения*. Как следствие, стационарное состояние ($t \rightarrow \infty$) достигается, когда система полностью опустошается.

Обе фазы описываются непрерывными процессами, для чего вводятся соответствующие системы дифференциальных уравнений с заданными граничными условиями. В результате решения таких систем дифференциальных уравнений устанавливаются замкнутые выражения, представляющие собой аппроксимацию для числа получивших фрагменты контента абонентов, а также для числа абонентов, успешно восстановивших целевой контент как при базовом, так и при кодированном способе его доставки. Следует подчеркнуть, что дан-

ная система достаточно сложна для непосредственного анализа, поэтому далее предлагается простой и практичный подход к ее моделированию, который предполагает разбиение общего интегрированного процесса на две составные части. Они характеризуют жизненный цикл целевого контента и его доставку получателям соответственно, и ниже рассматривается первая из них.

Д.1.1 Жизненный цикл контента

Система принадлежит к типу $M/M/\infty$ при условии пуассоновского характера поступления единиц контента и их экспоненциально распределенного времени нахождения в системе. Поэтому распределение числа единиц контента в такой системе описывается общеизвестным выражением, которое приведено в литературе по теории массового обслуживания. Можно предположить, что рассматриваемая простейшая система функционирует достаточно долгое время, чтобы достичь своего стационарного состояния. Тогда среднее количество имеющихся в ней одновременно различных единиц контента подчиняется выражению $E[N_c] = \lambda_c / \mu_c$. Следует также заметить, что приведенная здесь модель системы массового обслуживания может быть обобщена для более общего класса практических сценариев.

Вслед за описанием среднего числа одновременно существующих единиц активного контента, исследуется динамика процессов, связанных с некоторой типовой единицей контента. В частности, изучается существенно более сложный процесс доставки соответствующей меченой единицы контента. В этой связи принимается во внимание тот факт, что всего имеется $E[N_c]$ таких процессов, которые протекают в системе одновременно, т.е. интегрированная сеть связи делит свои радиоресурсы между $E[N_c]$ пользовательскими сессиями в среднем.

Д.1.2 Доставка контента получателям

Важно отметить, что в течение времени жизни некоторой меченой единицы контента заинтересованные в ее получении пользователи активируются

в произвольных местоположениях, так что скорость передачи данных между абонентским терминалом и обслуживающей его сотовой БС представляет собой случайную величину, имеющую вид распределения

$$f_r(r) = \begin{cases} \frac{2e^{\frac{r}{w_m}}}{\kappa_m w_m R^2} \left(\frac{p_{BS} G_m}{N_0 + I} \right)^{\frac{2}{\kappa_m}} \left[e^{\frac{r}{w_m}} - 1 \right]^{-\frac{2}{\kappa_m}}, & 0 \leq r \leq r_{\text{lim}}, \\ \frac{1}{R^2} \left(\frac{p_{BS} G_m}{(N_0 + I)(e^{\frac{r}{w_m}} - 1)} \right)^{\frac{2}{\kappa_m}}, & r = r_{\text{lim}}, \end{cases} \quad (\text{Д.1})$$

которое может быть легко получено из соответствующего распределения расстояний с учетом предельно достижимой скорости передачи данных. Здесь p_{BS} – уровень излучаемой мощности на стороне БС; w_m – полоса пропускания канала связи, а G_m и κ_m представляют собой параметры, определяющие процесс распространения радиосигнала. Следует заметить, что подобный подход в принципе применим для описания распределения скорости передачи данных в пределах максимальной дальности работы соединений D2D, но на практике $R_0 < d_0$ и для наглядности можно предположить, что скорость передачи данных r_0 сохраняется постоянной для всех D2D-партнеров.

Число пользователей, получающих в некоторый момент времени требуемую единицу контента, обозначается через $x(t)$, тогда как число пользователей, которые уже загрузили этот контент, обозначено как $y(t)$. Тогда для любого активного абонентского терминала из общего их числа $x(t)$ потенциальные D2D-партнеры расположены в радиусе R_0 от него. Для сохранения аналитичности также принимается во внимание среднее число соседей и предполагается, что меченый терминал наблюдает некоторую постоянную плотность расположенных поблизости терминалов-соседей, которая является типовой для рассматриваемой сети в целом. Соответственно, в среднем такой терминал может наблюдать $\frac{R_0^2}{R^2} [x(t) + y(t)]$ вероятных D2D-партнеров, а также сохранять канал связи со своей сотовой БС.

Применяя подход из работы [320], также вводится параметр η , характеризующий эффективность распределения контента. Тогда в предположении об идентичных пользователях итоговая скорость передачи определяется через индивидуальную скорость скачивания доли успешно полученного контента $\eta \cdot \frac{R_0^2}{R^2} x(t) + \frac{R_0^2}{R^2} y(t)$ вблизи рассматриваемого абонентского терминала. При этом во введенной модели скорость передачи данных поверх соединений D2D зависит от расстояния между их источником и получателем как $\mu_{i,j} = r_{i,j} / (Z \cdot E[N_c])$,

где Z – размер единицы контента; $E[N_c]$ – число одновременно активных процессов скачивания контента; $r_{i,j}$ – мгновенная скорость передачи данных.

В результате общая скорость получения контента поверх D2D и сотовых соединений соответственно составляет

$$\begin{aligned} r_{D2D} &= \frac{r_0}{Z} \left[\eta \cdot \frac{R_0^2}{R^2} x(t) + \frac{R_0^2}{R^2} y(t) \right], \\ r_{cell} &= \frac{r_{i,BS}}{ZE[N_c]}, \end{aligned} \quad (Д.2)$$

где r_0 – индивидуальная скорость передачи данных по D2D-соединениям внутри радиуса действия R_0 , а $r_{i,BS}$ – скорость скачивания контента через сотовое соединение с БС, усредненная по целевой области моделирования как $\int_0^{r_{lim}} r f_r(r) dr$. В предположении, что радиоресурсы системы связи делятся в равных долях между абонентскими терминалами как в сети D2D, так и в сотовой сети (благодаря функционированию «справедливого» алгоритма планирования в системе 3GPP LTE и механизмам случайного доступа при работе технологии WiFi-Direct), общая скорость скачивания через D2D в системе в целом пропорционально снижается как $r_0 \frac{R_0^2}{R^2} [\eta x(t) + y(t)]$.

Также важно напомнить, что время жизни единиц контента является достаточно коротким, поэтому процесс их распределения зачастую не может достичь своего стационарного состояния. Таким образом, далее изучается совместное развитие во времени процессов $x(t)$ и $y(t)$ (в переходном режиме) в предположении, что $\theta < \mu(1 - \eta)$. Ниже рассматриваются две фазы данных процессов: фаза роста в течение времени жизни контента $t \in [0; T]$, когда у абонентов возникают новые запросы на его скачивание, и фаза сокращения $t > T$, когда пользователи теряют интерес в получении этого контента и динамика обоих процессов спадает к нулевому уровню.

Целью последующего изложения является нахождение замкнутых выражений, которые задают динамику процессов $x(t)$ и $y(t)$ посредством решения двух вариаций для соответствующих систем дифференциальных уравнений, описывающих работу базового и кодированного режимов распределения контента:

$$\begin{cases} dx(t) = \lambda dt - \theta x(t) dt - \mu \eta x(t) - \mu y(t) dt - \mu_0 dt, \\ dy(t) = \mu \eta x(t) dt + \mu y(t) dt + \mu_0 dt - \theta y(t) dt, \\ \text{для начальных условий: } x(t_0) = x_0, \quad y(t_0) = y_0, \end{cases} \quad (Д.3)$$

где $\mu = \frac{r_0 R_0^2}{Z R^2}$ является константой, введенной для краткости, а $\mu_0 = \int_0^{r_{lim}} r f_r(r) dr / (Z \cdot E[N_c])$ представляет собой скорость передачи данных, усред-

ненную по области моделирования. Начальные условия x_0 и y_0 определяются фазой процесса.

Полезно подчеркнуть, что процесс $y(t)$ в обоих случаях развивается из точки $t_0 = Z/r_{BS}$, где первый пользователь из множества x завершает успешную загрузку контента. До этого момента динамику процесса x можно описать выражением $dx(t) = \lambda dt - \theta x(t)dt$ как

$$x(t) = \frac{\lambda}{\theta} (1 - e^{-\theta t}), \quad t \in [0; t_0], \quad (\text{Д.4})$$

которое отражает линейный характер роста λt при $\theta \rightarrow 0$. Тем самым для каждого режима распределения контента рассматриваемый процесс изучается на интервале $[t_0; T]$ или $[T; \infty)$. Начальные условия определяются при этом следующим образом: $x(0) = 0$, $y(t_0) = 1$ и $x(t_0) = x(t)$.

Д.1.3 Базовый режим распределения контента

При использовании базового (некодированного) режима (BT) получение единицы контента характеризуется сбором ее S фрагментов. Такой процесс распределения имеет достаточно низкую эффективность $\eta < 1$, что часто приводит к ситуации, когда D2D-соседи получают одни и те же фрагменты контента и далее не могут воспользоваться преимуществами прямых соединений. Более детальное обсуждение параметра η для популярных пиринговых систем BitTorrent приведено в [320], где предложена аппроксимация для вычисления его значения в виде функции от числа фрагментов.

Утверждение. В базовом режиме рассматривается динамика процессов изменения числа пользователей $x(t)$ и $y(t)$, которые заинтересованы в загрузке некоторого целевого контента C , имеющего время жизни T , и получили его частично или целиком соответственно.

1. Динамика данных процессов в течение фазы роста $t \in [t_0; T]$ характеризуется решением задачи Коши для начальных условий $x(t_0) = x_0$, $y(t_0) = y_0$ в виде

$$\begin{cases} x = \tilde{c}_1 e^{-\theta(t-t_0)} + \tilde{c}_2 e^{(\mu(1-\eta)-\theta)(t-t_0)} + c_3, \\ y = -\frac{\eta(\mu_0\theta + \lambda\mu)}{\theta(\mu(1-\eta)-\theta)} - \eta\tilde{c}_1 e^{-\theta(t-t_0)} - \tilde{c}_2 e^{(\mu(1-\eta)-\theta)(t-t_0)}, \end{cases} \quad (\text{Д.5})$$

где величины t_0 , x_0 и y_0 определяются из начальных условий, а c_3 задается как $\frac{\mu_0\theta + \lambda(\mu - \theta)}{\theta\mu(1-\eta) - \theta^2}$. Также

$$\begin{cases} \tilde{c}_1 = \frac{-\lambda/\theta + x_0 + y_0}{(1-\eta)}, \\ \tilde{c}_2 = \left(\frac{\lambda/\theta - \eta x_0 - y_0}{(1-\eta)} - c_3 \right). \end{cases}$$

2. Динамика данных процессов в течение фазы сокращения $t \in [T; \infty)$ характеризуется решением задачи Коши для начальных условий $x(T) = x_T, y(T) = y_T$, где x_T и y_T определяются как

$$\begin{cases} x = \frac{(\tilde{x}_T + \tilde{y}_T)}{(1-\eta)} e^{-\theta(t-T)} - \frac{\tilde{y}_T + \eta \tilde{x}_T}{(1-\eta)} e^{\theta_1(t-T)} + \frac{\mu_0}{\theta_1}, \\ y = -\eta \frac{(\tilde{x}_T + \tilde{y}_T)}{(1-\eta)} e^{-\theta(t-T)} + \frac{\tilde{y}_T + \eta \tilde{x}_T}{(1-\eta)} e^{\theta_1(t-T)} - \frac{\mu_0}{\theta_1}, \end{cases} \quad (\text{Д.6})$$

где $\theta_1 = (\mu(1-\eta) - \theta)$, $\tilde{x}_T = x_T - \frac{\mu_0}{\theta_1}$ и $\tilde{y}_T = y_T + \frac{\mu_0}{\theta_1}$.

Доказательство первой и второй частей данного утверждения приведено ниже в соответствующих подразделах.

Д.1.4 Кодированный режим распределения контента

При применении рассматриваемой системой режима с сетевым кодированием (НС) сотовая БС распределяет линейные комбинации фрагментов контента так, что соседние абонентские терминалы располагают полезными друг для друга данными с существенно более высокой вероятностью, что позволяет принять $\eta = 1$. Таким образом, анализ данной системы сводится к формулированию следующего утверждения.

Утверждение. В кодированном режиме рассматривается динамика процессов изменения числа пользователей $x(t)$ и $y(t)$, которые заинтересованы в загрузке некоторого целевого контента C , имеющего время жизни T , и получили его частично или целиком соответственно.

1. Динамика данных процессов в течение фазы роста $t \in [t_0; T]$ характеризуется решением задачи Коши для начальных условий $x(t_0) = x_0, y(t_0) = y_0$ в виде

$$\begin{cases} x = (\tilde{c}_1 + \tilde{c}_2 t) e^{-\theta(t-t_0)} - \frac{\mu_0\theta + \lambda(\mu - \theta)}{\theta^2}, \\ y = \frac{\lambda\mu + \mu_0\theta}{\theta^2} - (\eta\tilde{c}_1 + \frac{\tilde{c}_2}{\mu} + \tilde{c}_2 t) e^{-\theta(t-t_0)}, \end{cases} \quad (\text{Д.7})$$

где $\tilde{c}_1 = x_0 - c_3 + (\mu y_0 + \mu x_0 - \frac{\lambda\mu}{\theta})t_0$ и $\tilde{c}_2 = -\mu y_0 - \mu x_0 + \frac{\lambda\mu}{\theta}$. Величины t_0 , x_0 и y_0 восстанавливаются из соответствующих начальных условий.

2. Динамика данных процессов в течение фазы сокращения $t \in [T; \infty)$ характеризуется решением задачи Коши для начальных условий $x(T) = x_T, y(T) = y_T$, где x_T и y_T определяются как

$$\begin{cases} x = (\tilde{x}_T - \mu(\eta\tilde{x}_T + \tilde{y}_T)(t - T))e^{-\theta(t-T)} - \frac{\mu_0}{\theta}, \\ y = (\tilde{y}_T + \mu(\eta\tilde{x}_T + \tilde{y}_T)(t - T))e^{-\theta(t-T)} + \frac{\mu_0}{\theta}, \end{cases} \quad (\text{Д.8})$$

где $\tilde{x}_T = x_T + \frac{\mu_0}{\theta}$ и $\tilde{y}_T = y_T - \frac{\mu_0}{\theta}$.

Доказательство первой и второй частей данного утверждения приведено ниже в соответствующих подразделах.

В итоге вышеприведенные утверждения предоставляют замкнутые выражения для анализа базового режима распределения контента, а также для анализа кодированного режима в течение фаз роста и сокращения (т.е. до и после того, как контент начинает покидать систему соответственно). Полученные выражения позволяют исследовать динамику происходящих процессов при различных значениях их параметров, а также определить области стабильной работы. Однако приведенный здесь подход описывает динамику функционирования рассматриваемой системы при наличии ряда ограничений, таких как $\theta < \mu(1 - \eta)$ и $x(t), y(t) > 0$. Когда последнее из них нарушается, можно перевести число абонентов в значение $\max(x, 0)$ и переписать выражение для $dy(t) = -\theta y$ для соответствующих начальных условий.

На рисунке Д.1(а) приведено сравнение результатов предложенного выше математического подхода с данными имитационного моделирования системы в целом для некоторой отдельно взятой единицы контента в динамике при использовании базового режима распределения. Представленный график позволяет сравнить аналитическую оценку, описывающую динамику рассматриваемого процесса, со значениями, полученными при помощи имитационного моделирования в похожих условиях. Важно отметить, что представленные имитационные результаты соответствуют одной реализации наблюдаемой случайной величины без усреднения, что позволяет проиллюстрировать крайне высокую степень соответствия между ее ожидаемым и фактическим значением.

Основываясь на представленных результатах, можно сделать вывод о том, что разработанная математическая модель достаточно точно воспроизводит основные особенности, возникающие при распределении контента: формирование начального множества запрашивающих пользователей (сплошная линия), получение последнего фрагмента целевого контента (первая часть сплошной линии,

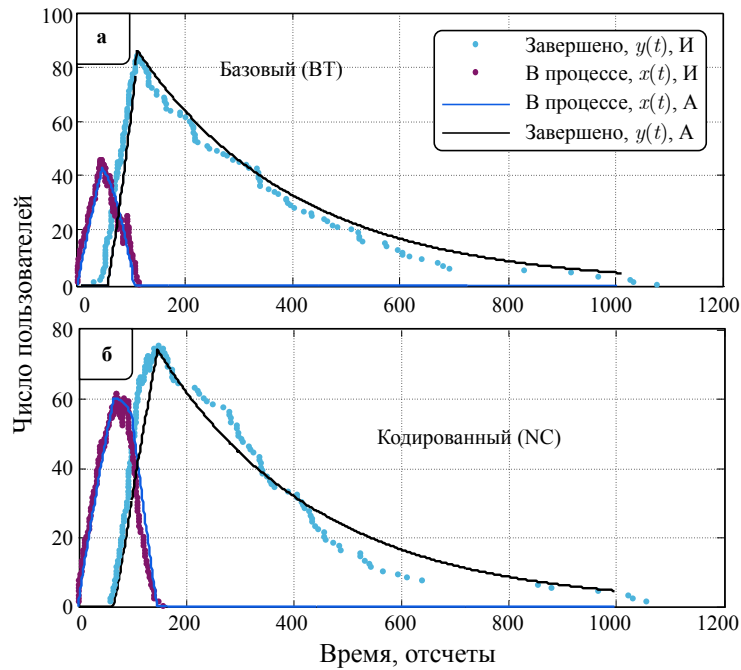


Рисунок Д.1 — Сравнение аналитических (А) и имитационных (И) результатов для режимов ВТ (а) и NC (б)

но только для режима ВТ), распределение контента, когда участники начального множества пользователей загружают последние оставшиеся фрагменты, а также процесс спада, когда абоненты начинают покидать систему. Подобным образом рисунок Д.1(б) описывает динамику процессов, происходящих в случае применения кодированного режима распределения контента.

С учетом полученных результатов можно сделать вывод о применимости предложенного подхода для анализа сложных систем распределения контента. Соответственно, работа в режиме NC характеризуется постепенным переходом в состояние успешного получения контента в отличие от функционирования в режиме ВТ. Причина подобного поведения заключается в том, что в режиме ВТ может находиться большое число пользователей (около 30 на графике), которые практически завершили загрузку своего целевого контента и ожидают его последние фрагменты, которые еще не поступили. В режиме NC такая ситуация исключена, поскольку понятие «последнего» фрагмента в принципе отсутствует и система переходит к непосредственному распределению контента гораздо раньше. Такой более ранний переход делает режим NC более желательным при высокой загрузке.

Д.2 Вспомогательные вычисления

В данном разделе приводятся вспомогательные вычисления, которые используются для получения результатов, составляющих суть сформулированных утверждений. Для наглядности последовательно решаются системы дифференциальных уравнений вначале для описания фазы роста (в течение времени жизни контента), $t \in [t_0; T]$, а затем – фазы сокращения (по мере того как контент покидает систему), $t \in [T; \infty)$. При этом предлагаются замкнутые выражения для $\eta < 1$ и $\eta = 1$ в обеих фазах. Общая структура предлагаемого подхода к решению поставленной задачи пояснена на рисунке Д.2, где динамика величин x и y определяет фазы роста и сокращения, порождая, таким образом, четыре технологические составляющие предлагаемого ниже способа вычисления.

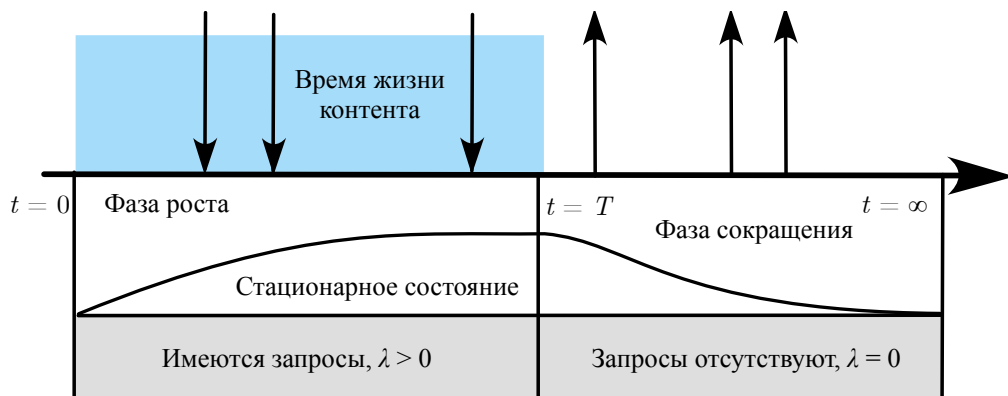


Рисунок Д.2 — Структура предложенного решения для двух фаз процесса: в течение времени жизни контента (фаза роста) и когда контент покидает систему (фаза сокращения)

Д.2.1 Описание фазы роста

Здесь рассматривается процесс распределения контента, представляющий собой совместную динамику изменения величин $x(t)$ и $y(t)$. Для нахождения искомого решения формулируется следующая задача Коши, которая рассмат-

ривается на интервале $[t_0; T]$:

$$\begin{cases} dx(t) = \lambda dt - \theta x(t)dt - \mu\eta x(t) - \mu y(t)dt - \mu_0 dt, \\ dy(t) = \mu\eta x(t)dt + \mu y(t)dt + \mu_0 dt - \theta y(t)dt, \\ \text{для начальных условий: } x(t_0) = x_0, \quad y(t_0) = y_0, \end{cases} \quad (\text{Д.9})$$

где значение $x_0 = x(t_0)$ для учета начальных условий может быть получено из соответствующего выражения.

Дальнейшее изложение приводит к решению рассматриваемой задачи при начальных условиях общего вида. Опуская индекс t для краткости, можно выразить переменную y из первого уравнения как

$$\begin{cases} y = \frac{1}{\mu} [\lambda - \theta x - \mu\eta x - x' - \mu_0], \\ y' = \mu\eta x + \mu y + \mu_0 - \theta y. \end{cases} \quad (\text{Д.10})$$

Далее записывается второе уравнение с учетом подстановки выражения для y

$$\begin{aligned} x'' + x'(\theta + \mu\eta + \theta - \mu) + x[\theta(\theta - \mu) + \theta\mu\eta] = \\ = -[\mu_0\theta + \lambda(\mu - \theta)]. \end{aligned} \quad (\text{Д.11})$$

Решение соответствующего однородного дифференциального уравнения представляет собой

$$x = c_1 e^{\alpha_1 t} + c_2 e^{\alpha_2 t},$$

где $\alpha_1 = -\theta$ и $\alpha_2 = (\mu(1 - \eta) - \theta)$ – корни характеристического уравнения.

В результате представляют интерес два отдельных случая: когда корни уравнения различны, т.е. $\alpha_1 \neq \alpha_2$, и когда они равны, т.е. $\alpha_1 = \alpha_2$.

Различные корни $\alpha_1 \neq \alpha_2$. Для этого случая искомое решение дифференциального уравнения может быть получено в виде

$$x = c_1 e^{-\theta t} + c_2 e^{(\mu(1-\eta)-\theta)t} + \frac{\mu_0\theta + \lambda(\mu - \theta)}{\theta\mu(1 - \eta) - \theta^2}.$$

Для решения задачи Коши с начальными условиями $x(0) = 0$ и $y(0) = 0$ вначале следует получить x' как

$$x' = -\theta c_1 e^{-\theta t} + (\mu(1 - \eta) - \theta)c_2 e^{(\mu(1-\eta)-\theta)t}.$$

Тогда соответствующее решение системы уравнений задается следующим образом

$$\begin{cases} x = c_1 e^{-\theta t} + c_2 e^{(\mu(1-\eta)-\theta)t} + \frac{\mu_0\theta + \lambda(\mu - \theta)}{\theta(\mu(1-\eta) - \theta)}, \\ y = -\frac{\eta(\mu_0\theta + \lambda\mu)}{\theta(\mu(1-\eta) - \theta)} - \eta c_1 e^{-\theta t} - c_2 e^{(\mu(1-\eta)-\theta)t}. \end{cases} \quad (\text{Д.12})$$

Принимая во внимание начальные условия $x(t_0) = x_0$ и $y(t_0) = y_0$, можно записать систему уравнений для коэффициентов c_1 и c_2 как

$$\begin{cases} c_1 + c_2 e^{\mu(1-\eta)t_0} = (x_0 - c_3) e^{\theta t_0}, \\ \eta c_1 + c_2 e^{\mu(1-\eta)t_0} = \left(\frac{\eta(\mu_0\theta - \lambda\mu)}{\theta(\mu(1-\eta) - \theta)} - y_0 \right) e^{\theta t_0}, \end{cases}$$

где c_3 представляет собой частное решение, приведенное выше,

$$\frac{\mu_0\theta + \lambda(\mu - \theta)}{\theta\mu(1 - \eta) - \theta^2}. \quad (\text{Д.13})$$

Тогда решение рассматриваемой системы уравнений не составляет трудностей, а искомые коэффициенты могут быть найдены в виде

$$\begin{cases} c_1 = \frac{-\lambda/\theta + x_0 + y_0}{(1-\eta)} e^{\theta t_0} = \tilde{c}_1 e^{\theta t_0}, \\ c_2 = \left(\frac{\lambda/\theta - \eta x_0 - y_0}{(1-\eta)} - c_3 \right) e^{(\theta - \mu(1-\eta))t_0} = \tilde{c}_2 e^{(\theta - \mu(1-\eta))t_0}. \end{cases} \quad (\text{Д.14})$$

Таким образом, итоговое решение можно записать в виде

$$\begin{cases} x = \tilde{c}_1 e^{-\theta(t-t_0)} + \tilde{c}_2 e^{(\mu(1-\eta)-\theta)(t-t_0)} + c_3, \\ y = -\frac{\eta(\mu_0\theta + \lambda\mu)}{\theta(\mu(1-\eta) - \theta)} - \eta \tilde{c}_1 e^{-\theta(t-t_0)} - \tilde{c}_2 e^{(\mu(1-\eta)-\theta)(t-t_0)}, \end{cases} \quad (\text{Д.15})$$

где константы \tilde{c}_1 , \tilde{c}_2 и c_3 были определены выше, а значения t_0 , x_0 и y_0 устанавливаются из начальных условий. В итоге данное выражение представляет собой искомое решение для описания фазы роста, когда $\alpha_1 \neq \alpha_2$.

Равные корни $\alpha_1 = \alpha_2$. Следует отметить, что $\alpha_1 = \alpha_2$, если и только если $\mu(1 - \eta) = 0$. Поскольку интенсивность поступления запросов в систему $\mu > 0$ по определению, данная ситуация полностью соответствует случаю линейного сетевого кодирования с $\eta = 1$. При этом коэффициент c_3 определяется как $-\frac{\mu_0\theta + \lambda(\mu - \theta)}{\theta^2}$. Тогда искомое решение уравнения можно получить в виде

$$x = (c_1 + c_2 t) e^{-\theta t} - \frac{\mu_0\theta + \lambda(\mu - \theta)}{\theta^2},$$

а соответствующая производная представляет собой

$$x' = -\theta(c_1 + c_2 t) e^{-\theta t} + c_2 e^{-\theta t}.$$

Решение системы уравнений можно записать как

$$\begin{cases} x = (c_1 + c_2 t) e^{-\theta t} + c_3, \\ y = \frac{\lambda\mu + \mu_0\theta}{\theta^2} - (c_1 + \frac{c_2}{\mu} + c_2 t) e^{-\theta t}. \end{cases} \quad (\text{Д.16})$$

Применяя начальные условия $x(t_0) = x_0$ и $y(t_0) = y_0$, можно записать систему уравнений для коэффициентов c_1 и c_2 как

$$\begin{cases} c_1 + c_2 t_0 = (x_0 - c_3) e^{\theta t_0}, \\ c_1 + \left(\frac{1}{\mu} + t_0\right) c_2 = \left(\frac{\lambda\mu + \mu_0\theta}{\theta^2} - y_0\right) e^{\theta t_0}, \end{cases} \quad (\text{Д.17})$$

что приводит к следующим значениям для искомым коэффициентов:

$$\begin{cases} c_1 = (x_0 - c_3 + (\mu y_0 + \mu x_0 - \frac{\lambda\mu}{\theta}) t_0) e^{\theta t_0} = \tilde{c}_1 e^{\theta t_0}, \\ c_2 = \left(-\mu y_0 - \mu x_0 + \frac{\lambda\mu}{\theta}\right) e^{\theta t_0} = \tilde{c}_2 e^{\theta t_0}. \end{cases} \quad (\text{Д.18})$$

Подставляя данные коэффициенты, можно получить итоговое решение для случая $\eta = 1$ в виде

$$\begin{cases} x = (\tilde{c}_1 + \tilde{c}_2 t) e^{-\theta(t-t_0)} - \frac{\mu_0\theta + \lambda(\mu-\theta)}{\theta^2}, \\ y = \frac{\lambda\mu + \mu_0\theta}{\theta^2} - (\eta\tilde{c}_1 + \frac{\tilde{c}_2}{\mu} + \tilde{c}_2 t) e^{-\theta(t-t_0)}, \end{cases} \quad (\text{Д.19})$$

где $\tilde{c}_1 = x_0 - c_3 + (\mu y_0 + \mu x_0 - \frac{\lambda\mu}{\theta}) t_0$ и $\tilde{c}_2 = -\mu y_0 - \mu x_0 + \frac{\lambda\mu}{\theta}$.

Д.2.2 Описание фазы сокращения

Здесь рассматривается ситуация, когда контент перестает быть востребованным после наступления момента времени T , и пользователи более не порождают новые запросы. Единственным дополнением, влияющим на динамику изменения наблюдаемых величин, является отсутствие новых поступлений в систему $\lambda = 0$. Таким образом, целесообразно получить решение на интервале $[T; \infty)$ с учетом новых начальных условий. Далее снова различаются случаи существования различных и равных корней характеристического уравнения.

Различные корни $\alpha_1 \neq \alpha_2$. Применяя новые начальные условия при решении данной задачи, возможно переписать рассматриваемую систему уравнений в виде

$$\begin{cases} x = c_1 e^{-\theta t} + c_2 e^{(\mu(1-\eta)-\theta)t} + \frac{\mu_0}{\mu(1-\eta)-\theta}, \\ y = -\frac{\mu_0}{\mu(1-\eta)-\theta} - \eta c_1 e^{-\theta t} - c_2 e^{(\mu(1-\eta)-\theta)t}, \\ \text{для начальных условий: } x(T) = x_T, \quad y(T) = y_T, \end{cases} \quad (\text{Д.20})$$

где x_T и y_T определяются динамикой исследуемых процессов в течение фазы роста.

По аналогии можно выразить коэффициенты c_1 и c_2 как

$$\begin{cases} (x_T - \frac{\mu_0}{\mu(1-\eta)-\theta})e^{\theta T} = \eta c_1 + c_2 e^{\mu(1-\eta)T}, \\ (y_T + \frac{\mu_0}{\mu(1-\eta)-\theta})e^{\theta T} = -\eta c_1 - c_2 e^{\mu(1-\eta)T}. \end{cases}$$

Искомые коэффициенты могут быть легко найдены следующим образом:

$$\begin{cases} c_1 = \frac{(\tilde{x}_T + \tilde{y}_T)}{(1-\eta)} e^{\theta T}, \\ c_2 = -\frac{\tilde{y}_T + \eta \tilde{x}_T}{(1-\eta)} e^{\theta T} e^{-\mu(1-\eta)T}, \end{cases}$$

где $\tilde{x}_T = x_T - \frac{\mu_0}{\mu(1-\eta)-\theta}$ и $\tilde{y}_T = y_T + \frac{\mu_0}{\mu(1-\eta)-\theta}$.

В результате решение для описания рассматриваемых процессов в течение фазы сокращения представляет собой

$$\begin{cases} x = \frac{(\tilde{x}_T + \tilde{y}_T)}{(1-\eta)} e^{-\theta(t-T)} - \frac{\tilde{y}_T + \eta \tilde{x}_T}{(1-\eta)} e^{\theta_1(t-T)} + \frac{\mu_0}{\theta_1}, \\ y = -\eta \frac{(\tilde{x}_T + \tilde{y}_T)}{(1-\eta)} e^{-\theta(t-T)} + \frac{\tilde{y}_T + \eta \tilde{x}_T}{(1-\eta)} e^{\theta_1(t-T)} - \frac{\mu_0}{\theta_1}, \end{cases} \quad (\text{Д.21})$$

где $\theta_1 = (\mu(1-\eta) - \theta)$.

Важно заметить, что приведенные выше выражения задают динамику изменения величин $x(t)$ и $y(t)$ после того, как контент начинает покидать систему, и до тех пор, пока данные величины не достигнут нулевого уровня. Начиная с момента времени, когда любая из этих величин достигнет нуля, выражение для другой меняется соответственно.

Равные корни $\alpha_1 = \alpha_2$. Наконец, рассматривается случай равных корней (т.е. $\eta = 1$) в течение фазы сокращения. Соответственно, исходная задача преобразуется в следующую систему уравнений:

$$\begin{cases} x = (c_1 + c_2 t) e^{-\theta t} - \frac{\mu_0}{\theta}, \\ y = \frac{\mu_0}{\theta} - (\eta c_1 + \frac{c_2}{\mu} + \eta c_2 t) e^{-\theta t}, \\ \text{для начальных условий: } x(T) = x_T, \quad y(T) = y_T. \end{cases} \quad (\text{Д.22})$$

С учетом начальных условий можно записать систему линейных дифференциальных уравнений для искомых коэффициентов в виде

$$\begin{cases} c_1 = x_T e^{\theta T} + \mu(x_T + y_T) e^{\theta T} T, \\ c_2 = -\mu(x_T + y_T) e^{\theta T}, \end{cases} \quad (\text{Д.23})$$

где $\tilde{x}_T = x_T + \frac{\mu_0}{\theta}$ и $\tilde{y}_T = y_T - \frac{\mu_0}{\theta}$.

В результате итоговое выражение представляет собой

$$\begin{cases} x = (\tilde{x}_T - \mu(\tilde{x}_T + \tilde{y}_T)(t - T)) e^{-\theta(t-T)} - \frac{\mu_0}{\theta}, \\ y = (\tilde{y}_T + \mu(\tilde{x}_T + \tilde{y}_T)(t - T)) e^{-\theta(t-T)} + \frac{\mu_0}{\theta}. \end{cases} \quad (\text{Д.24})$$

Таким образом, в данном приложении были последовательно получены выражения для аналитического описания процессов, происходящих в течение фазы роста (включая случаи $\eta < 1$ и $\eta = 1$), а также в течение фазы сокращения (также в случаях $\eta < 1$ и $\eta = 1$). При этом ситуация $\eta = 1$ соответствует случаю применения линейного сетевого кодирования для распределения контента в исследуемой системе.

Приложение Е

Документы, подтверждающие внедрение основных результатов диссертационной работы:

1. Российский университет дружбы народов (РУДН), Москва.
2. Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ), Санкт-Петербург.
3. Публичное акционерное общество «ГИПРОСВЯЗЬ» (ПАО «ГИПРОСВЯЗЬ»), Москва.
4. Научно-исследовательский институт радио (ФГУП НИИР), Москва.
5. Общество с ограниченной ответственностью «ЭВС» (ООО «ЭВС»), Санкт-Петербург.



**Российский университет
дружбы народов (РУДН)**

ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198
ОГРН 1027739189323; ОКПО 02066463; ИНН 7728073720

Телефон: +7495 434 53 00, факс: +7495 433 15 11
www.rudn.ru; rudn@rudn.ru

№ 17 рекогноска 20 18
0200-19-20/10

"УТВЕРЖДАЮ"

Первый проректор -
проректор по научной работе РУДН
доктор философских наук, профессор
Н.С. Кирабаев



АКТ

о внедрении результатов диссертационной работы

Андреева Сергея Дмитриевича на тему

«Разработка и исследование моделей множественного доступа и алгоритмов управления потоками трафика для гетерогенных беспроводных сетей» в исследованиях Российского университета дружбы народов

Комиссия в составе председателя Ю.В.Гайдамака, д.ф.-м.н., доцента, и членов комиссии И.А.Гудковой, к.ф.-м.н., доцента, и Э.Р.Зариповой, к.ф.-м.н., доцента кафедры прикладной информатики и теории вероятностей, установила, что научные результаты диссертации Андреева Сергея Дмитриевича «Разработка и исследование моделей множественного доступа и алгоритмов управления потоками трафика для гетерогенных беспроводных сетей», представленной на соискание ученой степени доктора технических наук были внедрены РУДН при выполнении «Плана мероприятий по реализации программы повышения конкурентоспособности («дорожная карта») РУДН (1 этап – 2016-2018 годы)» в рамках проекта 5-100.

Научные результаты диссертации С.Д.Андреева, а именно:

- модель гетерогенной сети с набором интегрированных между собой технологий радиодоступа, применимая для оценки параметров качества обслуживания;
- модель гетерогенной сети с возможностью установления прямых соединений между пользовательскими устройствами, применимая для оценки параметров качества обслуживания;
- модель сотовой сети с конкурентным доступом большого числа одновременно подключающихся устройств, применимая для изучения показателей эффективности системы связи;
- модель сети радиодоступа с конкурентным обслуживанием множества носимых устройств, применимая для оценки показателей эффективности системы радиодоступа на крайне высоких частотах

использованы при выполнении указанного выше плана мероприятий в части междисциплинарных исследований РУДН по пункту е) «Математическое и компьютерное моделирование сетей последующих поколений и высокотехнологичных систем».

Использование результатов диссертации С.Д.Андреева «Разработка и исследование моделей множественного доступа и алгоритмов управления потоками трафика для гетерогенных беспроводных сетей» позволило эффективно провести исследования по оптимизации работы механизмов в системах связи нового поколения 5G, а также создать в РУДН модельную сеть лаборатории опережающих исследований сетей 5G.

Председатель комиссии

Ю.В.Гайдамака

Члены комиссии:

И.А.Гудкова

Э.Р.Зарипова

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ
 ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
 ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
 ВЫСШЕГО ОБРАЗОВАНИЯ
 «САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
 УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ
 ИМ. ПРОФ. М.А. БОНЧ-БРУЕВИЧА»
 (СПбГУТ)

Юридический адрес: набережная реки Мойки,
 д. 61, Санкт-Петербург, 191186

Почтовый адрес: пр. Большевиков, д. 22, корп. 1,
 Санкт-Петербург, 193232
 Тел.(812) 3263156, Факс: (812) 3263159
 E-mail: rector@sut.ru
 ИНН 7808004760 КПП 784001001
 ОГРН 1027809197635 ОКТМО 40909000

25.12.2018 № _____
 на № _____ от _____

Утверждаю

Проректор по научной работе
 К.В.Дукельский



Акт

о внедрении научных результатов,

полученных Андреевым Сергеем Дмитриевичем в диссертационной работе "Разработка и исследование моделей множественного доступа и алгоритмов управления потоками трафика для гетерогенных беспроводных сетей" на соискание ученой степени доктора технических наук по специальности 05.12.13 – Системы, сети и устройства телекоммуникаций

Комиссия в составе декана факультета Инфокоммуникационных сетей и систем Л.Б.Бузюкова, заместителя заведующего кафедрой сетей связи и передачи данных М.А.Маколкиной и заведующей лабораторией кафедры сетей связи и передачи данных О.И.Ворожейкиной составила настоящий акт в том, что научные результаты, полученные в диссертации "Разработка и исследование моделей множественного доступа и алгоритмов управления потоками трафика для гетерогенных беспроводных сетей", использованы в учебном процессе Санкт-Петербургском государственном университете телекоммуникаций им. проф. М.А.Бонч-Бруевича:

1. При чтении лекций и проведении практических занятий по курсу «Самоорганизующиеся сети» (Рабочая Программа № 02.12.13/683, утверждена Первым проректором-проректором по учебной работе Г.М. Машковым 15.09.2015), раздел Программы:

Архитектура самоорганизующихся сетей связи. Сети связи пятого поколения как база для развития сетей связи. Сверхплотные сети и сети связи с ультра малой задержкой. Концепции Интернета Вещей, Тактильного Интернета и Интернета Навыков.

2. При чтении лекций и проведении практических занятий по курсу «Современные проблемы науки в области инфокоммуникаций» (Рабочая Программа № 02.12.15/783, утверждена Первым проректором-проректором по учебной работе Г.М. Машковым 21.09.2015), раздел Программы:

- Концепции развития сетей связи. Текущее состояние развития сетей. Прогнозы развития сетей связи.

В указанных дисциплинах используются следующие новые научные результаты, полученные С.Д.Андреевым в диссертационной работе:

- модель гетерогенной сети с набором интегрированных между собой технологий радиодоступа, учитывающая обслуживание потоков трафика (сессий) от множества пользователей и высокую плотность размещения узлов сетевой инфраструктуры,




- модель гетерогенной сети с двумя одновременно функционирующими технологиями радиодоступа и возможностью установления прямых соединений между соседними устройствами при высокой плотности,

- алгоритм распределения трафика пользователей в гетерогенной системе прямых соединений с большим количеством устройств, основанный на применении методов сетевого кодирования и учитывающий структуру передаваемых сессий.

Декан факультета ИКСС, канд. техн.наук, профессор

Зам. заведующего кафедрой ССиПД, канд.техн.наук

Зав. лабораторией кафедры ССиПД

 Л.Б.Бузюков
 М.А.Маколкина
 О.И.Ворожейкина

**ГИПРОСВЯЗЬ**

ОПЫТ МАСШТАБ ПЕРСПЕКТИВА

Публичное акционерное общество «ГИПРОСВЯЗЬ»

УТВЕРЖДАЮ

Заместитель генерального
директора ПАО «ГИПРОСВЯЗЬ»

А.Б. Васильев

« 24 » декабря 2018 г.

АКТ

о внедрении результатов диссертационной работы
Андреева Сергея Дмитриевича на тему
«Разработка и исследование моделей множественного доступа и алгоритмов
управления потоками трафика для гетерогенных беспроводных сетей» в
ПАО «ГИПРОСВЯЗЬ»

Настоящим актом подтверждаем, что научные результаты диссертационной работы Андреева Сергея Дмитриевича «Разработка и исследование моделей множественного доступа и алгоритмов управления потоками трафика для гетерогенных беспроводных сетей», представленной на соискание ученой степени доктора технических наук, внедрены в ПАО «Гипросвязь» при разработке «Методики оценки показателей эффективности для проектирования гетерогенных беспроводных сетей».

При разработке методики были использованы следующие новые научные результаты из диссертации С.Д. Андреева:

- Модель гетерогенной сети с набором интегрированных между собой технологий радиодоступа, которая, в отличие от известных, учитывает геометрию размещения узлов сети совместно с особенностями

обслуживания потоков трафика (сессий) пользователей во времени. Данная модель позволила определить, что вероятность блокировки сессии пользователя может быть снижена до 40% при наличии в рамках гетерогенной сети трех различных технологий радиодоступа.

- Метод управления информационными потоками, основанный на алгоритме с разделением пользовательской сессии для совмещенной сети радиодоступа, который отличается от известных тем, что позволяет вести передачу пользовательского трафика по нескольким технологиям радиодоступа одновременно с целью снижения нагрузки на сотовую сеть. Данный метод позволяет повысить скорость отправки данных пользователями на величину 50-60% при входной интенсивности 10-20 сессий в секунду.

В результате, при разработке методики были использованы результаты диссертации С.Д. Андреева в части моделей и методов перенаправления трафика пользователей между различными технологиями радиодоступа в рамках гетерогенной беспроводной сети с целью снижения нагрузки на сотовую сеть.

Председатель комиссии:

Начальник информационно-аналитического
отдела



А.А. Иванов

Члены комиссии:

Ведущий специалист



Е.Б. Юдина

Утверждаю
Заместитель генерального
директора ПАО «ГИПРОСВЯЗЬ»


А.Б. Васильев
« 09 декабря » 2018 г.



МЕТОДИКА ОЦЕНКИ ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ ДЛЯ ПРОЕКТИРОВАНИЯ ГЕТЕРОГЕННЫХ БЕСПРОВОДНЫХ СЕТЕЙ

Москва, 2018



ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ
 ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
 УНИТАРНОЕ ПРЕДПРИЯТИЕ
 ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
 НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ
 ИНСТИТУТ РАДИО
 (ФГУП НИИР)

Казакова ул., д. 16, Москва, 105064
 Телефон: (495) 647 18 30, для справок: (499) 261 63 70,
 Факс: (499) 261 00 90, E-mail: info@niir.ru
<http://www.niir.ru>
 ОКПО 01181481, ОГРН 1027700120766
 ИНН/КПП 7709025230/770901001

«УТВЕРЖДАЮ»

Первый заместитель генерального
 директора Федерального
 государственного унитарного
 предприятия «Научно-
 исследовательский институт
 радио», кандидат технических
 наук

М.Ю. Сподобаев

«10» января 2019 г.

АКТ

о внедрении результатов диссертационной работы
 Андреева Сергея Дмитриевича на тему «Разработка и
 исследование моделей множественного доступа и
 алгоритмов управления потоками трафика для
 гетерогенных беспроводных сетей», представленной на
 соискание ученой степени доктора технических наук

Комиссия в составе:

Председатель комиссии - директор НТЦ А ЭМС В.Э. Веерпалу, д.т.н.
 Члены комиссии - зам. директора НТЦ А ЭМС Е.Е. Девяткин, к.э.н.
 зам. директора НТЦ А ЭМС В.А. Корж, к.т.н.
 начальник лаборатории НТЦ А ЭМС Я.М. Гасс

установила, что при выполнении в Научно-исследовательском институте радио (ФГУП НИИР) работы «Реализация мероприятий по направлению «Информационная инфраструктура» плана мероприятий Программы «Цифровая экономика Российской Федерации» (определение принципов построения сетей 5G/IMT-2020 (архитектура сети)) М.: НИИР 2018, использованы следующие научные результаты диссертационной работы С.Д. Андреева:

1. Модель гетерогенной сети с возможностью установления прямых соединений между устройствами, отличающаяся от известных совместным учетом особенностей обслуживания потоков трафика во времени и геометрии размещения пользовательских устройств. Данная модель позволила определить характер увеличения среднего числа передающих пользователей с ростом входной интенсивности сессий.

2. Метод управления системой прямых соединений и способ его реализации в гетерогенной сети с интегрированными между собой технологиями радиодоступа, а также алгоритм распределения трафика с сетевым кодированием, учитывающий структуру контента в гетерогенной системе прямых соединений и позволяющий повысить вероятность успешного обслуживания пользователей.

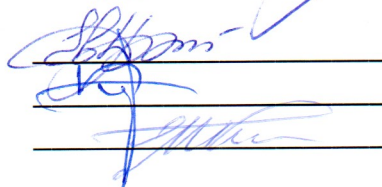
Использование указанных научных результатов диссертации С.Д. Андреева «Разработка и исследование моделей множественного доступа и алгоритмов управления потоками трафика для гетерогенных беспроводных сетей» дало возможность снизить нагрузку на сотовую сеть, значительно повышая при этом скорость передачи пользователей при выгрузке трафика на прямые соединения.

Председатель комиссии



В.Э. Веерпалу

Члены комиссии



Е.Е. Девяткин

В.А. Корж

Я.М. Гасс

Подписи В.Э. Веерпалу, Е.Е. Девяткина, В.А. Коржа, Я.М. Гасса заверяю.

Начальник отдела кадров ФГУП НИИР



Е.П. Буянова



Общество с ограниченной
ответственностью
“ Э В С ”

ИНН 7806371821

Россия, 195253, Санкт-Петербург,

Салтыковская дор., 18, лит. А, пом. 11Н

Тел.: (812) 606-6988; факс: (812) 226-1489

www.evs.ru

E-mail: mail@evs.ru

27.12.2018г. № 928

На № _____ от _____

АКТ

о внедрении научных результатов диссертационной работы
Андреева Сергея Дмитриевича «Разработка и исследование моделей
множественного доступа и алгоритмов управления потоками трафика
для гетерогенных беспроводных сетей» в ООО «ЭВС».

Данный акт подтверждает, что результаты диссертации «Разработка и исследование моделей множественного доступа и алгоритмов управления потоками трафика для гетерогенных беспроводных сетей» Андреева Сергея Дмитриевича, представленной на соискание ученой степени доктора технических наук, обладают актуальностью, представляют практический интерес и были внедрены в научно-исследовательской работе ООО «ЭВС».

Научные результаты диссертационной работы Андреева С.Д., в частности:

- модель сети радиодоступа для обслуживания большого числа устройств, отличающаяся от известных тем, что принимает во внимание особенности функционирования протокола случайного множественного доступа, стандартизованного для сотовой системы связи, и

позволяющая определить, что вероятность коллизии может быть снижена до 50% при использовании интервала отсрочки длительностью от 1 до 3 секунд;

- модель сети радиодоступа для обслуживания множества носимых устройств, отличающаяся от известных учетом особенностей работы протокола случайного множественного доступа, стандартизованного для локальной системы связи на крайне высоких частотах, и позволяющая определить, что энергетическая эффективность персональной сети может быть повышена до трех раз в случае направленной передачи.

Результаты диссертационной работы использованы в научно-исследовательских работах по совершенствованию линий связи для перспективных телевизионных систем безопасности, проводимых ООО «ЭВС».

Председатель комиссии:

Генеральный директор ООО «ЭВС»
к.т.н.



Н.В. Лебедев

Члены комиссии:

Начальник лаборатории
к.т.н.

А.Г. Стариков

Старший научный сотрудник
к.т.н.

С.Н. Ярышев

Ведущий инженер

С.Л. Шаров