

**ПРИОРИТЕТНЫЙ НАЦИОНАЛЬНЫЙ ПРОЕКТ «ОБРАЗОВАНИЕ»
РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ**

**Г.П. БАШАРИН, Ю.В. ГАЙДАМАКА,
К.Е. САМУЙЛОВ, Н.В. ЯРКИНА**

**УПРАВЛЕНИЕ КАЧЕСТВОМ
И ВЕРОЯТНОСТНЫЕ МОДЕЛИ
ФУНКЦИОНИРОВАНИЯ СЕТЕЙ СВЯЗИ
СЛЕДУЮЩЕГО ПОКОЛЕНИЯ**

Учебное пособие

Москва

2008

**«Создание комплекса инновационных образовательных программ
и формирование инновационной образовательной среды,
позволяющих эффективно реализовывать государственные интересы РФ
через систему экспорта образовательных услуг»**

Экспертное заключение –

доктор физико-математических наук, профессор *С.Я. Шоргин*

Башарин Г.П., Гайдамака Ю.В., Самуйлов К.Е., Яркина Н.В.

Управление качеством и вероятностные модели функционирования сетей связи следующего поколения: Учеб. пособие. – М.: РУДН, 2008. – 157 с.: ил.

В учебном пособии с помощью мультисервисных моделей теории телетрафика, относящихся к классу задач управления доступом в мультисервисных сетях на уровне соединений, изложен подход к построению и анализу вероятностных моделей для управления качеством обслуживания в сетях следующего поколения. Отдельная глава посвящена управлению доступом для мультисервисных систем массового обслуживания, в том числе вопросам определения оптимальной стратегии доступа.

Учебное пособие предназначено для студентов, обучающихся по магистерской программе «Управление инфокоммуникациями» по направлениям 010300 «Математика. Компьютерные науки», 010400 «Информационные технологии» и 010500 «Прикладная математика и информатика», а также для студентов, аспирантов и работников, специализирующихся в области инфокоммуникационных технологий.

Учебное пособие выполнено в рамках инновационной образовательной программы Российского университета дружбы народов, направление «Комплекс экспортноориентированных инновационных образовательных программ по приоритетным направлениям науки и технологий», и входит в состав учебно-методического комплекса, включающего описание курса, программу и электронный учебник.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	4
Глава 1. КЛАССИЧЕСКИЕ МУЛЬТИСЕРВИСНЫЕ МОДЕЛИ ЗВЕНА СЕТИ С ОДНОАДРЕСНЫМИ СОЕДИНЕНИЯМИ	9
§1.1. Мультисервисные модели Эрланга и Энгсета	9
§1.2. Мультисервисная модель Эрланга с явными потерями	11
Глава 2. ПРИМЕНЕНИЕ КЛАССИЧЕСКИХ МУЛЬТИСЕРВИСНЫХ МОДЕЛЕЙ К АНАЛИЗУ ФРАГМЕНТА ССПС	27
§2.1. Анализ фрагмента иерархической сети сотовой связи	27
§2.2. Анализ ВВХ микросоты с двумя типами каналов и учетом мобильности абонентов.....	42
Глава 3. УПРАВЛЕНИЕ ДОСТУПОМ ДЛЯ МУЛЬТИСЕРВИСНЫХ СМО	58
§3.1. Стратегии доступа	59
§3.2. Стратегия резервирования каналов	60
§3.3. Координатно-выпуклые стратегии	61
§3.4. Об оптимизации стратегии доступа.....	70
Глава 4. МОДЕЛЬ ЗВЕНА СЕТИ С ОДНОАДРЕСНЫМИ И МНОГОАДРЕСНЫМИ СОЕДИНЕНИЯМИ.....	75
§4.1. Модель мультисервисной сети с одноадресными и многоадресными соединениями	75
§4.2. Модель полнодоступного звена сети.....	87
§4.3. Модель звена с резервированием.....	101
РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА.....	120
СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ.....	122
ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ	129
ОПИСАНИЕ КУРСА И ПРОГРАММА.....	131

ВВЕДЕНИЕ

«Сети связи следующего поколения» (Next Generation Networks, NGN) – современная концепция, отражающая конвергенцию информационно-телекоммуникационных сетей в единую глобальную сеть. На сегодняшний день концепция NGN находится на этапе исследований и становления, причем чаще формулируются требования к сетям связи следующего поколения и задачи изучения, а не конкретные решения. Одной из особенностей NGN является возможность выбора пользователем определенного уровня качества предоставления услуги из нескольких уровней качества, предлагаемых сетью. Новые виды обслуживания, регулярно появляющиеся в современных информационно-телекоммуникационных сетях, делают задачу управления доступом к ресурсам сети для обеспечения качества услуги на различных уровнях сети одной из наиболее актуальных. Содержание курса обеспечивает слушателей необходимым объемом знаний для проведения исследований в этой новой области телекоммуникаций.

В учебном пособии с помощью мультисервисных моделей теории телетрафика, относящихся к классу задач управления доступом на уровне соединений, изложен подход к анализу показателей качества обслуживания в сетях следующего поколения. Отдельная глава посвящена управлению доступом для мультисервисных систем массового обслуживания, в том числе вопросам определения оптимальной стратегии доступа.

Настоящее учебное пособие предназначено для студентов, обучающихся по магистерской программе «Управление инфокоммуникациями» по направлению 010400 «Информационные технологии». В рамках инновационной образовательной программы, реализованной в РУДН в 2008–2009 гг. на кафедре систем

телекоммуникаций, разработан одноименный учебно-методический комплекс (УМК), включающий электронный учебник. Магистерская программа является авторской и включает в себя набор последовательно взаимосвязанных специальных дисциплин. По магистерской программе могут также обучаться лица, имеющие диплом бакалавра по направлениям 010300 «Математика. Компьютерные науки» и 010500 «Прикладная математика и информатика». Для эффективного обучения по магистерской программе учащимся рекомендуется в бакалавриате прослушать профиль специальных дисциплин по выбору в составе: «Основы формальных методов описания бизнес процессов»; «Модели для анализа качества обслуживания в сетях связи следующего поколения»; «Основы разработки корпоративных инфокоммуникационных систем»; «Основы управления инфокоммуникационными компаниями». Для этих дисциплин, в рамках инновационной образовательной программы в РУДН в 2008–2009 гг. также разработаны одноименные УМК и учебные пособия.

Целью курса является изучение принципов функционирования сетей связи следующего поколения; рассмотрение вопросов качества в NGN на различных уровнях; знакомство с методами анализа и расчета показателей качества отдельных элементов сетей, а также сети в целом; изучение управления доступом в мультисервисных системах массового обслуживания; исследование точных и приближенных методов анализа качества обслуживания в сетях связи следующего поколения. После успешного прохождения курса слушатели должны знать основные понятия и определения, относящиеся к концепции сетей связи следующего поколения; основные протоколы сетей связи следующего поколения; требования международных стандартов к показателям качества на различных уровнях сети NGN; методы разработки и анализа моделей телекоммуникационных систем сложной структуры; численные методы расчета (приближенные и точные) характеристик сети. Слушатели должны

уметь строить модели отдельных функциональных элементов NGN, а также модели сети в целом; при построении моделей выбирать стратегию доступа для обеспечения определенного уровня качества предоставления услуги с учетом нескольких уровней качества, предлагаемых сетью; проводить исследование построенных моделей, получать их вероятностные характеристики, требующиеся для анализа показателей качества; использовать изученные методы и принципы при разработке моделей и анализе качества обслуживания для реально существующих сетей.

Учебное пособие включает четыре главы. В главе 1 на примере мультисервисной модели Эрланга изложено обобщение основ классической теории телетрафика на мультисервисные модели. В главе 2 рассмотрены примеры применения мультисервисных моделей к анализу фрагмента сети сотовой подвижной связи. Глава 3 посвящена управлению доступом для мультисервисных систем массового обслуживания. В ней выделены основные типы стратегий доступа, а также описан подход к оптимизации стратегии доступа. В главе 4 две основные стратегии доступа применены при построении моделей звена мультисервисной сети с одноадресными и многоадресными соединениями.

В тексте приводятся ссылки на основную и дополнительную литературу. Главы книги разбиты на параграфы, которые нумеруются отдельно. Так, §2.3 означает «глава 2, параграф 3». В каждом параграфе нумерация формул начинается заново. Например, формула (4.12) означает «формула номер 12 внутри параграфа 4 текущей главы». При этом ссылка на формулу в пределах одной главы дается как есть, а при ссылке на формулу из другой главы к ней добавляется в начале номер соответствующей главы. Так, ссылка (1.4.12) означает «формула номер 12 внутри параграфа 4 главы 1». Нумерация рисунков, таблиц, теорем, лемм, утверждений и примеров сквозная в каждой главе.

СПИСОК ОСНОВНЫХ СОКРАЩЕНИЙ

Русскоязычные сокращения

БП	-	буферная память
БС	-	базовая станция
БЦК	-	базовый цифровой канал
ВВХ	-	вероятностно-временные характеристики
ИСС	-	интеллектуальная сеть связи
МСС	-	мультисервисная сеть связи
ОМП	-	обратимый марковский процесс
ПП	-	пуассоновский поток
ПРГ	-	процесс размножения и гибели
СВ	-	случайная величина
СЛАУ	-	система линейных алгебраических уравнений
СМО	-	система массового обслуживания
ССПС	-	сотовая сеть подвижной связи
СтМП	-	ступенчатый марковский процесс
СУГБ	-	система уравнений глобального баланса
СУЧБ	-	система уравнений частичного баланса
СУР	-	система уравнений равновесия
ТМО	-	теория массового обслуживания
ТТ	-	теория телетрафика
ЦЛ	-	цифровая линия
ЧНН	-	часы наибольшей нагрузки
ШПП	-	ширина полосы пропускания
ШЦЛ	-	широкополосная цифровая линия

Англоязычные сокращения

ATM	-	Asynchronous Transfer Mode асинхронный режим передачи
CP	-	Complete Partitioning стратегия полного разделения
CS	-	Complete Sharing полнодоступная стратегия
HDTV	-	High Definition Television телевидение высокой четкости
NGN	-	Next Generation Networks сети следующего поколения
MPEG	-	Moving Picture Experts Group Экспертная группа по вопросам движущегося изображения
PP	-	Partitioning Policy стратегия разделения
SMQMA	-	Sharing with Maximum Queue Length and Minimum Allocation неполнодоступная СМО с индивидуальными потолками
TP	-	Threshold Policy пороговая стратегия
TRP	-	Trunk Reservation Policy стратегия резервирования каналов

Глава 1. КЛАССИЧЕСКИЕ МУЛЬТИСЕРВИСНЫЕ МОДЕЛИ ЗВЕНА СЕТИ С ОДНОАДРЕСНЫМИ СОЕДИНЕНИЯМИ

§1.1. Мультисервисные модели Эрланга и Энгсета

В качестве примера классических мультисервисных моделей теории телетрафика (ТТ) можно назвать мультисервисные модели типа Эрланга и мультисервисные модели типа Энгсета с явными потерями [1].

Первая названная модель может быть использована для анализа звена мультисервисной сети связи (МСС), являющейся способом реализации концепции сети следующего поколения (Next Generation Networks, NGN), ключевыми особенностями которой являются пакетные технологии передачи и обеспечение функциональных возможностей «Triple Play Services» (коммерческой концепции, предполагающей предоставление услуг телефонии, телевидения и доступа в Интернет в виде одного коммерческого предложения). Далее в этой главе показан пример применения полnodоступной мультисервисной модели Эрланга к анализу широкополосной цифровой линии (ШЦЛ) с тремя типами нагрузки (голос, видео и данные). Начинается изложение с компактного описания математических моделей мультисервисных систем массового обслуживания (СМО) типа Эрланга, затем доказываются необходимые теоремы и излагается рекуррентный алгоритм вычисления макрохарактеристик указанных СМО. Понимание и освоение этого материала позволит грамотно организовать численный анализ макрохарактеристик действующих и проектируемых систем, а также разработку приближенных мультипликативных моделей для немультимпликативных систем. Более подробно и с большим количеством реальных содержательных примеров этот материал излагается в книгах В.С. Лагутина, С.Н. Степанова [5] и В.Б. Иверсена [8].

Мультисервисные модели типа Энгсета с явными потерями также успешно применяются для исследования функционирования звеньев МСС. Так, в [1, гл. 3] подробно рассмотрены две мультисервисные модели типа Энгсета – Энгсет-1 и Энгсет-2. В модели Энгсет-1 N близких по своим запросам абонентов могут запрашивать K различных услуг каждый, причем интенсивность запросов каждого абонента на услугу с одним и тем же номером примерно одинакова. Например, N семей одного многоквартирного дома или несколько близлежащих домов имеют доступ к интеллектуальной сети с помощью широкополосной ЦЛ емкостью V базовых цифровых каналов, причем $N \ll V$. В модели Энгсет-2 общее число N абонентов достаточно велико, $N \gg K$, и N_k абонентов ориентированы в основном на запрос только k -услуги. Обе модели могут быть применены при расчете мультисервисных систем мультимедийного доступа, например, при проектировании подключения к интеллектуальной сети связи (ИСС) с помощью ШЦЛ. В [1] объяснена необходимость рассмотрения двух случаев Энгсет-1 и Энгсет-2 и приведен рекуррентный алгоритм вычисления макрохарактеристик системы.

Анализ системы доступа к ИСС при помощи мультисервисных моделей типа Энгсета с явными потерями дает возможность провайдеру правильно выбрать параметры широкополосной ЦЛ и оценить качество обслуживания абонентов в часы наибольшей нагрузки (ЧНН). Напомним, что вопросы качества обслуживания являются одними из наиболее актуальных при предоставлении услуг связи в NGN. К параметрам качества здесь относятся такие характеристики, как вероятности π_k блокировки k -услуги, средняя загрузка ШЦЛ, ее пропускная способность TH_k для k -услуги, $k = \overline{1, K}$, и др. Это весьма непростая задача, тем более, что необходимо ее решить для некоторого реально прогнозируемого диапазона входных параметров – структурных и нагрузочных. Материал

главы 3 [1] дает много возможностей для самостоятельной теоретической, вычислительной и проектной работы слушателей и читателей.

§1.2. Мультисервисная модель Эрланга с явными потерями

1.2.1. Пример мультиплексирования в АТМ

Рассмотрим мультиплексирование высокоскоростной цифровой линии (ЦЛ) со скоростью C Мбит/с. Если принять, что один базовый цифровой канал (БЦК) имеет скорость 64 кбит/с, то, например, при $C = 2$ Мбит/с в широкополосной ЦЛ можно организовать $2048/64 = 32$ БЦК. Таким образом, базовый цифровой канал служит удачной для рассматриваемой модели единицей ширины полосы пропускания (ШПП).

Предположим, что к широкополосной ЦЛ организован доступ по схеме с явными потерями для K информационных потоков от разноскоростных источников нагрузки. Множество предоставляемых системой услуг удобно обозначить через $\mathbf{K} = \{1, \dots, K\}$, $|\mathbf{K}| = K$.

Пусть k -й источник в момент максимальной активности требует предоставления ему b_k единиц ШПП. Будем считать, что b_k – пиковая скорость k -источника или требуемое для него число БЦК. В системах передачи по технологии с асинхронным режимом передачи (Asynchronous Transfer Mode, АТМ) вся информация оцифровывается и упаковывается в ячейки по 53 байта, из которых 5 байт – заголовок, 8 байт – служебная информация, 40 байт – данные конкретного источника, передача которых и создает оплачиваемую абонентом нагрузку. Например, если голос оцифровывается в двоичный поток со скоростью порядка 64 кбит/с, то пиковая скорость передачи голосовой информации по соответствующей ЦЛ составит около 200 ячеек/с, т.е. в среднем каждые 5 мс в ЦЛ будет поступать одна ячейка.

На рис. 1.1 представлена упрощенная схема с $K = 3$ типами нагрузки – видео, данные и голос, причем в периоды активности каждого источника

генерируемые им ячейки поступают пачками. «БП» обозначает буферную память, входящую в состав устройства доступа к высокоскоростной ШЦЛ.

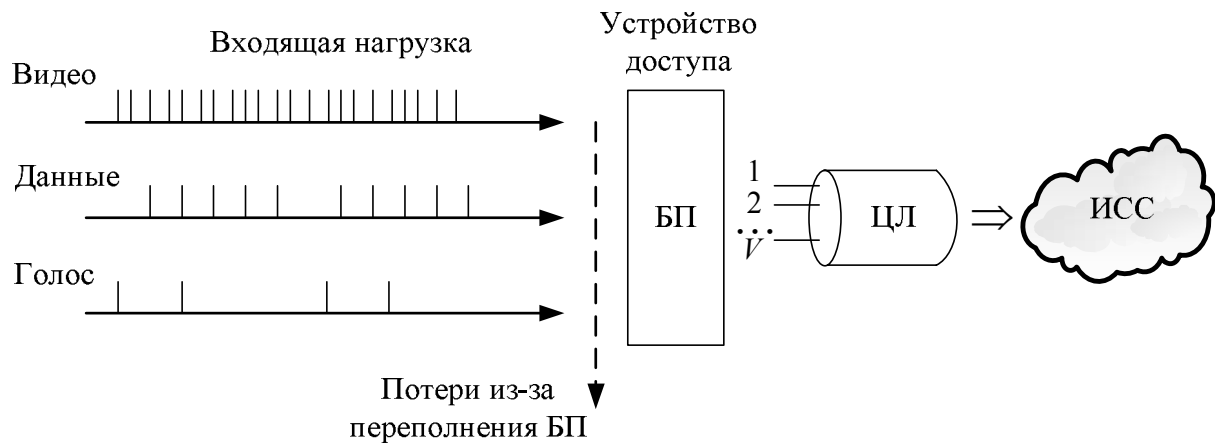


Рис. 1.1. Физическая модель функционирования широкополосной цифровой линии емкостью V БЦК

В этом случае стратегия доступа основана на пиковой скорости передачи каждого из K источников нагрузки.

1.2.2. Основные параметры мультисервисной модели Эрланга

Изучаемые нами многомерные модели достаточно сложны и поэтому требуют для своего описания развитой терминологии, позволяющей их описывать с различных позиций. Так, в ряде случаев удобно вести изложение как в физических терминах «ширины полосы пропускания», так и в терминах «базовых цифровых каналов». Кроме того, наряду с базовыми понятиями, терминами и моделями классической теории телетрафика целесообразно пользоваться эквивалентными или близкими им базовыми понятиями, терминами и методами теории массового обслуживания (ТМО).

Например, будем считать синонимами понятия « k -сообщение», « k -услуга» и « k -заявка»; «длительность передачи», «длительность занятия» и «длительность обслуживания» и т.д., отдавая предпочтение в зависимости от контекста тому или иному термину. Аналогичным образом

эквивалентны понятия «система передачи с явными потерями» и «система массового обслуживания без мест для ожидания».

Сделаем теперь следующие предположения о системе доступа к ЦЛ, оказываемых ею услугах и о создаваемой абонентами мультисервисной нагрузке.

- 1) На предоставляемую по широкополосной цифровой линии k -услугу поступает пуассоновский поток (ПП) k -сообщений с постоянной интенсивностью λ_k , $0 < \lambda_k < \infty$, $k = \overline{1, K}$. Все K ПП независимы в совокупности, причем k -услуга (k -сообщение) требует b_k единиц ШПП, т.е. b_k БЦК. Здесь b_k – целые числа, причем $b_k \in \{1, 2, \dots, V\}$, $k = \overline{1, K}$.
- 2) Если в момент поступления k -сообщения у ШЦЛ нет b_k свободных БЦК, то вновь поступившее k -сообщение получает отказ и теряется, не оказывая дополнительного влияния на интенсивность λ_k породившего его пуассоновского потока. Это означает, что эффект повторения получивших отказ сообщений мал и его можно не учитывать и считать, что система функционирует с явными потерями.
- 3) Длительность занятия k -услугой (передачей k -сообщения) имеет экспоненциальное распределение с интенсивностью μ_k , $0 < \mu_k < \infty$, $k = \overline{1, K}$, причем эти длительности не зависят друг от друга и от процессов поступления сообщений. По завершении обслуживания k -сообщение сразу освобождает всю занятую им ШПП, т.е. все занятые им b_k БЦК освобождаются одновременно.

Таким образом, в описании системы участвуют два структурных параметра – V и K , а также 3 K -мерных вектора, характеризующих предложенную нагрузку:

$$\boldsymbol{\lambda} := (\lambda_1, \dots, \lambda_K)^T, \quad \mathbf{b} := (b_1, \dots, b_K)^T, \quad \boldsymbol{\mu} := (\mu_1, \dots, \mu_K)^T,$$

т.е. всего $3K + 1$ независимых числовых параметра. Рассматриваемую модель K -сервисного доступа к ШЦЛ с явными потерями будем кодировать как

$$\mathbf{M} \left| \begin{array}{c} \mathbf{M} \\ \boldsymbol{\lambda}, \mathbf{b} \end{array} \right| \boldsymbol{\mu} \Big| V < \infty, \quad (2.1)$$

опуская 0 в четвертой позиции, когда из контекста ясно, что речь идет о системе с явными потерями.

Следуя традициям теории телетрафика, будем называть многомерную пуассоновско-экспоненциальную предложенную нагрузку этой модели *мультисервисной пуассоновской нагрузкой первого рода* и кодировать ее посредством тройки векторов $\langle \boldsymbol{\lambda}, \mathbf{b}, \boldsymbol{\mu} \rangle$.

Пусть $\rho_k := \frac{\lambda_k}{\mu_k}$ – интенсивность предложенной k -нагрузки безотносительно к требуемой ширине полосы, а $\rho'_k := \rho_k b_k$ – измеряемая в БЦК интенсивность предложенной k -нагрузки с учетом требуемой ширины полосы.

Оказывается, что, как и в моносервисном случае, мультисервисные модели типа (2.1) описываются ступенчатым марковским процессом (СтМП) с мультипликативным представлением стационарных (равновесных) вероятностей состояний. При этом многие вероятностно-временные характеристики (ВВХ) изучаемой системы зависят от нагрузки, задаваемой не тройкой, а парой векторов $\langle \boldsymbol{\rho}, \mathbf{b} \rangle$, а при выполнении условий 1)–3) автоматически выполняется и условие

$$4) \quad 0 < \rho_k b_k < \infty, \quad k = \overline{1, K}.$$

Переходим к построению пространства состояний системы. Пусть n_k – число k -услуг, т.е. число виртуальных каналов класса k , предоставляемых системой в некоторый момент времени $t > 0$. Тогда

$$n_k = 0, 1, \dots, \left\lfloor \frac{V}{b_k} \right\rfloor, \quad k = \overline{1, K},$$

где $\lfloor x \rfloor$ означает целую часть числа x . Вектор $\mathbf{n} = (n_1, \dots, n_K)^T$ описывает численность всех услуг, предоставляемых системой в момент t , а

$$U(\mathbf{n}) = \mathbf{b}^T \mathbf{n} = \sum_{k=1}^K b_k n_k \quad (2.2)$$

есть общее число занятых БЦК в состоянии \mathbf{n} . Величину (2.2) называют *мгновенным коэффициентом использования ШПП ШЦЛ* в состоянии \mathbf{n} , причем выполняется условие $0 \leq \mathbf{b}^T \mathbf{n} \leq V$. Очевидно, что $V - \mathbf{b}^T \mathbf{n}$ есть число свободных БЦК в состоянии \mathbf{n} .

Поэтому пространство \mathcal{S} всех возможных состояний системы имеет вид

$$\mathcal{S} := \{ \mathbf{n} : 0 \leq \mathbf{b}^T \mathbf{n} \leq V \}, \quad (2.3)$$

а *подпространства приема и блокировки* для k -сообщений имеют соответственно вид

$$\begin{aligned} \mathcal{S}_k &:= \{ \mathbf{n} : \mathbf{b}^T \mathbf{n} \leq V - b_k \}, \\ \overline{\mathcal{S}}_k &= \mathcal{S} \setminus \mathcal{S}_k = \{ \mathbf{n} : \mathbf{b}^T \mathbf{n} > V - b_k \}, \quad k = \overline{1, K}. \end{aligned} \quad (2.4)$$

Пусть $\lambda_k(\mathbf{n})$ – *интенсивность принятого*, а $\mu_k(\mathbf{n})$ – *интенсивность обслуженного потока* k -сообщений в состоянии \mathbf{n} . Тогда из предположений предыдущего раздела вытекает, что для $\forall k \in \mathcal{K}$

$$\lambda_k(\mathbf{n}) = \begin{cases} 0, & \mathbf{n} \in \overline{\mathcal{S}}_k \text{ (} k\text{-сообщение получает отказ),} \\ \lambda_k, & \mathbf{n} \in \mathcal{S}_k \text{ (} k\text{-сообщение принято),} \end{cases} \quad (2.5)$$

$$\mu_k(\mathbf{n}) = \mu_k n_k, \quad \mathbf{n} \in \mathcal{S}. \quad (2.6)$$

Последнее соотношение означает, что в состоянии \mathbf{n} суммарная интенсивность обслуживания всех n_k k -сообщений не зависит от числа n_l , $l \neq k$ сообщений других типов в системе. Если этого требуют запросы практики, то предположение (2.6) можно ослабить, но мы здесь этого делать не будем, концентрируя свое внимание на структурной сложности рассматриваемых СМО.

Упражнение 1.1. Рассмотрим ШЦЛ с параметрами $V = 8$, $K = 2$, $b_1 = 1$, $b_2 = 2$. Для данной системы

$$S = \{(n_1, n_2) : n_1 + 2n_2 \leq 8\},$$

$$S_1 = \{(n_1, n_2) : n_1 + 2n_2 \leq 7\}, \quad \bar{S}_1 = S \setminus S_1 = \{(n_1, n_2) : n_1 + 2n_2 = 8\},$$

$$S_2 = \{(n_1, n_2) : n_1 + 2n_2 \leq 6\}, \quad \bar{S}_2 = S \setminus S_2 = \{(n_1, n_2) : n_1 + 2n_2 = 7, 8\}.$$

Эти множества изображены на рис. 1.2. Черные точки обозначают подпространства S_1 и S_2 , а белые (полые) точки – подпространства \bar{S}_1 и \bar{S}_2 .

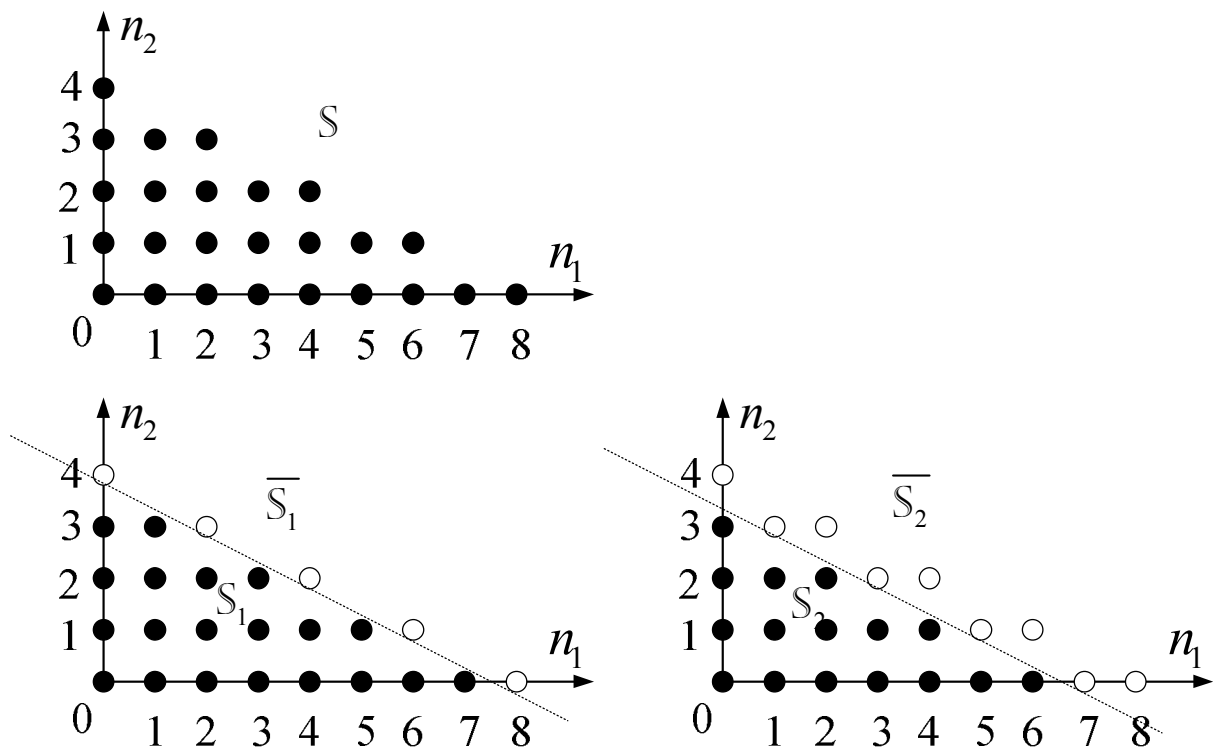


Рис. 1.2. Пространство состояний S и его подпространства S_1 , \bar{S}_1 , S_2 , \bar{S}_2

для $V = 8$, $K = 2$, $b_1 = 1$, $b_2 = 2$

Введем теперь K -мерный случайный процесс $\mathbf{X}(t) = (X_1(t), \dots, X_K(t))^T$, $t \geq 0$, где $X_k(t)$ – число k -услуг, оказываемых в момент времени t в мультисервисной модели Эрланга с явными потерями. Из предположений 1)–3) о параметрах модели следует, что $\mathbf{X}(t)$ – СтМП с пространством состояний S и матрицей интенсивностей переходов $\mathbf{A} = (a(\mathbf{m}, \mathbf{n}))_{\mathbf{m}, \mathbf{n} \in S}$. Примем дополнительно условие

5) Матрица \mathbf{A} неразложима (все состояния из S сообщаются).

Предположения 1)–5) имеют очевидный физический смысл, являются естественными для приложений и гарантируют существование стационарного распределения вероятностей, но не дают ответа на вопрос о том, какую нагрузку целесообразно предложить системе.

Теорема 1.1. Если для системы $\left. \begin{matrix} \mathbf{M} \\ \lambda, \mathbf{b} \end{matrix} \middle| \begin{matrix} \mathbf{M} \\ \mu \end{matrix} \right| V < \infty$ выполняются условия 1)–5), то ее описывает ступенчатый МП $\mathbf{X}(t)$, у которого существует стационарное (равновесное) распределение вероятностей. Это распределение не зависит от начального и является мультипликативным:

$$p(\mathbf{n}) = G^{-1} \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!}, \quad \mathbf{n} \in S. \quad (2.7)$$

Здесь $\rho_k = \frac{\lambda_k}{\mu_k}$, $k = \overline{1, K}$, и выполняется нормировочное условие

$$\sum_{\mathbf{n} \in S} p(\mathbf{n}) = 1, \quad (2.8)$$

так что нормирующая константа

$$G = \frac{1}{p(\mathbf{0})} = \left(\sum_{\mathbf{n} \in S} \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!} \right). \quad (2.9)$$

Доказательство. Вывод системы уравнений глобального баланса (СУГБ).

Пусть $\mathbf{A} = (a(\mathbf{m}, \mathbf{n}))_{\mathbf{m}, \mathbf{n} \in S}$ – матрица интенсивностей переходов СтМП $\mathbf{X}(t)$. При сделанных предположениях его равновесные вероятности удовлетворяют системе уравнений равновесия (СУР)

$$\mathbf{p}^T \mathbf{A} = \mathbf{0}^T \Leftrightarrow \sum_{\mathbf{m} \in S} p(\mathbf{m}) a(\mathbf{m}, \mathbf{n}) = 0, \mathbf{n} \in S \quad (2.10)$$

с нормирующим условием (2.8).

Пусть $\mathbf{e}_k = \left(\underbrace{0, \dots, 0}_{k-1}, 1, \underbrace{0, \dots, 0}_{K-k} \right)^T$ – K -мерный вектор, k -я компонента

которого равна 1, а остальные – нули. Для доказательства нам потребуется также функция Хевисайда

$$u(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leq 0 \end{cases}$$

и функция-индикатор

$$1(\text{событие } A) = \begin{cases} 1, & \text{если } A \text{ произошло,} \\ 0, & \text{в противном случае.} \end{cases} \quad (2.11)$$

Заметим, что за Δt попасть в состояние \mathbf{n} можно лишь из одного из соседних «снизу» состояний $\mathbf{n} - \mathbf{e}_k$, $n_k \geq 1$, за счет поступления нового k -сообщения, либо из соседних «сверху» состояний $\mathbf{n} + \mathbf{e}_k \in S$ за счет завершения передачи одного из k -сообщений, $k = \overline{1, K}$. Поэтому, перенося член $p(\mathbf{n})a(\mathbf{n}, \mathbf{n})$ в правую часть, СУР (2.10) можно записать в следующем виде:

$$\begin{aligned} & \sum_{k=1}^K p(\mathbf{n} - \mathbf{e}_k) u(n_k) a(\mathbf{n} - \mathbf{e}_k, \mathbf{n}) + \sum_{k=1}^K p(\mathbf{n} + \mathbf{e}_k) 1(\mathbf{n} \in S_k) a(\mathbf{n} + \mathbf{e}_k, \mathbf{n}) = \\ & = p(\mathbf{n})(-a(\mathbf{n}, \mathbf{n})), \mathbf{n} \in S. \end{aligned} \quad (2.12)$$

Система уравнений (2.12) после умножения обеих частей на Δt имеет простую физическую интерпретацию. Правая часть представляет

собой вероятность выхода из состояния \mathbf{n} за Δt в одно из соседних состояний, а левая часть представляет собой сумму вероятностей попадания в состояние \mathbf{n} за Δt из одного из соседних «снизу» или «сверху» состояний.

Так как сумма элементов любой строки матрицы \mathbf{A} равна 0, то

$$\begin{aligned} -a(\mathbf{n}, \mathbf{n}) &= \sum_{\mathbf{m} \in S \setminus \mathbf{n}} a(\mathbf{n}, \mathbf{m}) = \sum_{k=1}^K a(\mathbf{n}, \mathbf{n} - \mathbf{e}_k) u(n_k) + \sum_{k=1}^K a(\mathbf{n}, \mathbf{n} + \mathbf{e}_k) l(\mathbf{n} \in S_k) = \\ &= \sum_{k=1}^K \lambda_k l(\mathbf{n} \in S_k) + \sum_{k=1}^K \mu_k n_k, \quad \mathbf{n} \in S. \end{aligned} \quad (2.13)$$

Здесь $-a(\mathbf{n}, \mathbf{n})$ – положительная интенсивность выхода из состояния \mathbf{n} либо за счет поступления и приема одного нового сообщения, либо за счет завершения передачи одного из передаваемых сообщений. Поскольку

$$a(\mathbf{n} - \mathbf{e}_k, \mathbf{n}) = \lambda_k, \quad n_k \geq 1, \quad \mathbf{n} \in S;$$

$$a(\mathbf{n} + \mathbf{e}_k, \mathbf{n}) = (n_k + 1) \mu_k, \quad \mathbf{n} \in S_k,$$

то СУР (2.12) можно записать в виде

$$\begin{aligned} \sum_{k=1}^K p(\mathbf{n} - \mathbf{e}_k) u(n_k) \lambda_k + \sum_{k=1}^K p(\mathbf{n} + \mathbf{e}_k) l(\mathbf{n} \in S_k) (n_k + 1) \mu_k = \\ = p(\mathbf{n}) \left(\sum_{k=1}^K \lambda_k l(\mathbf{n} \in S_k) + \sum_{k=1}^K n_k \mu_k \right), \quad \mathbf{n} \in S. \end{aligned} \quad (2.14)$$

Отметим теперь, что (2.14) представляет собой СУГБ и имеет очевидную физическую интерпретацию. При наличии некоторого опыта СУГБ для представляющей практический интерес математической модели можно выписать сразу, минуя выполненные нами промежуточные выкладки.

Отметим еще, что порядок системы линейных алгебраических уравнений (СЛАУ) равен $|S|$, а ее ранг равен $|S| - 1$. Поэтому ее решение можно найти с точностью до произвольного множителя, который легко определить с помощью нормировочного условия (2.8).

Вывод и решение системы уравнений частичного баланса (СУЧБ)¹.

Как и ранее для однопоточковых СМО, выдвинем гипотезу о том, что наряду с глобальным балансом имеет место частичный баланс по каждой из k -услуг

$$p(\mathbf{n})\mu_k n_k = p(\mathbf{n} - \mathbf{e}_k)u(n_k)\lambda_k, \mathbf{n} \in S, k = \overline{1, K}; \quad (2.15a)$$

$$p(\mathbf{n} + \mathbf{e}_k)(n_k + 1)\mu_k = p(\mathbf{n})\lambda_k, \mathbf{n} \in S_k, k = \overline{1, K}. \quad (2.15b)$$

Физический смысл соотношений (2.15) иллюстрирует рис. 1.3.

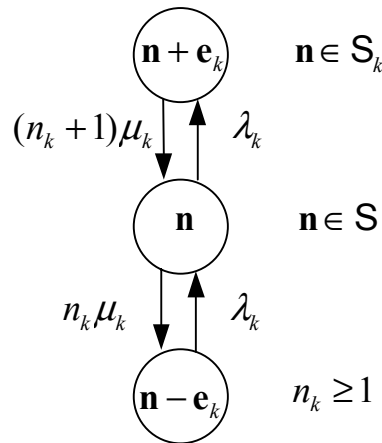


Рис. 1.3. Схема частичного баланса по k -услуге

Подсистема уравнений (2.15a) описывает баланс по k -услуге между состояниями \mathbf{n} и $\mathbf{n} - \mathbf{e}_k$:

$$P\{\mathbf{n} \xrightarrow{\Delta t} \mathbf{n} - \mathbf{e}_k\} = P\{\mathbf{n} - \mathbf{e}_k \xrightarrow{\Delta t} \mathbf{n}\} + o(\Delta t), \quad (2.16a)$$

а (2.15b) – между состояниями $\mathbf{n} + \mathbf{e}_k$ и \mathbf{n} :

$$P\{\mathbf{n} + \mathbf{e}_k \xrightarrow{\Delta t} \mathbf{n}\} = P\{\mathbf{n} \xrightarrow{\Delta t} \mathbf{n} + \mathbf{e}_k\} + o(\Delta t). \quad (2.16b)$$

Если решения СУЧБ–1 (2.15a) и СУЧБ–2 (2.15b) совпадают, то это их общее решение и будет решением СУГБ (2.14), т.к. мы просто приравняли попарно k -е слагаемые в левой и правой частях (2.14).

Из рекуррентных соотношений (2.15a) следует:

¹ Понятие частичного баланса для мультисервисных СМО соответствует понятию локального баланса для моносервисного случая. Однако слово «локальный» не вполне употребимо для многопоточковых систем. Поэтому для таких систем мы будем в дальнейшем использовать термин «частичный баланс».

$$p(\mathbf{n}) = p(\mathbf{n} - \mathbf{e}_k) \frac{\rho_k}{n_k} = p(\mathbf{n} - 2\mathbf{e}_k) \frac{\rho_k^2}{n_k(n_k - 1)} = \dots = p(\mathbf{n} - n_k \mathbf{e}_k) \frac{\rho_k^{n_k}}{n_k!}.$$

Продолжая рекурсию по всем остальным видам услуг, получим:

$$p(\mathbf{n}) = p(\mathbf{0}) \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!}, \quad \mathbf{n} \in S. \quad (2.17a)$$

Аналогичным образом из (2.15б) следует:

$$p(\mathbf{n} + \mathbf{e}_k) = p(\mathbf{n}) \frac{\rho_k}{n_k + 1} = p(\mathbf{n} - \mathbf{e}_k) \frac{\rho_k^2}{(n_k + 1)n_k} = \dots = p(\mathbf{0}) \frac{\rho_k}{n_k + 1} \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!}, \quad (2.17б)$$

$$\mathbf{n} \in S_k, \quad k = \overline{1, K},$$

что эквивалентно (2.17а).

Таким образом, СУЧБ-1 и СУЧБ-2 имеют общее решение, т.е. гипотеза о наличии баланса по каждой услуге оправдана, (2.17) совпадает с (2.7) и теорема 1.1 доказана. ■

1.2.3. Вероятность потерь и другие макрохарактеристики

Равновесное распределение (2.7) микросостояний в пространстве S дает прозрачную теоретическую базу для вывода необходимых вероятностно-временных макрохарактеристик ШЦЛ. Рассмотрим теперь основные из них.

Пусть $\mathbf{X} = (X_1, \dots, X_K)^T$ – случайный вектор с равновесным распределением

$$P\{\mathbf{X} = \mathbf{n}\} =: p(\mathbf{n}), \quad \mathbf{n} \in S. \quad (2.18)$$

Обозначим через π_k вероятность того, что вновь поступившее k -сообщение застанет ШЦЛ в макросостоянии \overline{S}_k и будет заблокировано:

$$\pi_k := P\{\overline{S}_k\} = \sum_{\mathbf{n} \in \overline{S}_k} p(\mathbf{n}). \quad (2.19a)$$

Величина π_k называется *вероятностью потерь по времени*, и в силу теоремы 1.1

$$\pi_k = \frac{1}{G} \sum_{\mathbf{n} \in \mathcal{S}_k} \prod_{j=1}^K \frac{\rho_j^{n_j}}{n_j!}. \quad (2.196)$$

Это название объясняется тем, что если при стационарном режиме в достаточно большом временном интервале $[t_0, t_0 + T)$ длины T суммарное время нахождения ШЦЛ в состоянии \overline{S}_k составляет T_k , то статистической оценкой для π_k служит величина $\hat{\pi}_k = \frac{T_k}{T}$, $k = \overline{1, K}$. Если в этом же интервале поступит всего N сообщений, из которых будет потеряно N_k k -сообщений, то величина $\tilde{\pi}_k = \frac{N_k}{N}$ будет являться оценкой *потерь по сообщениям*. Вопрос о точности этих оценок и их близости теоретическим значениям π_k в зависимости от параметра T или N решается с помощью методов математической статистики.

Поскольку k -сообщения поступают по пуассоновскому закону с интенсивностью λ_k , то за $[t_0, t_0 + T)$ поступает в среднем $\lambda_k T$ k -сообщений, из которых опять же в среднем теряется $\lambda_k T \pi_k$. Поэтому за определение *пропускной способности* (throughput) для k -сообщений можно принять величину:

$$TH_k := \lambda_k (1 - \pi_k), \quad k = \overline{1, K}. \quad (2.20)$$

Так как мгновенный коэффициент использования ШЦЛ в силу (2.2) является случайной величиной (СВ) $U(\mathbf{X}) = \mathbf{b}^T \mathbf{X}$, то за *коэффициент использования* (utility) ШЦЛ естественно принять среднее значение этой величины:

$$UTIL := EU(\mathbf{X}) = \mathbf{b}^T E\mathbf{X} = \sum_{k=1}^K b_k EX_k. \quad (2.21)$$

Здесь $b_k EX_k =: UTIL_k$ – среднее число БЦК, занятых обслуживанием k -сообщений, а $\eta := \frac{UTIL}{V}$ – средняя нагрузка, исполненная одним БЦК.

Мультипликативный характер равновесного распределения позволяет получить достаточно простой и эффективный алгоритм вычисления введенных выше и других вероятностно-временных макрохарактеристик ШЦЛ. С этой целью введем пространство $S(v)$ из таких состояний, в каждом из которых занято ровно v БЦК:

$$S(v) := \{\mathbf{n} \in S : \mathbf{b}^T \mathbf{n} = v\}, \bigcup_{v=0}^V S(v) = S. \quad (2.22)$$

Обозначим

$$q(v) := P\{U(\mathbf{X}) = v\} = \sum_{\mathbf{n} \in S(v)} p(\mathbf{n}), \sum_{v=0}^V q(v) = 1. \quad (2.23)$$

Очевидно, что мгновенный коэффициент использования $U(\mathbf{X})$ является случайной величиной с распределением (2.23), а его среднее значение

$$UTIL = \sum_{v=1}^V vq(v). \quad (2.24)$$

Введем теперь величину

$$R_k(v) := \sum_{\mathbf{n} \in S(v)} n_k p(\mathbf{n}). \quad (2.25)$$

Чтобы выяснить ее физический смысл, умножим каждое слагаемое на $q(v)/q(v)$, вынесем $q(v)$ за знак суммы и заметим, что $p(\mathbf{n})/q(v)$, $\mathbf{n} \in S(v)$ является условным распределением в пространстве $S(v)$. Поэтому

$$R_k(v) = q(v)E(X_k | v), \quad v = \overline{0, V}, \quad k = \overline{1, K}. \quad (2.26)$$

Поскольку $E(X_k | 0) = 0$, то $R_k(0) = 0$, а

$$\sum_{v=1}^V R_k(v) = \sum_{v=1}^V q(v)E(X_k | v) = EX_k, \quad k = \overline{1, K}.$$

Для удобства дальнейших записей примем, что v может быть отрицательным, но для сохранения здравого смысла потребуем, чтобы

$$q(v) = R_k(v) = 0, \quad v < 0. \quad (2.27)$$

Лемма 1.1. Макротяжность $q(v)$ удовлетворяет следующему рекуррентному соотношению:

$$vq(v) = \sum_{k=1}^K b_k \rho_k q(v - b_k), \quad v = \overline{0, V}. \quad (2.28)$$

Доказательство. Для доказательства запишем левую часть (2.28) более подробно:

$$vq(v) = \sum_{\mathbf{n} \in S(v)} p(\mathbf{n}) \sum_{k=1}^K b_k n_k = \sum_{k=1}^K b_k \sum_{\mathbf{n} \in S(v)} n_k p(\mathbf{n}) = \sum_{k=1}^K b_k R_k(v).$$

Поскольку все суммы конечные, то изменение порядка суммирования в двойной сумме законно. Подставляя теперь в (2.25) уравнения частичного баланса (2.15а), получим

$$R_k(v) = \sum_{\mathbf{n} \in S(v)} n_k p(\mathbf{n}) = \rho_k \sum_{\mathbf{n} \in S(v)} p(\mathbf{n} - \mathbf{e}_k) = \rho_k \sum_{\mathbf{n} \in S(v - b_k)} p(\mathbf{n}) = \rho_k q(v - b_k).$$

Отсюда следует (2.28). ■

Рекуррентные соотношения (2.28) позволяют получить эффективный рекуррентный алгоритм вычисления вероятностей $q(v)$, $v = \overline{0, V}$ занятия v БЦК, вероятностей π_k , $k = \overline{1, K}$ блокировок k -услуг и среднего числа занятых БЦК при стационарном режиме функционирования мультисервисной ШЦЛ. Алгоритм содержит следующие шаги.

Шаг 1. Вводим $2K + 1$ параметров V , b_k , ρ_k , $k = \overline{1, K}$ и резервируем $V + 1$ ячеек для $q(v)$ и K ячеек для π_k .

Шаг 2. Полагаем $q'(0) \leftarrow 1$, $q'(v) \leftarrow 0$, при $v < 0$.

Шаг 3. Рекуррентно по $v = \overline{1, V}$ вычисляем $q'(v) \leftarrow \frac{1}{v} \sum_{k=1}^K b_k \rho_k q'(v - b_k)$. Для

вычисления по этой формуле потребуется $2K$ операций умножения, $K - 1$ операций сложения и 1 операция деления, а всего на этом шаге потребуется выполнить $O(VK)$ арифметических

операций. Этот шаг дает ненормированное распределение вероятностей $q'(v)$, $v = \overline{0, V}$.

Шаг 4. Вычисляем нормирующую константу $G = \sum_{v=0}^V q'(v)$, что потребует $V + 1$ операций сложения.

Шаг 5. Переходим от ненормированного к нормированному распределению вероятностей: $q(v) \leftarrow q'(v) / G$, $v = \overline{0, V}$. Это потребует выполнения $V + 1$ операций деления.

Шаг 6. Получаем вероятности блокировок: $\pi_k \leftarrow \sum_{v=V-b_k+1}^V q(v)$, $k = \overline{1, K}$, что потребует $O(VK)$ операций сложения.

Шаг 7. Вычисляем среднее число занятых БЦК: $UTIL \leftarrow \sum_{v=1}^V vq(v)$. Это потребует выполнения V операций умножения и $V - 1$ операций сложения, т.е. всего $O(V)$ арифметических операций. ■

Таким образом, общее число используемых ячеек ОП составляет $O(V + 3K)$, а общее число арифметических операций – $O(VK)$. Поскольку число микросостояний в S и его подпространствах имеет комбинаторный характер и с ростом числа K услуг быстро растет, то по сравнению с прямым вычислением этих характеристик с помощью мультипликативного распределения $p(\mathbf{n})$, $\mathbf{n} \in S$ (2.7) вычисление необходимых для приложений макрохарактеристик по рекуррентным формулам дает огромный выигрыш. При этом структура рекуррентного алгоритма такова, что при его реализации исключены ошибки переполнения или обнуления, неизбежные при вычислениях с помощью (2.7).

Важным частным случаем изученной в этой главе модели является

СМО $\left. \begin{matrix} \mathbf{M} \\ \lambda, \mathbf{b} = \mathbf{1} \end{matrix} \right| \left. \begin{matrix} \mathbf{M} \\ \mu \end{matrix} \right| V < \infty$. В этой СМО заявки каждого из K видов требуют одинаковую ШПП, которую удобно принять за БЦК.

Поскольку при $\mathbf{b} = \mathbf{1}$ число занятых каналов в состоянии \mathbf{n} равно числу обслуживаемых заявок, то $U(\mathbf{n}) = \sum_{k=1}^K n_k := n_{\bullet}$. Поэтому

$S = \{\mathbf{n} : 0 \leq n_{\bullet} \leq V\}$, $S(v) = \{\mathbf{n} : n_{\bullet} = v\}$, $v = \overline{0, V}$, и из (2.28) сразу следует, что

$$vq(v) = q(v-1) \sum_{k=1}^K \rho_k = q(v-1) \rho_{\bullet}, \quad v = \overline{0, V}. \quad (2.29)$$

Следовательно, распределение числа занятых каналов в данной СМО имеет первое распределение Эрланга для полнодоступного пучка из V каналов, на который поступает пуассоновская нагрузка первого рода с интенсивностью ρ_{\bullet} :

$$q(v) = \frac{\rho_{\bullet}^v}{v!} q(0), \quad v = \overline{0, V}, \quad (2.30)$$

$$\pi_k = q(V), \quad k = \overline{1, K}. \quad (2.31)$$

Упражнение 1.1. Получите формулы (2.30) и (2.31), опираясь на теорему 1.1.

Упражнение 1.2. Рассмотрите случай $K = 1$, $\rho = \frac{\lambda_1}{\mu_1}$, $b_1 > 1$.

Глава 2. ПРИМЕНЕНИЕ КЛАССИЧЕСКИХ МУЛЬТИСЕРВИСНЫХ МОДЕЛЕЙ К АНАЛИЗУ ФРАГМЕНТА ССПС

§2.1. Анализ фрагмента иерархической сети сотовой связи

Расчет производительности систем сотовой подвижной связи (ССПС) традиционно строился на интуитивных и грубых оценках производительности, построенных, например, на основе классических формул Эрланга, которые были получены еще в 20-е годы прошлого века. Однако сети сотовой связи, как отдельный класс сетей телекоммуникаций, ставят перед теорией телетрафика большое количество задач новых типов, которые не возникают в традиционных стационарных кабельных сетях и требуют применения усовершенствованных методов теории телетрафика, таких, как многомерное распределение Эрланга, мультипликативные распределения разных типов, сложные модели нагрузки, а также эффективные численные методы вычисления необходимых для практики параметров. Естественно, что использование известных подходов не исключает и разработку новых методов теории телетрафика для оценки вероятностно-временных характеристик (ВВХ) систем сотовой связи.

На начальном этапе развертывания сети сотовой связи перед оператором стоит задача в кратчайшие сроки добиться максимальной зоны покрытия. Это означает, что прежде всего организуются *макросоты*, т.е. соты, обеспечивающие покрытие значительной территории. После достижения максимальной зоны покрытия следующей задачей становится оптимизация сети, которая заключается, в частности, в увеличении ее пропускной способности в местах значительной концентрации абонентов. Эта задача отчасти решается путем деления сот на секторы, внедрения повторного использования частот или непосредственного увеличения емкости сети за счет увеличения количества и уменьшения размера сот. Однако для такой оптимизации существуют определенные ограничения,

связанные с возможностью возникновения взаимных помех и возрастающей сложностью управления сетью. Например, при уменьшении размера сот или разделении их на секторы увеличивается частота возникновения хэндоверов, что значительно влияет на нагрузку и увеличивает затраты (см. соответствующие разделы в [14] и [18]).

Поэтому при решении задачи увеличения пропускной способности ССПС оптимальным представляется переход к *иерархической структуре сот*, при которой наряду с макросотами, которые обеспечивают широкую зону покрытия, появляются микросоты, которые размещаются внутри зоны покрытия макросот как бы под их «зонтиком» (поэтому иерархическую организацию сот еще иногда называют *зонтичной*). Передатчики микросот излучают гораздо меньшую мощность и располагаются, как правило, ниже уровня крыш зданий. Микросоты могут обеспечить высокую пропускную способность на отдельном участке сети, тем самым повышая производительность сети в местах значительного скопления абонентов.

При возникновении необходимости дальнейшего увеличения пропускной способности в отдельных частях сети могут быть введены пикосоты, по отношению к которым микросоты являются макросотами. На рис. 2.1 изображена типичная структура иерархической ССПС.

Иерархическое разделение сот, жестко связанное с размером их области покрытия, естественным образом влечет разделение их функциональности. Иерархическая сотовая структура трансформирует сотовую сеть в многоуровневую систему, где каждому уровню соответствуют специфические виды трафика:

- макросоты обеспечивают широкое покрытие и, как правило, берут на себя функции по обслуживанию быстродвижущихся абонентов, т.е. абонентов, передвигающихся в автомобилях или других транспортных средствах;

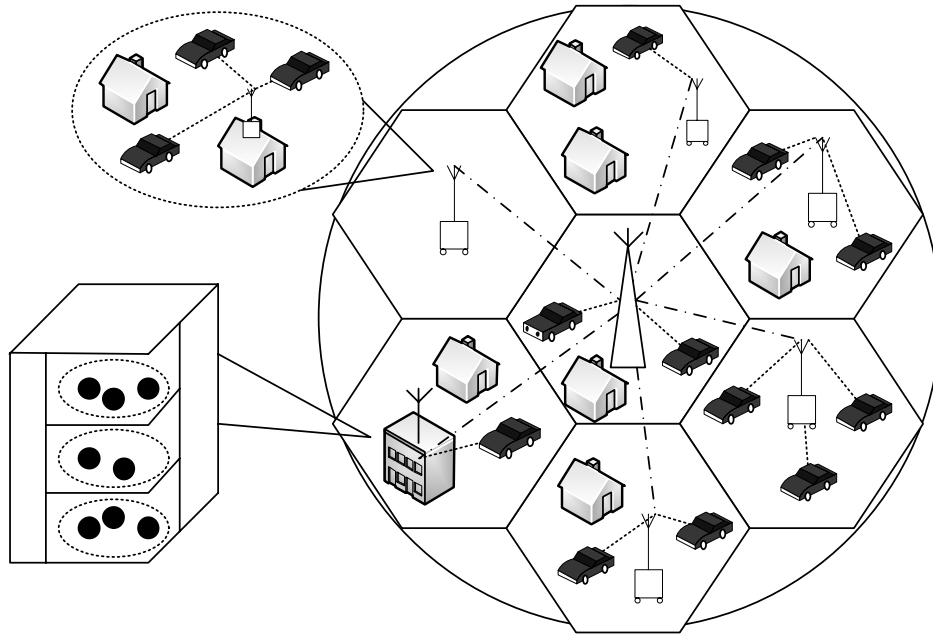


Рис. 2.1. Общий вид иерархической ССПС

- микросоты обеспечивают емкость на требуемых территориях, обслуживая медленно перемещающихся абонентов (как правило, это пешеходы или автомобили на городских улицах);
- пикосоты обеспечивают работу внутри зданий, где десятки абонентов могут быть сосредоточены на небольшой площади.

Кроме обслуживания абонентов с высокими скоростными характеристиками, макросота может также выступать в качестве дополнительного разделяемого ресурса по отношению к микросотам. Это означает, что в случае нехватки в микросоте каналов для обслуживания входящего нового или хэндовер-вызова, она может передать его на обслуживание в макросоту. Такая технология позволяет оперативно реагировать на изменение абонентской нагрузки в зоне покрытия макросоты. При использовании макросоты в качестве разделяемого ресурса должна применяться стратегия доступа, наиболее точно соответствующая динамике изменения нагрузки на микросоты. Макросота может являться и выделенным ресурсом, т.е. не обслуживать

быстродвижущихся абонентов, а только предоставлять свои каналы для обслуживания вызовов, заблокированных в микросотах.

Интуитивно понятно, что в целях наиболее эффективного использования емкости макросоты для обслуживания абонентских вызовов следует в первую очередь использовать емкость самих микросот таким образом, чтобы свободная емкость макросоты всегда оставалась максимальной. Для этого применяется метод *переупаковки каналов*, при использовании которого абонентские вызовы, занимающие каналы связи в макросоте, передаются обратно на обслуживание в микросоту при освобождении одного из ее каналов.

В этом разделе мы рассмотрим модель фрагмента (кластера) двухуровневой сети сотовой связи. Для нее будет построена математическая модель и выведены условия, при которых вероятностно-временные характеристики для этой модели могут быть получены с использованием изложенных ранее алгоритмов.

2.1.1. Модель кластера двухуровневой ССПС

Рассмотрим один кластер сети сотовой подвижной связи, состоящий из K микросот и одной макросоты, покрывающей все K микросот (рис. 2.2)¹. Будем считать, что в любой момент времени базовая станция (БС) микросоты k располагает набором из c_k частотных каналов для обслуживания входящих новых и хэндовер-вызовов, поступающих от абонентов, $k = \overline{1, K}$. Базовая станция макросоты располагает набором из C частотных каналов для обслуживания вызовов, заблокированных в микросотах.

Примем, что вновь возникающие вызовы в микросоте k образуют пуассоновский поток первого рода с интенсивностью λ_{0k} , $k = \overline{1, K}$,

¹ Башарин Г.П., Меркулов В.Е. Анализ пропускной способности в иерархических сетях сотовой подвижной связи // Электросвязь, 2003, № 4, с. 45–47.

$\sum_{k=1}^K \lambda_{0k} =: \lambda_0$. Положим $\theta_{0k} = \frac{\lambda_{0k}}{\lambda_0}$ – вероятность того, что новый вызов

поступит в микросоту k , $k = \overline{1, K}$. Узел 0 обозначает «внешнюю среду», так что $\theta_{00} := 0$.

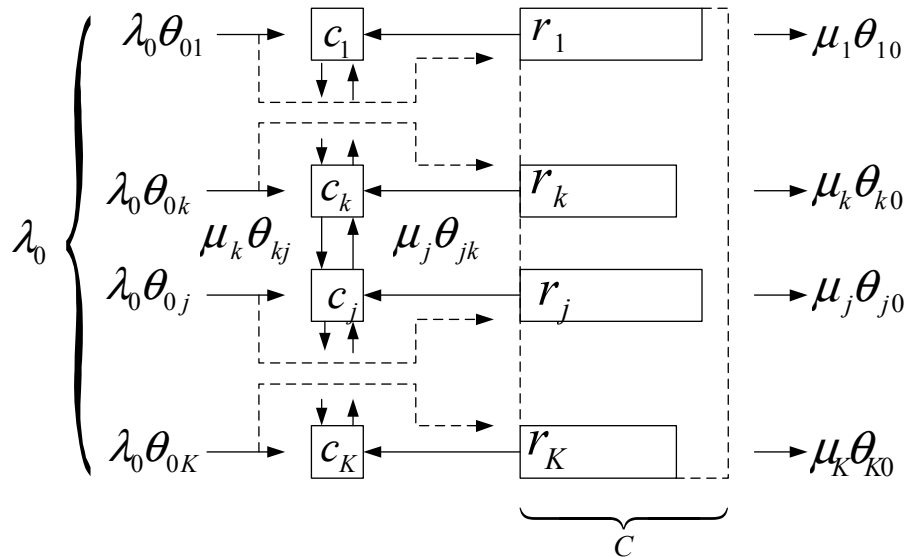


Рис. 2.2. Математическая модель кластера двухуровневой ССПС с одной макросотой, покрывающей K микросот

Вызов (как новый, так и хэндовер), принятый на обслуживание в k -микросоте, удерживает выделенный ему частотный канал случайное время, распределенное экспоненциально со средним значением $1/\mu_k$. Освобождение выделенного частотного канала происходит либо за счет окончания разговора с вероятностью $\theta_{k0} > 0$, либо за счет ухода абонента из зоны обслуживания базовой станции k в зону обслуживания одной из соседних с ней базовых станций с вероятностью $\theta_{kj} \geq 0$, $\theta_{kk} = 0$,

$$\sum_{j=1}^K \theta_{kj} + \theta_{k0} = 1, \quad k, j = \overline{1, K}, \quad j \neq k.$$

Маршрутная матрица $\Theta_0 = [\theta_{kj}]_{k, j = \overline{0, K}}$, где $\theta_{00} = 0$, в данном случае имеет вид, изображенный в таблице 2.1.

Таблица 2.1. Маршрутная матрица Θ_0

Θ_0	0	1	2	...	K	Σ
0	0	θ_{01}	θ_{02}	...	θ_{0K}	1
1	θ_{10}	0	θ_{12}	...	θ_{1K}	1
2	θ_{20}	θ_{21}	0	...	θ_{2K}	1
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
K	θ_{K0}	θ_{K1}	θ_{K2}	...	0	1

Новый или хэндовер-вызов, который не может быть принят на обслуживание базовой станцией микросоты из-за нехватки свободных частотных каналов, передается базовой станции макросоты. При этом устанавливается ограничение r_k на количество одновременно обслуживаемых макросотой абонентов k -микросоты. Если поступающий вызов, блокированный базовой станцией микросоты, также блокируется базовой станцией макросоты, то вызов теряется, не оказывая влияния ни на поступающие потоки вызовов, ни на функционирование системы.

Примем дополнительно, что в кластере реализована мгновенная переупаковка каналов. На языке модели это означает, что при освобождении одного из каналов в k -микросоте один из k -вызовов, обслуживаемых базовой станцией макросоты, мгновенно передается базовой станции k -микросоты, занимая в ней освободившийся канал. Канал базовой станции макросоты при этом освобождается.

Функционирование рассматриваемой системы описывается K -мерным марковским процессом $\mathbf{X}(t) = \{X_1(t), \dots, X_K(t)\}$, $t \geq 0$, с пространством состояний

$$S = \left\{ \mathbf{n} = (n_1, \dots, n_K) : \mathbf{0} \leq \mathbf{n} \leq \mathbf{c} + \mathbf{r}, \sum_{k=1}^K (n_k - c_k)^+ \leq C \right\},$$

где $(x)^+ = \begin{cases} x, & x > 0; \\ 0, & x \leq 0. \end{cases}$, а $X_k(t)$ – число обслуживаемых системой абонентов,

находящихся в зоне покрытия БС k , $k = \overline{1, K}$.

Из предположений о поступающей пуассоновской нагрузке и экспоненциальном времени нахождения в ячейке следует, что $\mathbf{X}(t)$ – СтМП с пространством состояний S и матрицей интенсивностей переходов $\mathbf{A} = [a(\mathbf{m}, \mathbf{n})]_{\mathbf{m}, \mathbf{n} \in S}$:

$$a(\mathbf{m}, \mathbf{n}) = \begin{cases} \lambda_0 \theta_{0k}, & \mathbf{n} = \mathbf{m} + \mathbf{e}_k, k = \overline{1, K}; \\ \mu_k n_k \theta_{k0}, & \mathbf{n} = \mathbf{m} - \mathbf{e}_k, k = \overline{1, K}; \\ \mu_k n_k \theta_{kj}, & \mathbf{n} = \mathbf{m} - \mathbf{e}_k + \mathbf{e}_j, k, j = \overline{1, K}, j \neq k; \\ 0, & \text{в остальных случаях,} \end{cases} \quad (1.1)$$

где $\mathbf{m}, \mathbf{n} \in S$, а \mathbf{e}_k – k -й столбец единичной матрицы размерности K .

2.1.2. Вывод условий мультипликативности

Введенная математическая модель во многом похожа на модель открытой экспоненциальной сети массового обслуживания. Отличие состоит в том, что в рассматриваемой модели емкости узлов (микросот) ограничены и, кроме этого, узлы используют общий ресурс (макросоту). Последняя особенность означает, что возможность изменения состояний отдельных узлов зависит от текущего состояния всех узлов сети. С другой стороны, изучаемую систему можно рассматривать как многопоточковую моносервисную неполнодоступную СМО с индивидуальными потолками (Sharing with Maximum Queue Length and Minimum Allocation, SMQMA) –

схема $\langle S^2, A_5 \rangle$ в [1, гл. 5]. Однако в отличие от этой классической модели в рассматриваемой системе принятые на обслуживание вызовы могут переходить с одного прибора на другой, что соответствует передаче вызовов на обслуживание из одной микросоты в другую. В обеих аналогиях указанные отличия влекут внесение изменений в матрицу интенсивностей переходов \mathbf{A} , которые могут повлиять на обратимость описывающего функционирование системы марковского процесса $\mathbf{X}(t)$. Поэтому, несмотря на близость рассматриваемой модели к классическим мультипликативным системам, факт наличия у нее мультипликативного стационарного распределения требует отдельного исследования. В этом разделе мы покажем, что для рассматриваемой модели стационарное распределение не всегда имеет мультипликативный вид, и выведем условия, при выполнении которых мультипликативное стационарное распределение существует.

Как и для открытых экспоненциальных сетей, введем вектор $\boldsymbol{\lambda}$, который является решением матричного уравнения

$$\boldsymbol{\lambda}^T (\mathbf{I} - \boldsymbol{\Theta}) = \boldsymbol{\lambda}_0 (\theta_{01}, \dots, \theta_{0K}), \quad (1.2)$$

где $\boldsymbol{\Theta} = [\theta_{ij}]_{i,j=1,\overline{K}}$ – главная подматрица маршрутной матрицы $\boldsymbol{\Theta}_0$, а \mathbf{I} – единичная матрица размерности K . Заметим, что в отличие от открытых экспоненциальных СеМО элементы вектора $\boldsymbol{\lambda}$ не являются интенсивностями суммарных потоков вызовов, поступающих на каждый узел (микросоту), поскольку в рассматриваемой системе существуют потери.

Лемма 2.1. *Справедливо следующее соотношение:*

$$\sum_{j=0}^K (\lambda_j \theta_{jk} - \lambda_k \theta_{kj}) = 0, \quad k = \overline{1, K}. \quad (1.3)$$

Доказательство. Заметим, что

$$\lambda_0 \theta_{0k} = \left[\lambda^T (\mathbf{I} - \Theta) \right]_k = \lambda_k - \sum_{j=1}^K \lambda_j \theta_{jk} \Leftrightarrow \lambda_k = \sum_{j=0}^K \lambda_j \theta_{jk}.$$

С другой стороны, из стохастического свойства матрицы Θ_0 следует, что

$$\lambda_k = \sum_{j=0}^K \lambda_k \theta_{kj}.$$

Утверждение леммы вытекает непосредственно из двух приведенных соотношений. ■

Введем функцию-индикатор

$$l(\mathbf{n}) = \begin{cases} 1, & \mathbf{n} \in S, \\ 0, & \mathbf{n} \notin S \end{cases}$$

и запишем систему уравнений глобального баланса (СУГБ) для рассматриваемой системы следующим образом:

$$\begin{aligned} p(\mathbf{n}) \left(\sum_{k=1}^K \lambda_0 \theta_{0k} l(\mathbf{n} + \mathbf{e}_k) + \sum_{k=1}^K \mu_k n_k \theta_{k0} + \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K \mu_k n_k \theta_{kj} l(\mathbf{n} + \mathbf{e}_j) \right) = \\ = \sum_{k=1}^K p(\mathbf{n} + \mathbf{e}_k) (n_k + 1) \mu_k \theta_{k0} (\mathbf{n} + \mathbf{e}_k) + \sum_{k=1}^K p(\mathbf{n} - \mathbf{e}_k) \lambda_0 \theta_{0k} (\mathbf{n} - \mathbf{e}_k) + \\ + \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p(\mathbf{n} - \mathbf{e}_k + \mathbf{e}_j) (n_j + 1) \mu_j \theta_{jk} l(\mathbf{n} - \mathbf{e}_k) l(\mathbf{n} + \mathbf{e}_j), \quad \mathbf{n} \in S. \end{aligned} \quad (1.4)$$

Следующая теорема дает ответ на вопрос, при каких условиях СУГБ (1.4) имеет решение мультипликативного вида.

Теорема 2.1. Пусть для элементов вектора λ (1.2) выполняются соотношения

$$\lambda_k \theta_{kj} = \lambda_j \theta_{jk}, \quad k = \overline{1, K}, \quad j \neq k. \quad (1.5)$$

Тогда

а) выполняются дополнительные условия

$$\lambda_0 \theta_{0k} = \lambda_k \theta_{k0}, \quad k = \overline{1, K}; \quad (1.6)$$

б) выполняются условия частичного баланса

$$\begin{cases} p(\mathbf{n} - \mathbf{e}_k) \lambda_0 \theta_{0k} 1(\mathbf{n} - \mathbf{e}_k) = p(\mathbf{n}) \mu_k n_k \theta_{k0}, \\ p(\mathbf{n} - \mathbf{e}_k + \mathbf{e}_j) \mu_j (n_j + 1) \theta_{jk} 1(\mathbf{n} - \mathbf{e}_k) = p(\mathbf{n}) \mu_k n_k \theta_{kj}, \\ k, j = \overline{1, K}, j \neq k, \mathbf{n} \in S, \end{cases} \quad (1.7)$$

и СУГБ (1.4) имеет единственное ненулевое решение мультипликативного вида

$$p(\mathbf{n}) = G \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!}, \quad \rho_k = \frac{\lambda_k}{\mu_k}, \quad k = \overline{1, K}, \quad G = \left(\sum_{\mathbf{n} \in S} \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!} \right)^{-1}; \quad (1.8)$$

в) соотношения (1.6) являются необходимым условием того, чтобы решение (1.8) удовлетворяло СУГБ (1.4).

Доказательство. Прежде всего заметим, что при выполнении условий (1.5) дополнительное условие (1.6) непосредственно следует из леммы 2.1.

Теперь подставим (1.8) в (1.4) и разделим полученное равенство на $p(\mathbf{n})$:

$$\begin{aligned} & \sum_{k=1}^K \lambda_0 \theta_{0k} 1(\mathbf{n} + \mathbf{e}_k) + \sum_{k=1}^K \mu_k n_k \theta_{k0} + \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K \mu_k n_k \theta_{kj} 1(\mathbf{n} + \mathbf{e}_j) = \\ & = \sum_{k=1}^K \lambda_k \theta_{k0} 1(\mathbf{n} + \mathbf{e}_k) + \sum_{k=1}^K \mu_k \frac{\lambda_0}{\lambda_k} \theta_{0k} + \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K \mu_k n_k \frac{\lambda_j}{\lambda_k} \theta_{jk} 1(\mathbf{n} + \mathbf{e}_k), \quad \mathbf{n} \in S. \end{aligned} \quad (1.9)$$

Приравнявая по отдельности соответствующие суммы в этом соотношении, получим:

$$\begin{cases} \sum_{k=1}^K (\lambda_0 \theta_{0k} - \lambda_k \theta_{k0}) 1(\mathbf{n} + \mathbf{e}_k) = 0; \\ \sum_{k=1}^K (\theta_{k0} - \frac{\lambda_0}{\lambda_k} \theta_{0k}) \mu_k n_k = 0; \\ \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K (\theta_{kj} - \frac{\lambda_j}{\lambda_k} \theta_{jk}) \mu_k n_k 1(\mathbf{n} + \mathbf{e}_k) = 0. \end{cases} \quad (1.10)$$

Очевидно, что выполнение условий (1.5)–(1.6) обращает все уравнения системы (1.10) в тождества, т.е. решение (1.8) в этом случае удовлетворяет СУГБ (1.4). Заметим также, что, умножая все соотношения (1.10) на $p(\mathbf{n})$ и производя обратные преобразования, мы получим условия частичного баланса (1.7), т.е. факт существования такого частичного баланса также доказан.

Пусть теперь решение (1.8) удовлетворяет СУГБ (1.4). Перепишем (1.9) следующим образом:

$$\sum_{k=1}^K 1(\mathbf{n} + \mathbf{e}_k)(\lambda_0 \theta_{0k} - \lambda_k \theta_{k0}) + \sum_{k=1}^K \sum_{\substack{j=0 \\ j \neq k}}^K \mu_k n_k 1(\mathbf{n} + \mathbf{e}_j)(\theta_{kj} - \frac{\lambda_j}{\lambda_k} \theta_{jk}) = 0, \quad \mathbf{n} \in S, \quad (1.11)$$

где \mathbf{e}_0 – нулевой вектор размерности K . Соотношение (1.11) должно выполняться для всех $\mathbf{n} \in S$. Пусть $\mathbf{n} = (0, \dots, 0) \in S$. Тогда из (1.11) следует необходимое условие

$$\sum_{k=1}^K (\lambda_0 \theta_{0k} - \lambda_k \theta_{k0}) = 0. \quad (1.12)$$

Теперь пусть $\mathbf{n} = (c_i + r_i) \mathbf{e}_i$ для некоторого $i \in \{1, \dots, K\}$. Очевидно, что такое состояние принадлежит пространству S и для него из формулы (1.11) следует, что

$$\sum_{\substack{k=1 \\ k \neq i}}^K (\lambda_0 \theta_{0k} - \lambda_k \theta_{k0}) + \mu_i (c_i + r_i) \sum_{\substack{j=0 \\ j \neq i}}^K (\theta_{ij} - \frac{\lambda_j}{\lambda_i} \theta_{ji}) = 0. \quad (1.13)$$

Второе слагаемое в соотношении (1.13) равно нулю в силу леммы 2.1. Учитывая, что из (1.12)

$$\sum_{\substack{k=1 \\ k \neq i}}^K (\lambda_0 \theta_{0k} - \lambda_k \theta_{k0}) = (\lambda_i \theta_{i0} - \lambda_0 \theta_{0i}),$$

получаем необходимое условие $(\lambda_i \theta_{i0} - \lambda_0 \theta_{0i}) = 0$ для всех $i = \overline{1, K}$. ■

Чтобы лучше понять физический смысл условий (1.5), рассмотрим пример модели кластера ССПС с количеством микросот $K = 2$. Для такой системы СУР (1.2) принимает вид

$$\begin{cases} \lambda_1 - \lambda_2 \theta_{21} = \lambda_0 \theta_{01}, \\ \lambda_2 - \lambda_1 \theta_{12} = \lambda_0 \theta_{02} \end{cases}$$

и имеет решение

$$\lambda_1 = \lambda_0 \frac{\theta_{01} + \theta_{02} \theta_{21}}{1 - \theta_{12} \theta_{21}}, \quad \lambda_2 = \lambda_0 \frac{\theta_{02} + \theta_{01} \theta_{12}}{1 - \theta_{12} \theta_{21}}.$$

С учетом этих соотношений условие $\lambda_1 \theta_{12} = \lambda_2 \theta_{21}$ после преобразований принимает вид $\theta_{01} \theta_{12} \theta_{20} = \theta_{02} \theta_{21} \theta_{10}$.

Таким образом, для кластера с двумя микросотами условия (1.5) эквивалентны тому, что в маршрутной матрице вероятность маршрута $0 \rightarrow 1 \rightarrow 2 \rightarrow 0$ равна вероятности обратного маршрута $0 \rightarrow 2 \rightarrow 1 \rightarrow 0$, т.е. вероятность появления вызова в первой микросоте с последующим переходом его во вторую микросоту и выходом из системы равна вероятности появления вызова вначале во второй микросоте, затем перехода его в первую и далее выхода из системы. Заметим, что аналогичным свойством очевидно обладают также маршруты $0 \rightarrow 1 \rightarrow 0$ и $0 \rightarrow 2 \rightarrow 0$, т.к. каждый из них является одним и тем же маршрутом в прямом и обратном направлениях.

Обобщая полученные результаты, можно сделать вывод, что рассматриваемая система имеет мультипликативное стационарное распределение в том случае, если вероятности перемещения абонента по любому маршруту, определенному в маршрутной матрице, равны при движении в прямом и обратном направлениях. Доказательство правомерности такого обобщения дает представленная ниже теорема 2.2.

Рассмотрим неориентированный граф G с множеством вершин $V = \{v_0, \dots, v_K\}$, соответствующих микросотам рассматриваемой системы

(включая «внешнюю среду») и матрицей смежности $\mathbf{M} = [m(v_i, v_j)]_{(v_i, v_j) \in V}$,
 $m(v_i, v_j) = u(\theta_{ij})$, $i, j = \overline{0, K}$, где $u(\cdot)$ – функция Хевисайда.

Будем называть *маршрутом абонента* любой замкнутый маршрут в графе \mathbf{G} , который проходит через вершину v_0 ровно один раз. Будем при этом называть v_0 началом и окончанием маршрута абонента.

Произведение $\theta_{0i_1} \theta_{i_1 i_2} \cdots \theta_{i_{n-1} i_n} \theta_{i_n 0}$ будем называть *вероятностью маршрута абонента* $(v_0 v_{i_1})(v_{i_1} v_{i_2}) \cdots (v_{i_{n-1}} v_{i_n})(v_{i_n} v_0)$.

Обратным маршрутом абонента по отношению к маршруту $(v_0 v_{i_1})(v_{i_1} v_{i_2}) \cdots (v_{i_{n-1}} v_{i_n})(v_{i_n} v_0)$ будем называть маршрут $(v_0 v_{i_n})(v_{i_n} v_{i_{n-1}}) \cdots (v_{i_2} v_{i_1})(v_{i_1} v_0)$.

Лемма 2.2. Пусть для любых $k, j \in \{1, \dots, K\}$, $j \neq k$, выполняется соотношение

$$\theta_{0k} \theta_{kj} \theta_{j0} = \theta_{0j} \theta_{jk} \theta_{k0}. \quad (1.14)$$

Тогда для любого маршрута $(v_0 v_{i_1})(v_{i_1} v_{i_2}) \cdots (v_{i_{n-1}} v_{i_n})(v_{i_n} v_0)$ его вероятность совпадает с вероятностью маршрута, обратного ему, т.е.

$$\theta_{0i_1} \theta_{i_1 i_2} \cdots \theta_{i_{n-1} i_n} \theta_{i_n 0} = \theta_{0i_n} \theta_{i_n i_{n-1}} \cdots \theta_{i_2 i_1} \theta_{i_1 0}. \quad (1.15)$$

Доказательство. Выберем произвольно $i \in \{1, \dots, K\}$, $i \neq j$. Умножая левую и правую части соотношения (1.14) соответственно на левую и правую части соотношения $\theta_{0j} \theta_{ji} \theta_{i0} = \theta_{0i} \theta_{ij} \theta_{j0}$, получим

$$\theta_{0k} \theta_{kj} \theta_{ji} \theta_{i0} = \theta_{0i} \theta_{ij} \theta_{jk} \theta_{k0}, \quad k, j, i = \overline{1, K}, \quad j \neq k, \quad i \neq j.$$

Действуя аналогичным образом, получим, что вероятности прохождения любого маршрута в прямом и обратном направлениях равны. ■

Теорема 2.2. Для выполнения условий (1.5) теоремы 2.1 необходимо и достаточно, чтобы выполнялись условия (1.14) леммы 2.2.

Доказательство. Докажем вначале необходимость. Если выполнены условия (1.5), то в силу теоремы 1.1 выполняются и условия (1.6). Подставляя (1.6) в (1.5), получим условия (1.14):

$$\lambda_0 \frac{\theta_{0k}}{\theta_{k0}} \theta_{kj} = \lambda_0 \frac{\theta_{0j}}{\theta_{j0}} \theta_{jk} \Rightarrow \theta_{0k} \theta_{kj} \theta_{j0} = \theta_{0j} \theta_{jk} \theta_{k0}, k, j = \overline{1, K}, j \neq k. \quad (1.16)$$

Теперь докажем достаточность. Пусть выполнены условия (1.14). Тогда в силу леммы 2.2 для любого маршрута $(v_0 v_{i_1})(v_{i_1} v_{i_2}) \cdots (v_{i_{n-1}} v_{i_n})(v_{i_n} v_0)$ выполнено соотношение (1.15).

Для доказательства достаточности воспользуемся математическим аппаратом теории сетей массового обслуживания. Введем величины h_{ij} – среднее количество посещения абонентом микросоты j при условии, что его маршрут начался в микросоте i , $i, j = \overline{1, K}$. Величины h_{ij} определяются соотношениями:

$$h_{ij} = [(\mathbf{I} - \Theta)^{-1}]_{ij} = \left[\sum_{n=0}^{\infty} \Theta^n \right]_{ij}.$$

Доказательство можно найти, например, в [1, гл. 4]. Таким образом, для любых $i, j = \overline{1, K}$, $i \neq j$,

$$h_{ij} = \theta_{ij} + \sum_{k=1}^K \theta_{ik} \theta_{kj} + \sum_{k=1}^K \sum_{l=1}^K \theta_{ik} \theta_{kl} \theta_{lj} + \dots \quad (1.17)$$

Домножая h_{ij} и h_{ji} соответственно на $\theta_{0i} \theta_{j0}$ и $\theta_{0j} \theta_{i0}$ и используя равенство (1.15), получим

$$\theta_{0i} h_{ij} \theta_{j0} = \theta_{0j} h_{ji} \theta_{i0}, i, j = \overline{1, K}, i \neq j. \quad (1.18)$$

Теперь введем величины h_i – среднее количество посещений абонентом микросоты i независимо от того, в какой микросоте был начат маршрут. Из теории СеМО (см. [1, гл. 4]) известно, что

$$h_j = \sum_{i=1}^K \theta_{0i} h_{ij} = \frac{\lambda_j}{\lambda_0}, j = \overline{1, K}.$$

Поэтому, суммируя (1.18) по $i = \overline{1, K}$, получим

$$\lambda_j \theta_{j0} = \lambda_0 \theta_{0j} \sum_{i=1}^K h_{ji} \theta_{i0}, i, j = \overline{1, K}, i \neq j. \quad (1.19)$$

Элементы суммы в правой части (1.19) представляют собой среднее число посещений абонентом каждой из микросот, при отправке из микросоты j , с последующим мгновенным выходом из системы. Поэтому вся сумма есть среднее количество выходов абонента из системы (при условии, что маршрут был начат в микросоте j) и равна 1.

Таким образом, имеют место условия (1.6). Прodelывая с использованием этих условий преобразования, обратные (1.16), получим выполнение условий (1.5). ■

Необходимо отметить, что в доказанной теореме речь идет именно о маршруте абонента, а не о маршруте его вызова, который может быть вынужденно прерван из-за блокировок. Типичные маршруты абонентов в зоне покрытия кластера ССПС, как правило, являются объективной информацией, в той или иной мере известной на момент проектирования сети. Следовательно, целесообразность предположения о выполнении условий (1.14) и наличии мультипликативного распределения может быть определена на достаточно раннем этапе построения ССПС.

Примерами систем двухуровневых ССПС, для которых выполняются условия (1.14) (и, следовательно, условия (1.5)), могут служить системы с равными по размеру микросотами, покрывающими территорию с большой плотностью абонентов, движение которых в разных направлениях можно принять равновероятным (крупный выставочный центр, бизнес-центр или торговая территория). Микросоты могут также располагаться вдоль крупных городских улиц, шоссе или железнодорожных путей, таких, на которых потоки пешеходов, машин или поездов можно считать равными в обоих направлениях. Таким образом, теоремы 2.1 и 2.2 дают возможность

применить классические вычислительные алгоритмы теории мультипликативных сетей для анализа целого класса двухуровневых сетей сотовой связи.

2.1.3. Анализ основных ВВХ

При выполнении условий (1.5) процедура анализа ВВХ рассматриваемой системы ничем не отличается от анализа неполнодоступной СМО с индивидуальными потолками. Известно большое количество методов и алгоритмов, которые могут быть использованы для вычисления необходимых ВВХ. К таким методам относятся модифицированный метод Бузена, эффективные сверточные алгоритмы, метод усеченных сверток, обобщенная рекурсия Кофмана-Робертса и др.

В частном случае, когда $c_k = 0$, $r_k = C$, $k = \overline{1, K}$, применима обычная рекурсия Кофмана-Робертса¹

$$P\{n_{\bullet} = m\} = \frac{1}{m} \sum_{k=1}^K \rho_k P\{n_{\bullet} = m - 1\}, m = \overline{0, c_{\bullet} + C}. \quad (1.20)$$

Заметим, что в силу условий (1.6) $\rho_k = \frac{\lambda_k}{\mu_k} = \frac{\lambda_0 \theta_{0k}}{\mu_k \theta_{k0}}$, т.е. вероятности

занятия ровно m каналов в системе, $m = \overline{0, c_{\bullet} + C}$, управляются только частичными балансами потоков вызовов между каждой из микросот и «внешней средой».

§2.2. Анализ ВВХ микросоты с двумя типами каналов и учетом мобильности абонентов

Этот параграф посвящен одной из проблем, возникших при переходе от систем 2G к системам 3G, существенной особенностью которых является многоуровневая архитектура с микро- и пикосотовым покрытием

¹ Первоначально этот метод был предложен Р. Форте – К. Гранджаном в 1964 г., а затем в 1981 г. вновь открыт Дж. С. Кофманом и Дж. У. Робертсом. Поэтому данный метод часто называют именем двух последних исследователей (см. [8, 15]).

[14]. При уменьшении размеров сот мобильность абонентов и случайные флуктуации их численности в сотах сильно влияют на пропускную способность и качество обслуживания в них. Поэтому для своевременного обнаружения перегрузок в сотах и оперативного перераспределения между ними резервных разговорных каналов необходимы новые алгоритмы контроля мобильности абонентов¹. Для этого в дополнение к процессам поступления вызовов от зарегистрированных в соте абонентов и их обслуживания необходимо рассматривать процессы прихода и ухода из соты пассивных (зарегистрированных) и активных (зарегистрированных и разговаривающих) абонентов. При приходе или уходе из соты активных абонентов происходит процедура межсотового хэндовера.

Таким образом, увеличивается число параметров, характеризующих процесс функционирования соты, и непосредственное применение классических формул Эрланга для вычисления ВВХ рассматриваемой системы сотовой связи невозможно.

2.2.1. Построение математической модели²

Примем, что сеть сотовой подвижной связи состоит из одинаковых сот, выделим некоторую соту (сота 1) и рассмотрим ее как СМО с двумя типами каналов. Каналы первого типа соответствуют каналам для регистрации пребывания абонентов в соте 1, и их число можно считать неограниченным. Каналы второго типа соответствуют разговорным каналам в соте 1, их число равно C .

В качестве правила установления соединения, дающего приоритет хэндовер-вызовам, используется выделение резервных каналов. Это означает, что базовой станции назначен некоторый целочисленный порог g , не превышающий C . Как только количество свободных разговорных

¹ Шорин О.А. Вероятность перегрузки сотовых систем связи с учетом подвижности абонентов // Электросвязь, 2004, № 5, с. 23–26.

² Башарин Г.П., Серебrenникова Н.В. Учет мобильности абонентов в микросоте с каналами двух типов // Электросвязь, 2007, № 11, с. 52–55.

каналов становится равным g , оставшиеся разговорные каналы выделяются для обслуживания только входящих хэндовер-вызовов. Запрос на установление соединения от уже зарегистрированного в соте абонента будет заблокирован. В таком режиме базовая станция функционирует до тех пор, пока количество свободных разговорных каналов не станет больше g . Обозначим $C_0 := C - g$ число каналов для обслуживания новых вызовов от абонентов, зарегистрированных в соте 1.

Примем, что поток поступления пассивных (только на регистрацию) и активных (с хэндовером на регистрацию и разговор одновременно) абонентов в соту 1 – пуассоновские и имеют постоянные интенсивности λ_0 и λ_H соответственно. Будем считать, что уход из соты 1 абонентов любого типа происходит с одинаковой интенсивностью ν мобильности. Кроме того, поток поступления запросов на установление соединения и предоставление разговорного канала от одного пассивного зарегистрированного в соте 1 абонента также является пуассоновским с постоянной интенсивностью ε . Длительность разговора в соте имеет экспоненциальное распределение с интенсивностью μ его успешного (без хэндовера) завершения и перехода активного абонента в пассивное состояние с освобождением разговорного канала и дальнейшим удержанием канала регистрации.

Процесс функционирования такой СМО при сделанных предположениях и двух структурных (C, g) и пяти нагрузочных параметрах описывается ступенчатым МП $(X(t), Y(t))$, $t \geq 0$, с двумерным пространством состояний $I = \{(i, j) : i = 0, 1, \dots; 0 \leq j \leq \min(i, C)\}$, показанным на рис. 2.3. Здесь i – число зарегистрированных в соте абонентов, j – число активных из них, т.е. число занятых каналов обслуживания.

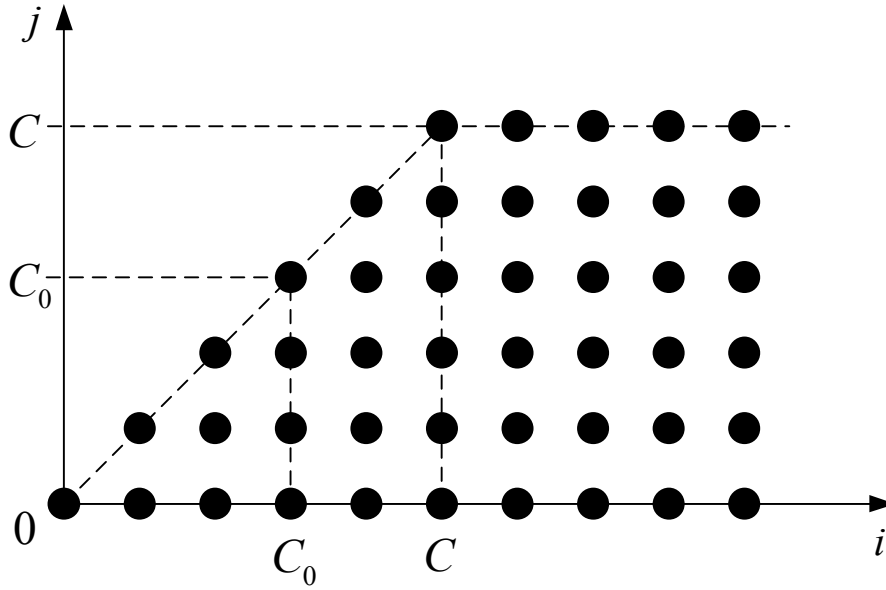


Рис. 2.3. Пространство I состояний МП $(X(t), Y(t))$, $t \geq 0$

2.2.2. Вывод СУГБ

Для получения ВВХ рассматриваемой системы необходимо найти равновесное распределение $P(i, j)$, $i, j \in I$, являющееся решением СУГБ.

Вид матрицы A интенсивностей переходов между состояниями системы в случае $C=3$ и $g=1$ представлен в таблице 2.2, причем для вычисления столбца « Σ » необходимо дополнительно выписать опущенные столбцы с ненулевыми элементами.

На рис. 2.4 приводится диаграмма интенсивностей переходов за время Δt между соседними состояниями процесса $(X(t), Y(t))$, с помощью которой можно получить компактную запись СУГБ:

$$\begin{aligned}
 & [\lambda_0 + \lambda_H + j\mu + j\nu + (i-j)\nu + (i-j)\epsilon u(C_0 - j)]P(i, j) = \\
 & = (i-j+1)\nu P(i+1, j) + (i-j+1)\epsilon u(C_0 + 1 - j)u(j)P(i, j-1) + \\
 & + \lambda_H u(j)P(i-1, j-1) + \lambda_0 u(i-j)P(i-1, j) + \\
 & + \lambda_H u(i-C)u(j-C+1)P(i-1, j) + (j+1)\nu u(C-j)P(i+1, j+1) + \\
 & + (j+1)\mu u(i-j)u(C-j)P(i, j+1), \quad i \geq 0, \quad 0 \leq j \leq \min(i, C),
 \end{aligned} \tag{2.1}$$

где $u(x)$ – функция Хевисайда.

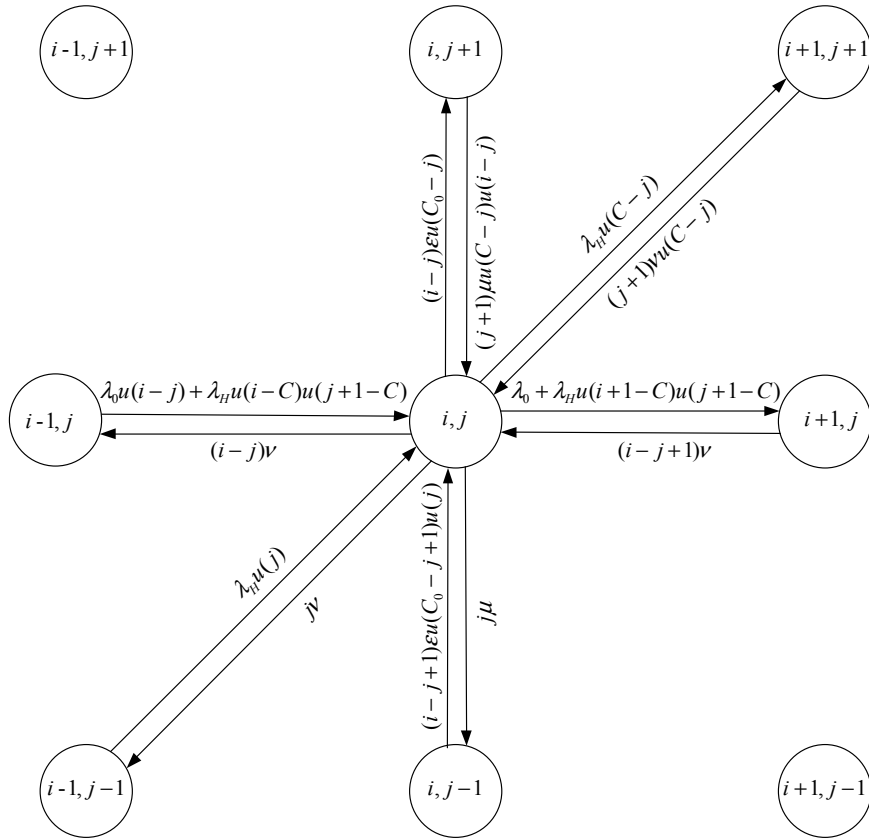


Рис. 2.4. Диаграмма интенсивностей переходов между состоянием $(i, j) \in I$ и соседними состояниями

Упражнение 2.1. Для наглядности рекомендуем записать систему (2.1) отдельно в каждом из непересекающихся подпространств (см. рис. 2.5):

$$I^{(1)} = \{(i, j) : i > 0; j = 0\},$$

$$I^{(2)} = \{(i, j) : i > 0; 0 < j < \min(i, C_0)\},$$

$$I^{(3)} = \{(i, j) : i > C_0; C_0 < j < \min(i, C)\},$$

$$I^{(4)} = \{(i, C_0) : i > C_0\},$$

$$I^{(5)} = \{(i, i) : 0 \leq i \leq C_0\},$$

$$I^{(6)} = \{(i, i) : C_0 < i \leq C\},$$

$$I^{(7)} = \{(i, C) : i > C\}, \quad I = \bigcup_{k=1}^7 I^{(k)}.$$

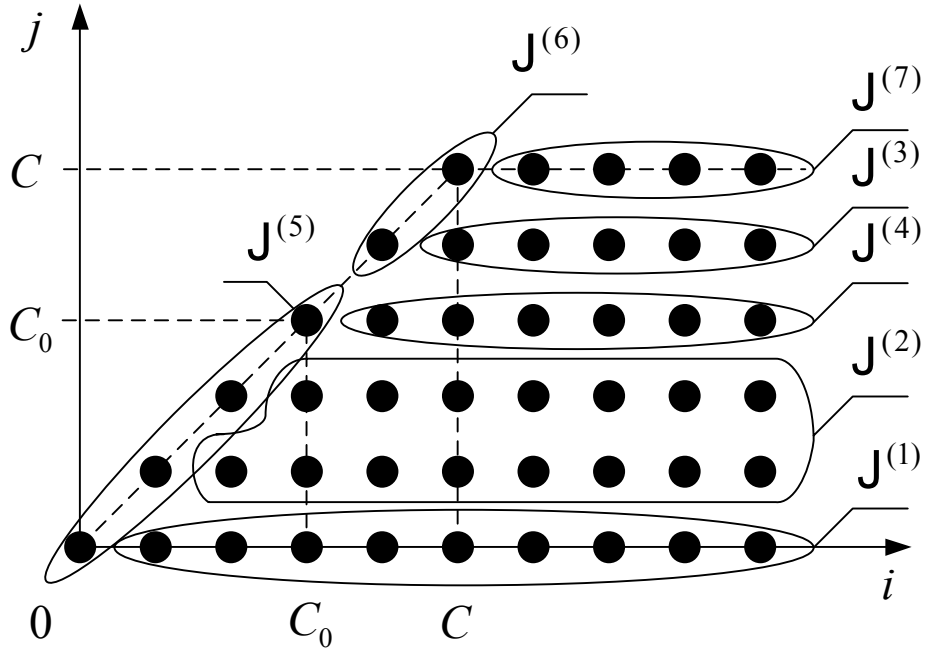


Рис. 2.5. Подпространства $l^{(k)}$, $k = \overline{1,7}$

Упражнение 2.2. Проверить, что если вместо функции Хевисайда $u(x)$ использовать функцию-индикатор $1(\square)$ события, указанного в скобках, то СУГБ (2.1) примет вид:

$$\begin{aligned}
 & [\lambda_0 + \lambda_H + j\mu + j\nu + (i-j)\nu + (i-j)\epsilon 1(j < C_0)] P(i, j) = \\
 & = (i-j+1)\nu P(i+1, j) + (i-j+1)\epsilon 1(0 < j < C_0 + 1) P(i, j-1) + \\
 & + \lambda_H 1(j > 0) P(i-1, j-1) + \lambda_0 1(i > j) P(i-1, j) + \\
 & + \lambda_H 1(i > C, j = C) P(i-1, j) + (j+1)\nu 1(j < C) P(i+1, j+1) + \\
 & + (j+1)\mu 1(j < i, j < C) P(i, j+1), \quad i \geq 0, \quad 0 \leq j \leq \min(i, C). \quad (2.1a)
 \end{aligned}$$

2.2.3. Вывод одномерных распределений

Для анализа ВВХ СМО рассмотрим два основных одномерных распределения в соте 1 – для числа i зарегистрированных и числа j активных абонентов:

$$P\{X = i\} = \sum_{j=0}^{\min(i, C)} P(i, j) =: P(i, \square), \quad i = 0, 1, \dots$$

$$P\{Y = j\} = \sum_{i=j}^{\infty} P(i, j) =: P(\square, j), \quad j = 0, 1, \dots, C.$$

Таблица 2.2. Вид матрицы интенсивностей переходов для $C = 3$ и $g = 1$

A	(0,0)	(1,0)	(2,0)	...	(1,1)	(2,1)	...	(2,2)	(3,2)	...	(3,3)	(4,3)	...	Σ
(0,0)	$-\lambda_0 - \lambda_H$	λ_0	0	...	λ_H	0	...	0	0	...	0	0	...	0
(1,0)	ν	$-\lambda_0 - \lambda_H -$ $-\nu - \varepsilon$	λ_0	...	ε	λ_H	...	0	0	...	0	0	...	0
(2,0)	0	2ν	$-\lambda_0 - \lambda_H -$ $-2\nu - 2\varepsilon$...	0	2ε	...	0	0	...	0	0	...	0
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots
(1,1)	ν	μ	0	...	$-\lambda_0 - \lambda_H -$ $-\nu - \mu$	λ_0	...	λ_H	0	...	0	0	...	0
(2,1)	0	ν	μ	...	ν	$-\lambda_0 - \lambda_H -$ $-\varepsilon - 2\nu - \mu$...	ε	λ_H	...	0	0	...	0
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots
(2,2)	0	0	0	...	2ν	2μ	...	$-\lambda_0 - \lambda_H -$ $-2\nu - 2\mu$	λ_0	...	λ_H	0	...	0
(3,2)	0	0	0	...	0	2ν	...	ν	$-\lambda_0 - \lambda_H -$ $-3\nu - 2\mu$...	0	λ_H	...	0
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots
(3,3)	0	0	0	...	0	0	...	3ν	3μ	...	$-\lambda_0 - \lambda_H -$ $-3\nu - 3\mu$	$\lambda_0 + \lambda_H$...	0
(4,3)	0	0	0	...	0	0	...	0	3ν	...	ν	$-\lambda_0 - \lambda_H -$ $-4\nu - 3\mu$...	0
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots

Суммируя СУГБ (2.1) по j и учитывая условие нормировки, получаем следующую СУЧБ для $P(i, \square)$:

$$\begin{cases} (\lambda_0 + \lambda_H)P(i, \square) = (i+1)\nu P(i+1, \square), i = 0, 1, \dots; \\ \sum_{i=0}^{\infty} P(i, \square) = 1. \end{cases} \quad (2.2)$$

Отметим, что (2.2) для распределения вероятностей одномерного процесса можно получить и непосредственно. Действительно, суммарный поток заявок на регистрацию при любом i является пуассоновским с интенсивностью $(\lambda_0 + \lambda_H)$, а интенсивность обслуживания одного занятого канала регистрации в состоянии $(i+1)$ равна $(i+1)\nu$ и не зависит от того, занимал ли при этом убывающий из соты 1 абонент одновременно и разговорный канал (см. рис. 2.6, $i = 0, 1, \dots$).

Вводя $\rho := \frac{\lambda_0 + \lambda_H}{\nu}$ – интенсивность предложенной и принятой в соте

1 нагрузки для регистрации прихода пассивных и активных абонентов, получаем из (2.2) для $P(i, \square)$ пуассоновское распределение с параметром $EX = \rho$:

$$P(i, \square) = \frac{\rho^i}{i!} \left(\sum_{k=0}^{\infty} \frac{\rho^k}{k!} \right)^{-1}, i = 0, 1, \dots \quad (2.3)$$

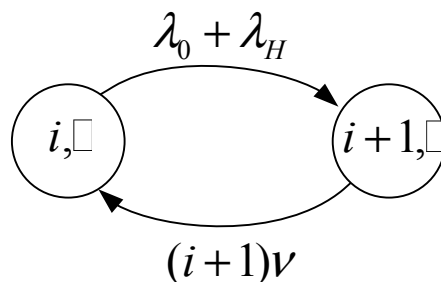


Рис. 2.6. Диаграмма интенсивностей переходов для процесса $X(t)$ при частичном балансе

Суммируя теперь СУГБ (2.1) по i и используя условие нормировки, получаем следующую СУР:

$$\left\{ \begin{array}{l} \lambda_H P(\square, j) + \varepsilon \sum_{i=j}^{\infty} (i-j) P(i, j) = (j+1)(\mu + \nu) P(\square, j+1), j = \overline{0, C_0 - 1}; \\ \lambda_H P(\square, j) = (j+1)(\mu + \nu) P(\square, j+1), j = \overline{C_0, C - 1}; \\ \sum_{j=0}^C P(\square, j) = 1. \end{array} \right. \quad (2.4)$$

Эта СУР также имеет очевидную физическую интерпретацию и может быть получена непосредственно, но она не является СУЧБ рекуррентного типа для $P(\square, j)$, $j = \overline{0, C}$, поскольку при $j = \overline{0, C_0 - 1}$ содержит также $P(i, j)$, $i \geq j$. Для дальнейшего упрощения рассмотрим условную вероятность $P(i|j)$, $i \geq j$, $j = \overline{0, C_0 - 1}$ того, что в соте 1 находятся $i \geq j$ зарегистрированных абонентов при условии, что j из них являются активными.

Из физических соображений построим приближение и будем считать, что, аналогично (2.2) для $P(i, \square)$, и для условных вероятностей $P(i|j)$ в диапазоне $0 \leq j \leq C_0 - 1$ также имеют место простые СУЧБ рекуррентного типа:

$$\left\{ \begin{array}{l} \lambda_0 P(i|j) \approx (i+1-j)\nu P(i+1|j), i \geq j; \\ \sum_{i=j}^{\infty} P(i|j) = 1. \end{array} \right. \quad (2.5)$$

Обозначая $\rho_0 := \frac{\lambda_0}{\nu}$ – интенсивность предложенной и принятой в соте

1 нагрузки для регистрации прихода только пассивных абонентов, получаем из (2.5):

$$P(i|j) \approx \frac{\rho_0^{i-j}}{(i-j)!} e^{-\rho_0}, i \geq j, 0 \leq j \leq C_0 - 1. \quad (2.6)$$

Отсюда среднее значение числа зарегистрированных абонентов в соте 1 при условии, что j из них активны, составляет

$$E(X | Y = j) \approx \sum_{i=j}^{\infty} iP(i | j) = j + \rho_0, \quad 0 \leq j \leq C_0 - 1.$$

Поэтому условное распределение (2.6) можно считать пуассоновским с параметром $EX = \rho_0$ и смещением j , что также имеет очевидную физическую интерпретацию, т.к. блокировок при регистрации нет.

Поскольку совместное распределение числа зарегистрированных и активных абонентов в соте 1 имеет вид

$$P(i, j) = P(\square, j)P(i | j), \quad i \geq j, \quad j = \overline{0, C}, \quad (2.7)$$

то, подставляя (2.6) и (2.7) в СУР (2.4) и упрощая, получим:

$$\begin{cases} (\lambda_H + \varepsilon\rho_0)P(\square, j) \approx (j+1)(\mu + \nu)P(\square, j+1), \quad j = \overline{0, C_0 - 1}; \\ \lambda_H P(\square, j) = (j+1)(\mu + \nu)P(\square, j+1), \quad j = \overline{C_0, C - 1}; \\ \sum_{j=0}^C P(\square, j) = 1. \end{cases} \quad (2.4a)$$

Вводя теперь интенсивности нагрузки

$$\rho_1 := \frac{\lambda_H + \varepsilon\rho_0}{\mu + \nu}, \quad \rho_2 := \frac{\lambda_H}{\mu + \nu}, \quad (2.8)$$

получим из (2.4a) равновесное распределение числа активных абонентов в соте 1:

$$\begin{cases} P(\square, j)G \approx \frac{\rho_1^j}{j!}, \quad j = \overline{0, C_0}; \\ P(\square, j)G \approx \frac{\rho_1^{C_0} \rho_2^{j-C_0}}{j!}, \quad j = \overline{C_0 + 1, C}; \end{cases}$$

$$G := \sum_{j=0}^{C_0} \frac{\rho_1^j}{j!} + \rho_1^{C_0} \sum_{j=C_0+1}^C \frac{\rho_2^{j-C_0}}{j!}. \quad (2.9)$$

Поэтому среднее число EY активных абонентов в соте 1, а также вероятность блокировок новых $\pi_0(C, g)$ и хэндовер-вызовов $\pi_H(C, g)$ могут быть вычислены соответственно по следующим формулам:

$$EY \approx \left[\rho_1 \sum_{j=0}^{C_0-1} \frac{\rho_1^j}{j!} + \rho_1^{C_0} \rho_2 \sum_{j=C_0}^{C-1} \frac{\rho_2^{j-C_0}}{j!} \right] G^{-1};$$

$$\pi_0(C, g) = \sum_{j=C_0}^C P(\square, j) \approx \rho_1^{C_0} \sum_{j=C_0}^C \left[\frac{\rho_2^{j-C_0}}{j!} G^{-1} \right]; \quad (2.10)$$

$$\pi_H(C, g) = P(\square, C) \approx \frac{\rho_1^{C_0} \rho_2^g}{C!} G^{-1}. \quad (2.11)$$

2.2.4. Эквивалентная физическая модель

Отметим, что по существу анализ исходной модели свелся к двум этапам. На первом было установлено, что распределение $P(i, \square)$, $i = 0, 1, \dots$

(2.3) является пуассоновским с параметром $\rho = \frac{\lambda_0 + \lambda_H}{\nu}$ и не зависит от

интенсивности μ успешного завершения разговора. На втором было

получено *квазиэрланговское* распределение $P(\square, j)$, $j = \overline{0, C}$ (2.9), в

котором интенсивность освобождения одного разговорного канала равна $(\mu + \nu)$, а интенсивность поступления пуассоновского потока заявок на

разговор равна $(\lambda_H + \varepsilon\rho_0)$ при $j = \overline{0, C_0}$ и λ_H при $j = \overline{C_0 + 1, C}$. Здесь $\varepsilon\rho_0$ –

средняя интенсивность предложенной разговорной нагрузки, создаваемой средним числом ρ_0 зарегистрированных в сети 1 пассивных абонентов,

инициирующих новые вызовы, что также физически очевидно.

После занятия C_0 разговорных каналов нагрузка интенсивности

$\frac{\varepsilon\rho_0}{\mu + \nu}$ блокируется с вероятностью $\pi_0(C, g)$ (2.10), а интенсивность

предложенной нагрузки на остальные g каналов уменьшается с ρ_1 до ρ_2

(2.8). Таким образом, после проведенного нами анализа достаточно

сложная СМО с двумерным пространством состояний I была аппроксимирована *квазиэрланговской* СМО с одномерным распределением

$P(\square, j), j = \overline{0, C}$ (2.9) числа занятых каналов, изображенной на рис. 2.7.

Такая модель рассмотрена, например, в [11, гл. 2].

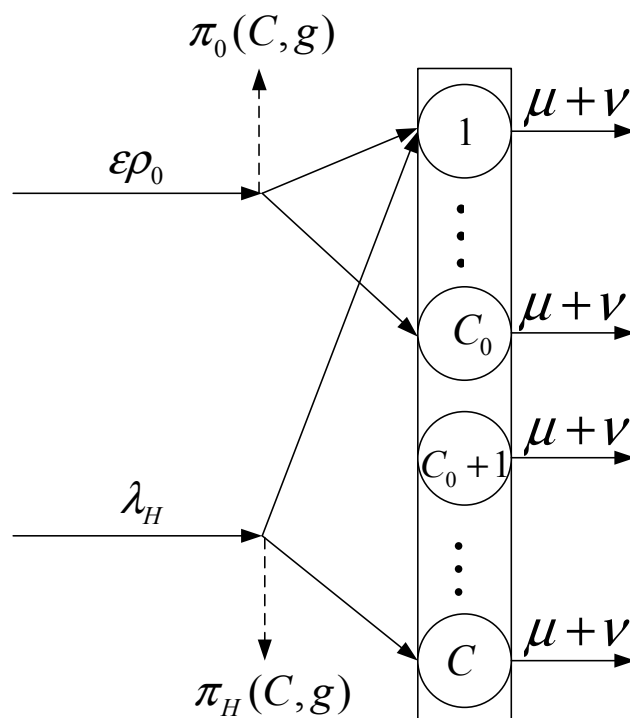


Рис. 2.7. Эквивалентная физическая модель в виде квазиэрланговской СМО

Напомним, что в классической модели Эрланга всего два параметра – число каналов и интенсивность предложенной нагрузки, а в рассматриваемой СМО после модификации структурные параметры не изменяются, а параметрами нагрузки в итоговых результатах являются ρ_1 и ρ_2 , причем ρ_0 и ε включены в ρ_1 . Поэтому численный анализ – особенно при решении проблем адаптивного управления резервными каналами с учетом скачков интенсивностей λ_0 и λ_H – остается достаточно громоздким и требует использования рекуррентных методов.

2.2.5. Рекуррентный алгоритм

При $g = 0$ для блокировок $\pi_0(C, 0)$ и $\pi_H(C, 0)$ верно, что

$$\pi_0(C, 0) = \pi_H(C, 0) = \frac{\rho_1^C}{C!} \left(\sum_{j=0}^C \frac{\rho_1^j}{j!} \right)^{-1}. \quad \text{Таким образом, можно записать}$$

$$\pi_0(C, 0) = \pi_H(C, 0) = E_C(\rho_1),$$

$$\text{где } E_C(\rho_1) = \frac{\rho_1^C}{C!} \left(\sum_{j=0}^C \frac{\rho_1^j}{j!} \right)^{-1} = \frac{\frac{\rho_1}{C} E_{C-1}(\rho_1)}{1 + \frac{\rho_1}{C} E_{C-1}(\rho_1)} \quad - \text{ B-формула Эрланга и ее}$$

запись, удобная для рекурсии по C . Воспользовавшись этой рекуррентной формулой для вычисления $E_k(\rho_1)$, получаем:

$$\pi_0(k, 0) = \frac{\frac{\rho_1}{k} \pi_0(k-1, 0)}{1 + \frac{\rho_1}{k} \pi_0(k-1, 0)}, \quad k = \overline{1, C}; \quad (2.10a)$$

$$\pi_H(k, 0) = \frac{\frac{\rho_1}{k} \pi_H(k-1, 0)}{1 + \frac{\rho_1}{k} \pi_H(k-1, 0)}, \quad k = \overline{1, C}; \quad (2.11a)$$

$$\pi_0(0, 0) = \pi_H(0, 0) = 1.$$

Непосредственное вычисление вероятностей блокировок по формулам (2.10) и (2.11) при достаточно больших значениях C может быть затруднено. Для вычисления $\pi_0(C, g)$ и $\pi_H(C, g)$ можно получить следующие рекуррентные формулы:

$$\pi_0(C_0 + l, l) = \frac{\frac{C}{\rho_2} \pi_0(C_0 + l - 1, l - 1) + \pi_H(C_0 + l - 1, l - 1)}{\pi_H(C_0 + l - 1, l - 1) + \frac{C}{\rho_2}}; \quad (2.10б)$$

$$\pi_H(C_0 + l, l) = \frac{\pi_H(C_0 + l - 1, l - 1)}{\pi_H(C_0 + l - 1, l - 1) + \frac{C}{\rho_2}}, \quad l = \overline{1, g}; \quad (2.11б)$$

где $\pi_0(C_0, 0) = \pi_H(C_0, 0) = E_{C_0}(\rho_1)$. Таким образом, формулы (2.11б) и (2.10б) позволяют эффективно проводить вычисления вероятностей блокировок.

2.2.6. Оптимизация числа каналов

Очевидно, что при увеличении числа резервных каналов и уменьшении числа общедоступных каналов вероятность блокировки $\pi_H(C, g)$ падает, а вероятность блокировки $\pi_0(C, g)$ – растет, и наоборот. Также понятно, что не всякое общее число каналов обеспечивает блокировки обоих типов ниже некоторого критического значения. Рассмотрим задачу определения оптимального числа резервных каналов, а также задачу определения оптимального общего числа каналов.

Задача нахождения оптимального числа резервных каналов состоит в следующем: при заданных интенсивностях $\lambda_0, \lambda_H, \nu, \varepsilon$ и μ , а также общем числе каналов C определить минимальное число g_{\min} резервных каналов, при котором

$$\begin{cases} g \rightarrow \min, \\ \pi_H(C, g) \leq \pi_H^0. \end{cases}$$

Решение рассматриваемой задачи оптимизации возможно путем перебора значений для g из диапазона $\{0, 1, \dots, C-1\}$ для нахождения такого g_{\min} , что $g_{\min} = \min\{g \mid \pi_H(C, g) \leq \pi_H^0\}$.

Задача определения оптимального общего числа каналов формулируется следующим образом: при заданных интенсивностях $\lambda_0, \lambda_H, \nu, \varepsilon$ и μ определить оптимальные общее число каналов C и число резервных каналов g , таких, что

$$\begin{cases} C \rightarrow \min, \\ \pi_0(C, g) \leq \pi_0^0, \\ \pi_H(C, g) \leq \pi_H^0. \end{cases}$$

Решение данной задачи зависит от значений π_0^0 и π_H^0 . Если $\pi_0^0 \leq \pi_H^0$, то искомое $C_{\min} = \min\{C | E_C(\rho_1) \leq \pi_0^0\}$. В этом случае оптимальное число резервных каналов $g_{\min} = 0$. Алгоритм решения задачи для случая $\pi_0^0 > \pi_H^0$ представлен ниже.

Шаг 1. Определить $C' = \min\{C | E_C(\rho_1) \leq \pi_H^0\}$ и $C'' = \min\{C | E_C(\rho_1) \leq \pi_0^0\}$.

Очевидно, что $C' \geq C''$. Перейти к шагу 2.

Шаг 2. $\bar{C} = \frac{C' + C''}{2}$, $g = 0$, $C = \bar{C}$. Перейти к шагу 3.

Шаг 3. $g := g + 1$, $C := C + 1$. Рекуррентно вычислить $\pi_0(C, g)$ и $\pi_H(C, g)$, используя (2.10а,б) и (2.11а,б) соответственно. Перейти к шагу 4.

Шаг 4. Если $\pi_0(C, g) \leq \pi_0^0$ и $\pi_H(C, g) \geq \pi_H^0$, то вернуться к шагу 3. Если $\pi_0(C, g) \geq \pi_0^0$ и $\pi_H(C, g) \geq \pi_H^0$, то $C'' := \bar{C}$, вернуться к шагу 2. Если $\pi_0(C, g) \geq \pi_0^0$ и $\pi_H(C, g) \leq \pi_H^0$, то $C'' := \bar{C}$, перейти к шагу 5. Если $\pi_0(C, g) \leq \pi_0^0$ и $\pi_H(C, g) \leq \pi_H^0$, то $C' := \bar{C}$, перейти к шагу 2.

Шаг 5. Найти $C^* = \min\{C | C \in [C'' + g, C' + g], \pi_0(C, g) \leq \pi_0^0\}$ и

$C^{**} = \min\{C | C \in [C'' + g - 1, C' + g - 1], \pi_H(C, g) \leq \pi_H^0\}$. Перейти к шагу 6.

Шаг 6. Если $C^* = C^{**}$, то $C_{\min} := C^*$, $g_{\min} := g$ или $g_{\min} := g - 1$. Если $C^* < C^{**}$, то $C_{\min} := C^*$, $g_{\min} := g$. Если $C^* > C^{**}$, то $C_{\min} := C^{**}$, $g_{\min} := g - 1$. ■

Оптимизационное управление числом полnodоступных и резервных каналов приобретает наибольшую важность в периоды ЧНН и при увеличении (особенно скачкообразном) нагрузки. В этих случаях задача оператора связи состоит в том, чтобы предотвратить ситуации перегрузок в сети или оперативно с ними справиться при их возникновении.

Предложенный в этом разделе метод анализа достаточно сложной модели микросоты с двумя типами каналов и учетом мобильности абонентов и хэндовер-вызовов (2 структурных и 5 параметров нагрузки) как более простой квазиэрланговской (рис. 2.7) с теми же структурными и только двумя параметрами нагрузки, естественным образом зависящими от исходных, позволяет решить эти важные практические задачи.

Глава 3. УПРАВЛЕНИЕ ДОСТУПОМ ДЛЯ МУЛЬТИСЕРВИСНЫХ СМО

В теории СМО первоначально предполагалось, что вновь поступившая заявка принимается на обслуживание, когда в системе имеется достаточно ресурсов для этого. В системах связи без мест для ожидания основными ресурсами являются каналы (приборы), а в системах с местами для ожидания – как каналы, так и емкость выделенной памяти (места для ожидания). Стратегия, при которой вновь прибывшая заявка принимается, когда объем требуемого ей ресурса меньше или равен объему свободного в этот момент ресурса, называется *полнодоступной*.

Эта стратегия проста в реализации, но обладает рядом недостатков.

- 1) Полнодоступная стратегия является несправедливой, т.к. некоторые классы заявок могут монополизировать каналный ресурс системы.
- 2) Полнодоступная стратегия может привести к слабому использованию ресурса.
- 3) Если доход (плата) за использование ресурса зависит от типа заявок, то полнодоступная стратегия может привести к уменьшению среднего дохода системы.

Поэтому естественным является построение таких стратегий доступа, которые могут ограничивать доступ в систему (СМО), даже если она обладает достаточным объемом свободного ресурса (числом свободных каналов). Для этого необходимо, во-первых, построить алгоритмы, которые эффективно оценивают производительность системы при заданной стратегии доступа, и, во-вторых, построить алгоритмы эффективного определения оптимальной стратегии доступа.

§3.1. Стратегии доступа

3.1.1. Основные определения

Рассмотрим мультисервисную СМО с блокировками поступающих заявок при отсутствии требуемого числа свободных каналов

$$\mathbf{M} \left| \begin{array}{c} \mathbf{M} \\ \lambda(\mathbf{n}), \mathbf{b} \end{array} \right| C \left| \begin{array}{c} 0 \\ \mu(\mathbf{n}) \end{array} \right|. \quad (1.1)$$

Ее функционирование описывает ступенчатый МП $\mathbf{X}(t)$, $t \geq 0$, в пространстве состояний $S = \{\mathbf{n} : 0 \leq \mathbf{b}^T \mathbf{n} \leq C\}$ [1, гл. 3].

Введем теперь стратегию доступа, определяемую посредством функции управления доступом $\mathbf{f} = (f_1, \dots, f_K)$, где

$$f_k(\mathbf{n}) = \begin{cases} 1, & \text{если в состоянии } \mathbf{n} \text{ } k\text{-заявка принимается в СМО, } k = \overline{1, K}; \\ 0, & \text{в противном случае;} \end{cases} \quad (1.2)$$

$$\mathbf{n} \in S(\mathbf{f}).$$

При сделанных предположениях СтМП $\mathbf{X}(t)$, $t \geq 0$, останется марковским при любой стратегии доступа \mathbf{f} типа (1.2), а $S(\mathbf{f})$ будет его множеством возвратных состояний, причем $\mathbf{0} \in S(\mathbf{f}) \subseteq S$.

Нетрудно показать, что ступенчатый МП $\mathbf{X}(t)$, $t \geq 0$, со стратегией доступа \mathbf{f} является стационарным, эргодическим и имеет финальные вероятности, которые удовлетворяют СУР

$$\mathbf{p}^T \mathbf{A} = \mathbf{0}^T, \quad \mathbf{p} > \mathbf{0}, \quad \dim \mathbf{p} = |S(\mathbf{f})| \quad (1.3)$$

и условию нормировки. Здесь \mathbf{A} – матрица интенсивностей переходов рассматриваемого процесса [1, прил. А]. Поэтому в приложениях вместо СП $\mathbf{X}(t)$, $t \geq 0$, при $t \rightarrow \infty$ можно говорить о случайном векторе \mathbf{X} с распределением

$$P\{\mathbf{X} = \mathbf{n}\} := p_{\mathbf{f}}(\mathbf{n}), \quad \sum_{\mathbf{n} \in S(\mathbf{f})} p_{\mathbf{f}}(\mathbf{n}) = 1. \quad (1.4)$$

В приложениях (1.4) часто называют равновесным распределением вероятностей для рассматриваемой СМО, что очевидно физически и упрощает изложение.

3.1.2. Важнейший частный случай

Опишем теперь с помощью функций доступа исторически первую, физически очевидную и с начала XX века известную *полнодоступную стратегию*.

Полнодоступная стратегия (Complete Sharing, CS) для СМО (1.1) задается с помощью K функций доступа

$$f_k(\mathbf{n}) = \begin{cases} 1, & \mathbf{b}^T \mathbf{n} \leq C - b_k, \quad k = \overline{1, K}; \\ 0, & \text{в противном случае.} \end{cases} \quad (1.5)$$

Здесь $\mathbf{n} \in S(\mathbf{f})$, и очевидно, что $S(\mathbf{f}) = S$. В главе 1 этого пособия (и более подробно – в [1, гл. 1–3]) мы рассматривали только эту стратегию, не фиксируя внимание на ее названии, а также вводили $U(\mathbf{n}) = \mathbf{b}^T \mathbf{n}$ – число каналов, используемых в состоянии \mathbf{n} , то есть по существу использовали (1.5). Вместе с тем в главе 2 этого пособия (или в [1, гл. 5 и 6]) мы наряду с полнодоступной рассматривали и несколько неполнодоступных стратегий, каждая из которых включает полнодоступную как частный случай.

Рассмотрим теперь одну простую в реализации и часто используемую в приложениях стратегию.

§3.2. Стратегия резервирования каналов

Стратегия резервирования каналов (Trunk Reservation Policy, TRP) для СМО (1.1) задается с помощью K дополнительных параметров резервирования $\mathbf{t} = (t_1, \dots, t_K)$ и K функций доступа

$$f_k(\mathbf{n}) = \begin{cases} 1, & \mathbf{b}^T \mathbf{n} \leq C - b_k - t_k, \quad k = \overline{1, K}; \\ 0, & \text{в противном случае,} \end{cases} \quad (2.1)$$

$\mathbf{n} \in S(\mathbf{f})$.

Обозначая через π_k вероятность блокировки при поступлении k -заявок, мы видим, что с ростом как b_k , так и t_k , при фиксированных значениях остальных параметров π_k также растет. Это означает, что t_k служит резервом не для k -заявок, а для заявок остальных типов.

Пример 3.1. Пусть $C=6$, $K=2$, $b_1=b_2=1$, $t_1=2$, $t_2=0$. Тогда

$$f_1(n_1, n_2) = \begin{cases} 1, & n_1 + n_2 \leq 3, \\ 0, & \text{в противном случае;} \end{cases}$$

$$f_2(n_1, n_2) = \begin{cases} 1, & n_1 + n_2 \leq 5, \\ 0, & \text{в противном случае.} \end{cases}$$

При таких функциях доступа пространство S , пространство $S(\mathbf{f})$ состояний СМО, а также множества $S_1(\mathbf{f})$ и $S_2(\mathbf{f})$ приема и множества $\bar{S}_1(\mathbf{f})$ и $\bar{S}_2(\mathbf{f})$ блокировки 1- и 2-заявок имеют вид:

$$S = \{(n_1, n_2) : 0 \leq n_1 + n_2 \leq 6\};$$

$$S(\mathbf{f}) = \{(n_1, n_2) \in S : 0 \leq n_1 + n_2 \leq 6, n_1 \leq 4\} = S \setminus \{(5, 0), (5, 1), (6, 0)\},$$

$$|S(\mathbf{f})| = 25;$$

$$S_1(\mathbf{f}) = \{(n_1, n_2) \in S(\mathbf{f}) : n_1 + n_2 \leq 3\}, \quad |S_1(\mathbf{f})| = 10;$$

$$\bar{S}_1(\mathbf{f}) = S(\mathbf{f}) \setminus S_1(\mathbf{f}) = \{(n_1, n_2) \in S(\mathbf{f}) : 3 < n_1 + n_2 \leq 6\}, \quad |\bar{S}_1(\mathbf{f})| = 15;$$

$$S_2(\mathbf{f}) = \{(n_1, n_2) \in S(\mathbf{f}) : n_1 + n_2 \leq 5\}; \quad |S_2(\mathbf{f})| = 20;$$

$$\bar{S}_2(\mathbf{f}) = S(\mathbf{f}) \setminus S_2(\mathbf{f}) = \{(n_1, n_2) \in S(\mathbf{f}) : n_1 + n_2 = 6\}, \quad |\bar{S}_2(\mathbf{f})| = 5.$$

На рис. 3.1 приводится пространство $S(\mathbf{f})$ и отмечены множества $\bar{S}_1(\mathbf{f})$ и $\bar{S}_2(\mathbf{f})$. Полые точки обозначают пространство $\bar{S}_1(\mathbf{f})$, а полые точки, обведенные серым цветом, – пространство $\bar{S}_2(\mathbf{f})$.

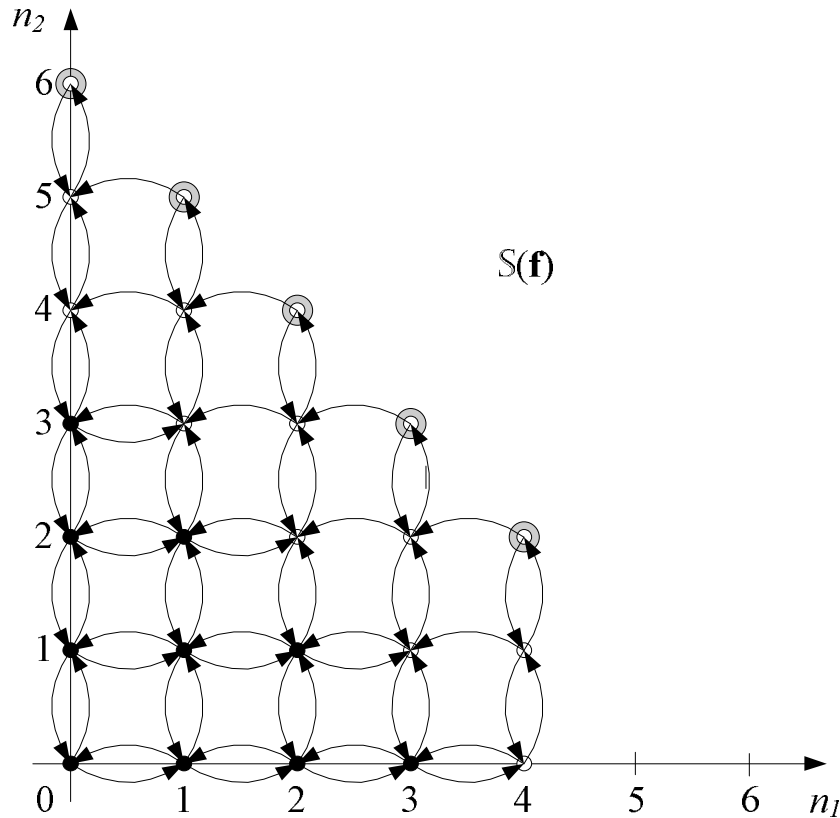


Рис. 3.1. Диаграмма интенсивностей переходов
для $C=6$, $K=2$, $b_1=b_2=1$, $t_1=2$, $t_2=0$

Мы видим, что не все переходы являются парными. Отсюда следует, что равновесное распределение не является мультипликативным. ■

Пример 3.2. Пусть $C=6$, $K=2$, $b_1=1$, $b_2=2$, $t_1=1$, $t_2=0$.

Тогда

$$f_1(n_1, n_2) = \begin{cases} 1, & n_1 + 2n_2 \leq 4, \\ 0, & \text{в противном случае;} \end{cases}$$

$$f_2(n_1, n_2) = \begin{cases} 1, & n_1 + 2n_2 \leq 4, \\ 0, & \text{в противном случае.} \end{cases}$$

В этом примере пространства S , $S(\mathbf{f})$, множества $S_1(\mathbf{f})$ и $S_2(\mathbf{f})$, $\bar{S}_1(\mathbf{f})$ и $\bar{S}_2(\mathbf{f})$ имеют вид:

$$S = \{(n_1, n_2) : 0 \leq n_1 + 2n_2 \leq 6\}, \quad |S| = 16;$$

$$S(\mathbf{f}) = \{(n_1, n_2) \in S : 0 \leq n_1 + 2n_2 \leq 6, n_1 \leq 5\} = S \setminus \{(6, 0)\}, \quad |S(\mathbf{f})| = 15;$$

$$S_1(\mathbf{f}) = \{(n_1, n_2) \in S(\mathbf{f}) : 0 \leq n_1 + 2n_2 \leq 4\} = S_2(\mathbf{f}), \quad |S_1(\mathbf{f})| = |S_2(\mathbf{f})| = 9;$$

$$\overline{S}_1(\mathbf{f}) = S(\mathbf{f}) \setminus S_1(\mathbf{f}) = \{(n_1, n_2) \in S(\mathbf{f}) : 4 < n_1 + 2n_2 \leq 6\} =$$

$$= \{(0, 3), (1, 2), (2, 2), (3, 1), (4, 1), (5, 0)\} = \overline{S}_2(\mathbf{f}), \quad |\overline{S}_1(\mathbf{f})| = |\overline{S}_2(\mathbf{f})| = 6.$$

На рис. 3.2 приводится пространство $S(\mathbf{f})$ и отмечены множества $\overline{S}_1(\mathbf{f})$ и $\overline{S}_2(\mathbf{f})$. Полые точки обозначают пространство $\overline{S}_1(\mathbf{f})$, а полые точки, обведенные серым цветом, – пространство $\overline{S}_2(\mathbf{f})$.

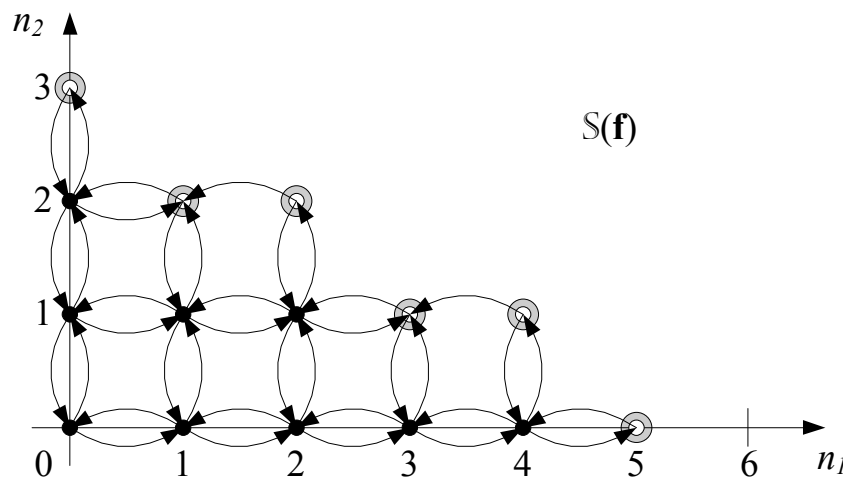


Рис. 3.2. Диаграмма интенсивностей переходов
для $C=6$, $K=2$, $b_1=1$, $b_2=2$, $t_1=1$, $t_2=0$

В этом примере $b_1 + t_1 = b_2 + t_2 = 2$ и $f_1(n_1, n_2) = f_2(n_1, n_2)$. Это должно привести к выравниванию вероятностей блокировок, но, как и в предыдущем примере, не все переходы являются парными. ■

§3.3. Координатно-выпуклые стратегии

3.3.1. Система уравнений глобального баланса

Рассмотрим теперь СУГБ для СМО (1.1) с дополнительным упрощающим условием

$$\lambda_k(\mathbf{n}) = \lambda_k(n_k), \quad \mu_k(\mathbf{n}) = \mu_k(n_k), \quad k = \overline{1, K}, \quad (3.1)$$

которой удовлетворяют равновесные вероятности $p_{\mathbf{f}}(\mathbf{n})$, $\mathbf{n} \in S(\mathbf{f})$ (1.4) при стратегии доступа \mathbf{f} . Кратко эту СМО можно обозначить как

$$\lambda_k(n_k), k = \overline{1, K}; \mathbf{b} \left| \begin{array}{c} \mathbf{M} \\ \mu_k(n_k), k = \overline{1, K} \end{array} \right| \mathbf{C} \left| \begin{array}{c} 0 \\ \mathbf{f} \end{array} \right|. \quad (3.2)$$

Выпишем систему уравнений глобального баланса с учетом (3.1):

$$\begin{aligned} & \sum_{k=1}^K p_{\mathbf{f}}(\mathbf{n}) [\lambda_k(n_k) f_k(\mathbf{n}) + \mu_k(n_k)] = \\ & = \sum_{k=1}^K p_{\mathbf{f}}(\mathbf{n} - \mathbf{e}_k) u(n_k) \lambda_k(n_k - 1) f_k(\mathbf{n} - \mathbf{e}_k) + \\ & + \sum_{k=1}^K p_{\mathbf{f}}(\mathbf{n} + \mathbf{e}_k) 1(\mathbf{n} + \mathbf{e}_k \in S(\mathbf{f})) \mu_k(n_k + 1), \quad \mathbf{n} \in S(\mathbf{f}). \end{aligned} \quad (3.3)$$

Эта система уравнений равновесия представляет собой СЛАУ и может быть решена одним из известных способов, например с помощью итераций Гаусса-Зейделя. Матрица коэффициентов системы (3.3) является сильно разреженной и обычно имеет четкую блочную структуру, например, часто является блочной трехдиагональной, т.е. квазикакобиевой. Поэтому в частных случаях и при небольших C и особенно малых K систему (3.3) можно решить численно, но в общем случае это нецелесообразно.

3.3.2. Система уравнений частичного баланса

Предположим, что имеет место не только глобальный баланс, но и частичный баланс по каждому из K входящих потоков, и выпишем две соответствующие СУЧБ:

$$p_{\mathbf{f}}(\mathbf{n}) \mu_k(n_k) = p_{\mathbf{f}}(\mathbf{n} - \mathbf{e}_k) \lambda_k(n_k - 1), \quad \mathbf{n} \in S(\mathbf{f}), \quad n_k \geq 1, \quad k = \overline{1, K}; \quad (3.4a)$$

$$p_{\mathbf{f}}(\mathbf{n}) \lambda_k(n_k) = p_{\mathbf{f}}(\mathbf{n} + \mathbf{e}_k) \mu_k(n_k + 1), \quad \mathbf{n} + \mathbf{e}_k \in S(\mathbf{f}), \quad k = \overline{1, K}. \quad (3.4b)$$

Это означает, что для графа интенсивностей переходов между \mathbf{n} и всеми соседними «сверху» и «снизу» состояниями $\mathbf{n} \pm \mathbf{e}_k$, $\mathbf{n} \in S(\mathbf{f})$, $k = \overline{1, K}$, (см. рис. 3.3) мы выделили подграф переходов в \mathbf{n} или из \mathbf{n} только за счет

k -заявок и предположили, что за Δt баланс существует как для переходов между \mathbf{n} и $\mathbf{n} - \mathbf{e}_k$, так и для \mathbf{n} и $\mathbf{n} + \mathbf{e}_k$ для каждого k в отдельности (см. рис. 3.4).

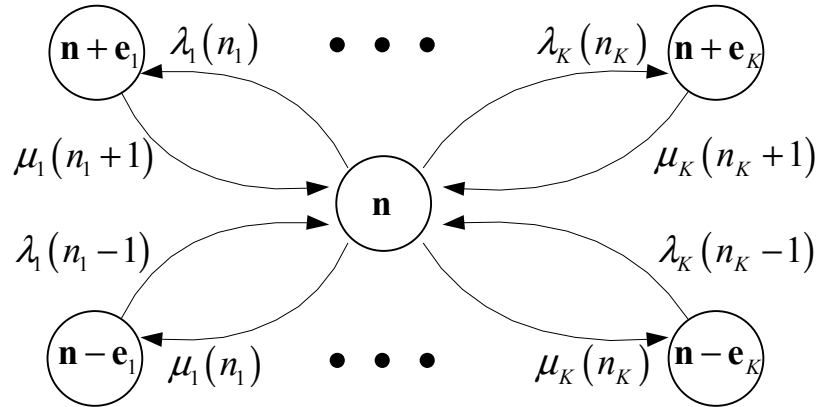


Рис. 3.3. Диаграмма интенсивностей переходов между \mathbf{n} и всеми соседними «сверху» и «снизу» состояниями $\mathbf{n} \pm \mathbf{e}_k$, $k = \overline{1, K}$, $\mathbf{n} \in S(\mathbf{f})$

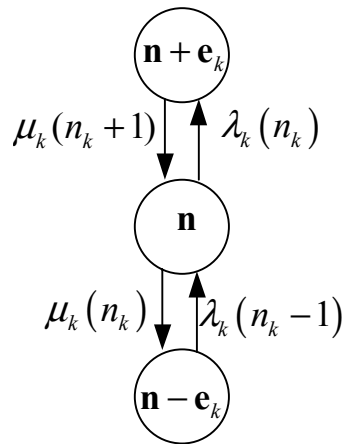


Рис. 3.4. Диаграмма интенсивностей переходов между \mathbf{n} и всеми соседними «сверху» и «снизу» состояниями $\mathbf{n} \pm \mathbf{e}_k$,

подграф для k -потока, $k = \overline{1, K}$, $\mathbf{n} \in S(\mathbf{f})$

Обозначая $\rho_k(n_k - 1) := \frac{\lambda_k(n_k - 1)}{\mu_k(n_k)}$ и используя при $n_k \geq 1$

последовательно рекурсивное соотношение (3.4а), получим:

$$\begin{aligned}
p_{\mathbf{f}}(\mathbf{n}) &= p_{\mathbf{f}}(\mathbf{n} - \mathbf{e}_k) \rho_k(n_k - 1) = \dots = \\
&= p_{\mathbf{f}}(\mathbf{n} - n_k \mathbf{e}_k) \rho_k(n_k - 1) \rho_k(n_k - 2) \dots \rho_k(0) = \\
&= \dots = p_{\mathbf{f}}(\mathbf{0}) \prod_{k=1}^K \rho_k(n_k - 1) \rho_k(n_k - 2) \dots \rho_k(0) = \\
&= p_{\mathbf{f}}(\mathbf{0}) \prod_{k=1}^K \prod_{m=0}^{n_k-1} \rho_k(m), \quad n_k \geq 1, \quad \mathbf{n} \in S(\mathbf{f}).
\end{aligned} \tag{3.5}$$

Из нормирующего условия следует, что

$$\frac{1}{p_{\mathbf{f}}(\mathbf{0})} = \sum_{\mathbf{n} \in S(\mathbf{f})} \prod_{k=1}^K \prod_{m=0}^{n_k-1} \rho_k(m). \tag{3.6}$$

Заметим теперь, что при замене $\mathbf{n} + \mathbf{e}_k$ на \mathbf{n} рекуррентное соотношение (3.4б) сводится к (3.4а), так что (3.5) будет общим решением как (3.4а), так и (3.4б) и, следовательно, будет также решением СУГБ (3.3).

3.3.3. Координатно-выпуклые множества

Координатно-выпуклые стратегии являются важным подмножеством всех стратегий, поскольку их равновесное распределение вероятностей является мультипликативным. Остановимся на этом немного подробнее и определим координатно-выпуклые множества.

Пусть $S(\mathbf{f})$ – пространство состояний для стратегии \mathbf{f} и $\exists \Omega \neq \emptyset$, $\Omega \subseteq S(\mathbf{f})$. Подмножество Ω называется *координатно-выпуклым*, если для него выполняется условие

$$\mathbf{n} \in \Omega, \quad n_k \geq 0 \Rightarrow \mathbf{n} - \mathbf{e}_k \in \Omega, \quad k = \overline{1, K}.$$

Координатно-выпуклая стратегия \mathbf{f} с ассоциированным координатно-выпуклым множеством Ω принимает входящую заявку, если соответствующий СтМП $\mathbf{X}_{\mathbf{f}}(t)$ после этого остается в Ω . Можно сказать, что стратегия \mathbf{f} называется координатно-выпуклой, если существует координатно-выпуклое множество $\Omega \subseteq S(\mathbf{f})$ такое, что

$$\forall \mathbf{n} \in S(\mathbf{f}), \quad f_k(\mathbf{n}) = 1 \Leftrightarrow \mathbf{n} + \mathbf{e}_k \in \Omega, \quad \forall k \in K.$$

Отметим, что рассмотренная нами ранее стратегия управления доступом в модели фрагмента ССПС ([1], §5.3) является координатно-выпуклой и, следовательно, ее равновесное распределение вероятностей является мультипликативным. Вместе с тем далеко не все стратегии являются координатно-выпуклыми. Например, не является координатно-выпуклой стратегия резервирования каналов, рассмотренная в §3.2. Это легко обнаружить на рис. 3.1 и рис. 3.2, где не все переходы являются парными.

3.3.4. Основные типы координатно-выпуклых стратегий

- 1) Наиболее простым и исторически первым примером таких стратегий является *полнодоступная* (Complete Sharing, CS). Для нее ассоциированным выпуклым множеством является само S .
- 2) *Стратегия полного разделения* (Complete Partitioning, CP) пучка из C приборов имеет место в том случае, если существует K целых положительных чисел C_1, \dots, C_K , таких, что $\sum_{k=1}^K C_k \leq C$ и

$$f_k(\mathbf{n}) = f_k(n_k) = \begin{cases} 1, & b_k n_k \leq C_k - b_k, \quad k = \overline{1, K}; \\ 0, & \text{в противном случае.} \end{cases}$$

Это означает, что пучок разделен на K независимых подпучков, причем k -заявка направляется только в полнодоступный подпучок из C_k приборов, $k = \overline{1, K}$.

Стратегия полного разделения является координатно-выпуклой, а ее ассоциированным координатно-выпуклым множеством является «прямоугольник»:

$$\Omega = \{0, \dots, D_1\} \times \dots \times \{0, \dots, D_K\}, \text{ где } D_k := \left\lfloor \frac{C_k}{b_k} \right\rfloor, \quad k = \overline{1, K}.$$

Поскольку одномерные случайные процессы размножения и гибели (ПРГ), описывающие обслуживание k -заявок на соответствующем подпучке из C_k приборов, независимы в совокупности, равновесное распределение $p_f(\mathbf{n})$ составного МП $\mathbf{X}_f(t)$, $t \geq 0$, имеет мультипликативный вид

$$p_f(\mathbf{n}) = p_1(n_1) \dots p_K(n_K),$$

где $p_k(n_k)$ – равновесное распределение для одномерного ПРГ с интенсивностями рождения $\lambda_k(n_k) \cdot 1(n_k < D_k)$ и гибели $\mu_k(n_k)$. Поэтому

$$p_k(n_k) = \frac{\prod_{m=0}^{n_k-1} \rho_k(m)}{\sum_{n_k=0}^{D_k} \prod_{m=0}^{n_k-1} \rho_k(m)}, \quad n_k = \overline{0, D_k}, \quad \rho_k(m) := \frac{\lambda_k(m)}{\mu_k(m+1)}.$$

3) *Стратегия разделения* (Partitioning Policy, PP) группы из K входящих потоков на $L < K$ непересекающихся подгрупп и одновременно пучка из C приборов на соответствующее количество L подгрупп означает, что разбиение $\zeta(K) = \{K_1, \dots, K_L\}$ имеет свойства $\prod_{l=1}^L K_l = K$, $C_1 + \dots + C_L \leq C$.

Здесь классы заявок из подгруппы K_l имеют доступ только к C_l приборам (единицам ресурса) и делят их полнодоступно. Если

$$k \in K_l, \text{ то } f_k(\mathbf{n}) = 1 \Leftrightarrow \sum_{m \in K_l} b_m n_m + b_k \leq C_l, \quad l = \overline{1, L}.$$

Упражнение 3.1. Доказать, что

$$\text{а) } L = 1 \Leftrightarrow C_1 = C \Rightarrow 3) \Rightarrow 1),$$

т.е. при $L = 1$ стратегия 3) совпадает со стратегией 1);

$$\text{б) при } L = K \Leftrightarrow 3) \Leftrightarrow 2),$$

т.е. при $L = K$ стратегия 3) совпадает со стратегией 2).

4) *Пороговая стратегия* (Threshold Policy, TP) с индивидуальными потолками d_1, \dots, d_K допускает новую k -заявку в СМО, если их общее число в СМО удовлетворяет прежнему условию $\mathbf{b}^T \mathbf{n} \leq C - b_k$ и, кроме того, ограничено число n_k уже принятых k -заявок: $n_k \leq d_k - 1, \quad k = \overline{1, K}$. Таким образом, $S_k = \{ \mathbf{n} : (\mathbf{b}^T \mathbf{n} \leq C - b_k) \wedge (0 \leq n_k \leq d_k - 1) \}$.

В отличие от стратегии полного разделения здесь, кроме случая $d_1 + \dots + d_K \leq C$, возможен и случай $d_1 + \dots + d_K > C$.

В качестве примера можно привести схему доступа к ШЦЛ K потоков по индивидуальным подпучкам из C_1, \dots, C_K каналов, где

$d_k := \left\lfloor \frac{C_k}{b_k} \right\rfloor, \quad k = \overline{1, K}$ (см. рис. 3.5 и 3.6). Эту схему можно обозначить

$$\begin{array}{c|c|c|c} \mathbf{M} & \mathbf{M} & C & 0 \\ \lambda, \mathbf{b} & \boldsymbol{\mu} & C & \end{array}.$$

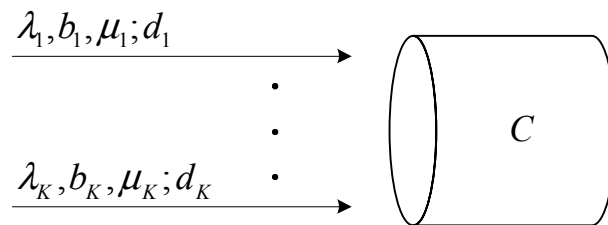


Рис. 3.5. Физическая модель ШЦЛ с индивидуальными потолками d_1, \dots, d_K

Здесь $f_k(\mathbf{n}) = f_k(n_k) = \begin{cases} 1, & (\mathbf{b}^T \mathbf{n} \leq C - b_k) \wedge (0 \leq n_k \leq d_k - 1); \\ 0, & \text{в противном случае;} \end{cases} = \begin{cases} 1, & \mathbf{n} \in S_k; \\ 0, & \mathbf{n} \in \bar{S}_k. \end{cases}$

Упражнение 3.2. Выпишите СУЧБ для состояний \mathbf{n} и $\mathbf{n} - \mathbf{e}_k$, а также для состояний \mathbf{n} и $\mathbf{n} + \mathbf{e}_k$, и убедитесь, что их общее решение является мультипликативным:

$$p(\mathbf{n}) = p(\mathbf{0}) \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!}, \quad \rho_k := \frac{\lambda_k}{\mu_k}.$$

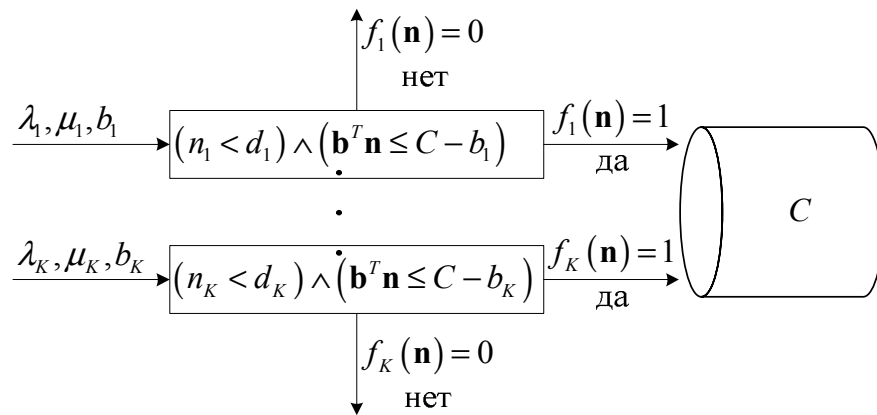


Рис. 3.6. Математическая модель ШЦЛ с индивидуальными потолками

В заключение этого раздела отметим, что множество пороговых стратегий включает как полностью доступную, так и стратегию полного разделения. Рис. 3.7 иллюстрирует взаимоотношение между множествами рассмотренных нами четырех типов координатно-выпуклых стратегий.



Рис. 3.7. Множество координатно-выпуклых стратегий

Упражнение 3.3. Приведите простые примеры к рис. 3.7.

§3.4. Об оптимизации стратегии доступа

3.4.1. Постановка задачи

Проблема построения оптимальной или хотя бы субоптимальной стратегии доступа является практически весьма важной, но теоретически

она достаточно сложна. Поэтому ограничимся только постановкой задачи и некоторыми примерами.

Рассмотрим СтМП $\mathbf{X}_f(t)$, $t \geq 0$, описывающий функционирование некоторой СМО с функцией доступа \mathbf{f} и пространством состояний $S(\mathbf{f})$.

Примем, что $w(\mathbf{n}) \left[\frac{\text{ден.ед.}}{\text{ед.вр.}} \right]$ – мгновенная интенсивность увеличения

дохода от функционирования СМО в состоянии $\mathbf{n} \in S(\mathbf{f})$. Тогда средняя интенсивность увеличения дохода при стратегии \mathbf{f} составит

$$W(\mathbf{f}) := \sum_{\mathbf{n} \in S(\mathbf{f})} w(\mathbf{n}) p_{\mathbf{f}}(\mathbf{n}). \quad (4.1)$$

Для упрощения выводов сделаем естественное предположение, что

$$w(\mathbf{n}) = \sum_{k=1}^K w_k(n_k), \quad (4.2)$$

т.е. общая интенсивность $w(\mathbf{n})$ обладает свойством аддитивности по отношению к интенсивностям $w_k(n_k)$ увеличения дохода от n_k заявок каждого из K классов.

3.4.2. Система с общей памятью и выделенными приборами

В качестве первого примера рассмотрим систему общей памяти емкостью C единиц, на которую поступает мультисервисная нагрузка,

$\mathbf{M} \mid \mathbf{M}$
 $\lambda, \mathbf{b} \mid \boldsymbol{\mu}$, но каждый из K классов заявок имеет свой выделенный прибор

(см. рис. 3.8).

Здесь k -заявка теряется, если при поступлении она застаёт свободными менее b_k единиц памяти, а в случае ее принятия занимает b_k единиц памяти на все время, пока k -й прибор не завершит ее обслуживание. Поэтому

$$S(\mathbf{f}) := \{ \mathbf{n} : \mathbf{b}^T \mathbf{n} \leq C \}; f_k(\mathbf{n}) = 1(\mathbf{b}^T \mathbf{n} + b_k \leq C) \quad (4.3)$$

– соответственно пространство всех состояний и функция доступа для k -заявок, $k = \overline{1, K}$.

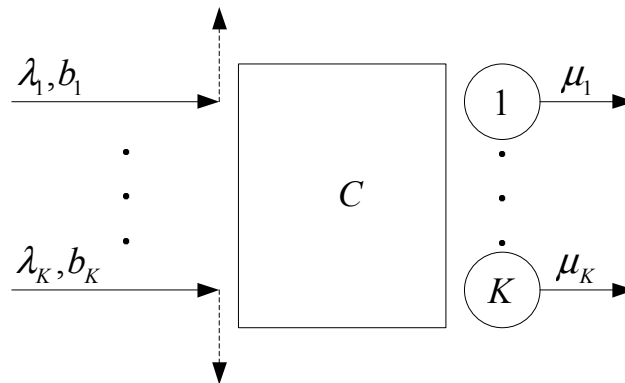


Рис. 3.8. СМО с общей памятью и K выделенными приборами

Эта мультисервисная СМО с очередью и блокировками может служить моделью мультипроцессорной вычислительной системы с разделяемой общей оперативной памятью емкостью C единиц. При $b_1 = \dots = b_K = 1$ ее можно использовать как модель коммутатора с буферизацией пакетов. В этом случае коммутатор имеет K выходных каналов, а класс пакета определен номером канала, на который он адресован. Если вновь поступивший пакет застал в СМО C пакетов, то он получает отказ и теряется. В этом случае все вероятности блокировок совпадают и их легко вычислить.

Обозначим интенсивность увеличения дохода при работе k -прибора через $w_k \left[\frac{\text{ден.ед.}}{\text{ед.вр.}} \right]$, тогда:

$$\begin{aligned}
 W(\mathbf{f}) &= \sum_{\mathbf{n} \in S(\mathbf{f})} p_{\mathbf{f}}(\mathbf{n}) \sum_{k=1}^K w_k u(n_k) = \sum_{k=1}^K w_k \sum_{\mathbf{n} \in S(\mathbf{f})} p_{\mathbf{f}}(\mathbf{n}) u(n_k) = \\
 &= \sum_{k=1}^K w_k \sum_{\mathbf{n} \in S(\mathbf{f})} p_{\mathbf{f}}\{n_k > 0\}.
 \end{aligned} \tag{4.4}$$

Таким образом, предложенный критерий является взвешенной оценкой среднего использования приборов. При $w_k = \mu_k$, $k = \overline{1, K}$, формула

(4.4) дает среднее значение пропускной способности (throughput) рассматриваемой СМО при стратегии \mathbf{f} .

3.4.3. Система без мест для ожидания

Рассмотрим теперь СМО с явными потерями типа $\begin{array}{c|c|c|c} \mathbf{M} & \mathbf{M} & \mathbf{C} & 0 \\ \lambda, \mathbf{b} & \boldsymbol{\mu} & \mathbf{f} & \end{array}$.

Примем, что обслуживание заявки начинается немедленно после ее поступления, т.е.

$$\mu_k(\mathbf{n}) = \mu_k(n_k) = n_k \mu_k, \quad (4.5)$$

и что $w_k \left[\frac{\text{ден.ед.}}{\text{ед.вр.}} \right]$ – интенсивность увеличения дохода при обслуживании

k -заявки, $k = \overline{1, K}$. Тогда $w(\mathbf{n}) = \sum_{k=1}^K w_k n_k$ – линейная функция

интенсивности доходов в состоянии \mathbf{n} , а

$$W(\mathbf{f}) = \sum_{\mathbf{n} \in S(\mathbf{f})} p_{\mathbf{f}}(\mathbf{n}) \sum_{k=1}^K w_k n_k \quad (4.6)$$

– средняя интенсивность дохода от функционирования системы при стратегии \mathbf{f} .

Отметим, что при $w_k = b_k$ (4.6) является средним значением использования приборов в рассматриваемой СМО, а при $w_k = \mu_k$ – средним значением пропускной способности при стратегии \mathbf{f} .

Рассмотрим теперь на эвристическом уровне два экстремальных случая – «легкий» и «тяжелый» трафик. В первом случае все интенсивности поступления являются достаточно малыми, так что практически всегда есть возможность принять позднее поступившие заявки. Поэтому блокировать вновь поступившую k -заявку в интересах будущей l -заявки даже при $w_l > w_k$, $l \neq k$, не имеет смысла. Поэтому для «легкого» трафика оптимальной или, по крайней мере, субоптимальной является полнодоступная стратегия.

Во втором случае все интенсивности поступающей нагрузки $\rho'_k := \rho_k b_k$ являются достаточно большими. Можно показать, что в этом случае оптимальной является стратегия полного разделения пучка из C приборов на подпучки емкостью $C_k = b_k s_k^*$, $k = \overline{1, K}$. Здесь (s_1^*, \dots, s_K^*) – оптимальное решение следующей детерминированной задачи об упаковке

[16]: найти $\max \sum_{k=1}^K w_k s_k$, при условии $\sum_{k=1}^K b_k s_k \leq C$, $s_k \in \{0, 1, \dots\}$, $k = \overline{1, K}$.

Введем теперь $\frac{w_k}{b_k} \left[\frac{\text{ден.ед.}}{\text{ед.вр.} \cdot \text{прибор}} \right]$ – интенсивность дохода от

одного прибора, занятого k -заявкой, и k^* – номер класса, для которого $\frac{w_k}{b_k}$

является максимальной при $k = \overline{1, K}$. Полагая $s_k = \begin{cases} \frac{w_k}{b_k}, & k = k^*, \\ 0, & k \neq k^*, \end{cases}$ получим

оптимальное решение для супертяжелого трафика: весь ресурс выделить только классу k^* -заявок.

Глава 4. МОДЕЛЬ ЗВЕНА СЕТИ С ОДНОАДРЕСНЫМИ И МНОГОАДРЕСНЫМИ СОЕДИНЕНИЯМИ

При анализе сети, в которой передача трафика осуществляется посредством как одноадресных, так и многоадресных соединений, естественно использовать комбинации методов, рассмотренных в [11, гл. 1 и 3]. Однако необходимо учитывать ряд особенностей, вызванных сложной комбинаторной структурой пространства состояний модели, а также функционирование модели при различных стратегиях доступа (см. главу 3). В этой главе предложены модели для полnodоступного звена и отдельного звена с резервированием ресурсов мультисервисной сети с двумя типами соединений. Построение моделей звена с различными стратегиями доступа, описанными в главе 3 пособия, предваряются построением в §4.1 модели МСС с многоадресной доставкой информации произвольной структуры с двумя типами соединений.

§4.1. Модель мультисервисной сети с одноадресными и многоадресными соединениями

4.1.1. Построение модели

Будем рассматривать мультисервисную сеть многоадресной передачи (*мультивещания*¹) произвольной топологии, состоящую из некоторого числа узлов, соединенных звеньями. Такая модель сети была построена в главе 3 [11]. Пусть L – общее число звеньев сети, а $\mathbb{L} = \{1, 2, \dots, L\}$ – множество всех звеньев, занумерованных произвольным образом. Обозначим C_l емкость l -звена. Напомним, что если за единицу емкости звена принять величину одной передаточной единицы (bandwidth unit), например, базовый цифровой канал 64 кбит/с, то C_l представляет собой пропускную способность соответствующего канала связи.

¹ Более подробно о технологии многоадресной передачи данных (multicast) или мультивещания см. главу 4 [11] или [15].

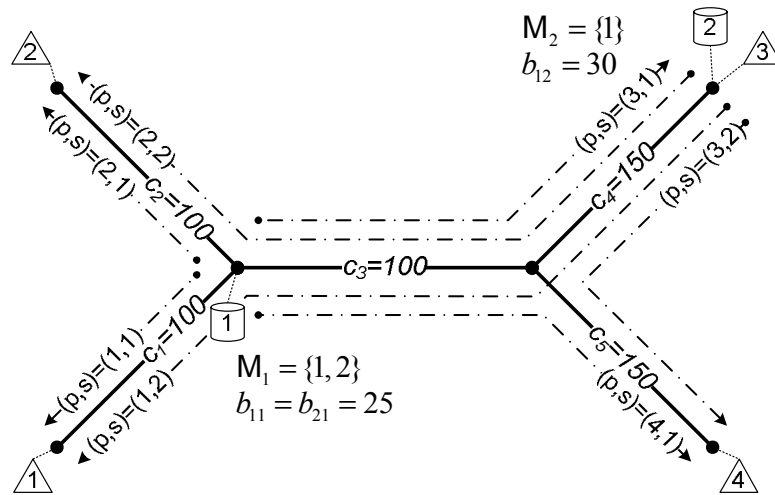
Пусть в сети имеется несколько *источников* многоадресной рассылки, каждый из которых предоставляет *пользователям* конечное число услуг. По запросам пользователей источник передает информацию, соответствующую содержанию услуги, например, транслирует выбранный пользователем телевизионный канал или данные видеоконференции. Если одна и та же услуга предоставляется одновременно нескольким пользователям, то информация на общих участках маршрутов передается без дублирования, то есть посредством единого многоадресного соединения. Таким образом, многоадресные соединения в сети могут быть установлены только между источником, который выступает в качестве отправителя данных, и одним или более пользователями в качестве получателей.

Обозначим $S = \{1, \dots, S\}$ множество всех источников в сети и $M_s = \{1, \dots, M_s\}$ – множество услуг, предоставляемых s -источником. Пусть b_{ms} – число единиц емкости звена, требуемое для предоставления услуги $m \in M_s$. Под *физическим путем* будем понимать набор звеньев сети между узлом подключения источника и узлом подключения пользователя. Обозначим $P_s = \{1, \dots, P_s\}$ множество физических путей от s -источника, $L_{ps} \subseteq L$ – множество всех звеньев p -пути к s -источнику, $P_s^l = \{p \in P_s : l \in L_{ps}\}$ – множество физических путей к s -источнику, проходящих через звено $l \in L$, и $S^l = \{s \in S : P_s^l \neq \emptyset\}$ – множество источников, предоставляющих услуги через l -звено.

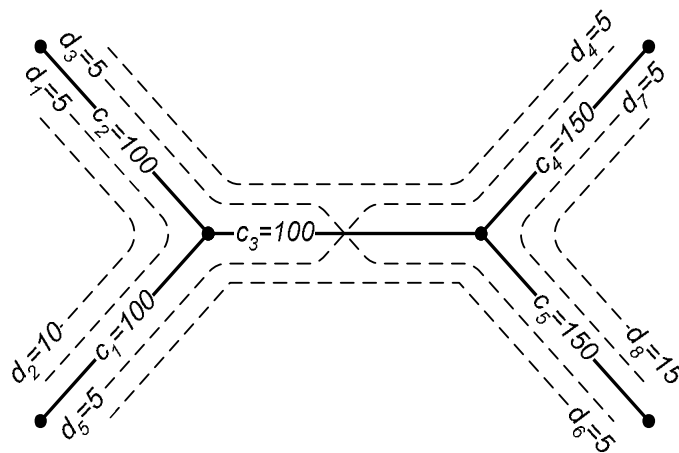
Одноадресные соединения, в отличие от многоадресных, могут быть установлены между двумя произвольными узлами сети. Обозначим $K = \{1, 2, \dots, K\}$ множество всех классов одноадресных соединений сети. Каждый класс соединений характеризуется двумя параметрами: маршрутом, то есть множеством звеньев сети, через которые

устанавливается соединение, и требованием к емкости звеньев маршрута. Пусть $L_k \subseteq L$ – маршрут, а d_k – требование к емкости всех звеньев маршрута соединения k -класса. Введем также множество $K^l = \{k \in K : l \in L_k\}$ классов одноадресных соединений, маршруты которых включают l -звено. Модель мультисервисной сети с одноадресными соединениями рассмотрена в [15], а модели отдельных звеньев сети с одноадресными соединениями – в [11, гл. 1].

На рис. 4.1 представлена схема мультисервисной сети с параметрами многоадресных и одноадресных соединений. Воспользуемся данным примером, для того чтобы пояснить введенные обозначения. Сеть состоит из пяти звеньев, следовательно, $L = \{1, 2, \dots, 5\}$; на каждом звене надписана его емкость. На рис. 4.1а показаны параметры многоадресных соединений сети. Сеть имеет два источника, изображенных на рисунке цилиндрами, то есть $S = \{1, 2\}$, и четыре узла подключения пользователей, отмеченных треугольниками. Около каждого источника на рисунке указано множество предоставляемых им услуг M_s , а также необходимое для предоставления каждой услуги число единиц емкости звеньев. Физические пути (штрихпунктирная линия) к первому источнику информации образуют множество $P_1 = \{1, 2, 3, 4\}$, и множества их звеньев имеют вид $L_{11} = \{1\}$, $L_{21} = \{2\}$, $L_{31} = \{3, 4\}$ и $L_{41} = \{3, 5\}$. Ко второму источнику ведут три физических пути, так как можно считать, что пользователь 3 подсоединен ко второму источнику напрямую, а не через звенья рассматриваемой сети, и $P_2 = \{1, 2, 3\}$, при этом $L_{12} = \{1, 3, 4\}$, $L_{22} = \{2, 3, 4\}$ и $L_{32} = \{4, 5\}$. Наконец, если рассматривать, например, третье звено, то через него проходят два физических пути к первому источнику: $P_1^3 = \{3, 4\}$, и два ко второму источнику: $P_2^3 = \{1, 2\}$. Деревья мультивещания от каждого из источников ко всем пользователям показаны на рис. 4.2.



а) многоадресные соединения



б) одноадресные соединения

Рис. 4.1. Схема мультисервисной сети с двумя типами соединений

На рис. 4.1б пунктирными линиями изображены маршруты классов одноадресных соединений. Рядом с каждым маршрутом указано требуемое для соединения число единиц емкости звеньев. Легко видеть, что всего имеется восемь классов одноадресных соединений, $K = \{1, 2, \dots, 8\}$, с параметрами $L_1 = L_2 = \{1, 2\}$, $L_3 = \{2, 3, 5\}$ и т. д. и $d_2 = 10$, $d_8 = 15$ и $d_k = 5$, $k \in K \setminus \{2, 8\}$. Вновь обратимся к звену 3: множество классов одноадресных соединений, маршруты которых проходят через это звено, имеет вид $K^3 = \{3, 4, 5, 6\}$.

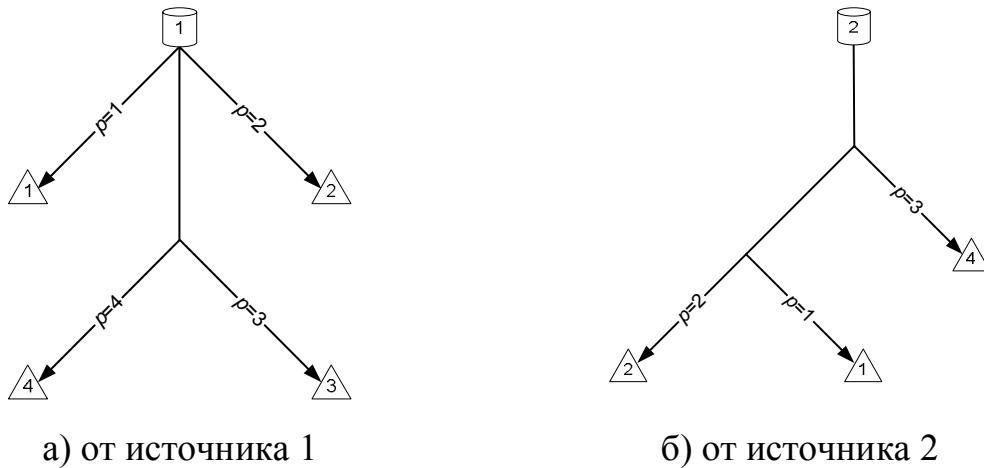


Рис. 4.2. Логические деревья мультивещания для сети на рис. 4.1

Вернемся к построению модели и сделаем предположения о характере запросов на установление соединений обоих типов и о продолжительности этих соединений. Тройку (m, p, s) , $m \in M_s$, $p \in P_s$, $s \in S$, будем называть *логическим путем* или (m, p, s) -путем. Состояние (m, p, s) -пути обозначим $x_{mps} \in \{0,1\}$: $x_{mps} = 1$ (в этом случае говорят, что путь «включен»), если s -источник передает по p -пути данные, соответствующие m -услуге, и $x_{mps} = 0$ (говорят, что путь «выключен») в противном случае. Условием включения логического пути по запросу пользователя является наличие для этого ресурсов на всех звеньях соответствующего маршрута, а именно: на каждом звене $l \in L_{ps}$ (m, s) -услуга либо уже предоставляется другому пользователю (свойство мультивещания), либо имеется b_{ms} свободных единиц емкости. При включении услуги на тех звеньях маршрута, через которые услуга ранее не предоставлялась, под передачу данных выделяется b_{ms} единиц емкости звена, освобождаемых после окончания предоставления услуги по всем проходящим через звено физическим путям. Если на момент поступления запроса хотя бы на одном из звеньев L_{ps} не оказывается свободных ресурсов, происходит блокировка установления соединения, путь остается в состоянии 0 и запрос пользователя теряется.

Для того чтобы пояснить вышесказанное, вновь обратимся к рис. 4.1: если, например, в момент поступления запроса от пользователя 3 на предоставление услуги 1 от источника 1 через звено 3 предоставляются услуги (2, 1) и (1, 2), а также 35 единиц емкости заняты одноадресными соединениями, то оставшихся $10 = 100 - 25 - 30 - 35$ единиц недостаточно для включения услуги и запрос будет потерян. Если же вместо услуги (2, 1) через звено 3 пользователю 4 предоставляется запрашиваемая услуга (1, 1), а на звене 4 найдутся свободные $b_{11} = 25$ единиц емкости, то запрос пользователя 3 будет удовлетворен и логический путь (1, 3, 1) перейдет в состояние 1, причем 25 единиц емкости будут выделены только на звене 4, так как на звене 3 ресурсы для предоставления услуги были выделены ранее.

Запрос пользователя на установление одноадресного соединения k -класса удовлетворяется при условии наличия на каждом звене маршрута свободных d_k единиц емкости. В этом случае указанное число единиц ресурса предоставляется данному соединению до его разъединения, после чего освобождается. При нехватке необходимых ресурсов на момент поступления запроса хотя бы на одном звене маршрута происходит блокировка установления соединения и запрос пользователя теряется. Состояние k -класса одноадресных соединений определяется числом установленных в сети соединений данного класса. Обозначим его $n_k \in \{0, 1, 2, \dots\}$. Например, если в момент поступления запроса на установление соединения класса 4 ($d_4 = 5$, $R_4 = \{1, 2\}$) в сети на рис. 4.1 имеется пять активных соединений класса 6 ($d_6 = 5$, $R_6 = \{1, 3\}$) и одно соединение класса 1 ($d_1 = 1$, $R_1 = \{1, 2\}$), запрос будет заблокирован, так как на звене 1 ($C_1 = 30$) окажутся свободными лишь 4 единицы емкости ($30 - 5 \times 5 - 1 \times 1 = 4$), что меньше требуемых пяти.

Будем предполагать, что логические пути и классы одноадресных соединений в сети функционируют независимо друг от друга. Пусть запросы на использование (m, p, s) -пути образуют пуассоновский поток интенсивности λ_{mps} , а время занятия пути не зависит от процесса поступления запросов и распределено по экспоненциальному закону со средним μ_{mps}^{-1} , $\rho_{mps} = \lambda_{mps} / \mu_{mps}$. Аналогично, пусть запросы на установление соединений k -класса образуют пуассоновский поток интенсивности ν_k , а времена занятия таких соединений не зависят от моментов их установления и распределены по экспоненциальному закону со средним κ_k^{-1} , $a_k = \frac{\nu_k}{\kappa_k}$.

4.1.2. Пространство состояний и равновесное распределение

Легко показать, что в сети со звеньями неограниченной емкости, то есть при $C_l = \infty$, $l \in L$, случайные процессы $\{X_{mps}(t), t \geq 0\}$, $m \in M_s$, $p \in P_s$, $s \in S$, и $\{N_k(t), t \geq 0\}$, $k \in K$, описывающие соответственно поведение (m, p, s) -пути и k -класса одноадресных соединений, являются обратимыми марковскими процессами (ОМП) [9] со стационарными распределениями

$$\pi_{mps}(x_{mps}) = P\{X_{mps}(t) = x_{mps}\} = \frac{\rho_{mps}^{x_{mps}}}{1 + \rho_{mps}}, x_{mps} \in \{0, 1\}, \quad (1.1)$$

и

$$p_k(n_k) = P\{N_k(t) = n_k\} = \frac{a_k^{n_k}}{n_k!} e^{-a_k}, n_k \in \{0, 1, 2, \dots\}. \quad (1.2)$$

Состояние всей сети определяется совокупностью состояний всех логических путей и классов одноадресных соединений. Рассмотрим составной случайный процесс

$$\left\{ \tilde{\mathbf{Z}}(t) = \left((X_{mps}(t))_{m \in M_s, p \in P_s, s \in S}, (N_k(t))_{k \in K} \right), t \geq 0 \right\},$$

описывающий поведение всех соединений сети при условии неограниченных емкостей звеньев. По построению данный процесс является ОМП на множестве

$$\tilde{\mathbf{Z}} = \tilde{\mathbf{X}} \times \tilde{\mathbf{N}}, \quad \tilde{\mathbf{X}} = \{0, 1\}^{\sum_{s \in S} M_s P_s}, \quad \tilde{\mathbf{N}} = \{0, 1, 2, \dots\}^K, \quad (1.3)$$

и, как следует из (1.1) и (1.2), имеет стационарное распределение мультипликативного вида

$$\pi(\mathbf{z}) = \pi(\mathbf{x}, \mathbf{n}) = G^{-1}(\tilde{\mathbf{Z}}) \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}} \prod_{k \in K} \frac{a_k^{n_k}}{n_k!}, \quad \mathbf{z} \in \tilde{\mathbf{Z}}, \quad (1.4)$$

где функция $G(\Omega)$ для любого множества $\Omega \subseteq \tilde{\mathbf{Z}}$ определяется соотношением

$$G(\Omega) = \sum_{\mathbf{z} \in \Omega} \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}} \prod_{k \in K} \frac{a_k^{n_k}}{n_k!}. \quad (1.5)$$

Из (1.5) следует, что нормирующая константа $G(\tilde{\mathbf{Z}})$ распределения вероятностей случайного процесса $\{\tilde{\mathbf{Z}}(t), t \geq 0\}$ имеет вид

$$G(\tilde{\mathbf{Z}}) = \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} (1 + \rho_{mps}) e^{\sum_{k \in K} a_k}.$$

Определим на каждом звене $l \in L$ для источников $s \in \mathbf{S}^l$ и состояний логических путей сети $\mathbf{x} \in \tilde{\mathbf{X}}$ функцию $y_{ms}^l(\mathbf{x})$, соответствующую состоянию (m, s) -услуги на l -звене:

$$y_{ms}^l(\mathbf{x}) := u \left(\sum_{p \in P_s^l} x_{mps} \right), \quad \text{где } u(x) \text{ – функция Хевисайда.}$$

Обозначим $\mathbf{y}^l(\mathbf{x}) := (y_{ms}^l(\mathbf{x}))_{m \in M_s, s \in \mathbf{S}^l}$ состояние услуг мультивещания на l -звене, когда логические пути сети находятся в состоянии $\mathbf{x} \in \tilde{\mathbf{X}}$, и

$\mathbf{z}^l(\mathbf{z}) := (y^l(\mathbf{x}), (n_k)_{k \in K^l})$ – состояние всех соединений на l -звене, когда сеть находится в состоянии $\mathbf{z} \in \tilde{\mathbf{Z}}$ (везде, где это не оговорено особо, вектор $\mathbf{z} \in \tilde{\mathbf{Z}}$ состоит из двух векторных компонент $\mathbf{x} \in \tilde{\mathbf{X}}$ и $\mathbf{n} \in \tilde{\mathbf{N}}$, а именно: $\mathbf{z} = (\mathbf{x}, \mathbf{n})$). Для $l \in \mathbf{L}$ введем величины

$$\begin{aligned} b_l(\mathbf{x}) &:= \sum_{s \in \mathbf{S}^l} \sum_{m \in \mathbf{M}_s} b_{ms} y_{ms}^l(\mathbf{x}), \quad \mathbf{x} \in \tilde{\mathbf{X}}, \\ d_l(\mathbf{n}) &:= \sum_{k \in K^l} d_k n_k, \quad \mathbf{n} \in \tilde{\mathbf{N}}, \end{aligned} \quad (1.6)$$

и

$$c_l(\mathbf{z}) := b_l(\mathbf{x}) + d_l(\mathbf{n}), \quad \mathbf{z} \in \tilde{\mathbf{Z}}, \quad (1.7)$$

представляющие собой число единиц емкости, занятых на l -звене соответственно многоадресными соединениями, одноадресными соединениями и соединениями обоих типов, когда сеть находится в состоянии $\mathbf{z} \in \tilde{\mathbf{Z}}$.

Пусть теперь $C_l < \infty$, $l \in \mathbf{L}$, и возможны блокировки установления многоадресных и одноадресных соединений. В этом случае пространство состояний модели примет вид

$$\mathbf{Z} = \{\mathbf{z} \in \tilde{\mathbf{Z}} : c_l(\mathbf{z}) \leq C_l, l \in \mathbf{L}\}. \quad (1.8)$$

Функционирование сети со звеньями ограниченной емкости описывает случайный процесс $\{\mathbf{Z}(t), t \geq 0\}$, являющийся сужением процесса $\{\tilde{\mathbf{Z}}(t), t \geq 0\}$ на множество \mathbf{Z} , заданное формулой (1.8). Из свойства сужения ОМП вытекает утверждение следующей теоремы.

Теорема 4.1. *Случайный процесс $\{\mathbf{Z}(t), t \geq 0\}$ является ОМП с распределением вероятностей мультипликативного вида*

$$\pi(\mathbf{z}) = G^{-1}(\mathbf{Z}) \prod_{s \in \mathbf{S}} \prod_{p \in \mathbf{P}_s} \prod_{m \in \mathbf{M}_s} \rho_{mps}^{x_{mps}} \prod_{k \in K} \frac{a_k^{n_k}}{n_k!}, \quad \mathbf{z} \in \mathbf{Z}, \quad (1.9)$$

где $G(\mathbf{Z})$ – нормирующая константа:

$$G(\mathbf{Z}) = \sum_{\mathbf{z} \in \mathbf{Z}} \prod_{s \in \mathbf{S}} \prod_{p \in \mathbf{P}_s} \prod_{m \in \mathbf{M}_s} \rho_{mps}^{x_{mps}} \prod_{k \in \mathbf{K}} \frac{a_k^{n_k}}{n_k!}, \quad (1.10)$$

и пространство состояний \mathbf{Z} задано формулой (1.8).

4.1.3. Вероятностные характеристики модели

Зная стационарное распределение вероятностей состояний системы, можно получить выражения для основных ее вероятностных характеристик. Выражения для многих важных характеристик сети, которые задаются вероятностью некоторого события, то есть подмножества Ω пространства состояний системы \mathbf{Z} , могут быть получены посредством функции $G(\Omega)$, определенной формулой (1.5), с помощью известного соотношения

$$P\{\mathbf{z} \in \Omega\} = \sum_{\mathbf{z} \in \Omega} \pi(\mathbf{z}) = \frac{G(\Omega)}{G(\mathbf{Z})}. \quad (1.11)$$

Для того чтобы установление многоадресного соединения оказалось заблокированным, помимо недостаточного числа свободных единиц емкости на каком-либо звене соответствующего физического пути, необходимо, чтобы запрашиваемая услуга не предоставлялась через это звено другим пользователям. Поэтому множество блокировок (m, p, s) -пути имеет вид

$$\mathbf{B}_{mps} = \left\{ \mathbf{z} \in \mathbf{Z} : \exists l \in \mathbf{L}_{ps} : y_{ms}^l(\mathbf{x}) = 0, c_l(\mathbf{z}) + b_{ms} > C_l \right\}. \quad (1.12)$$

Блокировка установления одноадресного соединения происходит в том случае, когда на момент поступления запроса пользователя хотя бы на одном из звеньев маршрута не оказывается необходимого числа свободных единиц емкости. Следовательно, множество блокировок одноадресных соединений k -класса имеет вид

$$\mathbf{B}_k = \left\{ \mathbf{z} \in \mathbf{Z} : \exists l \in \mathbf{L}_k : c_l(\mathbf{z}) + d_k > C_l \right\}. \quad (1.13)$$

Введем обозначения для вероятностей, заданных соотношениями (1.12) и (1.13) событий: $B_{mps} = P\{z \in B_{mps}\}$ и $B_k = P\{z \in B_k\}$. Величины B_{mps} и B_k представляют собой вероятности блокировки соответственно многоадресных и одноадресных соединений «по времени», и для нахождения их значений применима формула (1.11).

Для анализа функционирования многоадресных соединений, помимо вероятности потерь, интерес представляют вероятность того, что услуга предоставляется пользователю, и вероятность того, что услуга не предоставляется, но ресурсов достаточно, чтобы по запросу пользователя инициировать ее предоставление. Введем для любой тройки (m, p, s) , $m \in M_s$, $p \in P_s$, $s \in S$, события

$$F_{mps} = \{z \in Z : x_{mps} = 1\} \quad (1.14)$$

и

$$H_{mps} = \{z \in Z : x_{mps} = 0, \forall l \in L_{ps} \ c_l(z) + b_{ms} \leq C_l \vee y_{ms}^l(x) = 1\}. \quad (1.15)$$

Вычислить соответствующие вероятности $F_{mps} = P\{z \in F_{mps}\}$ и $H_{mps} = P\{z \in H_{mps}\}$ вновь можно по формуле (1.11). Первая из этих величин представляет собой вероятность того, что (m, p, s) -путь включен, вторая – вероятность того, что (m, p, s) -путь выключен, но в сети достаточно ресурсов для его включения.

Легко видеть, что для любого (m, p, s) -пути система множеств B_{mps} , F_{mps} , H_{mps} является разбиением пространства состояний Z , поэтому вероятности этих событий, как и в модели сети мультимедиа, связаны соотношением

$$B_{mps} + F_{mps} + H_{mps} = 1. \quad (1.16)$$

Упражнение 4.1. Докажите, что для любой тройки (m, p, s) , $m \in M_s$, $p \in P_s$, $s \in S$, выполняется соотношение

$$F_{mps} = \rho_{mps} H_{mps}. \quad (1.17)$$

Упражнение 4.2. Докажите, что для любого (m, p, s) -пути, $m \in M_s$, $p \in P_s$, $s \in S$, верно соотношение

$$F_{mps} = \frac{\rho_{mps}}{1 + \rho_{mps}} (1 - B_{mps}). \quad (1.18)$$

При исследовании одноадресных соединений дополнением подпространства блокировки B_k установления соединения k -класса является подпространство приема

$$\bar{B}_k = \{z \in Z : c_l(z) + d_k \leq C_l, l \in L_k\}, \quad (1.19)$$

объединяющее такие состояния сети, в которых запрос на установление соединений k -класса будет принят. Очевидно, что $\bar{B}_k = P\{z \in \bar{B}_k\} = 1 - B_k$.

Вычисление вероятностных характеристик модели непосредственно по формулам (1.9)–(1.11) подразумевает перебор пространства состояний

(1.8), мощность которого $|Z| \leq 2^{\sum_{s \in S} M_s P_s} N^K$, где N – наибольшее число одноадресных соединений одного класса, которые могут быть одновременно установлены в сети: $N = \max_{k \in K} N_k$,

$N_k = \max \{n_k \in \{0, 1, 2, \dots\} : z = (\mathbf{0}, \mathbf{n}) \in Z\}$. В общем случае для сети с одноадресными и многоадресными соединениями значение нормирующей константы $G(Z)$ можно получить путем прямого перебора пространства состояний в соответствии со следующим алгоритмом.

Алгоритм (расчет нормирующей константы методом прямого перебора пространства состояний):

$$G \leftarrow 0$$

для всех $z = (\mathbf{x}, \mathbf{n}) \in \{0, 1\}^{\sum_{s \in S} M_s P_s} \times \prod_{k \in K} \{0, 1, \dots, N_k\}$ выполнять

$$\text{если для всех } l \in L \quad \sum_{s \in S} \sum_{m \in M_s^l} b_{ms} \cdot u \left(\sum_{p \in P_s^l} x_{mps} \right) + \sum_{k \in K^l} d_k n_k \leq C_l$$

$$\text{то } G \leftarrow G + \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}} \prod_{k \in K} \frac{a_k^{n_k}}{n_k!}$$

возвратить G .

Обозначим $M = \max_{s \in S} M_s$ и $P = \max_{s \in S} P_s$. Тогда порядок временной сложности алгоритма оценивается выражением $O(2^{MPS} N^{K+1} MPS \max\{L, NK\})$ и, таким образом, экспоненциально зависит от основных параметров системы: M , P , S и K . По этой причине алгоритм применим лишь для ограниченного диапазона значений структурных параметров сети и для полноценного анализа модели требуется разработка более эффективных вычислительных методов [10].

§4.2. Модель полнодоступного звена сети

4.2.1. Постановка задачи

Предположим, что в модели сети, представленной в §4.1, все звенья, кроме некоторого звена l^* , имеют неограниченные ресурсы для обслуживания запросов пользователей, то есть $C_l = \infty$ для $l \in L \setminus \{l^*\}$. Задача анализа блокировок в такой системе сводится к анализу сети, состоящей из одного звена l^* , с одним источником s^* , который предоставляет услуги из множества $M = \bigcup_{s \in S^{l^*}} M_s$, и множеством классов одноадресных соединений K^{l^*} . Для удобства записи далее в этом параграфе индексы l^* и s^* опускаются.

Функционирование звена МСС с двумя типами соединений будем описывать с помощью многопоточковой мультисервисной системы массового обслуживания, схематично изображенной на рис. 4.3.

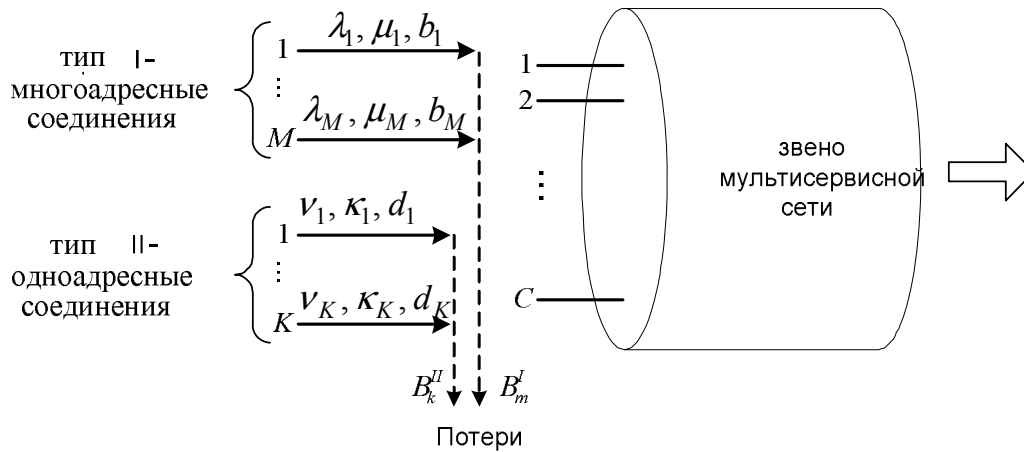


Рис. 4.3. Схема модели звена МСС

На полнодоступную систему, состоящую из $C = C_l^*$ приборов (единиц емкости звена сети) и не имеющую накопителя, поступают $M = |M|$ потоков заявок типа I и $K = |K|$ потоков типа II. Будем считать, что все $M + K$ поступающих в систему потоков являются пуассоновскими и независимы в совокупности. Первая группа потоков (I-потоки) моделирует поступление запросов на установление многоадресных соединений. Если на момент поступления (I, m)-заявки в системе нет ни одной заявки этого потока, то поступившая заявка принимается при условии наличия b_m свободных приборов и занимает их на случайное время, распределенное экспоненциально с параметром μ_m и не зависящее ни от длительности обслуживания заявок других потоков, ни от процессов поступления. Все поступившие в течение этого интервала времени (I, m)-заявки принимаются на обслуживание без выделения дополнительных приборов, а по истечении указанного интервала одновременно покидают систему и b_m приборов освобождаются. Потеря заявки типа I происходит только в том случае, если при ее поступлении в системе нет заявок того же потока, а также нет достаточного количества свободных приборов.

Обозначим $\rho_m := \frac{\lambda_m}{\mu_m}$, где $\lambda_1, \dots, \lambda_M$ – интенсивности входящих I-потоков.

Заметим, что, аналогично модели звена сети мультивещания, параметры ρ_1, \dots, ρ_M связаны с интенсивностями потоков запросов пользователей на включение соответствующих логических путей в сети соотношением

$$\rho_m = \prod_{p \in P^l} (1 + \rho_{mp}) - 1, \quad m = \overline{1, M}. \quad (2.1)$$

Описанная система соответствует полnodоступной стратегии CS (см. раздел 3.1.2).

Потоки второй группы (II-потоки) соответствуют потокам запросов пользователей на установление через звено l^* одноадресных соединений. Поступившая (II, k)-заявка принимается на обслуживание, если на момент ее прихода в системе имеется d_k свободных приборов. Принятая заявка занимает это число приборов на случайное время, распределенное по экспоненциальному закону с параметром κ_k и также не зависящее ни от длительности обслуживания заявок других потоков, ни от процессов поступления, после чего заявка покидает систему, освобождая d_k приборов. Если на момент поступления заявки достаточного количества свободных приборов не оказывается, заявка теряется. Интенсивности ν_1, \dots, ν_K входящих потоков этого типа совпадают с интенсивностями соответствующих потоков запросов пользователей, $a_k = \nu_k / \kappa_k$.

4.2.2. Пространство состояний и равновесное распределение

Положим $C = \infty$, в этом случае все поступившие в систему заявки принимаются на обслуживание и потери отсутствуют. Пусть случайный процесс $\{Y_m(t), t \geq 0\}$, $m = \overline{1, M}$, находится в состоянии 1, если в момент времени $t \geq 0$ в системе обслуживается хотя бы одна (I, m)-заявка, и в состоянии 0 в противном случае. Как было показано ранее, процесс $\{Y_m(t), t \geq 0\}$ является ОМП со стационарным распределением

$$\pi_m(y_m) = P\{Y_m(t) = y_m\} = \frac{\rho_m^{y_m}}{1 + \rho_m}, \quad y_m \in \{0, 1\}. \quad (2.2)$$

Введем также случайный процесс, характеризующий П-потоки. Пусть $N_k(t)$ – число (Π, k) -заявок в системе в момент времени $t \geq 0$, $k = \overline{1, K}$. Процесс $\{N_k(t), t \geq 0\}$ также является ОМП, а его стационарное распределение имеет вид

$$p_k(n_k) = P\{N_k(t) = n_k\} = \frac{a_k^{n_k}}{n_k!} e^{-a_k}, \quad n_k \in \{0, 1, 2, \dots\}. \quad (2.3)$$

Рассмотрим составной случайный процесс

$$\{\tilde{\mathbf{Z}}(t) = (Y_1(t), \dots, Y_M(t), N_1(t), \dots, N_K(t)), t \geq 0\}.$$

По построению $\{\tilde{\mathbf{Z}}(t), t \geq 0\}$ является ОМП на множестве $\tilde{\mathbf{Z}} = \tilde{\mathbf{Y}} \times \tilde{\mathbf{N}} = \{0, 1\}^M \times \{0, 1, 2, \dots\}^K$ и, как следует из формул (2.2) и (2.3), имеет стационарное распределение

$$\tilde{\pi}(\mathbf{z}) = G^{-1}(\tilde{\mathbf{Z}}) \prod_{m=1}^M \rho_m^{y_m} \prod_{k=1}^K \frac{a_k^{n_k}}{n_k!}, \quad \mathbf{z} = (\mathbf{y}, \mathbf{n}) \in \tilde{\mathbf{Z}}, \quad (2.4)$$

где функция $G(\Omega)$, аналогично §4.1, для любого множества $\Omega \subseteq \tilde{\mathbf{Z}}$ определяется соотношением

$$G(\Omega) = \sum_{\mathbf{z} \in \Omega} \prod_{m=1}^M \rho_m^{y_m} \prod_{k=1}^K \frac{a_k^{n_k}}{n_k!}, \quad (2.5)$$

следовательно, нормирующая константа $G(\tilde{\mathbf{Z}})$ распределения процесса $\{\tilde{\mathbf{Z}}(t), t \geq 0\}$ равна

$$G(\tilde{\mathbf{Z}}) = e^{\sum_{k=1}^K a_k} \prod_{m=1}^M (1 + \rho_m).$$

Процесс $\{\tilde{\mathbf{Z}}(t), t \geq 0\}$ с пространством состояний $\tilde{\mathbf{Z}}$ и распределением вероятностей (2.4) описывает состояние рассматриваемой системы для случая $C = \infty$.

Обозначим $c(\mathbf{z})$ число занятых приборов системы в состоянии $\mathbf{z} \in \tilde{Z}$ и заметим, что эта величина представима в виде

$$c(\mathbf{z}) = c(\mathbf{y}, \mathbf{n}) = b(\mathbf{y}) + d(\mathbf{n}) = \sum_{m=1}^M b_m y_m + \sum_{k=1}^K d_k n_k, \quad (2.6)$$

где $b(\mathbf{y})$ и $d(\mathbf{n})$ – число приборов, занятых в состоянии $\mathbf{z} = (\mathbf{y}, \mathbf{n})$ заявками I- и II-потоков соответственно. Пусть теперь $C < \infty$ и, следовательно, возможны потери заявок. Будем считать, что система функционирует с явными потерями (потерянные заявки обоих типов не оказывают влияние на интенсивность породившего их потока). В этом случае функционирование системы описывает случайный процесс $\{\mathbf{Z}(t), t \geq 0\}$, являющийся сужением процесса $\{\tilde{\mathbf{Z}}(t), t \geq 0\}$ на множество

$$Z = \{\mathbf{z} \in \tilde{Z} : c(\mathbf{z}) \leq C\}. \quad (2.7)$$

Как сужение обратимого процесса он также обратим, и, следовательно, справедлива следующая теорема.

Теорема 4.2. *Стационарное распределение вероятностей состояний процесса $\{\mathbf{Z}(t), t \geq 0\}$ имеет мультипликативный вид*

$$\pi(\mathbf{z}) = G^{-1}(Z) \prod_{m=1}^M \rho_m^{y_m} \prod_{k=1}^K \frac{a_k^{n_k}}{n_k!}, \quad \mathbf{z} \in Z, \quad (2.8)$$

где $G(Z)$ – нормирующая константа:

$$G(Z) = \sum_{\mathbf{z} \in Z} \prod_{m=1}^M \rho_m^{y_m} \prod_{k=1}^K \frac{a_k^{n_k}}{n_k!}. \quad (2.9)$$

4.2.3. Вероятностные характеристики

Как и для сети в целом, некоторые макрохарактеристики отдельного звена могут быть выражены с использованием функции $G(\Omega)$ от соответствующего подмножества пространства состояний посредством соотношения

$$P\{z \in \Omega\} = \frac{G(\Omega)}{G(Z)}. \quad (2.10)$$

К таким характеристикам, в частности, относятся вероятности потерь заявок потоков обоих типов, вероятность того, что (I, m) -заявка находится в системе, и вероятность того, что (I, m) -заявок в системе нет, но если такая заявка поступит, то будет принята на обслуживание. Напомним, что условием потери заявки I -потока, помимо недостаточного числа свободных приборов, является отсутствие в системе заявок данного потока. Следовательно, множество потерь (I, m) -заявок имеет вид

$$B_m^I = \{z \in Z : c(z) + b_m > C, y_m = 0\}. \quad (2.11)$$

Заявки II -потоков теряются в том случае, если в системе недостаточно свободных приборов для их обслуживания. Таким образом, множество потерь (II, k) -заявок имеет вид

$$B_k^{II} = \{z \in Z : c(z) + d_k > C\}. \quad (2.12)$$

Значения вероятностей $B_m^I = P\{z \in B_m^I\}$ и $B_k^{II} = P\{z \in B_k^{II}\}$ можно получить по формуле (2.10), они соответствуют вероятностям блокировки на звене многоадресных и одноадресных соединений, а в рассматриваемой системе дают вероятности потерь заявок по времени.

Множество таких состояний, когда (I, m) -заявка находится в системе (соответствующие m -услуге данные передаются через рассматриваемое звено), имеет вид

$$F_m = \{z \in Z : y_m = 1\}. \quad (2.13)$$

Множество таких состояний, когда (I, m) -заявок в системе нет, но если заявка поступит, то будет принята на обслуживание (m -услуга через звено не предоставляется, но ресурсов достаточно, чтобы по запросу пользователя инициировать ее предоставление), принимает вид

$$H_m = \{z \in Z : c(z) + b_m \leq C, y_m = 0\}. \quad (2.14)$$

Как и для модели сети произвольной топологии, для рассматриваемой модели верны соотношения (2.15)–(2.17).

$$B_m^I + F_m + H_m = 1. \quad (2.15)$$

Упражнение 4.3. Докажите, что для любого $m = \overline{1, M}$ выполняется соотношение

$$F_m = \rho_m H_m. \quad (2.16)$$

Упражнение 4.4. Докажите, что для любого $m = \overline{1, M}$ верно соотношение

$$F_m = \frac{\rho_m}{1 + \rho_m} (1 - B_m^I). \quad (2.17)$$

При анализе отдельного звена МСС большой интерес представляют характеристики дискретных СВ β , δ и γ , которые принимают значения $b(\mathbf{z})$, $d(\mathbf{z})$ и $c(\mathbf{z})$ соответственно. Здесь γ является случайной величиной занятых приборов в рассматриваемой системе и соответствует случайному числу занятых единиц емкости звена мультисервисной сети. Если за единицу емкости принять величину одной передаточной единицы, то γ представляет собой СВ ширины полосы пропускания, занятой на звене сети при обслуживании установленных через него соединений обоих типов. Аналогично интерпретируются СВ β и δ для многоадресных и одноадресных соединений соответственно. В частности, среднее значение занятой ширины полосы пропускания, то есть среднее число занятых приборов в рассматриваемой модели, можно найти как математическое ожидание $c^{(1)}$ СВ γ , а именно:

$$c^{(1)} = \sum_{\mathbf{z} \in Z} c(\mathbf{z}) \pi(\mathbf{z}). \quad (2.18)$$

Величина $c^{(1)}/C$ представляет собой коэффициент использования звена.

Обозначим $b^{(i)}$ и $d^{(i)}$ как начальные моменты СВ β и δ соответственно, тогда коэффициент корреляции $R_{\beta\delta}$ между этими СВ будет иметь вид

$$R_{\beta\delta} = \frac{\sum_{z \in Z} b(z)d(z)\pi(z) - b^{(1)}d^{(1)}}{\sqrt{\left(b^{(2)} - (b^{(1)})^2\right)\left(d^{(2)} - (d^{(1)})^2\right)}}. \quad (2.19)$$

Среднее число приборов, занятых I-заявками, можно найти как математическое ожидание функции $b(z)$:

$$b^{(1)} = \sum_{z \in Z} b(z)\pi(z). \quad (2.20)$$

Заметим, что, зная $c^{(1)}$ и $b^{(1)}$, значение величины $d^{(1)}$ легко получить из соотношения $c^{(1)} = b^{(1)} + d^{(1)}$.

Упражнение 4.5. Докажите, что для нахождения среднего числа приборов системы, занятых I-заявками, применима формула

$$b^{(1)} = \sum_{m=1}^M b_m F_m. \quad (2.21)$$

Доказательство.

$$\begin{aligned} \sum_{m=1}^M b_m F_m &= \sum_{m=1}^M b_m \sum_{z \in F_m} \pi(z) = \sum_{m=1}^M b_m \sum_{z \in F_m} y_m \pi(z) = \\ &= \sum_{m=1}^M b_m \left(\sum_{z \in F_m} y_m \pi(z) + \sum_{z \in Z \setminus F_m} y_m \pi(z) \right) = \\ &= \sum_{m=1}^M b_m \sum_{z \in Z} y_m \pi(z) = \sum_{z \in Z} \pi(z) \sum_{m=1}^M b_m y_m = \sum_{z \in Z} b(z) \pi(z) = b^{(1)}. \quad \blacksquare \end{aligned}$$

Упражнение 4.6. Докажите, что среднее число занятых I-заявками приборов в системе выражается формулой

$$b^{(1)} = \sum_{m=1}^M b_m \frac{\rho_m}{1 + \rho_m} (1 - B_m^I). \quad (2.22)$$

Доказательство. Утверждение (2.22) вытекает из соотношений (2.17) и (2.21). ■

4.2.4. Метод свертки для расчета вероятностных характеристик

Введем разбиение пространства состояний Z по числу занятых в системе приборов

$$Z(n) = \{z \in Z : c(z) = n\}, \quad n = \overline{1, C},$$

$$Z = \bigcup_{n=0}^C Z(n),$$

$$Z(i) \cap Z(j) = \emptyset, \quad i \neq j, \quad i, j = \overline{0, C},$$

и распределение числа занятых приборов

$$P(n) = P(Z(n)), \quad n = \overline{0, C}. \quad (2.23)$$

Тогда множество потерь (Π, k) -заявок можно представить в виде

$$B_k^{\Pi} = \bigcup_{n=C-d_k+1}^C Z(n), \quad k = \overline{1, K},$$

а вероятность их потерь определяется формулой

$$B_k^{\Pi} = \sum_{n=C-d_k+1}^C P(n), \quad k = \overline{1, K}.$$

Множество B_m^I потерь (I, m) -заявок имеет более сложную структуру по сравнению с множеством потерь заявок Π -потоков. Введем систему событий

$$Z_m(n) = \{z \in Z : c(z) = n, y_m = 0\}, \quad n = 0, \dots, C, \quad m = \overline{1, M},$$

и соответствующее распределение вероятностей

$$P_m(n) = P(Z_m(n)), \quad n = \overline{0, C}, \quad m = \overline{1, M}. \quad (2.24)$$

Теперь множество потерь заявок первого типа можно представить в виде

$$B_m^I = \bigcup_{n=C-b_m+1}^C Z_m(n), \quad m = \overline{1, M},$$

а вероятности потерь (I, m) -заявок вычисляются по формуле

$$B_m^I = \sum_{n=C-b_m+1}^C P_m(n), \quad m = \overline{1, M}.$$

Для случаев, когда в систему поступают заявки только одного типа, эффективные вычислительные алгоритмы для расчета распределений (2.23) и (2.24) нам уже известны. Так, если $M = 0$, то для вычисления $P(n)$ можно использовать сверточный алгоритм, если же $K = 0$, то $P(n)$ и $P_M(n)$ можно найти с помощью алгоритма, представленного в [11, разд. 3.2.3].

Для вычисления $P(n)$ в общем случае ($M > 0, K > 0$) естественно воспользоваться сочетанием этих двух алгоритмов на основе соотношения

$$Z(n) = \bigcup_{i=0}^n (Y(i) \times C(n-i)), \quad n = \overline{0, C},$$

где

$$Y(n) = \{y \in \tilde{Y} : b(y) = n\},$$

$$C(n) = \{n \in \tilde{N} : d(n) = n\}.$$

Легко показать, что для любых множеств $\Omega \subseteq \tilde{Y}$ и $\Upsilon \subseteq \tilde{N}$

$$G(\Omega \times \Upsilon) = G(\Omega)G(\Upsilon).$$

Тогда в вероятностной форме для $n = \overline{0, C}$ получим

$$P(n) = \frac{G(Z(n))}{G(Z)} = G^{-1}(Z) \sum_{i=0}^n G(Y(i))G(C(n-i)). \quad (2.25)$$

Аналогичное соотношение имеет место для $P_m(n)$:

$$Z_m(n) = \bigcup_{i=0}^n (Y_m(i) \times C(n-i)), \quad n = \overline{0, C},$$

где

$$Y_m(n) = \{y \in \tilde{Y} : b(y) = n, y_m = 0\}, \quad n = \overline{0, C},$$

или в вероятностной форме для $m = \overline{1, M}$ и $n = \overline{0, C}$

$$P_m(n) = \frac{G(Z_m(n))}{G(Z)} = G^{-1}(Z) \sum_{i=0}^n G(Y_m(i)) G(C(n-i)). \quad (2.26)$$

Лемма 4.1. Значение функции $f_m(i, n)$, удовлетворяющей для любого $n = \overline{0, C}$ соотношениям

$$G(Z(n)) = f_0(M, n),$$

$$G(Z_m(n)) = f_m(M, n), \quad m = \overline{1, M},$$

вычисляется по формуле

$$f_m(i, n) = \begin{cases} 0, & i = 0, \dots, M, \quad n < 0; \\ h(n), & i = 0, \quad n = \overline{0, C}; \\ f_m(i-1, n) + (1 - \delta(i, m)) \rho_i f_m(i-1, n - b_i), & \\ & i = 1, \dots, M, \quad n = \overline{0, C}, \end{cases} \quad (2.27)$$

где функция $h(n)$ задана формулой

$$h(n) = \begin{cases} 0, & n < 0; \\ 1, & n = 0; \\ \frac{1}{n} \sum_{k=1}^K d_k a_k h(n - d_k), & n = \overline{1, C}, \end{cases} \quad (2.28)$$

а $\delta(i, m)$ – символ Кронекера.

Доказательство. Функция $f_0(i, n)$ представляет собой свертку функций $g(m, n)$ (см. [11, гл. 3])

$$g(m, n) = \begin{cases} 0, & m = 0, n = \overline{1, C}; \\ 0, & m = \overline{0, M}, n < 0; \\ 1, & m = \overline{0, M}, n = 0; \\ g(m-1, n) + \rho_m g(m-1, n-b_m), & \\ & m = \overline{1, M}, n = \overline{1, C}, \end{cases} \quad (2.29)$$

и $h(n)$, заданной соотношениями (2.28):

$$f_0(i, n) = \sum_{j=0}^n g(i, n-j)h(j). \quad (2.30)$$

Множитель $(1 - \delta(i, m))$ в третьей строке (2.27) позволяет при переборе опускать m -ю компоненту вектора y , то есть учитывать условие $y_m = 0$ при расчете $P_m(n)$. Поскольку для аналогичных систем с одним типом потоков заявок выполняются соотношения

$$G(\mathbf{C}(n)) = h(n),$$

$$G(\mathbf{Y}(n)) = g(M, n),$$

$$G(\mathbf{Y}_M(n)) = g(M-1, n),$$

утверждение леммы следует из соотношений (2.25) и (2.26). ■

Следующие утверждения формулируют метод вычисления вероятностных характеристик звена сети с двумя типами соединений.

Утверждение 4.1. Нормирующая константа (2.9) вычисляется по формуле

$$G(\mathbf{Z}) = \sum_{n=0}^C f_0(M, n). \quad (2.31)$$

Утверждение 4.2. Вероятность потерь (I, m) -заявок вычисляется по формуле

$$B_m^I = \frac{\sum_{n=C-b_m+1}^C f_m(M, n)}{\sum_{n=0}^C f_0(M, n)}, \quad m = \overline{1, M}. \quad (2.32)$$

Вероятность потерь (II, k)-заявок вычисляется по формуле

$$B_k^{II} = \frac{\sum_{n=C-d_k+1}^C f_0(M, n)}{\sum_{n=0}^C f_0(M, n)}, \quad k = \overline{1, K}. \quad (2.33)$$

Вероятность того, что (I, m)-заявок в системе нет, но если такая заявка поступит, то будет принята на обслуживание, вычисляется по формуле

$$H_m = \frac{\sum_{n=0}^{C-b_m} f_m(M, n)}{\sum_{n=0}^C f_0(M, n)}, \quad m = \overline{1, M}. \quad (2.34)$$

Выражение для среднего числа занятых приборов имеет вид

$$c^{(1)} = \frac{\sum_{n=1}^C n \cdot f_0(M, n)}{\sum_{n=0}^C f_0(M, n)}. \quad (2.35)$$

Заметим, что $P(i, j) = P(Y(i) \times C(j))$, $i, j = \overline{0, C}$, $i + j \leq C$, представляет собой совместное распределение СВ β и δ , через которое удобно выразить коэффициент корреляции между ними и начальные моменты:

$$R_{\beta\delta} = \frac{\sum_{i=1}^{C-1} \sum_{j=1}^{C-i} ijP(i, j) - b^{(1)}d^{(1)}}{\sqrt{\left(b^{(2)} - (b^{(1)})^2\right)\left(d^{(2)} - (d^{(1)})^2\right)}},$$

где

$$b^{(i)} = \sum_{n=1}^C \left(n^i \sum_{j=0}^{C-n} P(n, j) \right),$$

$$d^{(i)} = \sum_{n=1}^C \left(n^i \sum_{j=0}^{C-n} P(j, n) \right).$$

Поскольку для $i, j = \overline{0, C}$, $i + j \leq C$ выполняется

$$P(i, j) = \frac{G(Y(i))G(C(j))}{G(Z)} = G^{-1}(Z)g(M, i)h(j), \quad (2.36)$$

верно следующее утверждение.

Утверждение 4.3. Коэффициент корреляции между СВ β и δ вычисляется по формуле

$$R_{\beta\delta} = \frac{\frac{\sum_{i=1}^{C-1} \sum_{j=1}^{C-i} ijg(M, i)h(j)}{\sum_{n=0}^C f_0(M, n)} - b^{(1)}d^{(1)}}{\sqrt{\left(b^{(2)} - (b^{(1)})^2\right)\left(d^{(2)} - (d^{(1)})^2\right)}}, \quad (2.37)$$

где

$$b^{(i)} = \left(\sum_{n=0}^C f_0(M, n) \right)^{-1} \sum_{n=1}^C \left(n^i g(M, n) \sum_{j=0}^{C-n} h(n) \right);$$

$$d^{(i)} = \left(\sum_{n=0}^C f_0(M, n) \right)^{-1} \sum_{n=1}^C \left(n^i h(n) \sum_{j=0}^{C-n} g(M, n) \right). \quad (2.38)$$

Полученные результаты позволяют реализовать эффективные алгоритмы для расчета вероятностных характеристик отдельного полнодоступного звена МСС. Добавим, что при реализации алгоритма расчета функции $f_m(i, n)$ для сокращения времени вычислений целесообразно добавить условие $f_m(i, n) = 1$ при $i = \overline{0, M}$ и $n = 0$.

§4.3. Модель звена с резервированием

В предыдущем параграфе разработана модель звена мультисервисной сети, где доступ к ресурсам звена организован по полнодоступной стратегии (см. раздел 3.1.2). Это означает, что для обслуживания поступившей заявки могут быть выделены любые из имеющихся S приборов при условии, что эти приборы не заняты обслуживанием других заявок. Такое правило действует для соответствующих одноадресным соединениям заявок II-поток и для заявок I-поток, пришедших в систему, где отсутствуют заявки того же потока. Напомним, что если на момент поступления заявки I-потока в системе обслуживается хотя бы одна заявка того же потока, то пришедшая заявка принимается на обслуживание без выделения дополнительных ресурсов. Как уже было отмечено, на практике в телекоммуникационных системах нередко применяется резервирование части ресурсов для некоторых классов трафика, являющихся приоритетными. В этом параграфе исследуются модель и метод анализа вероятных характеристик звена МСС с резервированием.

4.3.1. Звено с неполнодоступной стратегией разделения ресурсов

Применительно к звену МСС суть неполнодоступной стратегии разделения ресурсов состоит в том, что каждому классу (или группе классов) поступающего на звено трафика выделяется некоторая часть емкости звена, на которую не допускается трафик остальных классов (см. §3.2). При этом часть емкости звена отводится для совместного использования всеми классами. В терминах представленной в §4.2 модели неполнодоступную стратегию доступа определим следующим образом. Разобьем множества M и K соответственно на J^I и J^{II}

непересекающихся подмножеств: $M = \bigcup_{j=1}^{J^I} M_j$, $M_i \cap M_j = \emptyset$ при $i \neq j$, и

$K = \bigcup_{j=1}^{J^I} K_j$, $K_i \cap K_j = \emptyset$ при $i \neq j$. Пусть соответствующие подмножеству

M_j I-потoki заявок образуют нагрузочную группу с порогом резервирования r_j^I , $0 \leq r_j^I < C$. Аналогично пусть II-потoki заявок, соответствующие подмножеству K_j , образуют нагрузочную группу с порогом резервирования r_j^{II} , $0 \leq r_j^{II} < C$. Будем считать, что доступное

всем потокам количество приборов $C^0 = C - \sum_{j=1}^{J^I} r_j^I - \sum_{j=1}^{J^{II}} r_j^{II} > 0$. Количество

приборов, доступных заявкам потоков одной нагрузочной группы, складывается из значения порога резервирования данной группы и величины C^0 . Если поступившая заявка не находит ресурсов в доступной ей зоне, она теряется.

4.3.2. Пространство состояний и равновесное распределение

Очевидно, что, если каждая нагрузочная группа имеет доступ к неограниченному числу приборов, то есть при $r_j^I + C^0 = \infty$, $j = 1, \dots, J^I$, и $r_j^{II} + C^0 = \infty$, $j = 1, \dots, J^{II}$, потери заявок в неполнодоступной системе отсутствуют и ее функционирование описывает введенный в разделе 4.2.2 ОМП $\{\tilde{\mathbf{Z}}(t) = (Y_1(t), \dots, Y_M(t), N_1(t), \dots, N_K(t)), t \geq 0\}$ с множеством состояний $\tilde{\mathbf{Z}} = \tilde{\mathbf{Y}} \times \tilde{\mathbf{N}} = \{0, 1\}^M \times \{0, 1, 2, \dots\}^K$ и распределением вероятностей состояний мультипликативного вида (2.4).

Введем обозначения для количества приборов, занятых заявками потоков одной нагрузочной группы: $b_j(\mathbf{y}) = \sum_{m \in M_j} b_m y_m$, $j = 1, \dots, J^I$, для I-потоков и $d_j(\mathbf{n}) = \sum_{k \in K_j} d_k n_k$, $j = 1, \dots, J^{II}$, для II-потоков. Заметим, что

$b(\mathbf{y}) = \sum_{j=1}^{J^I} b_j(\mathbf{y})$ и $d(\mathbf{n}) = \sum_{j=1}^{J^{II}} d_j(\mathbf{n})$. Величина $\max\{r_j^I, b_j(\mathbf{y})\}$ равна числу

приборов, зарезервированных под обслуживание заявок I-потокв нагрузочной группы j и занятых заявками этих потоков в общей зоне,

если таковые имеются. Следовательно, сумма $\sum_{j=1}^{J^I} \max\{r_j^I, b_j(\mathbf{y})\}$ задает

число приборов, недоступных заявкам II-потокв, когда система находится в состоянии $\mathbf{z} = (\mathbf{y}, \mathbf{n})$. Аналогично величина $\max\{r_j^{II}, d_j(\mathbf{n})\}$ равна числу

приборов, зарезервированных под обслуживание заявок II-потокв нагрузочной группы j и занятых этими заявками в общей зоне, и сумма

$\sum_{j=1}^{J^{II}} \max\{r_j^{II}, d_j(\mathbf{n})\}$ равна числу приборов, недоступных заявкам I-потокв,

когда система находится в состоянии $\mathbf{z} = (\mathbf{y}, \mathbf{n})$. Таким образом, количество приборов, которые поступившая (I, m) -заявка застает недоступными по причине резервирования или занятости другими заявками (без учета того факта, что если в системе уже обслуживается заявка потока (I, m) , то поступившая заявка будет принята на обслуживание на те же приборы – эти приборы также считаются занятыми), равно

$$c_m^I(\mathbf{z}) = \sum_{j=1}^{J^I} \max\{r_j^I 1(m \notin M_j), b_j(\mathbf{y})\} + \sum_{j=1}^{J^{II}} \max\{r_j^{II}, d_j(\mathbf{n})\}, \quad \mathbf{z} \in \tilde{Z}, m \in M, \quad (3.1)$$

где $1(m \notin M_j)$ – функция-индикатор, принимающая значение 1, если

$m \notin M_j$, и 0, если, напротив, $m \in M_j$. Величина $C - c_m^I(\mathbf{z})$ равна количеству свободных приборов в доступной (I, m) -заявкам зоне, когда система находится в состоянии \mathbf{z} . Аналогичная величина для заявок II-потокв имеет вид

$$c_k^{\text{II}}(\mathbf{z}) = \sum_{j=1}^{J^I} \max\{r_j^I, b_j(\mathbf{y})\} + \sum_{j=1}^{J^{\text{II}}} \max\{r_j^{\text{II}} 1(k \notin K_j), d_j(\mathbf{n})\}, \quad \mathbf{z} \in \tilde{Z}, k \in K. \quad (3.2)$$

Здесь функция-индикатор $1(k \notin K_j)$ принимает значение 1, если $k \notin K_j$, и 0 в противном случае. Величина $C - c_k^{\text{II}}(\mathbf{z})$ равна количеству свободных приборов в доступной (Π, k) -заявкам зоне, когда система находится в состоянии \mathbf{z} . Введем также величину

$$c^0(\mathbf{z}) = \sum_{j=1}^{J^I} \max\{r_j^I, b_j(\mathbf{y})\} + \sum_{j=1}^{J^{\text{II}}} \max\{r_j^{\text{II}}, d_j(\mathbf{n})\}, \quad \mathbf{z} \in \tilde{Z}. \quad (3.3)$$

Заметим, что величина $C - c^0(\mathbf{z})$ равна количеству свободных приборов в общей зоне, когда система находится в состоянии \mathbf{z} .

Пусть теперь количество приборов ограничено и возможны потери заявок. Помимо ограничения на общее число занятых приборов $c(\mathbf{z}) \leq C$, на состояния недоступной системы накладываются ограничения, обусловленные резервированием, а именно: $b_j(\mathbf{y}) \leq C^0 + r_j^I$, $j = 1, \dots, J^I$, и $d_j(\mathbf{n}) \leq C^0 + r_j^{\text{II}}$, $j = 1, \dots, J^{\text{II}}$. Таким образом, пространство состояний недоступной системы равно

$$Z = \{\mathbf{z} \in \tilde{Z} : b_j(\mathbf{y}) \leq C^0 + r_j^I, j = 1, \dots, J^I, \quad d_j(\mathbf{n}) \leq C^0 + r_j^{\text{II}}, j = 1, \dots, J^{\text{II}}, \quad c(\mathbf{z}) \leq C\},$$

что можно представить с использованием обозначения (3.3) в виде

$$Z = \{\mathbf{z} \in \tilde{Z} : c^0(\mathbf{z}) \leq C\}. \quad (3.4)$$

Описывающий функционирование недоступной системы случайный процесс $\{\mathbf{Z}(t), t \geq 0\}$ представляет собой сужение процесса $\{\tilde{\mathbf{Z}}(t), t \geq 0\}$ на множество Z , заданное соотношением (3.4). Как сужение ОМП этот процесс также является ОМП, и для него справедлива теорема 4.2, если принять, что пространство состояний Z задано формулой (3.4).

4.3.3. Вероятностные характеристики

Найдем выражения для некоторых вероятностных характеристик неполнодоступной системы. (I, m) -заявка принимается на обслуживание в том случае, когда в системе уже обслуживается заявка того же потока или имеется b_m свободных приборов среди доступных заявке $C^0 + r_j^I$ приборов, $m \in M_j$. Таким образом, потеря (I, m) -заявки происходит в таких состояниях системы $\mathbf{z} = (\mathbf{y}, \mathbf{n}) \in Z$, где $y_m = 0$ и выполнено хотя бы одно из следующих условий:

$$b_j(\mathbf{y}) + b_m > C^0 + r_j^I$$

или

$$c(\mathbf{z}) + b_m > C.$$

Объединяя эти условия, мы можем записать множество потерь (I, m) -заявок с использованием выражения (3.1) в виде

$$B_m^I = \{ \mathbf{z} \in Z : c_m^I(\mathbf{z}) + b_m > C, y_m = 0 \}. \quad (3.5)$$

Для заявок II-потоков условием потери является недостаточное число свободных приборов в доступной зоне. Поэтому множество потерь (II, k) -заявок имеет вид

$$B_k^{II} = \{ \mathbf{z} \in Z : c_k^{II}(\mathbf{z}) + d_k > C \}. \quad (3.6)$$

Для заявок I-потоков введем множество таких состояний, что (I, m) -заявка находится в системе:

$$F_m = \{ \mathbf{z} \in Z : y_m = 1 \}; \quad (3.7)$$

и множество таких состояний, что (I, m) -заявок в системе нет, но если заявка поступит, то будет принята на обслуживание:

$$H_m = \{ \mathbf{z} \in Z : c_m^I(\mathbf{z}) + b_m \leq C, y_m = 0 \}. \quad (3.8)$$

Зная вид пространства состояний системы, а также вид множеств состояний, соответствующих искомым характеристикам (3.5)–(3.8), вероятности событий можно получить по формуле (2.10). Выражения для среднего числа занятых приборов (2.18) и коэффициента корреляции между СВ числа приборов, занятых заявками потоков разных типов (2.19), приведенные для полнодоступной системы в разделе 4.2.3, остаются верны для неполнодоступной системы с поправкой на вид пространства состояний. Однако, как уже было отмечено, вычисление вероятностных характеристик звена напрямую по представленным формулам допустимо лишь для небольшого диапазона значений структурных параметров модели – прежде всего для малой емкости звена. Поэтому далее мы вновь уделим особое внимание разработке эффективного метода расчета вероятностных характеристик. Будет рассмотрен актуальный для практики случай, когда резервирование ресурсов осуществляется для некоторого подмножества услуг мультивещания, например набора приоритетных телевизионных каналов или услуг конференц-связи.

4.3.4. Резервирование для подмножества услуг мультивещания

В модели звена с резервированием канальных ресурсов для подмножества услуг мультивещания структурные параметры принимают следующие значения. Имеются две нагрузочные группы I-потоков, то есть $J^I = 2$, причем ненулевой порог резервирования задан только для одной из них: $0 < r_1^I < C$ и $r_2^I = 0$. Вся совокупность II-потоков составляет одну нагрузочную группу с нулевым порогом резервирования: $J^{II} = 1$, $r_1^{II} = 0$. Для краткости записи обозначим $R = r_1^I$. Таким образом, резервирование осуществляется для услуг мультивещания из множества $M_1 \subseteq M$, или в терминах ТМО, имеет место резервирование приборов для заявок первых $M_1 = |M_1|$ I-потоков. Схема такой модели изображена на рис. 4.4.

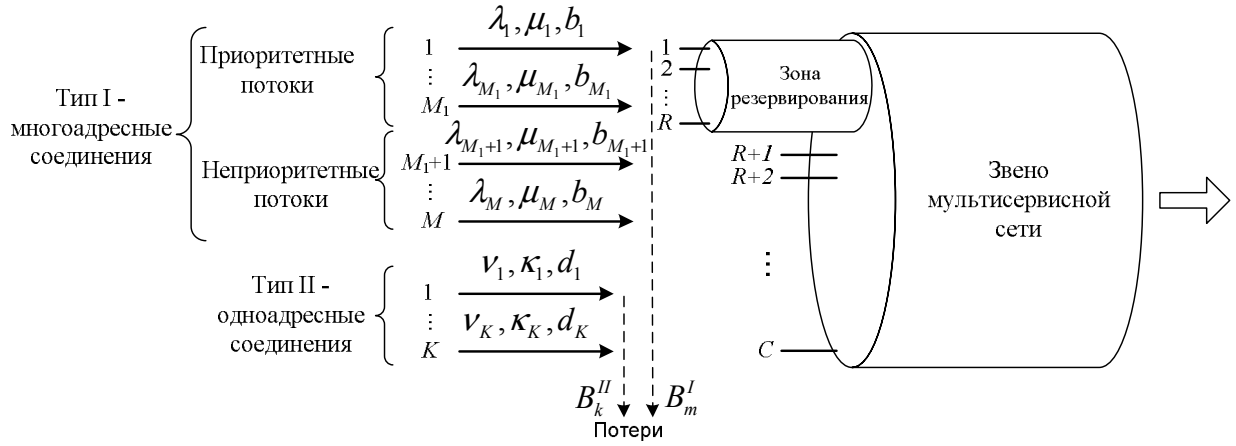


Рис. 4.4. Схема модели звена МСС с резервированием ресурсов для подмножества услуг мультивещания

Поскольку $r_2^I = r_1^{II} = 0$, для любого $\mathbf{z} = (\mathbf{y}, \mathbf{n}) \in \tilde{\mathbf{Z}}$ выполняются соотношения $\max\{r_2^I, b_2(\mathbf{y})\} = b_2(\mathbf{y})$ и $\max\{r_1^{II}, d(\mathbf{n})\} = d(\mathbf{n})$, следовательно, пространство состояний модели принимает вид

$$\mathbf{Z} = \{\mathbf{z} \in \tilde{\mathbf{Z}} : \max\{R, b_1(\mathbf{y})\} + b_2(\mathbf{y}) + d(\mathbf{n}) \leq C\}. \quad (3.9)$$

При $m \in M_1$ для любого состояния сети выполняется равенство $c_m^I(\mathbf{z}) = c(\mathbf{z})$ и множество потерь заявок приоритетного I-потока имеет вид $\mathbf{B}_m^I = \{\mathbf{z} \in \mathbf{Z} : c(\mathbf{z}) + b_m > C, y_m = 0\}$, $m \in M_1$.

При $m \in M_2$ для любого состояния сети $c_m^I(\mathbf{z}) = c^0(\mathbf{z})$, поэтому множество потерь заявок неприоритетного I-потока равно

$$\mathbf{B}_m^I = \{\mathbf{z} \in \mathbf{Z} : c^0(\mathbf{z}) + b_m > C, y_m = 0\}, m \in M_2.$$

Поскольку $c_k^{II}(\mathbf{z}) = c^0(\mathbf{z})$, для любого $k \in \mathbf{K}$, множество потерь (II, k)-заявок имеет вид

$$\mathbf{B}_k^{II} = \{\mathbf{z} \in \mathbf{Z} : c^0(\mathbf{z}) + d_k > C\}, k \in \mathbf{K}.$$

4.3.5. Метод свертки для расчета вероятностных характеристик

Введем разбиение пространства состояний (3.9) по числу занятых приборов:

$$Z(n) = \{\mathbf{z} \in Z : c(\mathbf{z}) = n\},$$

$$Z = \bigcup_{n=0}^C Z(n),$$

$$Z(i) \cap Z(j) = \emptyset, \quad i \neq j, \quad i, j = \overline{0, C},$$

и распределение числа занятых приборов

$$P(n) = P(Z(n)), \quad n = \overline{0, C}.$$

Также нам потребуется разбиение пространства состояний по значению функции $c^0(\mathbf{z})$:

$$Z^0(n) = \{\mathbf{z} \in Z : c^0(\mathbf{z}) = n\},$$

$$Z = \bigcup_{n=R}^C Z^0(n),$$

$$Z^0(i) \cap Z^0(j) = \emptyset, \quad i \neq j, \quad i, j = \overline{R, C},$$

и соответствующее распределение вероятностей

$$P^0(n) = P(Z^0(n)), \quad n = \overline{R, C}.$$

Отметим, что функция $c^0(\mathbf{z})$ принимает значения из множества $\{R, \dots, C\}$, поэтому $Z^0(n) = \emptyset$ и $P^0(n) = 0$ для всех $n = \overline{0, R-1}$.

Через вероятности событий $Z(n)$ и $Z^0(n)$ выражаются основные вероятностные характеристики системы. Среднее число занятых приборов выражается через вероятности событий $Z(n)$ следующим образом:

$$c^{(1)} = \sum_{n=1}^C nP(n). \quad (3.10)$$

Множество потерь (Π, k) -заявок можно представить в виде

$$B_k^{\Pi} = \bigcup_{n=C-d_k+1}^C Z^0(n), \quad k = \overline{1, K},$$

а вероятность их потери определяется формулой

$$B_k^{\Pi} = \sum_{n=C-d_k+1}^C P^0(n), \quad k = \overline{1, K}. \quad (3.11)$$

Множество B_m^I потерь (I, m) -заявок вновь имеет более сложную структуру по сравнению с множеством потерь заявок II-потоков. Введем условные распределения вероятностей

$$P_m(n) = P(Z(n) | y_m = 0), \quad m \in M_1, \quad n = \overline{0, C},$$

$$P_m^0(n) = P(Z^0(n) | y_m = 0), \quad m \in M_2, \quad n = \overline{0, C}.$$

Множества потерь заявок приоритетных I-потоков можно представить в виде

$$B_m^I = \bigcup_{n=C-b_m+1}^C Z(n) \cap \{z \in Z : y_m = 0\}, \quad m \in M_1.$$

Множества потерь заявок неприоритетных I-потоков – в виде

$$B_m^I = \bigcup_{n=C-b_m+1}^C Z^0(n) \cap \{z \in Z : y_m = 0\}, \quad m \in M_2.$$

Выражения для вероятностей потерь (I, m) -заявок имеют вид

$$B_m^I = \sum_{n=C-b_m+1}^C P_m(n), \quad m \in M_1, \quad (3.12)$$

и

$$B_m^I = \sum_{n=C-b_m+1}^C P_m^0(n), \quad m \in M_2. \quad (3.13)$$

Также через $Z(n)$ и $Z^0(n)$ можно записать выражение для множества H_m :

$$H_m = \begin{cases} \bigcup_{n=0}^{C-b_m} Z(n) \cap \{z \in Z : y_m = 0\}, & m \in M_1; \\ \bigcup_{n=0}^{C-b_m} Z^0(n) \cap \{z \in Z : y_m = 0\}, & m \in M_2. \end{cases}$$

Соответствующая вероятность равна

$$H_m = \begin{cases} \sum_{n=0}^{C-b_m} P_m(n), & m = \overline{1, M_1}; \\ \sum_{n=0}^{C-b_m} P_m^0(n), & m = \overline{M_1 + 1, M}. \end{cases} \quad (3.14)$$

Введем систему множеств, соответствующую разбиению множества состояний приоритетных услуг мультивещания по количеству занятых ими единиц ресурсов:

$$\underline{Y}(n) = \{ \mathbf{y} = (y_m)_{m \in M_1} : b_1(\mathbf{y}) = n \}, \quad n = \overline{0, C}. \quad (3.15)$$

Также введем систему множеств, соответствующую разбиению множества состояний неприоритетных соединений обоих типов по числу совместно занятых ими единиц ресурсов:

$$\bar{Z}(n) = \{ \mathbf{z} = ((y_m)_{m \in M_2}, \mathbf{n}) : b_2(\mathbf{y}) + d(\mathbf{n}) = n \}, \quad n = \overline{0, C}. \quad (3.16)$$

Заметим, что $\bar{Z}(n) = \emptyset$ при $n > C^0$. Введем совместное распределение числа приборов, занимаемых заявками приоритетных и неприоритетных потоков:

$$P(i, j) = P(\underline{Y}(i) \times \bar{Z}(j)), \quad i, j = 0, \dots, C, \quad i + j \leq C.$$

Здесь $P(i, j) = 0$ при $j > C^0$.

Найдем представления множеств $Z(n)$ и $Z^0(n)$ через множества систем (3.15) и (3.16). При $n = \overline{0, C^0}$

$$Z(n) = \bigcup_{i=0}^n (\underline{Y}(n-i) \times \bar{Z}(i)),$$

тогда как при $n = \overline{C^0, C}$

$$Z(n) = \bigcup_{i=0}^{C^0} (\underline{Y}(n-i) \times \bar{Z}(i)).$$

Следовательно, для $n = \overline{0, C}$ множество $Z(n)$ можно представить в виде

$$Z(n) = \bigcup_{i=0}^{\min\{n, C^0\}} (\underline{Y}(n-i) \times \bar{Z}(i)), \quad n = \overline{0, C},$$

или в вероятностной форме

$$P(n) = \sum_{i=0}^{\min\{n, C^0\}} P(n-i, i), \quad n = \overline{0, C}. \quad (3.17)$$

Множество $\bar{Z}(n)$, в свою очередь, можно представить в виде свертки

$$\bar{Z}(n) = \bigcup_{i=0}^n (\bar{Y}(i) \times C(n-i)), \quad n = \overline{0, C^0},$$

где

$$\bar{Y}(n) = \{\mathbf{y} = (y_m)_{m \in M_2} : b_2(\mathbf{y}) = n\}, \quad n = \overline{0, C^0},$$

и

$$C(n) = \{\mathbf{n} \in \tilde{N} : d(\mathbf{n}) = n\}, \quad n = \overline{0, C^0}.$$

Обозначив $P_{\bar{Z}}(i, j) = P(\bar{Y}(i) \times C(j))$ совместное распределение $\bar{Y}(n)$ и $C(n)$, получим в вероятностной форме

$$P(\bar{Z}(n)) = \sum_{i=0}^n P_{\bar{Z}}(i, n-i), \quad n = \overline{0, C^0}.$$

Разделение ресурсов в общей зоне между неприоритетными потоками происходит в соответствии с полнодоступной стратегией и описывается моделью звена, представленной в §4.2. При отсутствии резервирования, то есть если $M_1 = \emptyset$, $\bar{Z}(n) = Z(n)$ и вычислить вероятность события $Z(n)$ можно посредством свертки (2.27) функций (2.28) и (2.29):

$$P(Z(n)) = G^{-1}(Z) f_0(M, n) = G^{-1}(Z) \sum_{i=0}^n g(M, i) h(n-i).$$

Введем функцию $\tilde{f}_m(i, n)$, удовлетворяющую соотношениям

$$P(n) = G^{-1}(Z)\tilde{f}_0(1, n), \quad n = \overline{0, C}, \quad (3.18)$$

$$P_m(n) = G^{-1}(Z)\tilde{f}_m(1, n), \quad m = \overline{1, M_1}, \quad n = \overline{0, C}, \quad (3.19)$$

$$P(i, j) = G^{-1}(Z)g(M_1, i)\tilde{f}_0(M_1 + 1, j), \quad i, j = \overline{0, C}, \quad i + j \leq C, \quad (3.20)$$

где функция $g(m, n)$ задана формулой (2.29). В соответствии с (3.17) при $m = 0$ функция $\tilde{f}_0(1, n)$ удовлетворяет соотношению свертки

$$\tilde{f}_0(1, n) = \sum_{i=0}^{\min\{n, C^0\}} g(M_1, n-i)\tilde{f}_0(M_1 + 1, i),$$

причем при $i = 0, \dots, C^0$ величина $\tilde{f}_0(M_1 + 1, i)$ совпадает с $f_0(M_2, i)$, если при вычислении последней перенумеровать I-потoki таким образом, чтобы неприоритетные потоки стояли первыми. Таким образом, верна следующая лемма.

Лемма 4.2. Для любого $m = 0, \dots, M_1$ функция $\tilde{f}_m(i, n)$, удовлетворяющая соотношениям (3.18)–(3.20), вычисляется по формуле

$$\tilde{f}_m(i, n) = \begin{cases} 0, & i = \overline{1, M}, \quad n < 0; \\ 0, & n = \overline{C^0 + 1, C}; \\ h(n), & i > M, \quad n = \overline{0, C^0}; \\ \tilde{f}_m(i + 1, n) + (1 - \delta(m, i))\rho_i \tilde{f}_m(i + 1, n - b_i), & \\ & \text{в остальных случаях,} \end{cases} \quad (3.21)$$

где функция $h(n)$ задана формулой (2.28), а $\delta(m, i)$ – символ Кронекера.

Заметим, что в отличие от функции $f_m(i, n)$, при вычислении которой согласно формуле (2.27) перебор I-потокoв осуществляется в обратном порядке с i по 1, при расчете $\tilde{f}_m(i, n)$ по формуле (3.21) перебор I-потокoв производится в прямом порядке с i по M . Это сделано для того, чтобы сначала осуществлялся перебор приоритетных I-потокoв, занумерованных от 1 до M_1 , а затем всех неприоритетных потокoв.

Обратимся теперь к расчету вероятности события $Z^0(n)$. Множество $Z^0(n)$ при $n = \overline{R, C}$ можно представить в виде

$$Z^0(n) = \bigcup_{i=0}^R (\underline{Y}(i) \times \overline{Z}(n-R)) \cup \bigcup_{i=R+1}^n (\underline{Y}(i) \times \overline{Z}(n-i)).$$

Таким образом, можно записать

$$Z^0(n) = \bigcup_{i=0}^n (\underline{Y}(i) \times \overline{Z}(n - \max\{R, i\})), \quad n = \overline{R, C},$$

или в вероятностной форме:

$$P^0(n) = \sum_{i=0}^n P(i, n - \max\{R, i\}), \quad n = \overline{R, C}. \quad (3.22)$$

Введем функцию $\tilde{f}_m^n(i, j)$, удовлетворяющую следующим соотношениям:

$$P^0(n) = G^{-1}(Z) \tilde{f}_0^n(1, n), \quad n = \overline{R, C}, \quad (3.23)$$

$$P_m^0(n) = G^{-1}(Z) \tilde{f}_m^n(1, n), \quad m = \overline{M_1 + 1, M}, \quad n = \overline{R, C}. \quad (3.24)$$

Из соотношения (3.22) вытекает утверждение следующей леммы.

Лемма 4.3. Для любых $m = 0, \dots, M_1 + 1, M_1 + 2, \dots, M$ и $n = \overline{R, C}$ функция $\tilde{f}_m^n(i, j)$, удовлетворяющая соотношениям (3.23) и (3.24), вычисляется по формуле

$$\tilde{f}_m^n(i, j) = \begin{cases} 0, & i = \overline{1, M}, \quad j < 0; \\ \tilde{f}_m^n(i, n-R), & i > M_1, \quad j > n-R; \\ h(j), & j = \overline{0, C}; \\ \tilde{f}_m^n(i+1, j) + (1 - \delta(m, i)) \rho_i \tilde{f}_m^n(i+1, j-b_i), & \\ & \text{в остальных случаях,} \end{cases} \quad (3.25)$$

где функция $h(n)$ задана формулой (2.28), а $\delta(m, i)$ – символ Кронекера.

Проиллюстрируем предложенный метод вычисления вероятностей событий $Z(n)$ и $Z^0(n)$ на простом примере. Пусть имеется

неполнодоступная система рассматриваемого типа из 6 приборов. На систему поступают два приоритетных I-потока, один неприоритетный I-поток и два II-потока со следующими параметрами: $\mathbf{b} = (1, 2, 1)$, $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3)$, $\mathbf{d} = (1, 2)$, $\mathbf{a} = (a_1, a_2)$. Под заявки приоритетных потоков зарезервировано 2 прибора. Таким образом, структурные параметры системы имеют значения: $C = 6$, $R = 2$, $C^0 = 4$, $M_1 = \{1, 2\}$, $M_2 = \{3\}$, $M = 3$, $M_1 = 2$, $K = \{1, 2\}$, $K = 2$.

На рис. 4.5 разобран ход вычисления функции $\tilde{f}_0(m, n)$ при $m=1$ и $n=6$. На схеме видно, что при $m=3=M_1+1$ значения $\tilde{f}_0(3, 6)$ и $\tilde{f}_0(3, 5)$ обращаются в 0, так как второй аргумент функции $n > C^0$. Таким образом, суммирование осуществляется лишь по таким состояниям сети, в которых заявки неприоритетных потоков занимают C^0 или меньше, при этом общее число занятых приборов равно 6. Состояния сети, по которым в итоге производится суммирование, подписаны под соответствующими ветвями дерева.

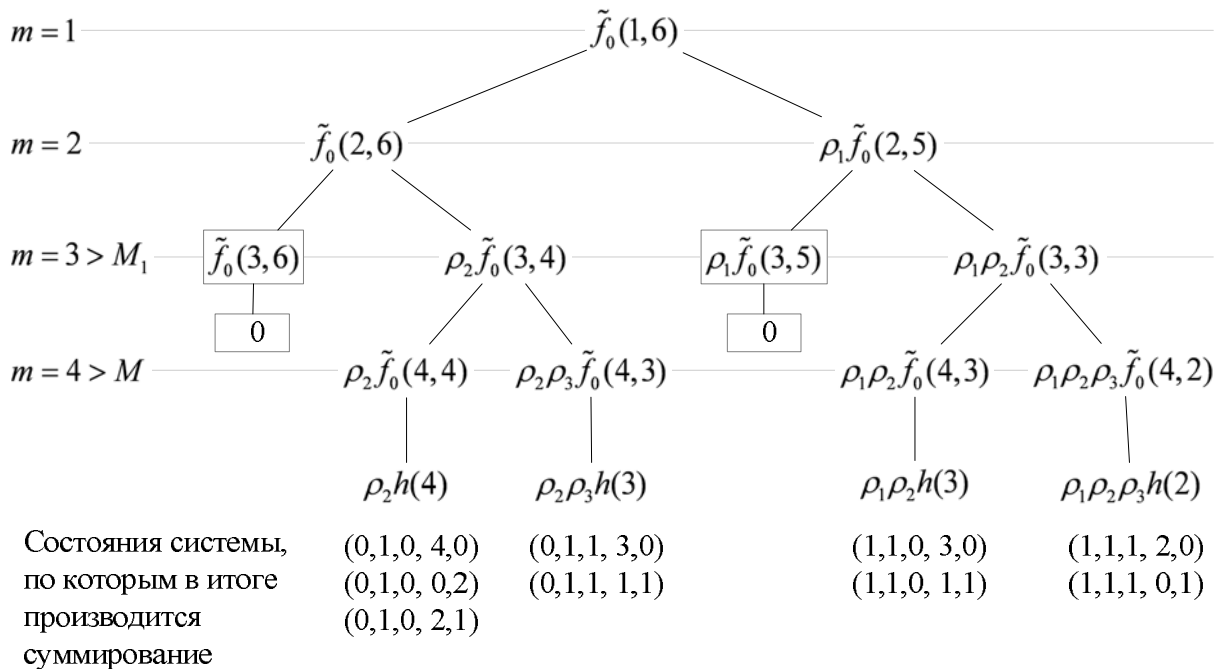
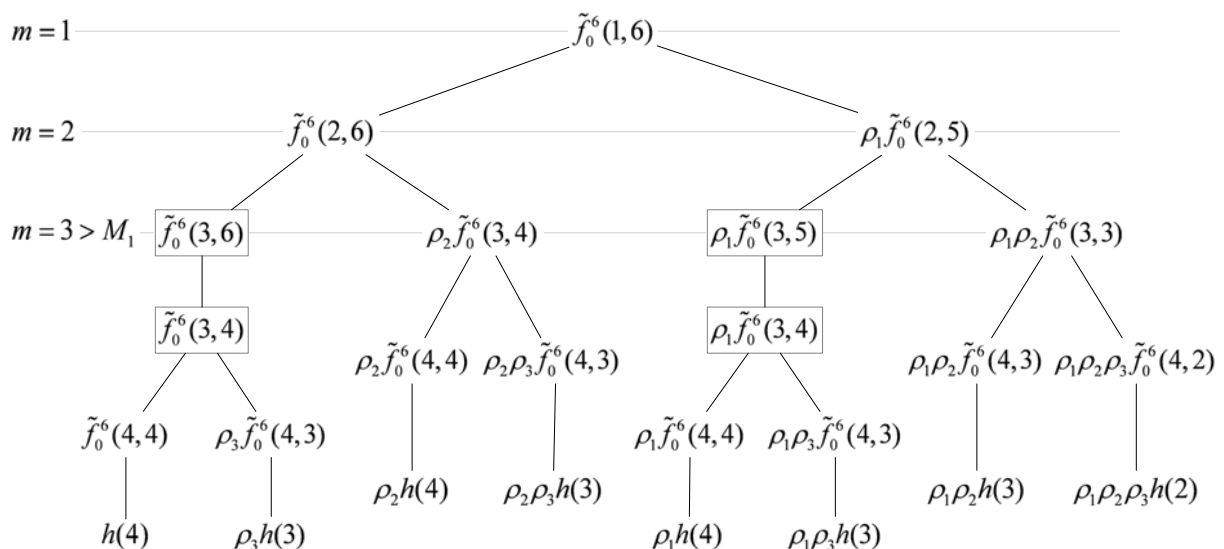


Рис. 4.5. Иллюстрация вычисления функции $\tilde{f}_0(m, n)$

На рис. 4.6 для рассматриваемого примера представлен ход вычисления функции $\tilde{f}_0^n(m, j)$ при тех же значениях аргументов: $m=1$ и $n=j=6$. Здесь, если при $m=3=M_1+1$ второй аргумент функции $j > n-R$, в качестве аргумента берется $j=n-R$. Таким образом, значения функции $\tilde{f}_0^6(3,6)$ и $\tilde{f}_0^6(3,5)$ преобразуются в $\tilde{f}_0^6(3,4)$. В итоге суммирование осуществляется по всем состояниям сети, в которых в свободной зоне занято $n-R$ приборов, при этом число приборов, занятых в резервной зоне заявками приоритетных потоков, может быть любым в пределах от 0 до R . Это видно из списков состояний, приведенных на рисунке под соответствующими ветвями дерева.

Функции $\tilde{f}_0(m, n)$ и $\tilde{f}_m^n(i, j)$ позволяют получить эффективные алгоритмы для вычисления вероятностных характеристик системы. Нетрудно убедиться, что из формул (3.10), (3.11) и (3.12)–(3.14), а также соотношений (3.18), (3.19), (3.23) и (3.24) вытекают следующие утверждения.



Состояния системы, по которым в итоге производится суммирование

(0,0,0, 4,0)	(0,0,1, 3,0)	(0,1,0, 4,0)	(0,1,1, 3,0)	(1,0,0, 4,0)	(1,0,1, 3,0)	(1,1,0, 3,0)	(1,1,1, 2,0)
(0,0,0, 0,2)	(0,0,1, 1,1)	(0,1,0, 0,2)	(0,1,1, 1,1)	(1,0,0, 0,2)	(1,0,1, 1,1)	(1,1,0, 1,1)	(1,1,1, 0,1)
(0,0,0, 2,1)		(0,1,0, 2,1)		(1,0,0, 2,1)			

Рис. 4.6. Иллюстрация вычисления функции $\tilde{f}_0^n(m, j)$

Утверждение 4.4. Нормирующая константа $G(Z)$, где множество Z имеет вид (3.9), вычисляется по формуле

$$G(Z) = \sum_{n=0}^C \tilde{f}_0(1, n). \quad (3.26)$$

Утверждение 4.5. Среднее число занятых приборов (3.10) вычисляется по формуле

$$c^{(1)} = \frac{\sum_{n=1}^C n \cdot \tilde{f}_0(1, n)}{\sum_{n=0}^C \tilde{f}_0(1, n)}. \quad (3.27)$$

Замечание. Здесь, как в модели звена МСС с полноступенчатой стратегией разделения ресурсов, величина $c^{(1)}/C$ представляет собой коэффициент использования звена.

Утверждение 4.6. Вероятностные характеристики (3.11)–(3.14) вычисляются по формулам:

$$B_k^{II} = \frac{\sum_{n=C-d_k+1}^C \tilde{f}_0^n(1, n)}{\sum_{n=0}^C \tilde{f}_0(1, n)}, \quad k = \overline{1, K}, \quad (3.28)$$

$$B_m^I = \begin{cases} \frac{\sum_{n=C-b_m+1}^C \tilde{f}_m(1, n)}{\sum_{n=0}^C \tilde{f}_0(1, n)}, & m = \overline{1, M_1}; \\ \frac{\sum_{n=C-b_m+1}^C \tilde{f}_m^n(1, n)}{\sum_{n=0}^C \tilde{f}_0(1, n)}, & m = \overline{M_1 + 1, M}, \end{cases} \quad (3.29)$$

$$H_m = \begin{cases} \frac{\sum_{n=0}^{C-b_m} \tilde{f}_m(1,n)}{C}, & m = \overline{1, M_1}; \\ \frac{\sum_{n=0}^C \tilde{f}_0(1,n)}{C} \\ \frac{\sum_{n=0}^{C-b_m} \tilde{f}_m^n(1,n)}{C}, & m = \overline{M_1 + 1, M}. \\ \frac{\sum_{n=0}^C \tilde{f}_0(1,n)}{C} \end{cases} \quad (3.30)$$

Полученные в данном параграфе результаты позволяют реализовать эффективные алгоритмы для расчета вероятностных характеристик отдельного звена мультисервисной сети с резервированием ресурсов для подмножества многоадресных соединений. Заметим, что важным частным случаем рассмотренной системы является звено, на котором резервируются ресурсы для всех многоадресных соединений, тогда как одноадресные соединения выступают как неприоритетные. Важность данного случая объясняется тем, что снижение качественных характеристик многоадресного соединения отражается на целой группе пользователей сети, поэтому таким соединениям целесообразно назначать высокий приоритет.

В заключение главы для иллюстрации рассмотрим пример выбора исходных данных (структурных и нагрузочных параметров) для анализа звена мультисервисной сети [15].

Пример 4.1. Рассмотрим передачу информации по звену МСС в рамках концепции «Triple Play Services» (коммерческой концепции, предполагающей предоставление услуг телефонии, телевидения и доступа в Интернет в виде одного коммерческого предложения)).

Пусть через звено Gigabit Ethernet устанавливаются одноадресные соединения для потоковой передачи данных со скоростью 1 Мбит/с, предоставления видео по требованию качества DVD со скоростью

5 Мбит/с, информационных голосовых услуг и телефонии со скоростью 8 кбит/с, а также многоадресные соединения для просмотра 25 каналов HDTV в формате MPEG-2 со скоростью передачи 30 Мбит/с. При этом пусть пять телеканалов пользуются большей популярностью у пользователя по сравнению с остальными, а доступ к еще пяти из менее популярных 20 каналов предлагается по более высокой цене.

Заметим, что максимальная пропускная способность звена для сетей Ethernet в среднем составляет около 60% от заявленной скорости 1 Мбит/с передачи данных. Примем 8 кбит/с за одну передаточную единицу. Тогда емкость звена C равна $600 \cdot 10^3 / 8 = 75000$ единиц. Источник мультивещания предоставляет 25 соответствующих телеканалам телевидения высокой четкости (High Definition Television, HDTV) услуг, $M = \{1, \dots, 25\}$, для которых $b_m = b = 3750$. Пусть последние пять каналов ($m = 21, \dots, 25$) являются более популярными и интенсивность потоков запросов на них выше. Пусть также первые пять каналов ($m = 1, \dots, 5$) являются приоритетными по причине более высокой цены. Вследствие этого возникает необходимость повышения качества обслуживания при предоставлении доступа к ним, что может быть достигнуто путем резервирования ресурсов звена. Поэтому, помимо полнодоступной системы, следует рассмотреть случаи, когда для первых пяти каналов на звене резервируются, например, 80 Мбит/с и 120 Мбит/с, что дает следующие значения параметров для модели звена с резервированием: $M_1 = 5$, $R = 10^4$ в первом случае и $R = 15 \cdot 10^3$ – во втором случае.

Помимо многоадресных соединений, на звене имеется 4 класса одноадресных соединений, $K = \{1, \dots, 4\}$, соответствующих потоковой передаче данных, видео по требованию, информационным голосовым услугам и телефонии. Параметры услуг мультивещания и классов одноадресных соединений представлены в таблице 4.1. Нагрузочные

параметры класса, соответствующего телефонии, при расчетах могут варьироваться.

Таблица 4.1. Параметры соединений
для численного анализа отдельного звена МСС

Услуги мультивещания				
Трафик	m	b_m	μ_m^{-1} , сек.	$\rho_m = \frac{\lambda_m}{\mu_m}$
HDTV, 30 Мбит/с	1,...,20	3750	3600	1
HDTV (популярные каналы), 30 Мбит/с	21,...,25	3750	3600	10
Классы одноадресных соединений				
Трафик	k	d_k	κ_k^{-1} , сек.	$a_k = \frac{V_k}{\kappa_k}$
Передача данных, 1 Мбит/с	1	125	900	1
Видео по требованию, 5 Мбит/с	2	625	4320	1,2
Информационные голосовые услуги, 8 кбит/с	3	1	180	50
Телефония, 8 кбит/с	4	1	180	$[0; 10^5]$

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Основная

- [1] Башарин Г.П. Лекции по математической теории телетрафика. – М.: Изд-во РУДН, 2007. – 268 с.
- [2] Корнышев Ю.Н., Пшеничников А.П., Харкевич А.Д. Теория телетрафика. Учебник для вузов. – М.: Радио и связь, 1996. – 272 с.
- [3] Крылов В.В., Самохвалова С.С. Теория телетрафика и ее приложения. – СПб.: «БХВ-Петербург», 2005. – 288 с.
- [4] Кучерявый А.Е., Цуприков А.Л. Сети связи следующего поколения. М.: ФГУП ЦНИИС, 2006. – 280 с.
- [5] Лагутин В.С., Степанов С.Н. Телетрафик мультисервисных сетей связи. – М.: Радио и связь, 2000. – 320 с.
- [6] Соколов Н.А. Телекоммуникационные сети. Монография в 4-х главах. – М.: Альварес Паблишинг, 2004.
- [7] Шнепс-Шнеппе М.А. Системы распределения информации. Методы расчета. Справочное пособие. – М.: Связь, 1979. – 344 с.
- [8] Iversen V.B. Teletraffic Engineering Handbook. – ITU-D, Nov. 2005. – 323 p.
- [9] Kelly F.P. Reversibility and stochastic network. – Chichester: Wiley, 1979. – 630 p.
- [10] Ross K.W. Multiservice loss models for broadband telecommunication networks. – London: Springer-Verlag, 1995. – 343 p.

Дополнительная

- [11] Башарин Г.П., Гайдамака Ю.В., Самуйлов К.Е. Яркина Н.В. Модели для анализа качества обслуживания в сетях связи следующего поколения. Уч. пособие для бакалавров. – М.: Изд-во РУДН, 2008. – 111 с.
- [12] Бабков В.Ю., Полинцев П.В., Устюжанин В.И. Качество услуг мобильной связи. Оценка, контроль и управление. – Горячая линия-Телеком, 2005. – 160 с.
- [13] Деарт В.Ю. Мультисервисные сети связи. – М.: Инсвязьиздат, 2007. – 166 с.: ил.
- [14] Маковеева М.М., Шинаков Ю.Г. Системы связи с подвижными объектами. Уч. пособие для вузов. – М.: Радио и связь, 2002. – 440 с.

- [15] Наумов В.А., Самуйлов К.Е., Яркина Н.В. Теория телетрафика мультисервисных сетей. – М.: Изд-во РУДН, 2007. – 192 с.
- [16] Таха Х.А. Введение в исследование операций. – М.: «Вильямс», 2007. – 912 с.
- [17] Телекоммуникационные системы и сети: Уч. пособие. В 3 томах. Том 3. – Мультисервисные сети / Величко В.В. и др. / Под ред. проф. Шувалова В.П. – М.: Горячая линия-Телеком, 2005. – 592 с.
- [18] Handbook of Wireless Networks and Mobile Computing. Ed. Ivan Stojmenovic. – New York: J. Wiley & Sons, 2002. – 630 p.

СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ

- a_k - интенсивность предложенной нагрузки, создаваемая запросами на установление соединения k -класса для одноадресных соединений
- $a_{n,m}$ - интенсивность перехода из состояния n в состояние m
- A - матрица интенсивностей переходов МП $X(t)$
- b_k - требование к ШПП для заявок k -го поступающего потока для одноадресных соединений
- b_{ms} - число единиц емкости звена, требуемое для предоставления услуги $m \in M_s$ для многоадресных соединений
- B_k - множество блокировок одноадресных соединений k -класса сети мультивещания
- B_{mps} - множество блокировок (m, p, s) -пути сети мультивещания
- B_{mps} - вероятность блокировки (m, p, s) -пути сети мультивещания
- B_m^I - множество потерь (I, m) -заявок для сети мультивещания
- B_m^I - вероятность потерь (I, m) -заявок для сети мультивещания
- B_k^{II} - множество потерь (II, k) -заявок для сети мультивещания
- B_k^{II} - вероятность потерь (II, k) -заявок для сети мультивещания
- $c^{(1)}$ - среднее число занятых приборов на отдельном звене сети мультивещания
- C - число частотных каналов на БС макросоты;
емкость отдельного звена сети мультивещания

C_l	- емкость l -звена сети мультивещания, $l \in L$
C^0	- емкость общедоступной (для одно- и многоадресных соединений) части на отдельном звене сети мультивещания
d_k	- требования к ШПП для заявок k -го класса для одноадресных соединений
\mathbf{f}	- функция управления доступом
F_{mps}	- вероятность того, что (m, p, s) -путь включен для многоадресных соединений
F_m	- вероятность того, что m -услуга для многоадресных соединений на отдельном звене сети мультивещания включена
G	- нормирующая константа
$G(\Omega)$	- нормирующая константа множества Ω
h_i	- среднее количество посещений абонентом микросоты i
H_{mps}	- вероятность того, что (m, p, s) -путь выключен, но в сети достаточно ресурсов для его включения для многоадресных соединений
H_m	- вероятность того, что m -услуга для многоадресных соединений на отдельном звене сети мультивещания выключена, но в сети достаточно ресурсов для ее включения
\mathbf{e}_k	- вектор, k -я компонента которого равна 1, а остальные – нули, $\mathbf{e}_k = (0^{k-1}, 1, 0^{K-k})^T$
EX	- среднее значение СВ X
\mathbf{I}	- единичная матрица
\mathbf{J}	- пространство состояний МП $X(t)$

- G** - неориентированный граф для модели фрагмента иерархической ССПС
- K** - число микросот для модели фрагмента иерархической ССПС, число потоков (услуг)
- K** - множество услуг (классов услуг для одноадресных соединений)
- L** - число звеньев сети мультивещания
- L** - множество звеньев сети мультивещания
- L_{ps} - множество всех звеньев физического пути $p \in P_s$ к источнику $s \in S$ сети мультивещания для многоадресных соединений
- (m, p, s) - логический путь, (m, p, s) -путь, $m \in M_s$, $p \in P_s$, $s \in S$, в сети мультивещания для многоадресных соединений
- M_s - множество услуг, предоставляемых источником $s \in S$ сети мультивещания для многоадресных соединений
- M** - матрица смежности
- n** - состояние системы, вектор численности всех услуг, предоставляемых системой в момент t , $\mathbf{n} = (n_1, \dots, n_K)^T$
- N** - число источников в модели Энгсета
- p** - физический путь в сети мультивещания, $p \in P_s$, для многоадресных соединений
- $\{p_n, n \in J\}$ - стационарное распределение МП $X(t) \in J$
- $P(i, \square)$ - стационарное распределения числа i зарегистрированных абонентов и числа j активных абонентов в соте иерархической ССПС, $i = 0, 1, \dots$
- $P(\square, j)$ - стационарное распределения числа j активных абонентов в соте иерархической ССПС, $j = 0, 1, \dots, C$

P_s	- множество всех физических путей к источнику $s \in S$ в сети мультивещания для многоадресных соединений
P_s^l	- множество физических путей к источнику $s \in S^l$, проходящих через звено $l \in L$ для многоадресных соединений
r_j^I	- порог резервирования (I, m) -заявок для сети мультивещания
r_j^{II}	- порог резервирования (II, k) -заявок для сети мультивещания
$R_{\beta\delta}$	- коэффициент корреляции между СВ β и δ
S	- число источников информации сети мультивещания
$S(\mathbf{f})$	- множество возвратных состояний СтМП $\mathbf{X}(t)$, $t \geq 0$,
S	- множество источников информации сети мультивещания
S^l	- множество источников информации, предоставляющих услуги через звено $l \in L$ для многоадресных соединений
S	- пространство состояний системы
S_k	- подпространство приема для k -сообщений, $k = \overline{1, K}$
$\overline{S_k}$	- подпространство блокировки для k -сообщений, $k = \overline{1, K}$
$\langle S, A \rangle$	- СМО со структурой ресурсов S и алгоритмом A их распределения между входящими потоками заявок
TH_k	- пропускная способность для k -сообщений
$u(x)$	- функция Хевисайда, $u(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$
$UTIL$	- коэффициент использования ШЦЛ
v	- число приборов в СМО
V	- емкость ШЦЛ в БЦК

x_{mps}	- состояние логического пути в сети мультивещания
$X(t)$	- случайный процесс, $t \geq 0$
$\mathbf{X}(t)$	- многомерный случайный процесс, $t \geq 0$
$\mathbf{Z}(t)$	- многомерный случайный процесс, $t \geq 0$, описывающий состояние сети с одноадресными и многоадресными соединениями
\tilde{Z}	- пространство состояний сети с одноадресными и многоадресными соединениями
γ	- СВ ширины полосы пропускания, занятой на звене сети мультивещания при обслуживании установленных через него соединений обоих типов
$\delta(i, j)$	- символ Кронекера, $\delta(i, j) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$
ε	- интенсивность ПП поступления запросов на установление соединения и предоставления разговорного канала от одного пассивного зарегистрированного в соте абонента
κ_k	- параметр экспоненциального обслуживания заявок на установление соединения k -класса для одноадресных соединений
λ_0	- интенсивность ПП активных абонентов в соту
λ_H	- интенсивность ПП пассивных абонентов в соту
λ_k	- интенсивность поступления заявок k -го потока, класса
λ_{mps}	- интенсивность ПП запросов на установление (m, p, s) -пути для многоадресных соединений, $m \in M_s$, $p \in P_s$, $s \in S$
$\lambda_k(\mathbf{n})$	- интенсивность принятого потока k -сообщений в

- состоянии \mathbf{n}
- μ_k - параметр экспоненциального обслуживания заявок k -го потока
 - μ_{mps} - параметр экспоненциального распределения ПП запросов на установление (m, p, s) -пути в сети мультивещания, $m \in M_s$, $p \in P_s$, $s \in S$ для многоадресных соединений
 - $\mu_k(\mathbf{n})$ - интенсивность обслуженного потока k -сообщений в состоянии \mathbf{n}
 - ν_k - интенсивность ПП запросов на установление соединения k -класса для одноадресных соединений
 - π_k - вероятность того, что вновь поступившее k -сообщение будет заблокировано
 - $\pi_0(C, g)$ - вероятность блокировки новых вызовов для модели фрагмента иерархической ССПС
 - $\pi_H(C, g)$ - вероятность блокировки хэндовер-вызовов для модели фрагмента иерархической ССПС
 - Θ_0 - маршрутная матрица
 - ρ - интенсивность предложенной нагрузки
 - ρ_k - интенсивность предложенной k -нагрузки
 - ρ'_k - измеряемая в БЦК интенсивность предложенной k -нагрузки с учетом требуемой ширины полосы
 - ρ_{mps} - нагрузка, создаваемая запросами на установление (m, p, s) -пути в сети мультивещания для многоадресных соединений
 - $1(\text{событие } A)$ - функция-индикатор,

$$1(\text{событие } A) = \begin{cases} 1, & \text{если } A \text{ произошло} \\ 0, & \text{в противном случае} \end{cases}$$

$\lfloor x \rfloor$

- целая часть числа x

■

- СИМВОЛ ОКОНЧАНИЯ ДОКАЗАТЕЛЬСТВА

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Абонент, 10, 11, 13, 27, 28, 30, 33, 38, 40, 42, 44, 57
 активный, 43, 44, 47, 51
 пассивный, 43, 44, 49, 50, 52
- Вероятность
 блокировки, 10, 21, 25, 51, 52, 54, 61, 63, 72, 85
 потери, 21, 85, 92, 95, 96, 98, 99, 108
 по времени, 21, 92
 по сообщениям, 22
- Вызов
 новый, 29, 30, 31, 32, 44, 51, 52
 хэндовер, 28, 29, 30, 31, 32, 43, 51, 57
- Дерево мультивещания, 77, 79
- Заявка, 12, 26, 58, 66, 69, 71, 73, 88, 99, 101, 114
- Интенсивность, 10, 13, 16, 19, 22, 30, 34, 44, 49, 52, 55, 68, 81, 88, 91, 118
 нагрузки, 14, 26, 49, 50, 51, 52, 53, 74
 увеличения дохода, 71, 72
- Источник, 11, 76, 79, 80, 82, 87, 118
- Коэффициент использования, 22, 93, 116
 мгновенный, 15, 22, 23
- Макросота, 27, 28, 29, 31, 32
- Маршрут, 38, 40, 41, 76, 77, 78, 80, 84
- Микросота, 29, 30, 38, 41
- Модель
 Энгсета, 9, 10
 Эрланга, 6, 9, 11, 12, 17, 53
- Мультивещание, 79, 82, 85, 89, 106, 110, 118, 119
- Нагрузка, 9, 11, 13, 17, 22, 26, 33, 53, 57, 71
- Пикосота, 28, 29, 42
- Подпространство
 блокировки, 15, 16, 86
 приема, 15, 16, 86
- Пользователь, 4, 76, 77, 79, 84, 85, 87, 89, 92, 117
- Пропускная способность, 10, 22, 27, 28, 43, 73, 75, 118
- Пространство состояний, 15, 16, 23, 33, 37, 44, 47, 52, 59, 61, 63, 66, 71, 72, 75, 81, 83, 85, 86, 89, 90, 95, 102, 104
- Пуассоновский поток, 13, 30, 44, 49, 51, 52, 81, 88
- Путь
 логический, 79, 80, 81, 82, 85, 89
 физический, 76, 77, 84
- Распределение
 квазиэрланговское, 52
 равновесное, 60, 62, 66, 67, 68, 81, 89, 102
 экспоненциальное, 13, 31, 33, 44, 81, 88
 Эрланга, 26, 27
- Сеть
 мультисервисная, 6, 9, 75, 77, 93, 101, 117
 следующего поколения, 4, 5, 9

Соединение

многоадресное, 6, 75, 76, 77, 78, 83, 86, 88, 92, 117

одноадресное, 6, 9, 75, 76, 78, 81, 89, 92, 93, 101, 117, 118

Стратегия

неполнодоступная, 101

полного разделения, 67, 70, 74

полнодоступная, 58, 60, 67, 70, 73, 75, 89, 101, 111, 116

разделения, 68

резервирования каналов, 60, 67, 118

Услуга, 4, 6, 10, 13, 15, 17, 20, 25, 76, 79, 82, 84, 87, 92, 106, 110, 118

Формула Эрланга, 27, 43, 54

Complete Partitioning, 67

Complete Sharing, 60

Next Generation Networks, 4, 5, 9, 10

Partitioning Policy, 68

Trunk Reservation Policy, 60

ОПИСАНИЕ КУРСА И ПРОГРАММА

1. Цели и задачи курса

Область знаний

Курс относится к области знаний «Информационно-телекоммуникационные системы», соответствующей одноименному приоритетному направлению развития науки и технологий, входящему в перечень, утвержденного Президентом Российской Федерации.

Уровень обучения и направления подготовки по действующему перечню

Курс является обязательной дисциплиной для студентов, обучающихся по магистерской программе «Управление инфокоммуникациями» по направлению 010400 «Информационные технологии».

Лица, желающие освоить данную программу специализированной подготовки магистра информационных технологий, должны иметь высшее профессиональное образование определенной ступени, подтвержденное документом государственного образца.

Лица, имеющие диплом бакалавра по направлениям 010300 «Математика. Компьютерные науки», 010400 «Информационные технологии», 010500 «Прикладная математика и информатика», зачисляются на специализированную магистерскую подготовку на конкурсной основе. Условия конкурсного отбора определяются вузом на основе государственного образовательного стандарта высшего профессионального образования бакалавра по данному направлению.

Для эффективного обучения на магистерской программе «Управление инфокоммуникациями» рекомендуется в бакалавриате прослушать профиль специальных дисциплин по выбору в составе следующих курсов:

«Основы формальных методов описания бизнес процессов»;

«Модели для анализа качества обслуживания в сетях связи следующего поколения»;

«Основы разработки корпоративных инфокоммуникационных систем»;

«Основы управления инфокоммуникационными компаниями».

Лица, желающие освоить программу специализированной подготовки магистра по данному направлению и имеющие высшее профессиональное образование, профиль которого не указан выше, допускаются к конкурсу по результатам сдачи экзаменов по дисциплинам, входящим в программу программы дополнительной профессиональной подготовки «Основы управления инфокоммуникациями», которая включает курсы:

«Введение в управление инфокоммуникациями»

«Введение в формальные методы описания бизнес-процессов»;

«Архитектура и принципы построения современных сетей и систем телекоммуникаций»;

«Корпоративные информационные системы».

Цели курса

- Изучение принципов функционирования сетей связи следующего поколения.
- Рассмотрение вопросов качества в NGN на различных уровнях.
- Знакомство с методами анализа и расчета показателей качества отдельных элементов сетей, а также сети в целом, в NGN.

- Исследование точных и приближенных методов анализа качества обслуживания в сетях связи следующего поколения.

Задачи курса

После успешного прохождения курса слушатели должны

знать:

- основные понятия и определения, относящиеся к концепции сетей связи следующего поколения;
- основные протоколы сетей связи следующего поколения;
- требования международных стандартов к показателям качества на различных уровнях сети NGN;
- методы разработки и анализа моделей телекоммуникационных систем сложной структуры;
- численные методы расчета (приближенные и точные) характеристик сети;

уметь:

- строить модели отдельных функциональных элементов NGN, а также модели сети в целом;
- проводить исследование построенных моделей, получать их вероятностные характеристики, требующиеся для анализа показателей качества;
- использовать изученные методы и принципы при разработке моделей и анализе качества обслуживания для реально существующих сетей.

2. Инновационность курса

По содержанию.

«Сети связи следующего поколения» (NGN) – современная концепция, отражающая конвергенцию информационно-телекоммуникационных сетей в единую глобальную сеть. На сегодняшний день концепция NGN находится на этапе исследований и становления, причем чаще формулируются требования к NGN и задачи изучения, а не конкретные решения. Одной из особенностей NGN является возможность выбора пользователем определенного уровня качества предоставления услуги из нескольких уровней качества, предлагаемых сетью. Новые виды обслуживания, регулярно появляющиеся в современных информационно-телекоммуникационных сетях, делают проблему обеспечения качества услуги на различных уровнях сети одной из наиболее актуальных. Это привело к необходимости рассмотрения вопросов качества с новых позиций.

Содержание курса обеспечивает слушателей необходимым объемом знаний для проведения исследований в этой новой области телекоммуникаций.

По методике преподавания и организации учебного процесса.

Методика преподавания основана на применении современных информационных технологий. Учебно-методический комплекс с одноименным названием помимо традиционных методических материалов включает электронный учебник, интегрированный в инфокоммуникационную среду типа eLearning. Эти средства позволяют организацию и проведение лабораторных занятий в виде виртуального класса, где студенты работают под руководством преподавателя в асинхронном режиме. Такой режим позволяет осуществлять эффективный контроль уровня знаний за счет постоянного наблюдения за степенью

освоения курса учащимися и за ходом выполнения промежуточных видов контроля знаний.

По литературе.

В настоящее время существует недостаток учебной литературы по вопросам качества в сетях связи следующего поколения, как на русском, так и на английском языке.

3. Структура курса

Общие положения

Трудоемкость курса: 4 кредита.

Аудиторные занятия:

лекции – 2 часа в неделю;

лабораторные занятия – 2 часа в неделю;

Самостоятельная работа студента: 2 часа в неделю.

Содержание курса, объем знаний, общие требования к промежуточному и итоговому контролю знаний определяются программой курса, график обучения определяется календарным планом, а оценка освоения программы курса студентом – методикой оценки уровня знаний.

Содержание курса

Темы лекций

Тема 1. Концепция сетей связи следующего поколения (NGN – Next Generation Networks).

1.1. Организации, занимающиеся стандартами NGN. Архитектура, протоколы и услуги сетей связи следующего поколения.

1.2. Понятие качества в NGN. Качество восприятия, качество обслуживания, качество функционирования сети. Параметры оценки качества и требования международных стандартов.

1.3. Принципы построения моделей для анализа качества.

Тема 2. Математические модели телекоммуникационных систем сложной структуры.

2.1. Пример классической мультисервисной модели: модель Эрланга с явными потерями.

2.2. Общий подход к построению моделей телекоммуникационных систем сложной структуры в виде системы массового обслуживания (СМО) $\langle S, A \rangle$ с ресурсами некоторой структуры S и алгоритмом A их распределения между входящими потоками заявок.

2.3. Математическая модель буферизации в узле коммутации пакетов в виде СМО $\langle S^i, A_u \rangle$, $i = \overline{1, 5}$.

2.4. Основные параметры модели фрагмента системы спутниковой связи $\langle S^2, A_5 \rangle$.

Тема 3. Модели сетей сотовой подвижной связи (ССПС).

3.1. Принципы функционирования ССПС. Организация каналов связи. Соты, кластеры и эстафетная передача обслуживания (хэндовер).

3.2. Модель кластера двухуровневой ССПС: условия мультипликативности, анализ основных вероятностно-временных характеристик (ВВХ).

3.3. Модель микросоты с двумя типами каналов и учетом мобильности абонентов: построение пространства состояний, вывод СУР.

- 3.4. Анализ модели микросоты с помощью квазиэрланговской СМО с одномерным распределением числа занятых каналов.
- 3.5. Рекуррентный алгоритм расчета вероятностей блокировок. Оптимизация числа каналов.

Тема 4. Управление доступом для мультисервисных СМО.

- 4.1. Стратегии доступа: основные определения. Стратегия резервирования каналов.
- 4.2. Координатно-выпуклые стратегии. Системы уравнений глобального (СУГБ) и частичного (СУЧБ) балансов. Основные типы координатно-выпуклых стратегий.
- 4.3. Об оптимизации стратегии доступа.

Тема 5. Особенности анализа мультисервисных сетей связи (МСС).

- 5.1. Мультисервисные сети с одноадресными и многоадресными соединениями. Понятие мультивещания. Дерево мультивещания.
- 5.2. Модель сети мультивещания с одноадресными и многоадресными соединениями: пространство состояний, равновесное распределение вероятностей.
- 5.3. Модель полнодоступного звена сети мультивещания: пространство состояний и равновесное распределение, вероятностные характеристики.
- 5.4. Модель звена сети мультивещания с резервированием: пространство состояний и равновесное распределение, вероятностные характеристики.

Темы семинарских занятий

Тема 1. Основные протоколы NGN: H.323, SIP, MGCP, MEGACO/H.248, BICC, Sigtran. Мультимедийная IP-ориентированная подсистема связи (IP-based Multimedia Subsystem, IMS) как ядро NGN.

Тема 2. Эволюция мобильных сетей. Базовые характеристики сетей разных поколений. Сети 3G и 4G.

Тема 3. Методы определения момента инициации хэндовера. Процедуры жесткой и мягкой эстафетной передачи. Упрощающие предположения для построения моделей оценки качества. Параметры качества: вероятности блокировок нового вызова, хэндовера, многократного хэндовера.

Тема 4. Обратимые марковские процессы. Необходимое и достаточное условия обратимости. Обратимость усеченного марковского процесса.

Тема 5. Четыре основные координатно-выпуклые стратегии и связь между ними. Примеры оптимизации для мультисервисных СМО.

Тема 6. Статическое и динамическое мультивещание. Дерево мультивещания.

Тема 7. Основные алгоритмы MCC: RPB/TRPB, RPM, СВТ.

Тема 8. Основные протоколы MCC: IGMP, BGMP, СВТ, DVMRP, MOSPF, PIM.

Тема 9. Метод расчёта отдельного звена сети мультивещания: алгоритм свертки.

Требования к контролю знаний

В процессе чтения курса предусмотрен один промежуточный контроль знаний и итоговый контроль знаний. Оценка знаний студента по каждому виду контроля осуществляется в соответствии с методикой оценки знаний.

Промежуточный контроль знаний.

Контроль уровня знаний осуществляется в виде письменной контрольной работы, включающей 2 вопроса по темам № 1, № 2 и № 3 содержания курса.

Промежуточный контроль знаний.

Примерный перечень вопросов контрольной работы № 1:

1. Назовите организации, занимающиеся разработкой стандартов для NGN. Расскажите об истории их работы над стандартами NGN.
2. Поясните значение термина «Triple Play Services».
3. Каковы ключевые особенности сетей связи следующего поколения?
4. Перечислите услуги, которые можно предоставлять пользователям с помощью сети NGN.
5. Опишите многоуровневую архитектуру NGN, назовите основные функциональные элементы сети.
6. Каково назначение программного коммутатора Softswitch?
7. Расскажите о мультимедийной IP-ориентированной подсистеме связи. Какая роль отводится подсистеме IMS в NGN?
8. Назовите показатели, которыми оценивается качество на уровнях пользователя, приложений, сети.

9. Перечислите классы обслуживания QoS, рекомендованные сектором стандартизации электросвязи Международного союза электросвязи (МСЭ-Т).
10. Перечислите параметры качества обслуживания QoS.
11. Дайте краткую характеристику основных протоколов NGN: H.323, SIP, MGCP, MEGACO/H.248, BICC, Sigtran.
12. Выпишите перечень предположений для мультисервисной модели Эрланга.
13. Постройте пространство состояний и выпишите теорему о мультипликативности стационарного распределения вероятностей (без доказательства) для мультисервисной модели Эрланга.
14. Для мультисервисной модели Эрланга выпишите СУГБ.
15. Для мультисервисной модели Эрланга выпишите и решите СУЧБ, получите выражения для стационарных вероятностей.
16. Определите следующие показатели качества для мультисервисной модели Эрланга: вероятность потерь по времени, пропускная способность, коэффициент использования ШЦЛ.
17. Изложите рекуррентный алгоритм вычисления макровероятностей $q(v)$ и других макрохарактеристик для мультисервисной модели Эрланга.
18. Опишите общий подход к построению моделей телекоммуникационных систем сложной структуры в виде СМО $\langle S, A \rangle$.
19. Постройте модель буферной памяти в узле коммутации пакетов. Опишите пять способов управления буферной памятью.
20. На примере СМО $\langle S^1, A_2 \rangle$ изложите алгоритм свертки для расчета вероятностно-временных характеристик системы.

21. Назовите ключевые особенности спутниковых систем связи, а также различные способы разделения пропускной способности ретранслятора.
22. Постройте модель спутниковой системы связи в виде СМО $\langle S^2, A_5 \rangle$. Выпишите систему уравнений глобального баланса (СУГБ).
23. Получите равновесное распределение для СМО $\langle S^2, A_5 \rangle$ и докажите его мультипликативность.
24. Для СМО $\langle S^2, A_5 \rangle$ оцените вероятность потерь и другие макрохарактеристики.
25. Назовите поколения ССПС, приведите примеры соответствующих стандартов сетей. Расскажите о принципах организации связи в ССПС.
26. Определите понятие иерархической сотовой системы связи, понятия соты, кластера и хэндовера. Перечислите виды сот иерархической ССПС.
27. Постройте модель кластера двухуровневой ССПС. Получите равновесное распределение для этой модели.
28. Докажите мультипликативность равновесного распределения для модели кластера двухуровневой ССПС.
29. Дайте определение обратимого марковского процесса. Выпишите необходимое и достаточное условия обратимости. При каких условиях усеченный марковский процесс будет обладать свойством обратимости?

Примерный перечень задач контрольной работы №1:

1. Для модели ШЦЛ с параметрами $V = 8$, $K = 2$, $b_1 = 1$, $b_2 = 2$ выписать подпространства S_1 и S_2 , а также подпространства \bar{S}_1 и \bar{S}_2 .

2. Для мультисервисной модели Эрланга с параметрами $V = 6$, $K = 2$, $b_1 = 1$, $b_2 = 2$, $\rho_1 = 3$, $\rho_2 = 2$ вычислить по рекуррентному алгоритму макровероятности $q(v)$, $v = \overline{0, V}$, а затем UTIL, π_1 , π_2 .
3. Для мультисервисной модели Эрланга получите формулы для вычисления макровероятности $q(v)$ и вероятности потерь π_k при $K = 1$, $\rho = \frac{\lambda_1}{\mu_1}$, $b_1 > 1$.
4. Для мультисервисной модели Эрланга при $K = 2$, $c = 10$, $b_1 = 1$, $b_2 = 2$, $\lambda_2^* = 1$, $\mu_1 = \mu_2 = 1$ для $\lambda_1^* = 2$ и $\lambda_1^* = 3$ для каждого $N = 5, 10, 20, 50, 100$ вычислить $q(c)$, $c = \overline{0, C}$, среднее число занятых каналов UTIL, вероятность блокировки π_1 и π_2 заявок обоих типов.

Итоговый контроль знаний.

Контроль уровня знаний осуществляется в виде письменной контрольной работы, включающей 2 вопроса по темам № 4 и № 5 содержания курса.

Примерный перечень вопросов:

1. Назовите стратегии управления доступом для СМО

\mathbf{M}	\mathbf{M}	C	0
$\lambda(\mathbf{n}), \mathbf{b}$	$\mu(\mathbf{n})$.
2. Дайте характеристику четырех координатно-выпуклых стратегий, опишите связи между ними.
3. Дайте краткую характеристику доходности для СМО с явными потерями. Приведите примеры оптимизации для мультисервисных СМО.
4. Определите понятие МСС, одноадресного и многоадресного соединения.

5. Что такое дерево мультивещания? Приведите примеры услуг, образующих динамические и статические деревья мультивещания. В чем разница между этими деревьями?
6. Расскажите об основных алгоритмах, используемых в МСС: RPB/TRPB, RPM, СВТ.
7. Назовите и опишите основные протоколы МСС: IGMP, VGMP, СВТ, PIM.
8. Постройте модель МСС с одноадресными соединениями, выпишите пространство состояний, получите выражения для стационарных вероятностей.
9. Получите выражения для вероятностей блокировок установления соединения в МСС с одноадресными соединениями. Приведите примеры расчета.
10. Опишите метод расчёта отдельного звена МСС с одноадресными соединениями с помощью алгоритма Кауфмана-Робертса.
11. Для модели сети мультивещания постройте пространство состояний, выпишите основные множества для определения вероятностных характеристик модели.
12. Постройте модель отдельного звена сети мультивещания. Выпишите СУР, получите ее решение.
13. Изложите алгоритм свертки для расчета вероятностных характеристик отдельного звена сети мультивещания.

Примерный перечень задач контрольной работы № 2:

1. Доказать, что стратегия доступа РР (Partitioning Policy) при $L=1$ представляет собой полнодоступную стратегию СС (Complete

Sharing), а при $L = K$ - стратегию полного разделения СР (Complete Partitioning).

- Для математической модели ШЦЛ с индивидуальными потолками выписать СУЧБ для состояний \mathbf{n} и $\mathbf{n} - \mathbf{e}_k$, а также для состояний \mathbf{n} и $\mathbf{n} + \mathbf{e}_k$. Убедиться, что их общее решение является мультипликативным:

$$p(\mathbf{n}) = p(\mathbf{0}) \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!}, \quad \rho_k := \frac{\lambda_k}{\mu_k}.$$

- Привести простые примеры, иллюстрирующие связи между четырьмя координатно-выпуклыми стратегиями.
- Доказать, что для отдельного звена сети с многоадресными и одноадресными соединениями $F_m = \rho_m H_m \quad \forall m = \overline{1, M}$.
- Доказать, что для отдельного звена сети с многоадресными и одноадресными соединениями $F_m = \frac{\rho_m}{1 + \rho_m} (1 - B_m) \quad \forall m = \overline{1, M}$.
- Показать, что для модели отдельного звена сети с многоадресными и одноадресными соединениями среднее число приборов, занятых I-заявками, определяется по формуле $b^{(1)} = \sum_{m=1}^M b_m F_m$.

- Доказать, что для модели отдельного звена сети с многоадресными и одноадресными соединениями среднее число приборов, занятых I-заявками, имеет вид $b^{(1)} = \sum_{m=1}^M b_m \frac{\rho_m}{1 + \rho_m} (1 - B_m^I)$.

Литература

Обязательная литература

- [1] Башарин Г.П. Лекции по математической теории телетрафика. – М.: Изд-во РУДН, 2007. – 268 с.
- [2] Корнышев Ю.Н., Пшеничников А.П., Харкевич А.Д. Теория телетрафика. Учебник для вузов. – М.: Радио и связь, 1996. – 272 с.
- [3] Крылов В.В., Самохвалова С.С. Теория телетрафика и ее приложения. – СПб.: «БХВ-Петербург», 2005. – 288 с.
- [4] Кучерявый А.Е., Цуприков А.Л. Сети связи следующего поколения. М.: ФГУП ЦНИИС, 2006. – 280 с.
- [5] Лагутин В.С., Степанов С.Н. Телетрафик мультисервисных сетей связи. – М.: Радио и связь, 2000. – 320 с.
- [6] Соколов Н.А. Телекоммуникационные сети. Монография в 4-х главах. – М.: Альварес Паблишинг, 2004.
- [7] Шнепс-Шнеппе М.А. Системы распределения информации. Методы расчета. Справочное пособие. – М.: Связь, 1979. – 344 с.
- [8] Iversen V. B. Teletraffic engineering and network planning. – ITU-D, May 2006. – 590 p.
<http://oldwww.com.dtu.dk/education/34340/telenook.pdf>
- [9] Kelly F.P. Reversibility and stochastic network. – Chichester: Wiley, 1979. – 630 p.
- [10] Ross K.W. Multiservice loss models for broadband telecommunication networks. – London: Springer-Verlag, 1995. – 343 p.

Дополнительная литература и источники Интернет

- [11] Башарин Г.П., Бочаров П.П., Коган Я.А. Анализ очередей в вычислительных сетях. Теория и методы расчета. // М.: Наука, Гл. ред. физ.-мат. лит., 1989. – 336 с.
- [12] Башарин Г.П., Гайдамака Ю.В., Самуйлов К.Е. Яркина Н.В. Модели для анализа качества обслуживания в сетях связи следующего поколения. Уч. пособие для бакалавров. – М.: Изд-во РУДН, 2008. – 111 с.
- [13] Бабков В.Ю., Полынцев П.В., Устюжанин В.И. Качество услуг мобильной связи. Оценка, контроль и управление. – Горячая линия – Телеком, 2005. – 160 с.

- [14] Деарт В.Ю. Мультисервисные сети связи. - М.: Инсвязьиздат, 2007. – 166 с.: ил.
- [15] Маковеева М.М., Шинаков Ю.Г. Системы связи с подвижными объектами. Уч. пособие для вузов. – М.: Радио и связь, 2002. – 440 с.
- [16] Наумов В.А., Самуйлов К.Е., Яркина Н.В. Теория телетрафика мультисервисных сетей. – М.: Изд-во РУДН, 2007. – 192 с.
- [17] Таха Х.А. Введение в исследование операций. — М.: «Вильямс», 2007. – 912 с.
- [18] Телекоммуникационные системы и сети: Уч. пособие. В 3 томах. Том 3. – Мультисервисные сети / Величко В.В. и др. /под ред. проф. Шувалова В.П. – М.: Горячая линия-Телеком, 2005. – 592 с.
- [19] Handbook of Wireless Networks and Mobile Computing. Ed. Ivan Stojmenovic. – New York: J. Wiley & Sons, 2002. – 630 p.
- [20] <http://www.itu.int>
- [21] <http://www.minsvyaz.ru>

Аннотированное содержание курса.

Первый модуль трудоемкостью 2 кредита составляют:

- теоретический материал, излагаемый в лекциях 1 – 10 календарного плана курса;
- содержание семинарских занятий в течение 20 академических часов.

В конце модуля проводится промежуточный контроль знаний.

Второй модуль трудоемкостью 2 кредита составляют:

- теоретический материал, излагаемый в лекциях 11 – 18 календарного плана курса;
- содержание семинарских занятий в течение 16 академических часов.

В конце модуля проводится итоговый контроль знаний.

Календарный план курса

Виды и содержание учебных занятий				
Неде- ля	Лекции	Число часов	Лабораторные занятия	Число часов
1	Концепция качества обслуживания в сетях связи следующего поколения (NGN – Next Generation Networks). Архитектура, протоколы и услуги сетей связи следующего поколения.	2	Основные протоколы сетей связи следующего поколения: H.323, SIP.	2
2	Понятие качества в NGN. Параметры оценки качества и требования международных стандартов.	2	Основные протоколы сетей связи следующего поколения: MGCP, MEGACO/H.248.	2
3	Принципы построения моделей для анализа качества.	2	Основные протоколы сетей связи следующего поколения: BICC, Sigtran.	2
4	Классические мультисервисные модели: модель Эрланга с явными потерями.	2	Эволюция мобильных сетей. Базовые характеристики сетей разных поколений. Сети 3G и 4G.	2

Виды и содержание учебных занятий				
Неде- ля	Лекции	Число часов	Лабораторные занятия	Число часов
5	Общий подход к построению моделей телекоммуникационных систем сложной структуры в виде СМО $\langle S, A \rangle$.	2	Вероятность потерь и другие макро-характеристики СМО $\langle S^2, A_5 \rangle$.	2
6	Математическая модель буферизации в узле коммутации пакетов в виде СМО $\langle S^1, A_u \rangle$, $u = \overline{1,5}$.	2	Пять способов организации доступа к спутниковым каналам.	2
7	Математическая модель фрагмента системы спутниковой связи в виде СМО $\langle S^2, A_5 \rangle$. Основные параметры. Равновесное распределение для СМО $\langle S^2, A_5 \rangle$.	2	Обратимые марковские процессы.	2
8	Модель кластера двухуровневой ССПС: условия мультипликативности.	2	Организация каналов связи. Соты, кластеры и эстафетная передача обслуживания (хэндовер).	2
9	Модель микросоты с двумя типами каналов и учетом мобильности абонентов:	2	Анализ основных ВВХ модели кластера двухуровневой ССПС.	2

Виды и содержание учебных занятий				
Неде- ля	Лекции	Число часов	Лабораторные занятия	Число часов
	построение пространства состояний, вывод СУР.			
10	Анализ модели микросоты с помощью квазиэрланговской СМО с одномерным распределением числа занятых каналов.	2	Рекуррентный алгоритм расчета вероятностей блокировок. Оптимизация числа каналов.	2
11	Промежуточный контроль знаний (Контрольная работа № 1)			2
12	Стратегии доступа для мультисервисных СМО: основные определения. Стратегия резервирования каналов.	2	Четыре основные координатно-выпуклые стратегии и связь между ними.	2
13	Координатно-выпуклые стратегии. Системы уравнений глобального (СУГБ) и частичного (СУЧБ) балансов. Основные типы координатно-выпуклых стратегий.	2	Примеры оптимизации для мультисервисных СМО.	2
14	Об оптимизации стратегии доступа.	2	Алгоритмы RPB и TRPB.	2

Виды и содержание учебных занятий				
Неде- ля	Лекции	Число часов	Лабораторные занятия	Число часов
15	Сети с одноадресными и многоадресными соединениями. Понятие мультивещания. Дерево мультивещания.	2	Алгоритмы RPF, CBT.	2
16	Модель сети мультивещания с одноадресными и многоадресными соединениями: пространство состояний, равновесное распределение вероятностей.	2	Протоколы IGMP, BGP.	2
17	Модель полнодоступного звена сети мультивещания: пространство состояний и равновесное распределение, вероятностные характеристики.	2	Протоколы CBT, PIM-SM, PIM-DM.	2
18	Модель звена сети мультивещания с резервированием: пространство состояний и равновесное распределение, вероятностные характеристики.	2	Метод свёртки для расчета вероятностных характеристик звена сети мультивещания.	2

Виды и содержание учебных занятий				
Неде- ля	Лекции	Число часов	Лабораторные занятия	Число часов
19	Консультации по подготовке к итоговому контролю знаний.	2	Консультации по подготовке к итоговому контролю знаний.	2
20	Итоговый контроль знаний (Контрольная работа № 2)			2

4. Описание системы контроля знаний

Шкала балльно-рейтинговой системы.

Баллы за семестр	Баллы за итоговый контроль знаний	Общая сумма баллов	Итоговая оценка
61 – 80	Автоматическая оценка.		
	Если студент решил не проходить итоговый контроль знаний, то	86 – 100	5
	начисляются дополнительные баллы - по 1 баллу за каждый балл свыше 60.	70 – 84	4
		62 – 68	3
31 – 80	0 – 20	86 – 100	5
		69 – 85	4
		51 – 68	3
		31 – 50	2
0 – 30	Нет	0 – 30	2

Соответствие систем оценок (используемых ранее оценок итоговой академической успеваемости, оценок ECTS и балльно-рейтинговой системы (БРС) оценок текущей успеваемости)

Баллы БРС	Традиционные оценки в РФ	Баллы для перевода оценок	Оценки	Оценки ECTS
86 - 100	5	95 - 100	5+	A
		86 - 94	5	B
69 - 85	4	69 - 85	4	C
51 - 68	3	61 - 68	3+	D
		51 - 60	3	E
0 - 50	2	31 - 50	2+	FX
		0 - 30	2	F
51 – 100	Зачет		Зачет	Passed

Порядок начисления баллов.

1. Порядок начисления баллов за семестр.

Общая оценка работы в семестре. Посещаемость занятий, активность работы на семинарских занятиях: 0 – 10 баллов

Промежуточный контроль знаний: 0 – 30 баллов

Контрольная работа № 1.

Вопрос 1: 0 – 15 баллов

Вопрос: 2 0 – 15 баллов

Работа над рефератом: 0 – 40 баллов

Содержание реферата: 0 – 30 баллов.

Качество презентации реферата: 0 – 10 баллов.

2. Порядок начисления баллов за итоговый контроль знаний.

Контрольная работа № 2: 0 – 20 баллов

Вопрос 1: 0 – 10 баллов

Вопрос: 2 0 – 10 баллов

Пример применения методики оценки знаний

Студент посетил 80% занятий, однако на семинарских занятиях не принимал участие в обсуждениях.

Набранные баллы: 6 баллов.

Контрольная работа № 1: студент письменно отвечал на следующие вопросы:

Вопрос 1. Выпишите перечень предположений для мультисервисной модели Эрланга.

Ответ неполный: указаны три из пяти предположений.

Набранные баллы: 9 баллов.

Вопрос 2. На примере СМО $\langle S^1, A_2 \rangle$ постройте алгоритм свертки для расчета вероятностно-временных характеристик системы

Ответ полный, содержит описание системы, пространства состояний модели, однако, при выводе формулы ненормированной вероятности $g(m, n)$ допущена ошибка.

Набранные баллы: 5 баллов.

Работа над рефератом.

Тема. Многоадресная доставка в сетях сотовой подвижной связи. Протоколы MoM (Mobile Multicast) и MMA (Multicast by Multicast Agent).

Реферат содержит обзор концепции сетей с многоадресной доставкой информации. Подробно описаны принципы организации сотовой сети подвижной связи, пояснены особенности протоколов МоМ и ММА, проведен сравнительный анализ этих протоколов.

Недостатком является небольшое число (три) источников, использованных при написании реферата. Объем реферата составил 20 страниц, реферат содержит иллюстрации и схемы, оформлен в соответствии с требованиями к написанию учебно-научных материалов. При написании реферата студент активно использовал возможности виртуального кабинета преподавателя, задавал вопросы, выкладывал промежуточные версии реферата.

Набранные баллы: 25 баллов.

Студент подготовил в электронном виде презентацию по содержанию реферата, сделал 10-минутный доклад, четко отвечал на вопросы преподавателя и других слушателей.

Набранные баллы: 10 баллов.

Таким образом, в течение семестра студент набрал следующие баллы.

Посещаемость занятий и активность: 6 баллов

Промежуточный контроль знаний № 1: $9 + 5 = 14$ баллов

Работа над рефератом: $25 + 10 = 35$ баллов

Баллы за семестр $N =$: 55 баллов

Для оценки работы в семестре применяется вторая строка шкалы балльно-рейтинговой системы, поскольку $31 \leq N \leq 80$.

Студент не имеет права получить автоматическую итоговую оценку и обязан проходить итоговый контроль знаний.

1. Начисление баллов за итоговый контроль знаний.

На контрольной работе № 2 (итоговый контроль знаний) студент письменно отвечал на следующие вопросы:

Вопрос 1. Назовите стратегии управления доступом для СМО

$\lambda(\mathbf{n}), \mathbf{b} \mid \mathbf{M} \mid \mu(\mathbf{n}) \mid C \mid 0$. Дайте характеристику четырех

координатно-выпуклых стратегий, опишите связи между ними.

Ответ на вопрос полный, перечислены и правильно описаны все стратегии доступа, показано, при каких значениях параметров, различные стратегии совпадают.

Набранные баллы: 10 баллов.

Вопрос 2. Получите выражения для вероятностных характеристик модель отдельного звена: вероятности блокировки одноадресного соединения в МСС с многоадресными и одноадресными соединениями, вероятности блокировки логического пути, вероятностей B_m^I, B_k^II, F_m^I и H_m^I .

Ответ на вопрос полный, без замечаний.

Набранные баллы: 10 баллов.

Таким образом, в течение семестра и за итоговый контроль знаний студент набрал следующие баллы.

Баллы за семестр $N =$ 55 баллов

Итоговый контроль знаний $M =$ 10 + 10 = 20 баллов

Общая сумма баллов $N + M =$ 55 + 20 = 75 баллов

Итоговая оценка по 5-балльной шкале: 4 (хорошо).

Итоговая оценка шкале ECTS: C.

Академическая этика, соблюдение авторских прав.

Все имеющиеся в тексте всех компонент УМК ссылки на литературные источники и источники Интернет являются актуальными, тщательно выверены и снабжены «адресами». Не включены в тексты выдержки из работ других авторов без ссылки на соответствующий источник, не

пересказаны работы других авторов близко к их тексту и без ссылки на соответствующий источник. В УМК не использованы чужие идеи без указания первоисточников. Это распространяется на литературные источники (монографии, учебники, статьи и пр.) и источники Интернет, для которых необходимых случаях указан полный адрес соответствующего сайта.