

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

---

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

**МОДЕЛИ ДЛЯ АНАЛИЗА  
КАЧЕСТВА ОБСЛУЖИВАНИЯ  
В СЕТЯХ 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 технических наук, профессор *С.Н. Степанов*

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

Модели для анализа качества обслуживания в сетях связи следующего поколения: Учеб. пособие. – М.: РУДН, 2008. – 137 с.: ил.

Излагается подход к анализу показателей качества обслуживания в сетях следующего поколения посредством классических моделей теории телетрафика. Важной частью учебного пособия является компактный и достаточно полный обзор современной концепции качества обслуживания в сетях следующего поколения, который вынесен в Приложение.

Учебное пособие предназначено для студентов бакалавриата, обучающихся по направлениям 010300 «Математика. Компьютерные науки», 010400 «Информационные технологии» или 010500 «Прикладная математика и информатика». Одноименный курс входит в состав модуля «Управление инфокоммуникациями» профиля специализации в бакалавриате и является дисциплиной по выбору. Студенты, выбравшие данный профиль, должны также прослушать следующие дисциплины: «Основы формальных методов описания бизнес-процессов», «Основы разработки корпоративных инфокоммуникационных систем», «Основы управления инфокоммуникационными компаниями».

*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
СПИСОК ОСНОВНЫХ СОКРАЩЕНИЙ.....	7
Глава 1. КЛАССИЧЕСКИЕ МОДЕЛИ ЗВЕНА СЕТИ С ОДНОАДРЕСНЫМИ СОЕДИНЕНИЯМИ .....	9
§1.1. Первая модель Эрланга .....	9
§1.2. Нагрузка и ее характеристики .....	18
§1.3. Модель Эрланга с ожиданием и блокировками .....	25
§1.4. Модель Энгсета.....	34
§1.5. Равновесное распределение и система уравнений равновесия (приложение к главе 1).....	46
Глава 2. ПРИМЕНЕНИЕ КЛАССИЧЕСКИХ МОДЕЛЕЙ К АНАЛИЗУ ОДНОЙ СОТЫ ССПС .....	48
§2.1. Физическая модель процесса обслуживания в соте.....	48
§2.2. Полнодоступная модель с потерями.....	50
§2.3. Неполнодоступная модель с потерями.....	52
Глава 3. МОДЕЛЬ ЗВЕНА СЕТИ С МНОГОАДРЕСНЫМИ СОЕДИНЕНИЯМИ .....	57
§ 3.1. Модель сети с многоадресными соединениями.....	57
§ 3.2. Модель отдельного звена .....	71
Приложение. КОНЦЕПЦИЯ КАЧЕСТВА ОБСЛУЖИВАНИЯ В СЕТЯХ NGN.....	81
§П.1. Современная концепция качества обслуживания в сетях связи ..	83
§П.2. Характеристики качества функционирования сети .....	91
§П.3. Требования к качеству обслуживания типовых услуг NGN .....	94
РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА.....	103
СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ.....	105
ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ .....	109
ОПИСАНИЕ КУРСА И ПРОГРАММА.....	112

## **ВВЕДЕНИЕ**

Современная концепция «сети связи следующего поколения» (Next Generation Network, NGN) отражает конвергенцию информационно-телекоммуникационных сетей в единую глобальную сеть. Движение в этом направлении только началось, на сегодняшний день международные стандарты содержат, главным образом, не конкретные решения, а лишь требования к сетям следующего поколения, из которых вытекает множество задач для исследования и изучения. Одним из важнейших направлений исследований является анализ качества обслуживания в сетях следующего поколения. Усложнение конфигураций сетей по сравнению с телефонными сетями общего пользования, а также появление новых видов обслуживания, предусматривающих возможность выбора пользователем услуги с заранее заданным уровнем качества, привело к необходимости построения новых моделей для анализа качества обслуживания в сети. В рамках курса слушатели получают необходимый объем знаний для проведения исследований в этой новой области, которую сейчас также называют теорией телетрафика мультисервисных сетей.

В учебном пособии излагается подход к анализу показателей качества обслуживания в сетях следующего поколения посредством классических моделей теории телетрафика, которые относятся к классу задач управления доступом в мультисервисных сетях на уровне соединений. Важной частью учебного пособия является Приложение, в котором содержится компактный и достаточно полный обзор современной концепции качества обслуживания в сетях следующего поколения.

Учебное пособие предназначено для студентов бакалавриата, обучающихся по направлениям 010300 «Математика. Компьютерные науки», 010400 «Информационные технологии» или 010500 «Прикладная математика и информатика». Одноименный курс входит в состав модуля «Управление инфокоммуникациями» профиля специализации в

бакалавриате и является дисциплиной по выбору студента. Студенты, выбравшие данный профиль, должны также прослушать следующие дисциплины: «Основы формальных методов описания бизнес-процессов»; «Основы разработки корпоративных инфокоммуникационных систем»; «Основы управления инфокоммуникационными компаниями».

Целью курса является ознакомление слушателей с принципами функционирования сетей связи следующего поколения, изучение вопросов качества обслуживания в этих сетях, освоение методов построения вероятностных моделей для анализа качества обслуживания в терминах теории массового обслуживания и теории марковских процессов, изучение методов анализа и расчета показателей качества обслуживания. После успешного прохождения курса слушатели должны знать: основные понятия и определения, относящиеся к концепции сетей связи следующего поколения; принципы построения сетей связи следующего поколения; требования международных стандартов к показателям качества обслуживания – QoS-параметрам. Слушатели должны уметь с помощью аппарата теории случайных процессов, теории массового обслуживания и теории телетрафика строить простые модели узлов и звеньев NGN, для построенных моделей составлять и решать системы уравнений равновесия (СУР), получать вероятностные характеристики моделей, связанные с показателями качества обслуживания; применять численные методы при анализе полученных характеристик моделей.

Лица, успешно окончившие бакалавриат и прослушавшие перечисленные курсы, могут продолжить обучение в магистратуре по направлению 010400 «Информационные технологии» по специализации «Управление инфокоммуникациями».

Учебное пособие включает три главы и Приложение. Глава 1 посвящена классическим моносервисным моделям Эрланга и Энгсета и служит введением к последующим главам. В главе 2 рассмотрен пример

применения классических моносервисных моделей к анализу модели сети сотовой подвижной связи, а глава 3 посвящена анализу модели звена мультисервисной сети связи с многоадресными соединениями. В Приложении к учебному пособию содержится обзор современной концепции качества обслуживания в перспективных сетях, определены основные показатели качества и требования к качеству обслуживания типовых услуг NGN.

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, определений и примеров сквозная в каждой главе.

# СПИСОК ОСНОВНЫХ СОКРАЩЕНИЙ

## Русскоязычные сокращения

МВОС	–	модель взаимодействия открытых систем
МСС	–	мультисервисная сеть связи
МСЭ	–	Международный союз электросвязи
ОМП	–	обратимый марковский процесс
ПНН	–	период времени наибольшей нагрузки
ПП	–	пуассоновский поток
ПРГ	–	процесс размножения и гибели
СВ	–	случайная величина
СДОП	–	сеть передачи данных общего пользования
СЛАУ	–	система линейных алгебраических уравнений
СМО	–	система массового обслуживания
ССПС	–	сотовая сеть подвижной связи
СтМП	–	ступенчатый марковский процесс
СУГБ	–	система уравнений глобального баланса
СУЛБ	–	система уравнений локального баланса
СУР	–	система уравнений равновесия
ТМО	–	теория массового обслуживания
ТТ	–	теория телетрафика
ТфОП	–	телефонная сеть общего пользования
ЧНН	–	часы наибольшей нагрузки
ШПП	–	ширина полосы пропускания

## **Англоязычные сокращения**

3GPP	–	3rd Generation Partnership Project
ATM	–	Asynchronous Transfer Mode
DiffServ	–	Differentiated Services
ETSI	–	European Telecommunications Standards Institute
FCFS	–	first come – first served
HDTV	–	High Definition Television
IMS	–	Internet Protocol-based Multimedia Subsystem
IP	–	Internet Protocol
IPDV	–	IP packet delay variation
IPER	–	IP packet error ratio
IPLR	–	IP packet loss ratio
IPTD	–	IP packet transfer delay
IPTV	–	Internet Protocol Television
LCFS	–	last come – first served
MOS	–	Mean Opinion Score
MP	–	Measurement point
MPLS	–	Multiprotocol Label Switching
NGN	–	Next Generation Network
NP	–	Network Performance
QoE	–	Quality of Experience
QoS	–	Quality of Service
SIP	–	Session Initiation Protocol
SNMP	–	Simple Network Management Protocol
TIPHON	–	Telecommunications and Internet Protocol Harmonization Over Networks
TISPAN	–	Telecoms & Internet converged Services & Protocols for Advanced Networks



# Глава 1. КЛАССИЧЕСКИЕ МОДЕЛИ ЗВЕНА СЕТИ С ОДНОАДРЕСНЫМИ СОЕДИНЕНИЯМИ

## §1.1. Первая модель Эрланга

### 1.1.1. Постановка задачи

Рассмотрим многолинейный пучок из  $\nu$  полнодоступных идентичных линий (обслуживающих устройств, каналов и т. п.), работающих параллельно и не имеющих мест для ожидания. Примем, что на этот пучок поступает пуассоновский поток (ПП) однотипных заявок с постоянной интенсивностью  $\lambda$  заявок в единицу времени. Пусть  $E_n = \{n\}$  – состояние рассматриваемой системы массового обслуживания (СМО), когда занято  $n$  линий,  $n \in J = \{0, 1, \dots, \nu\}$ .

Поскольку мест для ожидания нет, то в состоянии  $E_\nu$  поступающие заявки блокируются, т. е. теряются и не оказывают влияния ни на характер поступающего потока, ни на его интенсивность. С точки зрения приложений это упрощающее предположение означает, что либо вероятность  $p_\nu$  потерь достаточно мала (например, порядка нескольких процентов), либо что потерянная заявка направляется на другую СМО, так что влиянием повторных попыток установить соединение можно пренебречь. При этом порядок занятия свободных линий пучка может быть любым – *случайным* или *упорядоченным*.

Каждая принятая заявка занимает одну линию на все время обслуживания, которое имеет экспоненциальное распределение с параметром, не зависящим ни от поступающего потока, ни от состояния СМО, ни от длительности обслуживания других заявок. Такую СМО часто называют *моделью Эрланга с явными потерями* или *первой моделью Эрланга* в честь датского математика и инженера А. К. Эрланга (1878–

1929), работы которого в 1908–1918 гг. положили начало теории телетрафика (ТТ) и теории массового обслуживания (ТМО).

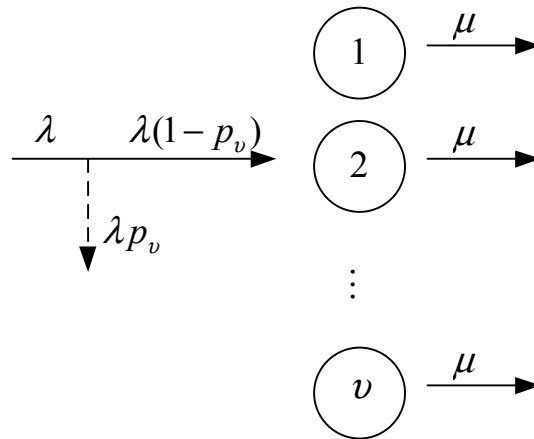


Рис. 1.1. СМО  $M \left| M \right| v$ , где  $p_v$  – вероятность блокировки

Напомним теперь идею обозначений Кендалла, согласно которой СМО простой структуры задается посредством следующего описания:



Символ  $M$  на первом месте означает пуассоновский поток, а на втором – экспоненциальное обслуживание. Рассматриваемую СМО можно

кодировать как  $M \left| M \right| v \mid r=0$  или как  $M \mid M \mid v$ ,  $1 \leq v \leq \infty$ , опуская

четвертую позицию при  $r=0$ . Значения интенсивностей  $\lambda$   $\frac{\text{заявок}}{\text{ед.вр.}}$  и

$\mu$   $\frac{\text{заявок}}{\text{ед.вр.}}$  указывать не обязательно (рис. 1.1), причем физический смысл

абстрактного термина «заявка» может интерпретироваться по-разному в каждом конкретном случае: вызов, сообщение, пакет, слот и т. д. Каждая

заявка на все время обслуживания в моносервисной СМО занимает одно из  $v$  идентичных обслуживающих устройств, которые можно называть также

приборами, линиями, каналами и т. д.

Для пуассоновского потока среднее число заявок, поступивших за интервал длины  $T$ , равно  $\lambda T$ , а средняя длительность экспоненциального обслуживания составляет  $1/\mu$ . Поэтому за время  $T=1/\mu$  поступит в среднем  $\rho := \lambda/\mu$  заявок. Величина  $\rho$  не зависит от выбранной единицы времени и называется *интенсивностью поступающей нагрузки*. На практике  $\rho$  измеряется в *эрлангах* (Эрл), причем, например,  $\rho = 3.5$  Эрл означает, что либо среднее значение числа занятых в СМО линий в некоторый фиксированный или произвольный момент времени равно 3.5, либо в среднем за рассматриваемый интервал занято 3.5 линии. Более подробно важный для теории и приложений параметр  $\rho$  рассматривается в §1.2.

### ***1.1.2. Построение процесса размножения и гибели***

Напомним некоторые важные для дальнейшего изложения факты из курса теории вероятностей и ТМО. Пусть  $X(t)$  – число заявок, обслуживаемых  $\nu$ -линейным пучком в момент  $t$ ,  $t \geq 0$ . Так как принятая заявка на все время обслуживания занимает одну линию, то число обслуживаемых заявок совпадает с числом занятых линий, т. е. *пространством состояний* для случайного процесса  $X(t)$ ,  $t \geq 0$ , служит множество  $J = \{0, 1, \dots, \nu\}$ ,  $1 \leq \nu \leq \infty$ .

Поскольку по предположению поступающий поток является пуассоновским с параметром  $\lambda$ , то при любых  $t_0 \geq 0$ ,  $\tau \geq 0$  число  $Y(\tau)$  заявок, предложенных СМО на интервале  $(t_0, t_0 + \tau)$ , обладает пуассоновским распределением с параметром  $\lambda\tau$ :

$$P\{Y(\tau) = m\} = e^{-\lambda\tau} \frac{(\lambda\tau)^m}{m!}, \quad m = 0, 1, \dots \quad (1.1)$$

Это распределение не зависит ни от того, будут ли поступившие заявки приняты на обслуживание или получают отказ, ни от того, как долго будут обслуживаться принятые заявки. В частности, из (1.1) следует, что

$$P\{\text{нет поступлений заявок на } (t_0, t_0 + \tau)\} = P\{Y(\tau) = 0\} = e^{-\lambda\tau}, \quad (1.2)$$

т. е. расстояние между поступающими заявками имеет экспоненциальное распределение и не зависит ни от  $t_0$ , ни от состояния  $X(t_0)$ , ни от возможных уходов заявок из СМО на интервале  $(t_0, t_0 + \tau)$ .

Далее, пусть  $X(t_0) = n$ ,  $n = \overline{1, \nu}$ , а  $z_1, \dots, z_n$  – остаточные длительности обслуживания этих заявок. Тогда в силу отсутствия последействия у экспоненциального распределения длительности обслуживания заявок и их независимости как от потока, так и друг от друга, получим:

$$\begin{aligned} P\{\text{ни одна заявка не покинет СМО на } (t_0, t_0 + \tau)\} &= P\{\min(z_1, \dots, z_n) > \tau\} = \\ &= P\{z_1 > \tau, \dots, z_n > \tau\} = P\{z_1 > \tau\} \cdots P\{z_n > \tau\} = e^{-n\mu\tau}. \end{aligned} \quad (1.3)$$

Используя эти факты, легко убедиться в том, что случайный процесс  $X(t)$ ,  $t \geq 0$ , является *однородным марковским процессом*. Действительно, вероятностный прогноз будущего состояния  $X(t_0 + \tau)$  при известном прошлом  $X(s)$ ,  $s < t_0$ , зависит лишь от настоящего состояния  $X(t_0)$  и от длины  $\tau$  интервала прогнозирования, но не зависит от  $t_0$ , т. е. от расположения этого интервала (см. приложение А в [1]).

При этом  $X(t)$ ,  $t \geq 0$ , – ступенчатый марковский процесс (СтМП) типа процесса размножения и гибели (ПРГ), так как за малый интервал времени  $(t_0, t_0 + \Delta)$  с точностью до  $o(\Delta)$  возможно только сохранение «статус-кво» или переход в соседнее «снизу» или «сверху» состояние (см. приложение Б в [1]). Обозначая функцию Хевисайда

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

получим следующие вероятности переходов за интервал длины  $\Delta$ :

$$\begin{aligned}
p_{n,n}(\Delta) &:= P\{X(t_0 + \Delta) = n \mid X(t_0) = n\} = \\
&= P\{\text{за } (t_0, t_0 + \Delta) \text{ не поступила ни одна заявка} \\
&\text{и не освободилась ни одна линия} \mid X(t_0) = n\} = e^{-\lambda u(v-n)\Delta} e^{-n\mu\Delta} = \\
&= [1 - \lambda u(v-n)\Delta + o(\Delta)] \times [1 - n\mu\Delta + o(\Delta)] = \\
&= 1 - \lambda u(v-n)\Delta - n\mu\Delta + o(\Delta), \quad n = \overline{0, v}; \tag{1.4a}
\end{aligned}$$

$$\begin{aligned}
p_{n,n-1}(\Delta) &:= P\{X(t_0 + \Delta) = n - 1 \mid X(t_0) = n\} = \\
&= P\{\text{за } (t_0, t_0 + \Delta) \text{ не поступила ни одна заявка} \\
&\text{и освободилась одна из } n \text{ занятых линий} \mid X(t_0) = n\} = \\
&= e^{-\lambda\Delta} \binom{n}{1} (1 - e^{-\mu\Delta}) e^{-(n-1)\mu\Delta} = n\mu\Delta + o(\Delta), \quad n = \overline{1, v}; \tag{1.4б}
\end{aligned}$$

$$\begin{aligned}
p_{n,n+1}(\Delta) &:= P\{X(t_0 + \Delta) = n + 1 \mid X(t_0) = n\} = \\
&= P\{\text{за } (t_0, t_0 + \Delta) \text{ поступила одна заявка} \\
&\text{и не освободилась ни одна линия} \mid X(t_0) = n\} = \\
&= (1 - e^{-\lambda\Delta}) e^{-n\mu\Delta} = \lambda\Delta + o(\Delta), \quad n = \overline{0, v-1}; \tag{1.4в}
\end{aligned}$$

$$p_{n,m}(\Delta) = 0, \quad |m - n| \geq 2, \quad n, m = \overline{0, v}. \tag{1.4г}$$

Из формул (1.4) следует, что при  $n = \overline{0, v}$  интенсивности потоков занятия и освобождения линий имеют следующий вид:

$$\begin{aligned}
a_{n,n-1} &= \mu_n = n\mu; \quad a_{n,n+1} = \lambda_n = \lambda u(v-n); \\
a_n &:= -a_{n,n} = \lambda u(v-n) + n\mu; \tag{1.5}
\end{aligned}$$

$$a_{n,m} = 0, \quad |m - n| \geq 2,$$

т.е.  $X(t)$ ,  $t \geq 0$ , является ПРГ (рис. 1.2).

В силу формул (1.2), (1.3) и (1.5) длительность  $\tau_n$  «сидения» в состоянии  $n$  имеет экспоненциальное распределение со средним значением  $1/a_n$ :

$$P\{\tau_n > t\} = e^{-(\lambda u(v-n) + n\mu)t}, \quad t \geq 0, \quad n = \overline{0, v}.$$

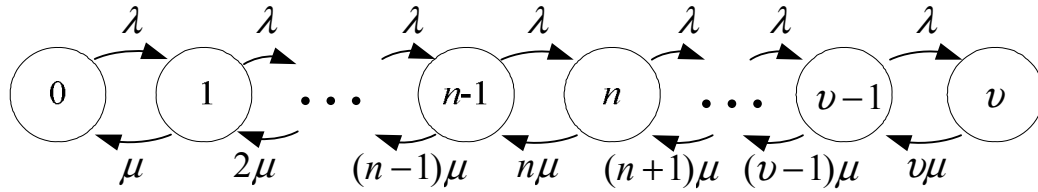


Рис. 1.2. Диаграмма интенсивностей переходов для СМО  $M \left| \begin{array}{c} M \\ \lambda \\ \mu \end{array} \right| v$

Следовательно, моментами выхода из состояния  $n$  управляет пуассоновский поток с интенсивностью  $a_n$ , а самими переходами в соседние снизу или сверху состояния управляют соответственно переходные вероятности

$$q_{n,n-1} = \frac{n\mu}{\lambda + n\mu}, \quad n = \overline{1, v}; \quad q_{n,n+1} = \frac{\lambda}{\lambda + n\mu}, \quad n = \overline{0, v-1}; \quad (1.6a)$$

причем

$$q_{n,n} = q_{n,m} = 0, \quad |m - n| \geq 2, \quad n, m = \overline{0, v}. \quad (1.6b)$$

Формулы (1.6) физически очевидны и поясняют смысл условия  $0 < \lambda, \mu < \infty$ .

Таким образом, мы показали, что функционирование полностью доступного пучка из  $v$  линий с явными потерями в стационарном режиме можно описать с помощью ПРГ  $X(t)$ ,  $t \geq 0$  с пространством состояний  $J = \{0, 1, \dots, v\}$ , трехдиагональной матрицей интенсивностей переходов  $\mathbf{A}$  (1.5) и двухдиагональной стохастической матрицей  $\mathbf{Q}$  (1.6).

**Упражнение 1.1.** Запишите формулы (1.5) в виде матрицы  $\mathbf{A}$ , а (1.6) – в виде матрицы  $\mathbf{Q}$  (см. таблицы Б1 и Б2 в приложении Б в [1]).

### 1.1.3. Распределение и первая формула Эрланга

Искомое равновесное распределение  $\mathbf{p}^T = (p_0, \dots, p_v)$ , где  $P\{X = n\} =: p_n$ ,  $n = \overline{0, v}$ , для ПРГ с матрицей  $\mathbf{A}$  (1.5) удовлетворяет

системе уравнений глобального баланса (СУГБ)  $\mathbf{p}^T \mathbf{A} = \mathbf{0}^T$ , которую можно записать в виде (см. также приложение Б в [1]):

$$\begin{aligned} -\lambda p_0 + \mu p_1 &= 0, \\ \lambda p_{n-1} - (\lambda + n\mu)p_n + (n+1)\mu p_{n+1} &= 0, \quad n = \overline{1, v-1}, \\ \lambda p_{v-1} - v\mu p_v &= 0. \end{aligned} \quad (1.7)$$

Суммируя теперь уравнения (1.7) от нулевого до  $(n-1)$ -го включительно, получим следующую систему уравнений локального баланса (СУЛБ)

$$\lambda p_{n-1} = n\mu p_n, \quad n = \overline{1, v}. \quad (1.8)$$

Физическую интерпретацию (1.8) можно дать с помощью рис. 1.2: левая часть – это взвешенный поток интенсивностей переходов от состояния  $(n-1)$  к состоянию  $n$ , а правая часть – в обратном направлении.

Обозначая  $\rho = \lambda/\mu$ ,  $0 < \rho < \infty$ , и используя нормировочное условие

$$\sum_{n=0}^v p_n = 1, \text{ получим из (1.8):}$$

$$p_n =: p_n(\rho) = \frac{\rho^n}{\sum_{m=0}^v \frac{\rho^m}{m!}}, \quad n = \overline{0, v}. \quad (1.9)$$

Для простоты вычислений в дальнейшем будем полагать, что

$$p_n(0) = \begin{cases} 1, & n = 0, \\ 0, & n = \overline{1, v} \end{cases} = u(1-n), \quad n = \overline{0, v}.$$

Распределение (1.9) числа занятых линий полнодоступного пучка называется *первым распределением Эрланга*.

Важная для проектирования реальных систем связи величина  $p_v(\rho)$  является *вероятностью блокировок* предложенного потока или *вероятностью потерь по времени*, ее обычно обозначают

$$E_v(\rho) := \frac{\rho^v}{\sum_{m=0}^v \frac{\rho^m}{m!}}, \quad 0 \leq \rho < \infty, \quad (1.10)$$

и называют *первой формулой Эрланга* или *V-формулой Эрланга*. Эта формула и другие характеристики полнодоступного пучка неоднократно табулировалась<sup>1</sup>, а сама модель Эрланга уже почти столетие широко применяется при проектировании телекоммуникационных систем и сетей различных поколений, включая самые современные.

Легко видеть, что справедлива следующая

**Теорема 1.1.** *Если  $v \rightarrow \infty$ , то при  $0 \leq \rho < \infty$*

$$p_n(\rho) \xrightarrow{v \rightarrow \infty} e^{-\rho} \frac{\rho^n}{n!}, \quad n = 0, 1, \dots \quad (1.11)$$

Следовательно, распределение  $p_n(\rho)$ ,  $n = \overline{0, v}$ , (1.9) при  $v < \infty$  является *усеченным распределением Пуассона* с параметром  $\rho$  и, в принципе, может быть получено с помощью таблицы распределения Пуассона. Заметим, что поскольку после Второй мировой войны распределением Эрланга стали называть распределение суммы  $n$  независимых случайных величин с одинаковым экспоненциальным распределением каждая, то в настоящее время (1.9) называют также *усеченным распределением Пуассона*.

Поскольку вероятность блокировок  $E_v(\rho)$  широко используется на практике, а с ростом  $v$  величины  $\rho^v$  при  $\rho > 1$  и  $v!$  очень быстро возрастают, то при вычислении по формуле (1.10) могут произойти как переполнение, так и исчезновение величины  $\rho^v/v!$ . Поэтому вычисление  $E_v(\rho)$  производится с помощью следующей рекуррентной процедуры:

---

<sup>1</sup> Башарин Г.П. Таблицы вероятностей и средних квадратических отклонений потерь на полнодоступном пучке линий. – М.: Изд-во АН СССР, 1962. – 128 с.



$$E_0(\rho) = 1, E_v(\rho) = \frac{\rho E_{v-1}(\rho)}{v + \rho E_{v-1}(\rho)}, v = 1, 2, \dots \quad (1.12)$$

Действительно, поделив числитель и знаменатель дроби в формуле (1.10) на  $\sum_{m=0}^{v-1} \rho^m / m!$ , получаем:

$$E_v(\rho) = \frac{\left[ \frac{\rho^v}{v!} \right]}{\left[ \sum_{m=0}^{v-1} \frac{\rho^m}{m!} \right]} \left[ \frac{\left( \sum_{m=0}^{v-1} \frac{\rho^m}{m!} + \frac{\rho^v}{v!} \right)}{\sum_{m=0}^{v-1} \frac{\rho^m}{m!}} \right]^{-1} = \frac{\frac{\rho}{v} E_{v-1}(\rho)}{1 + \frac{\rho}{v} E_{v-1}(\rho)}.$$

Обозначив  $I_v(\rho) = [E_v(\rho)]^{-1}$ , получим из (1.12) следующую рекуррентную формулу:

$$I_0(\rho) = 1, I_v(\rho) = 1 + \frac{v}{\rho} I_{v-1}(\rho), v = 1, 2, \dots \quad (1.13)$$

Обе рекуррентные процедуры дают устойчивый результат, но каждая из них имеет свою предпочтительную область изменения параметров  $\rho$  и  $v$ .

**Упражнение 1.2.** К АТС подключено  $10^4$  абонентов, каждый из которых создает нагрузку в 0,04 Эрл, причем 10% вызовов являются междугородными. Требуется найти количество  $v$  каналов для междугородных вызовов при условии, что вероятность их блокировки не превысит 1%.

**Решение.** Полагая  $N = 10^4$ , получим для интенсивностей  $\rho_1$  общей и  $\rho_2$  междугородной нагрузки на АТС следующие значения:

$$\rho_1 = 10^4 \cdot 0,04 = 400 \text{ Эрл};$$

$$\rho_2 = 0,1\rho_1 = 40 \text{ Эрл}.$$

Согласно формуле (1.12) и условию задачи

$$\frac{\rho_2 E_{v-1}(\rho_2)}{v + \rho_2 E_{v-1}(\rho_2)} \leq 0,01.$$

Отсюда сразу следует, что  $v \geq 99 \rho_2 E_{v-1}(\rho_2)$ .

Используя таблицы для  $E_v(\rho)$  или рекуррентную процедуру (1.12), получим, что при  $\rho_2 = 40$  условие задачи будет выполнено, если  $v \geq 53$  канала.

**Упражнение 1.3.** Докажите, что  $\frac{\rho}{v} [1 - E_v(\rho)] = \frac{E_v(\rho)}{E_{v-1}(\rho)}$ .

## §1.2. Нагрузка и ее характеристики

### 1.2.1. Определение и виды нагрузки

**Определение 1.1.** *Обслуживаемой (принятой) некоторой СМО в момент времени  $t$  мгновенной нагрузкой* называется случайный процесс  $X(t)$ ,  $t \geq 0$ , представляющий собой число одновременно обслуживаемых в этой системе заявок или число одновременно занятых ресурсов (приборов).

**Определение 1.2.** *Интенсивностью обслуживаемой в момент  $t$  системой  $S$  мгновенной нагрузки* называется среднее число занятых приборов (единиц ресурсов) этой системы, т. е.  $EX(t; S) \geq 0$ ,  $t \geq 0$ .

Интенсивность нагрузки измеряется в эрлангах, причем 1 Эрл соответствует интенсивности нагрузки, при которой за среднюю длительность занятия прибора поступает одна заявка.

Если процесс  $X(t)$  стационарный и эргодический, то можно показать, что

$$EX(t) \xrightarrow{t \rightarrow \infty} EX = \bar{\lambda} h, \quad (2.1)$$

где  $\bar{\lambda}$  – средняя интенсивность принятого к обслуживанию потока,  $h$  – средняя длительность обслуживания. Поэтому  $\bar{\lambda} h$  – средняя

интенсивность нагрузки на  $S$ , которую в приложениях для краткости часто называют просто «нагрузка».

**Примечание.** Если каждая заявка при обслуживании занимает не единицу рассматриваемого ресурса, а  $b \left[ \frac{\text{ед.ресурса}}{\text{заявку}} \right]$ , то это учитывается в начале или после расчетов. В этой главе все заявки принадлежат к одному типу (моносервисные СМО), так что можно считать  $b = 1$ . Если же на СМО поступают потоки заявок  $K$  типов (мультисервисные СМО) с независимыми пуассоновскими распределениями и независимыми экспоненциальными длительностями обслуживания, то наряду с параметрами  $\lambda_1, \dots, \lambda_K, \mu_1, \dots, \mu_K$  приходится вводить и параметры  $b_1, \dots, b_K$ . Для мультисервисной полнодоступной  $v$ -линейной СМО с такой пуассоновско-экспоненциальной нагрузкой и потерями при блокировке удобно использовать обозначение

$$M \left| \begin{array}{c} M \\ \mu \end{array} \right| v \mid 0 . \quad (2.2)$$

Проиллюстрируем определение 1.2 с помощью СМО  $M \left| \begin{array}{c} M \\ \mu \end{array} \right| v < \infty \mid 0$ .

Пусть случайная величина (СВ)  $X$  имеет первое распределение Эрланга (1.9), т. е. распределение СВ  $X$  является равновесным распределением для соответствующего СтМП и описывает функционирование этой СМО в стационарном (равновесном) режиме. Тогда из СУЛБ (1.8) получим

$$EX = \sum_{n=1}^v n p_n(\rho) = \frac{\lambda}{\mu} \sum_{n=1}^v p_{n-1}(\rho) = \rho [1 - E_v(\rho)]. \quad (2.3)$$

В силу определения 1.2 эту величину естественно принять за интенсивность *обслуженной* нагрузки:

$$Y_{\text{обсл}} = \rho [1 - E_v(\rho)]. \quad (2.4)$$

Для  $\nu = \infty$  и любого  $\rho$ ,  $0 < \rho < \infty$ , интенсивность *предложенной* нагрузки равна интенсивности обслуженной:

$$Y_{\text{предл}} = \rho \sum_{n=1}^{\infty} p_{n-1}(\rho) = \sum_{n=1}^{\infty} n p_n(\rho) = \rho = Y_{\text{обсл}}. \quad (2.5)$$

Очевидно, что при  $\nu < \infty$  и отсутствии мест для ожидания часть *предложенной* (поступающей) нагрузки теряется. Пусть  $Y_{\text{пот}}$  – интенсивность потерянной нагрузки. Тогда

$$Y_{\text{предл}} = Y_{\text{обсл}} + Y_{\text{пот}}, \quad (2.6)$$

причем физически очевидно, что это соотношение справедливо при очень широких предположениях.

Для первой модели Эрланга

$$Y_{\text{пот}} = Y_{\text{предл}} E_{\nu}(\rho). \quad (2.7)$$

### 1.2.2. О статистической оценке характеристик нагрузки

На практике обычно измеряется обслуживаемая системой  $S$  с емкостью  $\nu$  нагрузка на интервале  $[t_0, t_0 + T)$ . Пусть  $n(t)$ ,  $t \in [t_0, t_0 + T)$  некоторая реализация ПРГ  $X(t)$ ,  $t \geq 0$ , описывающего *мгновенную обслуживаемую нагрузку* в рассматриваемой системе. Величину

$$Z(t_0, t_0 + T) := \int_{t_0}^{t_0+T} n(t) dt \quad (2.8)$$

будем называть *работой* по обслуживанию заявок, выполняемой системой  $S$  на рассматриваемом интервале. Величина работы равна общей продолжительности занятий линий (каналов, входов, выходов или других обслуживающих устройств) системы  $S$  на  $[t_0, t_0 + T)$  и измеряется в эрланг-часах (Эрл-ч), эрланг-минутах (Эрл-мин), эрланг-секундах (Эрл-с) и т. д., причем в нашей стране раньше эрланг-час назывался «часозанятие». Работу в 1 Эрл-ч выполняет линия, непрерывно занятая в течение 1 часа;

2 Эрл-ч – 2 линии, непрерывно занятые в течение 1 часа, или 1 линия, но занятая непрерывно в течение 2 часов.

**Пример 1.1.** Пусть  $\lambda = 15 \frac{\text{заявок}}{\text{мин.}}$ ,  $h = \frac{1}{\mu} = 3 \frac{\text{мин.}}{\text{заявку}}$ . Тогда интенсивность предложенной нагрузки равна  $Y_{\text{предл}} = \rho = 45$  Эрл. Объем нагрузки, или работы, за 4 часа составляет  $4 \text{ ч} \times 45 \text{ Эрл} = 180$  Эрл-ч, за 30 минут –  $0,5 \text{ ч} \times 45 \text{ Эрл} = 22,5$  Эрл-ч.

**Упражнение 1.4.** Приведите несколько примеров подсчета объема нагрузки, аналогичных примеру 1.1, когда за единицу времени естественно принять минуты, секунды и миллисекунды.

Если работу системы  $S$  на интервале  $[t_0, t_0 + T)$  представить графически как функцию от  $t$ , то получим быстро возрастающую ломаную линию, вершины которой соответствуют моментам изменения мгновенной нагрузки  $n(t)$ .

В силу (2.8) работу, выполняемую системой  $S$  на непересекающихся отрезках времени, можно суммировать, т. е. работа системы  $S$  обладает свойством *аддитивности*.

Средняя величина работы (2.8) в единицу времени

$$Y(t_0, t_0 + T) := \frac{1}{T} \int_{t_0}^{t_0+T} n(t) dt \quad (2.9)$$

служит статистической оценкой для среднего значения интенсивности обслуживаемой нагрузки (2.1), причем

$$EX = \bar{\lambda}h \approx Y(t_0, t_0 + T). \quad (2.10)$$

На рис. 1.3 (см. рис. 2.1 [11]) приводится пример результатов измерения величины  $n(t)$  обслуживаемой нагрузки на пяти последовательных интервалах длиной  $T$  каждый (кривая  $C$ ). Кроме того,

приводятся (кривая  $D$ ) результаты измерения на каждом из этих интервалов величины

$$Y_i = \frac{1}{T} \int_{(i-1)T}^{iT} n(t) dt, \quad i = \overline{1,5}. \quad (2.11)$$

Величина  $Y$  служит оценкой среднего числа занятых линий и, тем самым, оценкой среднего значения интенсивности обслуживаемой нагрузки на интервале  $[(i-1)T, iT)$ . Аналогичным образом можно оценить дисперсию и другие характеристики обслуживаемой нагрузки. Подробнее эти интересные вопросы математической статистики применительно к СМО рассматриваются в книгах [4, 5, 10, 11].

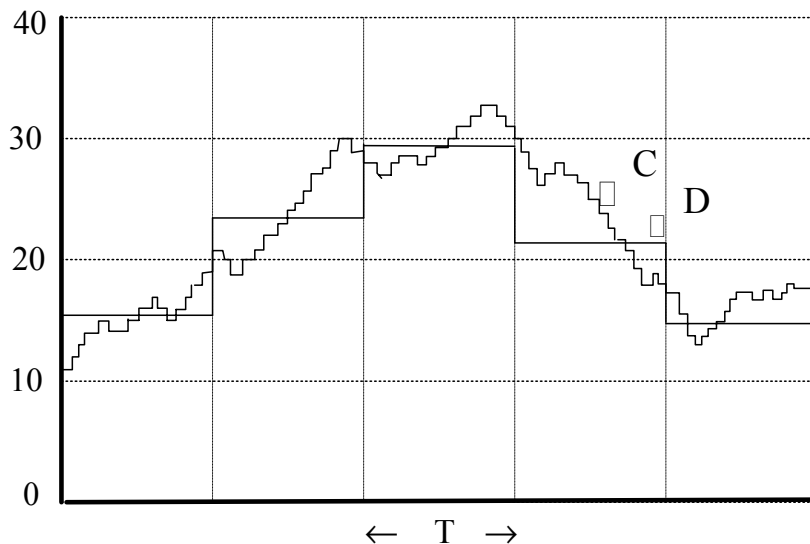


Рис. 1.3. Мгновенная обслуживаемая нагрузка  $n(t)$

На выбор  $T$  влияют многие факторы, включая специфику и параметры самой системы  $S$ , применяемой аппаратуры измерений и требуемую точность статистических оценок. Исторически в телефонии обычно принимают  $T = 1$  час и измеряют обслуживаемую нагрузку в течение суток, в разные дни недели – рабочие, выходные и праздничные – в течение года. На основе этих измерений можно определить часы наибольшей нагрузки (ЧНН) для различных суток в течение года. При проектировании или модернизации современных телекоммуникационных

систем требуемые характеристики нагрузки могут также рассчитываться на основе измерений типа (2.11) с соответствующим  $T$ , после чего определяется период времени наибольшей нагрузки (ПНН) в интересующем операторе временном интервале. Применяется также и экспертная оценка характеристик предполагаемых предложенной и обслуженной нагрузок.

### 1.2.3. О порядке занятия свободных приборов

Рассмотрим СМО  $M \left| \begin{matrix} M \\ \lambda \end{matrix} \right| \begin{matrix} M \\ \mu \end{matrix} \left| v \right| 0$  при случайном и упорядоченном

занятии (искании) свободных приборов и пусть  $Y_{\text{обсл}}^{\text{сл}}(i) =: a_i^{\text{сл}}$  и  $Y_{\text{обсл}}^{\text{уп}}(i) =: a_i^{\text{уп}}$ ,  $i = \overline{1, v}$ , – среднее использование (загрузка) прибора №  $i$ , или интенсивность обслуженной им нагрузки, в первом и втором случаях соответственно. Очевидно, что в силу (2.9)

$$a_i^{\text{сл}} = \frac{Y_{\text{обсл}}}{v} = \frac{\rho(1 - E_v(\rho))}{v}, \quad i = \overline{1, v}. \quad (2.12)$$

На рис. 1.4 видно, что при упорядоченном искании на прибор №  $i$  поступает нагрузка, потерянная на приборе №  $(i-1)$ , с интенсивностью  $\rho E_{i-1}(\rho)$ , а на прибор №  $(i+1)$  – потерянная на приборе №  $i$  с интенсивностью  $\rho E_i(\rho)$ . Поэтому интенсивность обслуженной прибором №  $i$  нагрузки равна

$$a_i^{\text{уп}} = \rho(E_{i-1}(\rho) - E_i(\rho)), \quad i = \overline{1, v}. \quad (2.13)$$

Поскольку  $E_0(\rho) = 1$ , то  $a_1^{\text{уп}} = \rho[1 - E_1(\rho)]$ . Так как интенсивность обслуженной  $v$  приборами нагрузки равна сумме интенсивностей нагрузок, обслуженных каждым из  $v$  приборов, то в силу (2.12) и (2.13)

$$Y_{\text{обсл}}^{\text{уп}} = \sum_{n=1}^v a_n^{\text{уп}} = \rho \left[ \sum_{i=1}^v E_{i-1}(\rho) - \sum_{i=1}^v E_i(\rho) \right] = \rho[1 - E_v(\rho)] = Y_{\text{обсл}}^{\text{сл}} = Y_{\text{обсл}}. \quad (2.14)$$

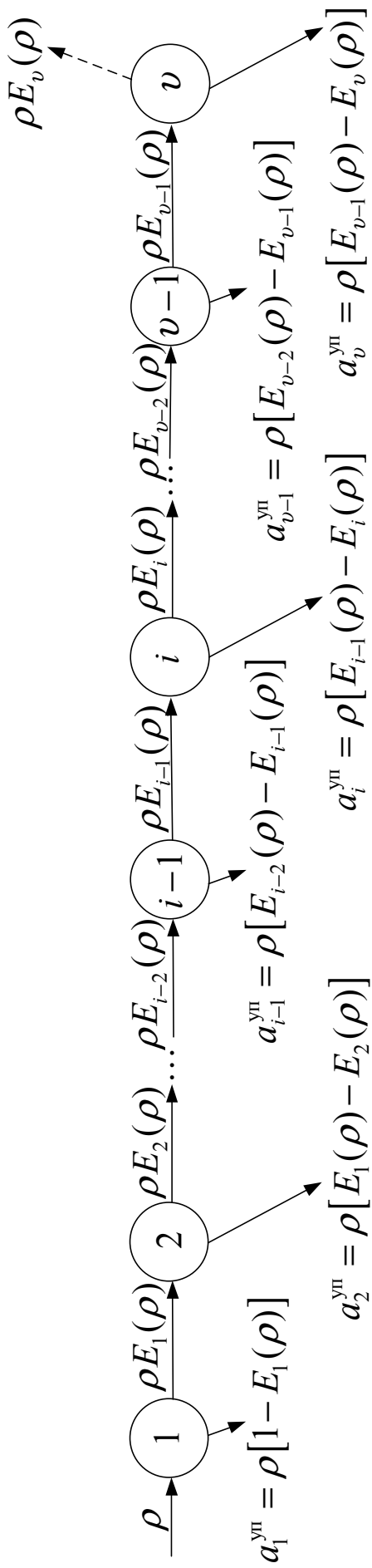


Рис. 1.4. Схема расчета пучка из  $v$  упорядоченных приборов



Таким образом, как вероятность потерь, так и суммарная интенсивность обслуженной пучком из  $\nu$  приборов нагрузка не зависит от вида искания свободных приборов, а загрузка  $i$ -прибора – зависит.

Эта задача возникла в 1930-е годы при планировании работы искателей свободных выходов или свободных линий в декадно-шаговых АТС. Аналогичные задачи возникают в координатных и электронных АТС, а также в других системах.

### §1.3. Модель Эрланга с ожиданием и блокировками

#### 1.3.1. Постановка задачи

Рассмотрим  $\nu$ -линейную СМО с  $r$  местами для ожидания, на которую поступает ПП однородных заявок с постоянной интенсивностью  $\lambda$ , а длительности обслуживания заявок независимы и имеют идентичное экспоненциальное распределение с интенсивностью  $\mu$ :

$$M \left| \begin{array}{c} M \\ \lambda \end{array} \right| \left| \begin{array}{c} M \\ \mu \end{array} \right| 1 \leq \nu < \infty \mid 1 \leq r \leq \infty. \quad (3.1)$$

Алгоритм обслуживания с ожиданием состоит в том, что если в момент поступления заявки есть свободные линии, то заявка занимает одну из них в установленном порядке. Если же все  $\nu$  линий заняты, то заявка занимает комплект ожидания (устройство или зону в памяти), т. е. принимается в очередь до получения доступа к одной из освободившихся линий пучка. При этом может быть рассмотрен как тот случай, когда заявка освобождает место в очереди (комплект ожидания) после начала обслуживания, так и тот, когда заявка удерживает за собой место в очереди до завершения обслуживания. В первом случае емкость системы  $R := \nu + r$ , а во втором  $R = r \geq \nu$ . Для простоты мы будем рассматривать лишь первый случай.

Порядок выбора заявок из очереди на обслуживание при появлении свободной линии может быть различным: в порядке поступления (first

come – first served, FCFS), в обратном порядке (last come – first served, LCFS), случайно (Random) и др. Если с точки зрения проектировщика или оператора вероятность блокировки  $p_{v+r}$  столь мала, что ею можно пренебречь, то анализируют самый простой случай  $r = \infty$ , когда блокировок нет. При  $r < \infty$  заявка, заставшая в СМО  $v$  еще обслуживаемых и  $r$  ожидающих заявок, т. е. полностью занятую систему, получает отказ и теряется, не оказывая дальнейшего влияния на поступающий поток (рис. 1.5).

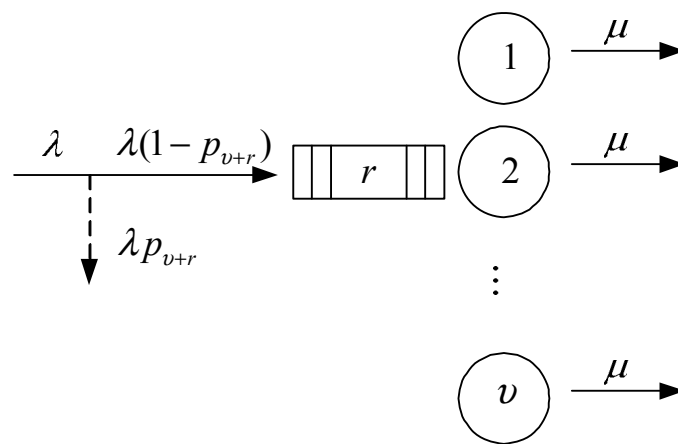


Рис. 1.5. СМО  $M \left| \begin{matrix} M \\ \lambda \end{matrix} \right| v \left| \begin{matrix} M \\ \mu \end{matrix} \right| r < \infty$ , где  $p_{v+r}$  – вероятность блокировки

Эту вторую модель Эрланга в теории телетрафика называют также *системой Эрланга с неявными потерями*, поскольку затянувшееся ожидание часто приводит к отказу ожидающей заявки (абонента) от требуемого соединения или к повторению вызовов, что значительно усложняет модель и соответствующие расчеты.

### 1.3.2. Второе распределение Эрланга

Пусть  $R := v + r$  – общая емкость СМО;  $X(t)$ ,  $t \geq 0$ , – случайный процесс, описывающий общее число заявок в СМО в момент  $t$ ;  $J = \{0, 1, \dots, v, v+1, \dots, R\}$  – пространство состояний системы. В силу сделанных предположений о пуассоновском характере входящего потока

заявок и экспоненциальном распределении длительности их обслуживания, процесс  $X(t)$ ,  $t \geq 0$ , является ступенчатым марковским типа ПРГ с интенсивностями переходов из состояния  $n$  соответственно «вверх», «вниз» и сохранения «статус-кво» ( $n = \overline{0, R}$ )

$$\begin{aligned} a_{n,n+1} &=: \lambda_n = \lambda u(R - n), \\ a_{n,n-1} &=: \mu_n = \mu \min(n, v), \\ a_{n,n} &= -\lambda_n - \mu_n. \end{aligned} \tag{3.2}$$

Для строгого вывода этих соотношений рекомендуем читателям проделать рассуждения, аналогичные тем, которые приведены в разделе 1.1.2 для первой модели Эрланга. Физический смысл интенсивностей переходов «вверх» и «вниз» хорошо поясняет рис. 1.6.

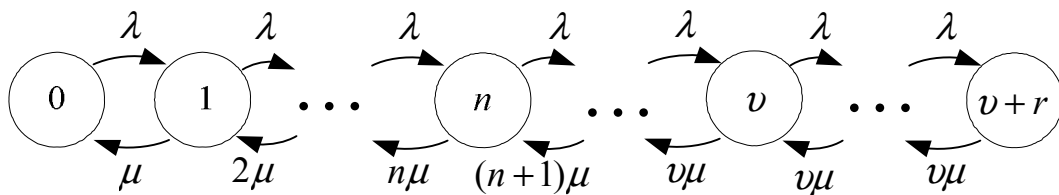


Рис. 1.6. Диаграмма интенсивностей переходов для СМО  $M \left| \begin{array}{c} M \\ \lambda \end{array} \right| \mu \left| \begin{array}{c} v \\ v+r \end{array} \right| r$

Действительно, интенсивность  $\lambda_n$  перехода «вверх» в состоянии  $n = \overline{0, R-1}$  является постоянной и равна интенсивности предложенного потока, а в состоянии  $R$  все заявки предложенного потока блокируются, так что интенсивность  $\lambda_n$  принятого потока равна 0. Интенсивность  $\mu_n$  переходов «вниз» в состоянии  $n = \overline{0, v}$  равна  $n\mu$ , а затем при  $n = \overline{v+1, R}$  интенсивность  $\mu_n$  остается постоянной и равна  $v\mu$ .

В силу (3.2) матрица  $A$  интенсивностей переходов для рассматриваемого ПРГ  $X(t)$ ,  $t \geq 0$ , является трехдиагональной и имеет вид, изображенный в таблице 1.1.

Таблица 1.1. Матрица интенсивностей переходов для СМО  $M|M|v|r$

$A$	0	1	...	$n$	...	$v$	$v+1$	...	$v+r-1$	$R$	$\Sigma$
0	$-\lambda$	$\lambda$									0
1	$\mu$	$-\lambda-\mu$	$\ddots$								0
$\vdots$		$\ddots$	$\ddots$								$\vdots$
$n-1$		$\ddots$	$\lambda$								0
$n$			$-\lambda-n\mu$	$\ddots$							0
$n+1$				$(n+1)\mu$	$\ddots$						0
$\vdots$				$\ddots$							$\vdots$
$v-1$						$\lambda$					0
$v$						$-\lambda-v\mu$	$\lambda$				0
$v+1$						$v\mu$	$-\lambda-v\mu$	$\ddots$			0
$\vdots$							$\ddots$	$\ddots$			$\vdots$
$v+r-1$								$-\lambda-v\mu$	$\lambda$		0
$R$								$v\mu$	$-v\mu$		0

Искомое равновесное распределение  $\mathbf{p}^T = (p_0, \dots, p_R)$  удовлетворяет СУГБ

$$\mathbf{p}^T \mathbf{A} = \mathbf{0}^T, \quad (3.3)$$

которая, в силу (3.2) и вида матрицы  $\mathbf{A}$ , принимает следующий вид:

$$\begin{aligned} -\lambda p_0 + \mu p_1 &= 0; \\ \lambda p_{n-1} - (\lambda + n\mu)p_n + (n+1)\mu p_{n+1} &= 0, \quad n = \overline{1, v-1}; \\ \lambda p_{n-1} - (\lambda + \nu\mu)p_n + \nu\mu p_{n+1} &= 0, \quad n = \overline{\nu, \nu+r-1}; \\ \lambda p_{R-1} - \nu\mu p_R &= 0. \end{aligned} \quad (3.4)$$

**Упражнение 1.5.** Выведите уравнения (3.4) при помощи табл. 1.1.

Используя функцию Хевисайда, запишем СУГБ (3.4) более компактно:

$$\begin{aligned} \lambda p_{n-1} u(n) - [\lambda u(R-n) + \mu \min(n, \nu)] p_n + \\ + \mu \min(n+1, \nu) p_{n+1} u(R-n) = 0, \quad n = \overline{0, R}. \end{aligned} \quad (3.4a)$$

Суммируя теперь уравнения (3.4) от нулевого до  $(n-1)$ -го включительно, получим следующую СУЛБ:

$$\begin{aligned} \lambda p_{n-1} &= n\mu p_n, \quad n = \overline{1, \nu}; \\ \lambda p_{n-1} &= \nu\mu p_n, \quad n = \overline{\nu, R}, \end{aligned} \quad (3.5)$$

или, что эквивалентно,

$$\lambda p_{n-1} = \min(n, \nu)\mu p_n, \quad n = \overline{1, R}. \quad (3.5a)$$

Физическую интерпретацию СУЛБ можно дать с помощью рис. 1.6: левая часть – это взвешенный поток интенсивностей переходов от состояния  $(n-1)$  к состоянию  $n$ , а правая часть – в обратном направлении.

Очевидно, что эта интерпретация справедлива как при  $n = \overline{1, \nu}$ , так и при  $n = \overline{\nu+1, R}$ .

Обозначая  $\rho = \lambda/\mu$ , запишем рекуррентное соотношение (3.5а) в виде  $p_n = \frac{\rho}{\min(n, \nu)} p_{n-1}$ ,  $n = \overline{1, R}$ . Поэтому

$$p_n = \begin{cases} \frac{\rho^n}{n!} p_0, n = \overline{0, \nu}; \\ \left(\frac{\rho}{\nu}\right)^{n-\nu} p_\nu = \frac{\rho^n}{\nu! \nu^{n-\nu}} p_0, n = \overline{\nu, R}. \end{cases} \quad (3.6)$$

Из условия нормировки и (3.6) следует, что

$$\begin{aligned} \frac{1}{p_0} &= \sum_{n=0}^{\nu-1} \frac{\rho^n}{n!} + \frac{\rho^\nu}{\nu!} \sum_{n=\nu}^R \left(\frac{\rho}{\nu}\right)^{n-\nu} = \sum_{n=0}^{\nu-1} \frac{\rho^n}{n!} + \frac{\rho^\nu}{\nu!} \sum_{m=0}^r \left(\frac{\rho}{\nu}\right)^m = \\ &= \sum_{n=0}^{\nu-1} \frac{\rho^n}{n!} + \frac{\rho^\nu}{\nu!} \frac{1 - \left(\frac{\rho}{\nu}\right)^{r+1}}{1 - \frac{\rho}{\nu}}. \end{aligned} \quad (3.7)$$

Итак, нами доказана следующая

**Теорема 1.2.** *Равновесное распределение числа заявок в системе*

*Эрланга с ожиданием  $M \left| \begin{matrix} M \\ \lambda \end{matrix} \right| \nu \left| r \right.$  при  $r < \infty$  и любом  $\rho < \infty$  определяется*

*формулами (3.6)–(3.7). При  $r = R = \infty$  эти формулы справедливы, если выполняется условие*

$$\rho < \nu, \quad (3.8)$$

*причем формула (3.7) упрощается:*

$$\frac{1}{p_0} = \sum_{n=0}^{\nu-1} \frac{\rho^n}{n!} + \frac{\rho^\nu}{\nu!} \frac{1}{1 - \frac{\rho}{\nu}}. \quad (3.7a)$$

Распределение (3.6)–(3.7) называют *вторым распределением Эрланга*.

**Следствие.** Для однолинейной системы массового обслуживания

$M \left| \begin{matrix} M \\ \lambda \end{matrix} \right| \mu \left| \begin{matrix} M \\ \mu \end{matrix} \right| 1 \mid r$  из (3.6)–(3.7) при  $r < \infty$  и любом  $\rho < \infty$  получаем:

$$p_n = \rho^n p_0, \quad n = \overline{0, R}; \quad \frac{1}{p_0} = \frac{1 - \rho^{r+2}}{1 - \rho}. \quad (3.9)$$

Поэтому вероятность блокировки

$$\pi := p_R = \rho^R \frac{1 - \rho}{1 - \rho^{R+2}}. \quad (3.9a)$$

При  $r = R = \infty$  блокировки отсутствуют, но (3.9) справедливо, если  $\rho < 1$ , причем

$$p_n = \rho^n (1 - \rho), \quad n = 0, 1, \dots \quad (3.10)$$

Таким образом, при  $v = 1$  и  $r = \infty$  второе распределение Эрланга является *геометрическим* с параметром  $\rho < 1$ .

Для проверки полученных результатов применим соотношения (3.6)–(3.7) к системе без мест для ожидания ( $r = 0, R = v$ ):

$$p_n = \frac{\rho^n}{n!} p_0, \quad n = \overline{0, v}; \quad \frac{1}{p_0} = \sum_{n=0}^v \frac{\rho^n}{n!}. \quad (3.11)$$

Поскольку (3.11) совпадает с (1.9), то контроль выполняется, причем, как и следовало ожидать, первое распределение Эрланга является частным случаем второго.

### Упражнение 1.6.

а. Доказать, что в СМО  $M \left| \begin{matrix} M \\ \lambda \end{matrix} \right| \mu \left| \begin{matrix} M \\ \mu \end{matrix} \right| v \mid r$  при  $r < \infty$  и любом  $\rho < \infty$

средняя длина очереди  $\bar{q}$  имеет вид

$$\bar{q} = p_v \frac{\rho}{v} \frac{1 + r \left( \frac{\rho}{v} \right)^{r+1} - (r+1) \left( \frac{\rho}{v} \right)^r}{\left( 1 - \frac{\rho}{v} \right)^2}.$$

б. При  $\rho < v$  получить формулу для  $\bar{q}$  при  $r = \infty$  и изучить скорость роста  $\bar{q}$  при  $\frac{\rho}{v} = 1 - \varepsilon$ . Рассмотреть несколько численных примеров при  $\varepsilon = 10^{-1}, 10^{-2}, 10^{-3}$ .

### 1.3.3. Оценка интенсивности принятой нагрузки и вторая формула Эрланга

Рассмотрим снова случайную величину  $X$  – число заявок в СМО, имеющую второе распределение Эрланга (3.6), и пусть случайная величина  $Y$  – число обслуживаемых, а случайная величина  $Z$  – число ожидающих заявок.

Тогда очевидно, что

$$Y = \min(X, v), \quad Z = \begin{cases} X - v, & X \geq v, \\ 0, & X < v \end{cases} =: (X - v)^+, \quad (3.12)$$

$$X = Y + Z. \quad (3.13)$$

Пусть  $\pi := p_R$  – вероятность блокировки, т. е. потерь по времени.

**Теорема 1.3.** При  $r < \infty$  и любом  $\rho < \infty$  средняя интенсивность принятого потока равна средней интенсивности обслуженного потока, а при  $r = \infty$  и  $\rho < v$  равна интенсивности предложенного потока.

**Доказательство.** Суммируя уравнения локального баланса (3.5) по  $n$  от 1 до  $R$ , получим:

$$\lambda(1 - \pi) = \mu \left[ \sum_{n=1}^v n p_n + v \sum_{n=v+1}^R p_n \right]. \quad (3.14)$$

Левая часть представляет собой интенсивность принятого при  $r < \infty \Leftrightarrow \pi > 0$  и предложенного при  $r = R = \infty$ ,  $\rho < v \Leftrightarrow \pi = 0$  потоков. Вместе с тем правая часть в обоих случаях представляет собой среднюю интенсивность обслуженного потока. ■

Равенство (3.14) допускает еще одну полезную интерпретацию.

**Следствие 1.**



$$EY = \begin{cases} \rho(1-\pi), r < \infty, \rho < \infty, \\ \rho, r = \infty, \rho < v. \end{cases} \quad (3.15)$$

Это означает, что среднее число занятых приборов равно средней интенсивности принятой при  $r < \infty$  и средней интенсивности предложенной при  $r = \infty$  нагрузки. Поскольку  $EY \leq v$ , то (3.15) влечет

**Следствие 2.**

$$\rho(1-\pi) \leq v. \quad (3.16)$$

Данное соотношение означает, что средняя интенсивность как принятой, так и предложенной нагрузки не превосходит число  $v$  имеющихся в СМО обслуживающих приборов, что физически очевидно.

Рассмотрим теперь случайную величину  $\omega$  – длительность ожидания заявкой начала обслуживания при  $r = \infty$ ,  $\rho < v$ . Очевидно, что в силу формулы (3.6)

$$P\{\omega > 0\} = \sum_{n=v}^{\infty} p_n = p_0 \frac{\rho^v}{v!} \sum_{n \geq v} \left(\frac{\rho}{v}\right)^{n-v} = \frac{\frac{\rho^v}{v!} \frac{v}{v-\rho}}{\sum_{n=0}^{v-1} \frac{\rho^n}{n!} + \frac{\rho^v}{v!} \frac{v}{v-\rho}}. \quad (3.17)$$

В теории телетрафика формулу (3.17) называют *второй формулой Эрланга* или *C-формулой Эрланга*, поскольку ее можно интерпретировать как *вероятность неявных потерь*. Выразим теперь  $P\{\omega > 0\}$  через  $E_v(\rho)$  – вероятность явных потерь в первой модели Эрланга. Поделив числитель и

знаменатель (3.17) на  $\sum_{n=0}^v \frac{\rho^n}{n!}$  и произведя очевидные упрощения, получим:

$$P\{\omega > 0\} = \frac{\frac{\rho^v}{v!} \frac{v}{v-\rho} \left(\sum_{n=0}^v \frac{\rho^n}{n!}\right)^{-1}}{\left(\sum_{n=0}^{v-1} \frac{\rho^n}{n!} + \frac{\rho^v}{v!} \frac{v}{v-\rho}\right) \left(\sum_{n=0}^v \frac{\rho^n}{n!}\right)^{-1}} =$$

$$= \frac{vE_v(\rho)}{(v-\rho)[1-E_v(\rho)]+vE_v(\rho)} = \frac{vE_v(\rho)}{v-\rho[1-E_v(\rho)]}. \quad (3.17a)$$

Эта формула позволяет вычислять значения  $P\{\omega > 0\}$  как функции от  $v$  и  $\rho$ ,  $\rho < v$  в требуемом диапазоне с помощью ранее полученных значений  $E_v(\rho)$  или с использованием рекуррентных соотношений (1.12) для  $E_v(\rho)$ .

## §1.4. Модель Энгсета

### 1.4.1. Описание модели Энгсета

В первой модели Эрланга  $M \left| \begin{array}{c} M \\ \lambda \end{array} \right| \mu \left| v \right| 0$  предложенный поток заявок является пуассоновским с постоянной интенсивностью поступления заявок, которая не зависит от числа обслуживаемых в этот момент в СМО заявок.

В модели Энгсета<sup>1</sup> также рассматривается полнодоступный пучок линий без мест для ожидания, но вместо одного пуассоновского потока заявок рассматривается  $N$ ,  $N > v$ , малоинтенсивных независимых источников заявок, каждый из которых может находиться в одном из двух состояний: «0» – свободен или «1» – занят сам и занимает одну из линий. Каждый из  $N$  источников в свободном состоянии может с интенсивностью  $\varepsilon$  сгенерировать одну заявку, которая мгновенно займет одну из свободных линий, если они имеются, а источник перейдет в состояние «1», либо предложенная источником заявка будет потеряна с вероятностью  $p_v$  занятия всех  $v$  линий пучка, а источник останется в состоянии «0» (рис. 1.7 и 1.8).

---

<sup>1</sup> Т. Engset (1869–1943) – норвежский математик, один из основателей теории телетрафика. В 1918 г. опубликовал статью, в которой впервые рассмотрен полнодоступный пучок с конечным числом источников нагрузки.

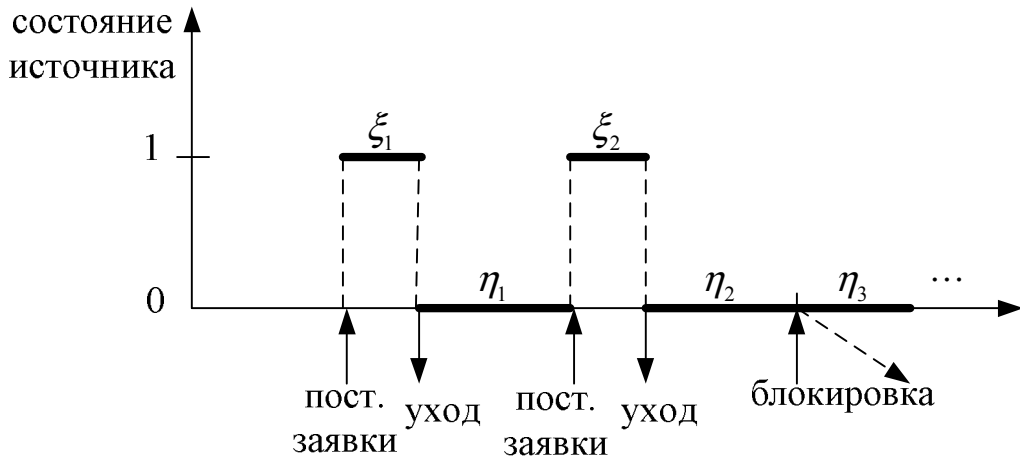


Рис. 1.7. Смена состояний отдельного источника,  $E\xi_i = \frac{1}{\mu}$ ,  $E\eta_i = \frac{1}{\varepsilon}$ ,  
 $i = 1, 2, \dots$

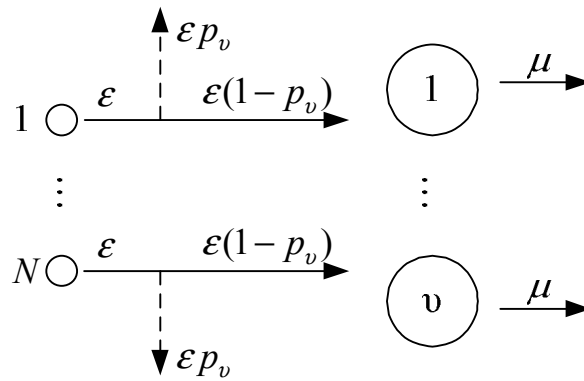


Рис. 1.8. СМО  $M \left| M \right| v | 0$ , где  $p_v$  – вероятность блокировок при  $N > v$

Длительность пребывания источника в состоянии «1» совпадает с длительностью занятия его заявкой линии и имеет экспоненциальное распределение со средним значением  $1/\mu$ , т. е. с интенсивностью  $\mu$  завершения обслуживания и перехода источника из состояния «1» в состояние «0». Отметим, что на практике чаще всего рассматривается случай  $\mu \ll \varepsilon$ .

### 1.4.2. Построение процесса размножения и гибели

Легко убедиться, что функционирование модели Энгсета без мест для ожидания описывается посредством ПРГ  $X(t)$ ,  $t \geq 0$ , параметры которого определяются числом  $j$  занятых линий в системе, причем  $j \in J = \{0, 1, \dots, \min(N, \nu)\}$ . Если  $N \geq \nu$ , то  $J = \{0, 1, \dots, \nu\}$ , т. е. пространство состояний для модели Энгсета при  $r = 0$  совпадает с пространством состояний для модели Эрланга при  $r = 0$ . Модель Энгсета с местами для ожидания при условии  $N > \nu + r$  мы рассматривать не будем, рекомендуя заинтересованному читателю рассмотреть ее самостоятельно.

Если  $j$  из  $\nu$  линий заняты, то интенсивность  $\mu_j$  освобождения одной из них, как и для первой модели Эрланга, равна

$$\mu_j = j\mu, \quad j = \overline{0, \nu}. \quad (4.1)$$

Если  $j$  из  $\nu$  линий заняты, то  $N - j$  источников нагрузки свободны, так что суммарная интенсивность  $\Lambda_j$  предложенного потока заявок зависит от  $j$ , совпадает с интенсивностью  $\lambda_j$  принятого потока при  $j = \overline{0, \nu - 1}$  и равна интенсивности  $\Lambda_\nu$  предложенного, но заблокированного потока при  $N > \nu$  (рис. 1.9):

$$\begin{aligned} \Lambda_j &= (N - j)\varepsilon, \quad j = \overline{0, \nu}; \\ \Lambda_j &= \lambda_j, \quad j = \overline{0, \nu - 1}. \end{aligned} \quad (4.2)$$

Отметим, что при  $N = \nu$  в состоянии  $\nu$  нет свободных источников, так что  $\Lambda_\nu = 0$ , блокировки отсутствуют, потери по вызовам равны 0, хотя вероятность занятия всех линий положительна. Этот факт иллюстрирует принципиальное различие моделей Эрланга и Энгсета (подробнее см. в разделе 1.4.5).

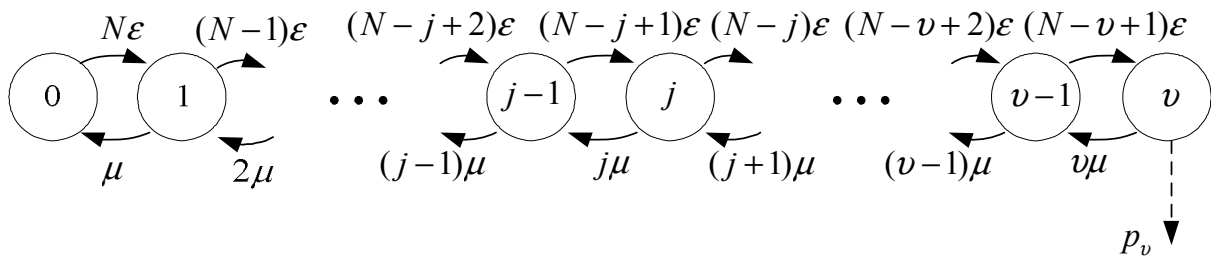


Рис. 1.9. Диаграмма интенсивностей переходов для СМО  $M \left| \begin{matrix} M \\ N, \varepsilon \end{matrix} \right| \mu \left| v \right| 0$ ,

где  $p_v$  – вероятность блокировок при  $N > v$

Заметим теперь, что интенсивность  $\Lambda_j$  предложенного пуассоновского потока заявок в модели Эрланга не зависит от ее состояния  $j$  и равна  $\lambda$  при  $j = \overline{0, v}$ , причем  $\Lambda_j$  совпадает с интенсивностью принятого потока при  $j = \overline{0, v-1}$  и с интенсивностью заблокированного потока при  $j = v$ . Хотя в модели Энгсета разделение предложенного потока заявок на принятый и заблокированный потоки происходит аналогичным образом, но интенсивность этих потоков в силу (4.2) зависит от  $j$ , т. е. от состояния СМО. Поэтому предложенный пуассоновский поток с постоянной интенсивностью  $\lambda$ , не зависящей от параметров СМО, называют еще *чисто случайным*, а предложенный поток с параметрами  $\langle N, \varepsilon \rangle$ , интенсивность которого зависит еще от параметров  $v$  и  $\mu$  модели Энгсета, называют *поток Энгсета*, или *квазислучайным*.

Если распределение длительности обслуживания вызовов является экспоненциальным, то создаваемая пуассоновским потоком нагрузка называется *пуассоновской нагрузкой первого рода*, или *чисто случайной*, а потоком Энгсета – *пуассоновской нагрузкой второго рода*, или *квазислучайной*.

В классической телефонии обычно применяется модель Энгсета при  $N/v$  порядка нескольких единиц, а при  $N/v$  порядка 10 и более единиц применяют модель Эрланга с  $\lambda = N\varepsilon$ .

### 1.4.3. Распределение Энгсета числа занятых линий для случая $N > v$

В силу (3.1)–(3.2) матрица интенсивностей переходов для рассматриваемого ПРГ  $X(t)$ ,  $t \geq 0$  имеет вид, представленный в таблице 1.2.

Равновесное распределение  $\mathbf{p}^T = (p_0, \dots, p_v)$  удовлетворяет СУГБ

$$\mathbf{p}^T \mathbf{A} = \mathbf{0}^T, \quad (4.3)$$

которая, в силу вида матрицы  $\mathbf{A}$ , принимает вид:

$$\begin{aligned} -p_0 N \varepsilon + \mu p_1 &= 0, \\ (N - j + 1) \varepsilon p_{j-1} - [(N - j) \varepsilon + j \mu] p_j + (j + 1) \mu p_{j+1} &= 0, \quad j = \overline{1, v-1}, \\ (N - v + 1) \varepsilon p_{v-1} - v \mu p_v &= 0. \end{aligned} \quad (4.4)$$

Поскольку матрица  $\mathbf{A}$  трехдиагональная, а сумма элементов в каждой строке равна 0, то суммируя уравнения (4.4) от нулевого до  $(j-1)$ -го включительно, получим систему уравнений локального баланса

$$(N - j + 1) \varepsilon p_{j-1} = j \mu p_j, \quad j = \overline{1, v}. \quad (4.5)$$

Физическую интерпретацию этого рекуррентного соотношения можно дать с помощью рис. 1.9: левая часть (4.5) – это взвешенный поток интенсивностей перехода из состояния  $j-1$  в состояние  $j$ , а правая часть – в обратном направлении. Очевидно, что эта интерпретация справедлива при любом  $j$ ,  $j = \overline{1, v}$ .

Пусть теперь  $\rho_1 := \varepsilon / \mu$  – интенсивность нагрузки от одного свободного источника. Тогда (4.5) влечет

$$p_j = \frac{N - j + 1}{j} \rho_1 p_{j-1} = \dots = \frac{(N - j + 1)(N - j + 2) \cdots (N - 1) N}{j!} \rho_1^j p_0, \quad j = \overline{1, v}.$$

Таблица 1.2. Матрица интенсивностей переходов для СМО  $M \left| \begin{matrix} M \\ N, \varepsilon \end{matrix} \right. \mu \left| \begin{matrix} M \\ \mu \end{matrix} \right. v | 0$

$A$	0	1	...	$j$	...	$v-1$	$v$	$\Sigma$
0	$-N\varepsilon$	$N\varepsilon$						0
1	$\mu$	$-(N-1)\varepsilon - \mu$				<b>0</b>		0
$\vdots$		$\ddots$						$\vdots$
$j-1$				$(N-j+1)\varepsilon$				0
$j$				$-(N-j)\varepsilon - j\mu$				0
$j+1$				$(j+1)\mu$	$\ddots$			0
$\vdots$					$\ddots$	$\ddots$		$\vdots$
$v-1$					$\ddots$	$-(N-v+1)\varepsilon - (v-1)\mu$	$(N-v+1)\varepsilon$	0
$v$						$v\mu$	$-v\mu$	0

Записывая эту формулу с помощью сочетаний и используя нормировочное условие  $\sum_{j=0}^v p_j = 1$ , получим распределение Энгсета с параметрами  $\langle v; N, \rho_1 \rangle$ :

$$p_j(v; N, \rho_1) := p_j = \frac{\binom{N}{j} \rho_1^j}{\sum_{m=0}^v \binom{N}{m} \rho_1^m}, \quad j = \overline{0, v}, \quad N \geq v. \quad (4.6)$$

Здесь  $p_j$  – вероятность того, что в СМО Энгсета занято  $j$  каналов из  $v$ , т. е. при большом  $T$  состоятельной статистической оценкой  $p_j$  может служить доля времени пребывания СМО в состоянии  $j$  на интервале  $(t_0, t_0 + T)$ . Поэтому  $p_v$  при  $N > v$  называют также *вероятностью блокировки по времени* в СМО Энгсета.

Рассмотрим теперь систему Энгсета при  $v = N = 1$ , т. е. с параметрами  $\langle 1; 1, \rho_1 \rangle$ . Для нее в силу (4.6) вероятность занятия единственного канала

$$p_1(1; 1, \rho_1) = \frac{\rho_1}{1 + \rho_1} =: \alpha, \quad (4.7a)$$

а вероятность того, что канал свободен

$$p_0(1; 1, \rho_1) = \frac{1}{1 + \rho_1} =: 1 - \alpha. \quad (4.7b)$$

Отсюда следует, что при  $0 \leq \alpha < 1$

$$\rho_1 = \frac{\alpha}{1 - \alpha}. \quad (4.8)$$

Подставляя теперь (4.8) в (4.6) и умножая числитель и знаменатель (4.6) на  $(1 - \alpha)^N$ , получим *усеченное биномиальное распределение*:



$$p_j = \frac{\binom{N}{j} \alpha^j (1-\alpha)^{N-j}}{\sum_{m=0}^v \binom{N}{m} \alpha^m (1-\alpha)^{N-m}}, \quad j = \overline{0, v}. \quad (4.6a)$$

Физический смысл этих результатов становится более ясным при сопоставлении со случаем  $N \leq v$ .

#### 1.4.4. Биномиальное распределение числа занятых линий для случая $N \leq v$

Рассмотрим теперь интересный теоретически случай  $N \leq v$ , когда число каналов превосходит число источников нагрузки. На практике этот редкий случай избытка каналов может произойти, например, если развитие общедоступной сети происходит быстрее, чем подключение к ней новых абонентов или если информация от источников слабой нагрузки является очень ответственной.

Если в модели Энгсета  $N > v$ , а пространство состояний системы  $J = \{0, 1, \dots, v\}$ , то в этом разделе  $N \leq v$ , так что  $J = \{0, 1, \dots, N\}$ . Оба эти случая охватывает пространство состояний  $J = \{0, 1, \dots, \min(v, N)\}$  и, повторяя рассуждения предыдущего раздела с заменой  $v$  на  $\min(v, N)$ , получим следующую формулу для вероятности числа занятых каналов в стационарном режиме:

$$p_j(v; N, \rho_1) =: p_j = \frac{\binom{N}{j} \rho_1^j}{\sum_{m=0}^{\min(v, N)} \binom{N}{m} \rho_1^m}, \quad j = \overline{0, \min(v, N)}. \quad (4.9)$$

При  $N > v$  (4.9) совпадает с распределением Энгсета (4.6), которое в силу (4.6a) является одновременно *усеченным биномиальным распределением*, а при  $N \leq v$ , (4.9) является биномиальным

распределением с параметрами  $N, \alpha$ . Действительно, справедлива следующая

**Теорема 1.4.** Для СМО  $M \left| \begin{matrix} M \\ N, \varepsilon \end{matrix} \right| \mu \left| \begin{matrix} M \\ \mu \end{matrix} \right| 1 \leq N \leq \nu \mid 0$  равновесное

распределение числа  $X$  занятых каналов является биномиальным с числом  $N$  опытов и вероятностью успеха  $\alpha$  в каждом из них.

**Доказательство.** Утверждение теоремы сразу следует из формул (4.9) и (4.7):

$$P\{X = j\} := p_j = \frac{\binom{N}{j} \rho_1^j}{\sum_{m=0}^N \binom{N}{m} \rho_1^m} = \frac{\binom{N}{j} \rho_1^j}{(1 + \rho_1)^N} = \binom{N}{j} \alpha^j (1 - \alpha)^{N-j}, \quad j = \overline{0, N}. \quad (4.10)$$

■

Этот интересный результат можно пояснить с помощью следующих элементарных рассуждений. Разделим все  $N$  источников на две группы – занятые и свободные. Вероятность того, что некоторый источник занят, равна  $\alpha$ , причем она не зависит от состояния всех других источников, поскольку блокировки отсутствуют и любой поступающей заявке предоставляется свободный канал. Поэтому распределение числа занятых источников совпадает с распределением числа занятых каналов и описывается биномиальным распределением  $Bi(N, \alpha)$  (4.10).

Легко показать, что среднее значение и дисперсия числа  $X$  занятых линий соответственно равны

$$EX = N\alpha, \quad DX = N\alpha(1 - \alpha). \quad (4.11)$$

Рассмотрим теперь предельный случай, когда число  $N$  источников нагрузки и число приборов  $\nu$ ,  $\nu \geq N$ , неограниченно возрастают, а интенсивность  $\varepsilon$  каждого из источников стремится к 0, причем

$$\lim_{N \rightarrow \infty, \varepsilon \rightarrow 0} N\varepsilon = \lambda. \quad (4.12)$$

Отметим, что при этом

$$\varepsilon \rightarrow 0 \Leftrightarrow \rho_1 = \frac{\varepsilon}{\mu} \rightarrow 0 \Leftrightarrow \alpha = \frac{\rho_1}{1 + \rho_1} \rightarrow 0,$$

$$N\varepsilon \rightarrow \lambda \Leftrightarrow N\alpha = N \frac{\frac{\varepsilon}{\mu}}{1 + \frac{\varepsilon}{\mu}} = \frac{N\varepsilon}{\mu + \varepsilon} \rightarrow \frac{\lambda}{\mu} =: \rho. \quad (4.12a)$$

Как и в первой модели Эрланга, здесь  $\rho$  – интенсивность предложенной нагрузки для бесконечнолинейного пучка, а после предельного перехода распределение Энгсета (4.10), как и первое распределение Эрланга (1.8), становится распределением Пуассона с параметром  $\rho$ :

$$P\{X = j\} = p_j = e^{-\rho} \frac{\rho^j}{j!}, \quad j = 0, 1, \dots \quad (4.13)$$

Вместе с тем предельный переход (4.12a) в распределении Энгсета (4.6) в силу  $N\rho_1 \rightarrow \frac{\lambda}{\mu} = \rho$  дает распределение Эрланга (1.9).

#### ***1.4.5. Связь между вероятностями блокировок по времени, по вызовам и по нагрузке***

В теории телетрафика часто рассматривают блокировки по времени и по вызовам, поскольку вторые проще измерять экспериментально. Поэтому остановимся вкратце на этом вопросе на примере СМО Энгсета при  $N > v$ .

Вероятность блокировки по времени в силу (4.6) имеет вид:

$$E_{v,N}(\rho_1) = p_v = \frac{\binom{N}{v} \rho_1^v}{\sum_{j=0}^v \binom{N}{j} \rho_1^j}. \quad (4.14)$$

За определение вероятности блокировки по вызовам в силу (4.2) и (4.6) естественно принять величину

$$\begin{aligned}
 B_{v,N}(\rho_1) &:= \lim_{T \rightarrow \infty} \frac{\text{Ср. число заблокир. заявок на } [t_0, t_0 + T)}{\text{Ср. число предл. заявок на } [t_0, t_0 + T)} = \\
 &= \lim_{T \rightarrow \infty} \frac{(N-v)\varepsilon T p_v}{\sum_{j=0}^v (N-j)\varepsilon T p_j} = \frac{(N-v) \binom{N}{v} \rho_1^v}{\sum_{j=0}^v (N-j) \binom{N}{j} \rho_1^j} = \frac{\binom{N-1}{v} \rho_1^v}{\sum_{j=0}^v \binom{N-1}{j} \rho_1^j}. \tag{4.15}
 \end{aligned}$$

Здесь мы воспользовались соотношением  $\frac{(N-j)}{N} \binom{N}{j} = \binom{N-1}{j}$ .

Таким образом,

$$B_{v,N}(\rho_1) = E_{v,N-1}(\rho_1). \tag{4.16}$$

Физическая интерпретация этого равенства состоит в том, что вероятность отвергнуть вызов от одного случайно выбранного источника равна вероятности того, что остальные  $N-1$  источников заняли все  $v$  каналов.

Так как с ростом  $N$  растут и потери по времени, то очевидно, что

$$B_{v,N}(\rho_1) = E_{v,N-1}(\rho_1) < E_{v,N}(\rho_1). \tag{4.17}$$

Неравенство  $B_{v,N}(\rho_1) < E_{v,N}(\rho_1)$  можно также легко как непосредственно доказать аналитически, так и интерпретировать физически.

Рассмотрим теперь потери по нагрузке. Поскольку  $\rho_1 = \varepsilon/\mu$  – интенсивность нагрузки от одного свободного источника, то величины

$$a_j := (N-j)\rho_1, \quad j = \overline{0, v}, \tag{4.18}$$

$$y_j := \begin{cases} (N-j)\rho_1, & j = \overline{0, v-1}, \\ 0, & j = v, \end{cases} \tag{4.19}$$

представляют собой соответственно интенсивности предложенной и принятой нагрузки в состоянии  $j$ ,  $j = \overline{0, v}$ . Очевидно, что интенсивность предложенной нагрузки совпадает с интенсивностью принятой (обслуженной) нагрузки при  $j = \overline{0, v-1}$  и с интенсивностью потерянной нагрузки при  $j = v$ .

Рассмотрим теперь три случайные величины  $X$ ,  $A$  и  $Y$ :

$$P\{X = j\} = P\{A = a_j\} = P\{Y = y_j\} = p_j, \quad j = \overline{0, v},$$

с распределением Энгсета (4.6) и запишем СУЛБ (4.5) в виде

$$y_{j-1}p_{j-1} = jp_j, \quad j = \overline{1, v}. \quad (4.20)$$

Отсюда сразу следует равенство средних значений

$$\bar{Y} = \bar{X}, \quad (4.21)$$

причем факт равенства интенсивностей принятой и обслуженной нагрузок физически также очевиден. Вычислим теперь средние значения (4.21) с помощью (4.6):

$$\begin{aligned} \bar{Y} = \bar{X} &= \sum_{j=1}^v jp_j = p_0 \sum_{j=1}^v j \binom{N}{j} \rho_1^j = \\ &= N\rho_1 p_0 \sum_{j=1}^v \binom{N-1}{j-1} \rho_1^{j-1} = N\rho_1 p_0 \sum_{s=0}^{v-1} \binom{N-1}{s} \rho_1^s. \end{aligned} \quad (4.22)$$

Вероятностью блокировки по нагрузке назовем отношение интенсивности блокированной к интенсивности поступающей нагрузки:

$$C_{v,N}(\rho_1) = \frac{\bar{A} - \bar{Y}}{\bar{A}} = \frac{(N-v)\rho_1 p_v}{\sum_{j=0}^v (N-j)\rho_1 p_j} = B_{v,N}(\rho_1). \quad (4.23)$$

Таким образом, справедливо соотношение

$$C_{v,N}(\rho_1) = B_{v,N}(\rho_1) = E_{v,N-1}(\rho_1) < E_{v,N}(\rho_1). \quad (4.24)$$

Повторяя теперь рассуждения этого раздела для первой модели Эрланга, получим, что для нее вероятности блокировок по времени, по вызовам и по нагрузке совпадают:

$$E_v(\rho) = B_v(\rho) = C_v(\rho). \quad (4.25)$$

Это утверждение следует также и из (4.24), когда число  $N$  источников нагрузки растет, а интенсивность  $\rho_1$  каждого из них в свободном состоянии уменьшается, так что выполняются условия (4.12)–(4.12a).

Практические рекомендации о том, когда предложенную пуассоновскую нагрузку второго рода с параметрами  $N, \rho_1$  можно аппроксимировать пуассоновской нагрузкой первого рода с параметром  $\rho$  можно найти в учебниках по инженерной ТТ.

### §1.5. Равновесное распределение и система уравнений равновесия (приложение к главе 1)

Вероятностный вектор  $\boldsymbol{\pi}^T = (\pi_i)_{i \in J}$  с положительными элементами представляет собой *равновесное распределение* некоторого случайного вектора  $x$  на  $J$ , который отражает фундаментальные статистические свойства ступенчатого марковского процесса  $\{X(t), t \geq 0\}$ , проявляющиеся на достаточно длинном отрезке времени. Действительно, можно показать, что равновесное распределение  $\boldsymbol{\pi}$  содержит всю необходимую для приложений информацию о наличии *предельных вероятностей, эргодичности* (сходимости доли времени пребывания в состоянии  $i$  к  $\pi_i, i \in J$ ) у рассматриваемого СтМП:

$$P\{X(t) = i\} \xrightarrow{t \rightarrow \infty} \pi_i, \quad \forall i \in J;$$

$$P\left\{\frac{1}{T} \int_0^T 1(X(t) = i) dt \xrightarrow{T \rightarrow \infty} \pi_i\right\} = 1, \quad \forall i \in J.$$

Поскольку  $\{X(t), t \geq 0\}$  является математической моделью некоторой реальной СМО  $\langle S, A \rangle$ , функционирующей в состоянии равновесия, то это оправдывает название *равновесное распределение* для  $\pi$  и позволяет использовать физические соображения и результаты реальных измерений для поддержки гипотезы о наличии равновесия или ее частичного или полного опровержения. Здесь  $S$  – структура ресурсов СМО,  $A$  – алгоритм их распределения между входящими потоками заявок [1]. Далее для обозначения системы иногда будет использоваться только символ  $S$ .

Важно отметить, что матрица  $A$  для соответствующего СтМП содержит всю информацию о нем, которая обычно необходима в приложениях. При этом некоторые ограничения на элементы матрицы  $A$  и ее неразложимость физически оправданы и обычно выполняются в инженерных приложениях. В этих предположениях существуют ненулевые решения вырожденной системы линейных алгебраических уравнений (СЛАУ)

$$\sum_{i \in J} x_i a_{ij} = 0, \quad j \in J \quad \Leftrightarrow \quad \mathbf{x}^T \mathbf{A} = \mathbf{0}^T, \quad (5.1)$$

причем все  $x_i$  имеют один знак и отличны от 0, а если  $|J| = \infty$ , то

$$\sum_{i \in J} |x_i| < \infty. \text{ Поэтому, полагая}$$

$$\pi_i := \frac{x_i}{\sum_{j \in J} x_j}, \quad i \in J,$$

запишем СЛАУ (2.1) в виде *системы уравнений равновесия* (СУР)

$$\sum_{i \in J} \pi_i a_{ij} = 0, \quad j \in J \quad \Leftrightarrow \quad \boldsymbol{\pi}^T \mathbf{A} = \mathbf{0}^T$$

с нормировочным условием  $\sum_{i \in J} \pi_i = 1$ .

## Глава 2. ПРИМЕНЕНИЕ КЛАССИЧЕСКИХ МОДЕЛЕЙ К АНАЛИЗУ ОДНОЙ СОТЫ ССПС

### §2.1. Физическая модель процесса обслуживания в соте

Одним из подходов к проектированию сетей является применение принципа декомпозиции, согласно которому сеть разбивается на блоки, после чего для отдельного блока выполняется анализ и расчет необходимых характеристик. Будем рассматривать сотовую сеть подвижной связи (ССПС), территория обслуживания которой покрыта сотами правильной шестиугольной формы. Допустим, что все соты сети идентичны по размеру, числу радиоканалов и требованиям к обслуживанию вызовов. Предположим также, что в любой момент времени абоненты сети, находящиеся в пределах территории обслуживания, распределены по ней равномерно. В сделанных предположениях отдельным блоком для анализа процесса функционирования сети может служить одна сота ССПС.

Особенностью систем беспроводной подвижной связи является мобильность абонента, которая влечет необходимость передачи обслуживания текущего соединения мобильного абонента из одной соты в другую без прекращения связи. Если сигнал между базовой станцией и мобильным абонентом становится слабым, необходимо передать обслуживание вызова базовой станции соседней соты, т.е. выполнить процедуру *хэндовера* – смену радиоканала без разрыва текущего соединения. Таким образом, в каждой соте ССПС возникают вызовы двух типов: так называемые *новые вызовы* (new calls), к возникновению которых привела инициация соединения абонентом, находящимся на территории данной соты, и *хэндовер-вызовы* (handover calls).

Новый или хэндовер-вызов, обслуживаемый базовой станцией фиксированной соты, может либо успешно закончиться на территории



данной соты по причине окончания разговора мобильным абонентом, либо потребовать передачи обслуживания на базовую станцию соседней соты. Приближение мобильного абонента к границе соты сопровождается ухудшением качества связи. При падении мощности сигнала от базовой станции данной соты ниже определенного уровня начинается поиск свободного радиоканала на базовой станции соты, в направлении которой движется мобильный абонент. Далее возможны три варианта:

- а) передача обслуживания текущего соединения на один из свободных радиоканалов базовой станции соседней соты;
- б) успешное завершение обслуживания текущего соединения по причине окончания разговора мобильным абонентом во время нахождения в зоне хэндовера;
- в) вынужденный разрыв текущего соединения (*forced call termination*) на территории соседней соты – *блокировка хэндовера*, которая произойдет, если в момент пересечения абонентом границы соты передача обслуживания текущего соединения базовой станции соседней соты невозможна (например, на базовой станции соседней соты нет свободных радиоканалов).

Вероятность  $B_2$  вынужденного разрыва текущего соединения (вероятность блокировки хэндовер-вызова) является одним из основных параметров качества обслуживания (Quality of Service, QoS-параметров), нормированных в рекомендациях МСЭ. Еще один QoS-параметр – вероятность  $B_1$  отказа в обслуживании при инициации соединения (вероятность блокировки нового вызова).

Далее в главе 2 исследуются две модели обслуживания вызовов в одной соте ССПС [15]. Заметим, что психологически абоненту ССПС легче принять отказ в обслуживании при первой попытке установления соединения, чем столкнуться с разрывом уже установленного соединения в процессе разговора, поэтому в ССПС применяются механизмы,

обеспечивающие приоритет хэндовер-вызовов по отношению к новым вызовам, например резервирование каналов для обслуживания хэндовер-вызовов. Эти детали учитываются при построении и анализе модели §2.3.

Введём упрощающие предположения, позволяющие перейти от описанной ранее физической модели к математической модели. Общими для обеих моделей являются следующие предположения:

1. Потоки новых и хэндовер-вызовов являются независимыми пуассоновскими потоками с постоянными интенсивностями  $\lambda_1$  и  $\lambda_2$  соответственно.
2. Если в момент поступления нового или хэндовер-вызова на базовой станции нет свободных радиоканалов, то вызов теряется, не оказывая дополнительного влияния на интенсивность породившего его пуассоновского потока.
3. Длительность обслуживания базовой станцией как нового, так и хэндовер-вызова, является экспоненциально распределенной СВ с параметром  $\mu$ .
4. Число радиоканалов в соте равно  $C$ .

Далее в данной главе для обеих моделей получены аналитические формулы для расчета вероятностных характеристик, определенных выше. При заданных ограничениях на QoS-параметры эти формулы дают возможность рассчитать структурные параметры модели, такие как общее число радиоканалов в соте и число радиоканалов, зарезервированных для обслуживания хэндовер-вызовов.

## **§2.2. Полнодоступная модель с потерями**

В предположениях 1–4 математической моделью процесса обслуживания вызовов в соте ССПС может служить  $C$ -линейная (по числу радиоканалов в соте) полнодоступная СМО, на которую поступают два пуассоновских потока заявок. Поток 1-заявок, соответствующий потоку

новых вызовов, имеет интенсивность  $\lambda_1$ , а поток 2-заявок (хэндовер-вызовы) – интенсивность  $\lambda_2$ . Если в момент поступления заявки любого типа в СМО имеется хотя бы один свободный прибор, заявка поступает на обслуживание и занимает один прибор на все время обслуживания. Длительности обслуживания как 1-заявок, так и 2-заявок, являются независимыми случайными величинами, имеющими экспоненциальное распределение с параметром  $\mu$ . Если в момент поступления заявки любого типа в СМО нет свободных приборов, заявка теряется.

Таким образом, поток вызовов, создающих нагрузку на базовую станцию сотовой, является суперпозицией независимых пуассоновских потоков интенсивностей  $\lambda_1$  и  $\lambda_2$ , а следовательно – пуассоновским потоком интенсивности

$$\lambda = \lambda_1 + \lambda_2. \quad (2.1)$$

Тогда для анализа процесса обслуживания вызовов в сети ССПС можно применить первую модель Эрланга  $M | M | C | 0$  (см. §1.1).

Схематически модель системы показана на рис. 2.1.

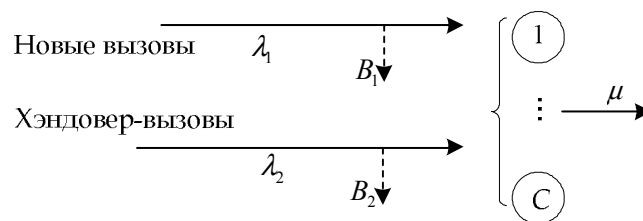


Рис. 2.1. Двухпоточковая полнодоступная СМО с потерями

Интересующими нас характеристиками являются вероятность потери 1-заявки, соответствующая вероятности  $B_1$  блокировки нового вызова, и вероятность потери 2-заявки, соответствующая вероятности  $B_2$  блокировки хэндовер-вызова.

Пусть  $X(t)$  – число заявок в СМО в момент  $t$ ,  $t \geq 0$ . Пространством состояний для случайного процесса  $X(t)$  служит множество

$J = \{0, 1, \dots, C\}$ ,  $1 \leq C \leq \infty$ . Обозначим  $p_n(t)$  вероятность того, что в системе в момент  $t$  было  $n$  заявок:

$$p_n(t) = P\{X(t) = n\}, \quad n = \overline{0, C}. \quad (2.2)$$

Согласно (1.1.9) стационарное распределение вероятностей имеет вид:

$$p_n = \frac{\rho^n}{n!} \left[ \sum_{i=0}^C \frac{\rho^i}{i!} \right]^{-1}, \quad n = \overline{0, C}. \quad (2.3)$$

где  $\rho = \lambda/\mu$  – суммарная нагрузка на систему, создаваемая новыми и хэндовер-вызовами, а  $\lambda$  определяется формулой (2.1).

Потеря как 1-заявок, так и 2-заявок в рассмотренной СМО произойдет в случае, когда в момент  $\tau$  поступления заявки любого типа в СМО нет свободных приборов, т. е. в случае  $X(\tau) = C$ . Следовательно, вероятности блокировки нового вызова и блокировки хэндовер-вызова совпадают и определяются формулой

$$B_1 = B_2 = \frac{\rho^C}{C!} \left[ \sum_{n=0}^C \frac{\rho^n}{n!} \right]^{-1}. \quad (2.4)$$

### §2.3. Неполнодоступная модель с потерями

Сделаем еще одно предположение.

5. Для обеспечения приоритета хэндовер-вызовов по отношению к новым вызовам применяется стратегия доступа с резервированием: на базовой станции соты  $g$  радиоканалов предназначены для обслуживания как новых, так хэндовер-вызовов, а остальные  $C - g$  радиоканалов зарезервированы только для обслуживания хэндовер-вызовов.

Более подробно о стратегиях доступа и управлении доступом см. главу 3 [14].

В предположениях 1–5 математической моделью процесса обслуживания вызовов в сети ССПС может служить  $S$ -линейная неполнодоступная СМО, на которую поступают два пуассоновских потока заявок. Поток 1-заявок, соответствующий потоку новых вызовов, имеет интенсивность  $\lambda_1$ , а поток 2-заявок (хэндовер-вызовы) – интенсивность  $\lambda_2$ . Если в момент поступления 1-заявки в СМО число свободных приборов больше, чем  $C - g$ ,  $0 \leq g \leq C$ , 1-заявка поступает на обслуживание и занимает один прибор на все время обслуживания, в противном случае 1-заявка теряется. Если в момент поступления 2-заявки в СМО есть хотя бы один свободный прибор, 2-заявка поступает на обслуживание и занимает один прибор на все время обслуживания, в противном случае 2-заявка теряется. Длительности обслуживания как 1-заявок, так и 2-заявок, являются независимыми случайными величинами, имеющими экспоненциальное распределение с параметром  $\mu$ . Схематически модель системы показана на рис. 2.2.

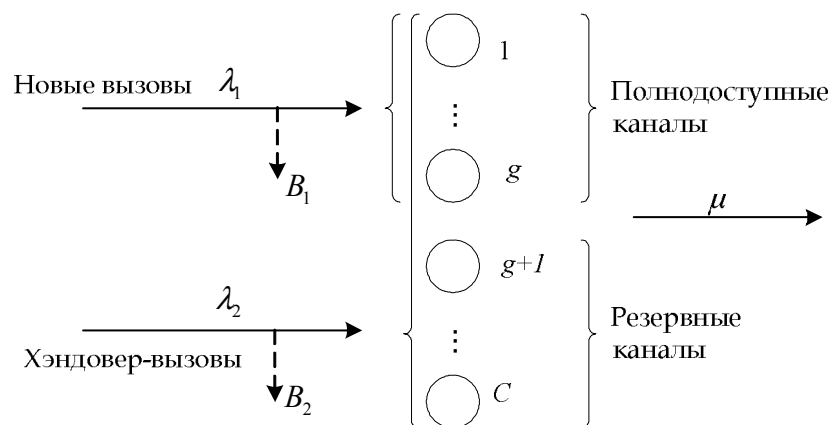


Рис. 2.2. Двухпоточковая неполнодоступная СМО с потерями

На рисунке изображена модель со стягиванием – при освобождении одного из приборов в полнодоступной части вызов, обслуживающийся на одном из резервных приборов, мгновенно переходит на обслуживание в полнодоступную часть, занимая в ней освободившийся прибор. Прибор в резервной части при этом освобождается.

Интересующими нас характеристиками вновь являются вероятность потери 1-заявки, соответствующая вероятности  $B_1$  блокировки нового вызова, и вероятность потери 2-заявки, соответствующая вероятности  $B_2$  блокировки хэндовер-вызова.

Обозначим  $X(t)$  число заявок в СМО в момент времени  $t$ ,  $t \geq 0$ . Случайный процесс  $\{X(t), t \geq 0\}$  также является ПРГ. Пространство состояний МП  $X(t)$  имеет вид  $J = \{0, 1, \dots, C\}$ ,  $1 \leq C \leq \infty$ . На рис. 2.3 представлена диаграмма интенсивностей переходов МП  $X(t)$ .

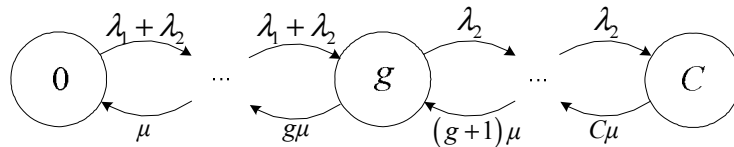


Рис. 2.3. Диаграмма интенсивностей переходов МП  $X(t)$

При  $C < \infty$ , а также при  $C = \infty$  и  $\rho < C$  МП  $\{X(t), t \geq 0\}$  – эргодический, следовательно, существуют его стационарные вероятности  $\{p_n, n = \overline{0, C}\}$ .

Стационарные вероятности  $\{p_n, n \geq 0\}$  удовлетворяют системе уравнений равновесия, выведенной с помощью принципа глобального баланса:

$$\begin{cases} (\lambda_1 + \lambda_2) p_0 = \mu p_1; \\ (n\mu + \lambda_1 + \lambda_2) p_n = (\lambda_1 + \lambda_2) p_{n-1} + (n+1)\mu p_{n+1}, & 1 \leq n \leq g-1; \\ (g\mu + \lambda_2) p_g = (\lambda_1 + \lambda_2) p_{g-1} + (g+1)\mu p_{g+1}; \\ (n\mu + \lambda_2) p_n = \lambda_2 p_{n-1} + (n+1)\mu p_{n+1}, & g+1 \leq n \leq C-1; \\ C\mu p_C = \lambda_2 p_{C-1}. \end{cases} \quad (3.1)$$

При решении СУР воспользуемся принципом локального баланса. Выпишем систему уравнений локального баланса:

$$\begin{cases} (\lambda_1 + \lambda_2) p_{n-1} = n\mu p_n, & 1 \leq n \leq g; \\ \lambda_2 p_{n-1} = n\mu p_n, & g+1 \leq n \leq C. \end{cases} \quad (3.2)$$

Из (3.2) следует, что

$$p_n = \begin{cases} \frac{(\lambda_1 + \lambda_2)^n}{n! \mu^n} p_0, & 1 \leq n \leq g; \\ \frac{(\lambda_1 + \lambda_2)^g \lambda_2^{n-g}}{n! \mu^n} p_0, & g+1 \leq n \leq C. \end{cases} \quad (3.3)$$

Получаем выражения для стационарных вероятностей МП  $\{X(t), t \geq 0\}$  в виде

$$p_n = \begin{cases} \frac{\rho^n}{n!} p_0, & 1 \leq n \leq g; \\ \frac{\rho^g \rho_2^{n-g}}{n!} p_0, & g+1 \leq n \leq C, \end{cases} \quad (3.4)$$

где  $\rho_1 = \lambda_1/\mu$ ,  $\rho_2 = \lambda_2/\mu$ , а  $p_0$  определяется из нормировочного условия

$\sum_{n=0}^C p_n = 1$  и имеет вид

$$p_0 = \left[ \sum_{n=0}^g \frac{\rho^n}{n!} + \sum_{n=g+1}^C \frac{\rho^g \rho_2^{n-g}}{n!} \right]^{-1}. \quad (3.5)$$

Потеря 1-заявки произойдет в случае, когда в момент  $\tau$  поступления 1-заявки в СМО занято не менее, чем  $g$  приборов, т. е.  $X(\tau) \in \{g, g+1, \dots, C\}$ . Отсюда следует, что вероятность блокировки нового вызова определяется формулой

$$B_1 = \left( \sum_{n=g}^C \frac{\rho^g \rho_2^{n-g}}{n!} \right) \left[ \sum_{n=0}^g \frac{\rho^n}{n!} + \sum_{n=g+1}^C \frac{\rho^g \rho_2^{n-g}}{n!} \right]^{-1}. \quad (3.6)$$

Потеря 2-заявки в рассмотренной СМО произойдет в случае, когда в момент  $\tau$  поступления 2-заявки в СМО нет свободных приборов, т. е.

$X(\tau) = C$ . Отсюда следует, что вероятность блокировки хэндовер-вызова определяется формулой

$$B_2 = \frac{\rho^g \rho_2^{C-g}}{C!} \left[ \sum_{n=0}^g \frac{\rho^n}{n!} + \sum_{n=g+1}^C \frac{\rho^g \rho_2^{n-g}}{n!} \right]^{-1}. \quad (3.7)$$

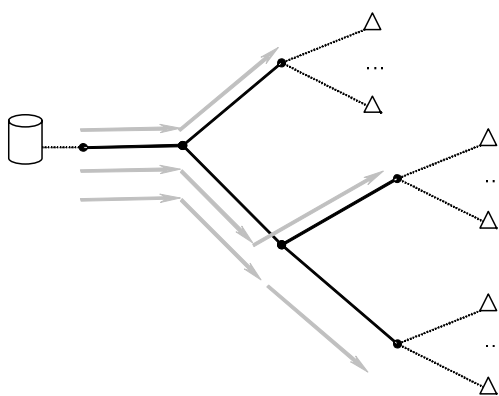
Оценивая характеристики  $B_1$  и  $B_2$  для недоступной модели с приоритетом хэндовер-вызовов, можно сказать, что вероятность  $B_2$  блокировок хэндовер-вызовов, как этого и следовало ожидать, меньше, чем вероятность  $B_1$  блокировок новых вызовов в соте.



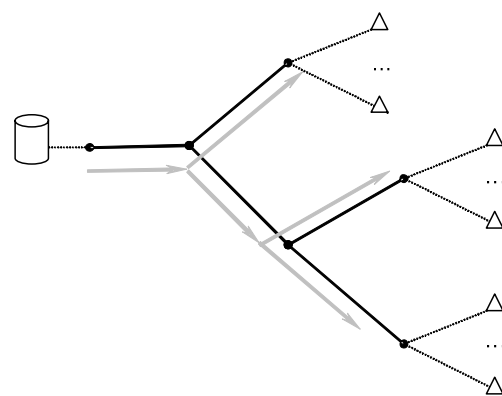
## Глава 3. МОДЕЛЬ ЗВЕНА СЕТИ С МНОГОАДРЕСНЫМИ СОЕДИНЕНИЯМИ

### § 3.1. Модель сети с многоадресными соединениями

Многоадресная передача данных (multicast), или *мультивещание* – технология доставки информации по принципу «точка–много точек», позволяющая более эффективно по сравнению с одноадресной передачей использовать ресурсы сети при рассылке идентичных данных нескольким пользователям. При традиционных способах передачи информации для каждой пары источник–получатель устанавливается соединение «точка–точка» (unicast). В случае передачи одинаковых данных нескольким получателям (например, рассылка обновлений программного обеспечения или телетрансляция в СДОП) это приводит к тому, что по общим для нескольких соединений звеньям сети передается несколько копий одной и той же информации. При мультивещании пользователи совместно используют пропускную способность общих звеньев маршрутов без дублирования информации. Этот основной принцип мультивещания проиллюстрирован на рис. 3.1.



а) традиционная маршрутизация



б) многоадресная маршрутизация

Рис. 3.1. Основной принцип мультивещания

Первые спецификации многоадресной передачи появились в середине 1980-х гг., однако востребована эта технология стала только в конце 1990-х гг. при организации на базе мультисервисных сетей связи

(МСС), включая Интернет, телевизионного вещания и видеотрансляций, поскольку предоставление таких услуг предполагает обеспечение возможности передавать большим группам абонентов трафика существенного объема – десятков видеопотоков, соответствующих различным телевизионным каналам, со скоростью не ниже 2–4 Мбит/с в случае телевидения традиционного качества и около 25–34 Мбит/с в случае телевидения высокой четкости (High Definition Television, HDTV).

Различают статическое и динамическое мультивещание. При статическом мультивещании на первом этапе источник формирует группу пользователей, запросивших предоставление какой-либо услуги (так называемую группу мультивещания). На втором этапе источник посылает информацию, соответствующую содержанию услуги, всем пользователям из этой группы. Примером может служить рассылка обновлений программного обеспечения группе пользователей, подписавшихся на постоянное получение новых версий. Рассылка инициируется источником (push-технология) при появлении в источнике очередной обновленной версии программы. С момента начала рассылки информации происходит формирование очередной группы мультивещания. Все пользователи, приславшие запрос на предоставление данной услуги после начала рассылки источником текущей версии документа, войдут в группу мультивещания уже для рассылки следующей версии.

При динамическом мультивещании источник начинает передачу информации, соответствующей содержанию услуги, по запросу пользователя. Момент окончания предоставления услуги также определяется пользователем. Примером динамического мультивещания служит теле- и радиовещание в мультисервисной сети.

Анализ функционирования сети динамического мультивещания, которому посвящена настоящая глава, относится к классу задач анализа систем с разделением ресурсов в мультисервисных сетях, однако в этом

случае, в отличие от сетей с одноадресными соединениями, известные для СМО с потерями результаты невозможно применять напрямую. Отдельной задачей является численный анализ, осуществление которого ввиду комбинаторного характера пространства состояний системы и его большой размерности невозможно без разработки специальных вычислительных алгоритмов.

### **3.1.1. Основные понятия**

Будем рассматривать мультисервисную сеть многоадресной передачи с несколькими источниками информации (отправителями данных), которые, как и пользователи (получатели данных), могут быть подключены к любому узлу сети [16]. Каждый *источник* предоставляет конечное число *услуг* и по запросам пользователей передает информацию, соответствующую содержанию услуги. При этом если одна и та же услуга предоставляется одновременно нескольким пользователям, то информация передается без дублирования. Из-за этого в сети мультивещания соединения имеют структуру дерева, корнем которого является узел подключения источника, а листовыми вершинами – узлы подключения пользователей. Такое дерево мы будем называть *деревом мультивещания* (multicasting tree). Поскольку каждый пользователь сети может обращаться к источнику независимо от других пользователей, структура многоадресных соединений динамична: сделавший запрос на предоставление услуги пользователь будет подключен к активному дереву мультивещания, если эта услуга уже предоставляется другим пользователям, или инициализирует новое дерево мультивещания (имеющее пока лишь одну ветвь), если эта услуга никому не предоставлялась.

Рассмотрим фрагмент мультисервисной сети заданной структуры, состоящей из  $L$  звеньев (рис. 3.2). Последовательность звеньев сети от

узла подключения пользователя до источника будем называть *физическим путем*. Обозначим

$L = \{1, \dots, L\}$  – множество звеньев сети;

$S = \{1, \dots, S\}$  – множество источников информации;

$P_s = \{1, \dots, P_s\}$  – множество всех физических путей к источнику  $s \in S$  ;

$M_s = \{1, \dots, M_s\}$  – множество услуг, предоставляемых источником  $s \in S$  ;

$L_{ps} \subseteq L$  – множество всех звеньев физического пути  $p \in P_s$  к источнику  $s \in S$  ;

$S^l = \{s \in S : P_s^l \neq \emptyset\}$  – множество источников информации, предоставляющих услуги через звено  $l \in L$  ;

$P_s^l = \{p \in P_s : l \in L_{ps}\}$  – множество физических путей к источнику  $s \in S^l$ , проходящих через звено  $l \in L$  ;

$b_{ms}$  – число единиц емкости звена, требуемое для предоставления услуги  $m \in M_s$  ;

$C_l$  – емкость звена  $l \in L$  .

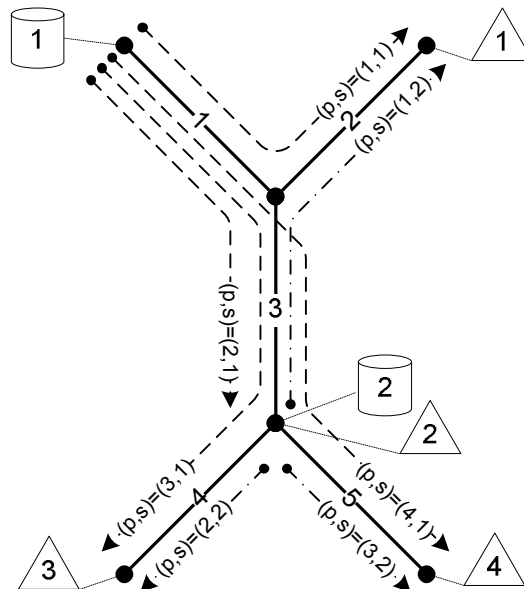


Рис. 3.2. Пример сети мультивещания

**Упражнение 3.1.** Для примера сети многоадресной передачи, схема которой изображена на рис. 3.2, выпишите множества  $L$ ,  $S$ ,  $P_s$ ,  $M_s$ ,  $L_{ps}$ , где  $s \in S$ , а также множества  $S^l$ ,  $P_s^l$ , где  $l \in L$ ,  $s \in S^l$ . Постройте дерево мультивещания от каждого из источников.

**Решение.** Сеть состоит из пяти звеньев, следовательно  $L = \{1, \dots, 5\}$ . В ней имеются два источника информации, изображенные на рисунке цилиндрами,  $S = \{1, 2\}$ . К первому ведут физические пути от каждого из четырех пользователей (пользователи показаны треугольниками), т.е.  $P_1 = \{1, \dots, 4\}$ , при этом  $L_{11} = \{2, 1\}$ ,  $L_{21} = \{3, 1\}$ ,  $L_{31} = \{4, 3, 1\}$  и  $L_{41} = \{5, 3, 1\}$  (рис. 3.3а).

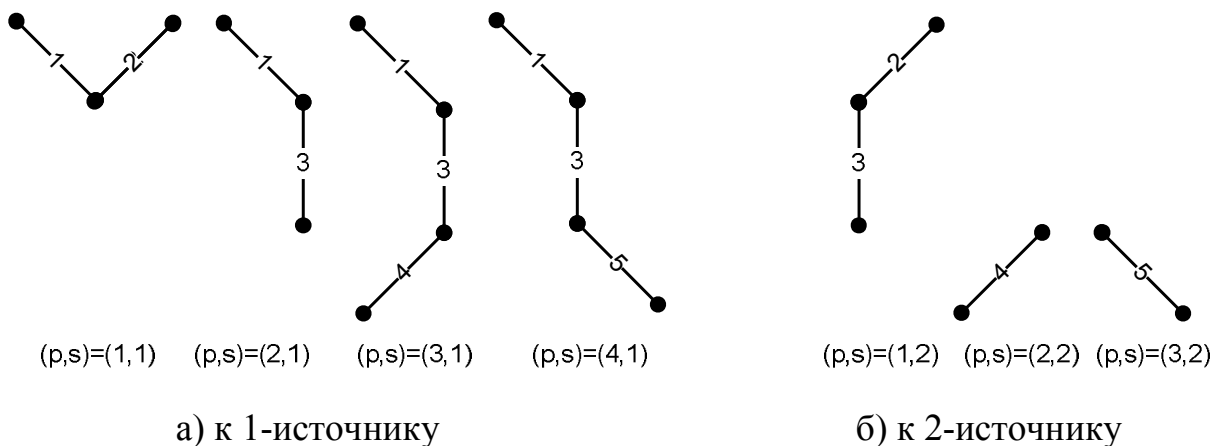


Рис. 3.3. Физические пути в сети на рис. 3.2

Ко второму источнику имеются три физических пути, так как можно считать, что пользователь 2 подсоединен ко второму источнику напрямую, а не через звенья рассматриваемой сети, т.е.  $P_2 = \{1, 2, 3\}$  и  $L_{12} = \{2, 3\}$ ,  $L_{22} = \{4\}$ ,  $L_{32} = \{5\}$  (рис. 3.3б). Наконец, если рассматривать, например, третье звено, то через него проходят два физических пути к первому источнику от третьего и четвертого пользователей:  $P_1^3 = \{3, 4\}$ , и один физический путь ко второму источнику:  $P_2^3 = \{1\}$ .

Деревья мультивещания от каждого из источников ко всем пользователям показаны на рис. 3.4.

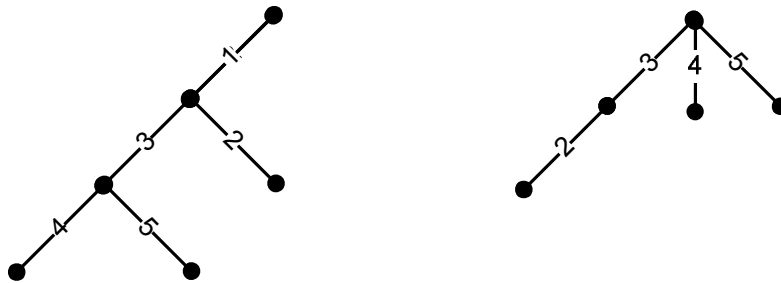


Рис. 3.4. Примеры деревьев мультивещания

Пару  $(m, s)$ ,  $m \in M_s$ ,  $s \in S$ , будем называть  $(m, s)$ -услугой. Пусть для предоставления  $(m, s)$ -услуги на всех звеньях физического пути  $p \in P_s$  требуется  $b_{ms}$  единиц емкости звена. 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 логического пути обозначим  $x_{mps}$ , при этом пусть  $x_{mps} = 1$ , когда  $(m, s)$ -услуга предоставляется пользователю, которого связывает с источником  $s$  физический путь  $p$ ; в противном случае  $x_{mps} = 0$ . Теперь детальное состояние всех логических путей сети можно описать набором матриц  $\mathbf{x} = (\mathbf{X}_1, \dots, \mathbf{X}_s, \dots, \mathbf{X}_S)$ , где матрица

$$\mathbf{X}_s = \begin{pmatrix} x_{11s} & \dots & x_{1P_s s} \\ \dots & & \dots \\ x_{M_s 1s} & \dots & x_{M_s P_s s} \end{pmatrix}$$

определяет детальное состояние всех логических путей к источнику  $s \in S$ . Множество всех возможных матриц такого вида обозначим  $\tilde{\mathcal{X}}$ .

**Упражнение 3.2.** Покажите, что мощность множества  $\tilde{\mathcal{X}}$  определяется формулой  $|\tilde{\mathcal{X}}| = 2^{\sum_{s \in S} M_s P_s}$ .

### 3.1.2. Пространство состояний модели и равновесное распределение вероятностей

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

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

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

$$c_l(\mathbf{x}) = \sum_{s \in S^l} \sum_{m \in M_s} b_{ms} y_{ms}^l(\mathbf{x}), \quad \mathbf{x} \in \tilde{X}, \quad (1.1)$$

задающую количество занятых единиц емкости  $l$ -звена, если сеть находится в состоянии  $\mathbf{x}$ . Тогда пространство состояний логических путей сети можно выразить следующим образом:

$$X = \{ \mathbf{x} : c_l(\mathbf{x}) \leq C_l, \quad l \in L \}. \quad (1.2)$$

Функционирование путей в сети мультивещания описывается случайным процессом ввиду случайного характера потока запросов пользователей на предоставление услуг. Кроме того, случайным является и время занятия пути, поскольку завершение предоставления услуги может произойти не только по причине истечения времени передачи соответствующей информации, но и по причине преждевременного завершения сеанса связи со стороны пользователей. Построение математической модели сети будем проводить в предположении о независимости функционирования физических путей, т. е. не будем учитывать некоторые особенности протоколов мультивещания в части

реализации функций установления, поддержания и разъединения соединений (сигнализации).

Построение модели начнем с описания поведения логического пути. На рис. 3.5 изображена схема функционирования  $(m, p, s)$ -пути, на которой показаны его «жизненные циклы», состоящие из двух периодов:  $(m, p, s)$ -путь включен ( $x_{mps} = 1$ ) и  $(m, p, s)$ -путь выключен ( $x_{mps} = 0$ ).

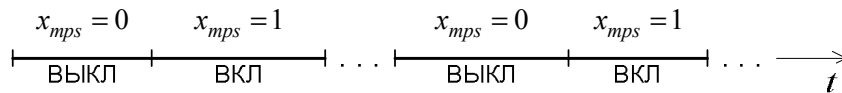


Рис. 3.5. Схема функционирования логического пути

Более детально «жизненный цикл» логического пути показан на рис. 3.6.

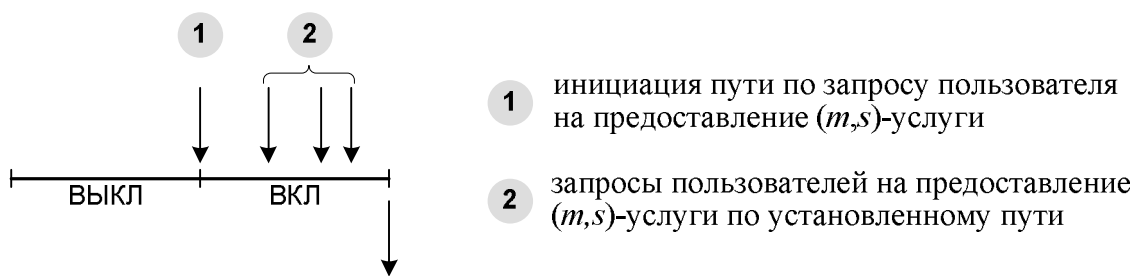


Рис. 3.6. «Жизненный цикл» логического пути

Будем рассматривать систему в простейших предположениях о поведении логического пути. Пусть все логические пути функционируют независимо друг от друга, поток запросов на установление  $(m, p, s)$ -пути является пуассоновским с параметром  $\lambda_{mps}$ , а время предоставления услуги (время занятия  $(m, p, s)$ -пути) является случайной величиной, распределенной по экспоненциальному закону с параметром  $\mu_{mps}$ .

Поведение  $(m, p, s)$ -пути опишем в терминах СМО

$M \mid M \mid 1 \mid 0$  с «прозрачными заявками». В этой СМО заявка (запрос  $\lambda_{mps}$   $\mu_{mps}$ )



пользователя), поступившая в период простоя прибора (период, когда логический путь выключен), обслуживается в течение некоторого случайного интервала времени (период, когда логический путь включен). Все последующие заявки, поступившие в СМО на этом интервале времени, не теряются и не становятся в очередь на обслуживание, а получают его вместе с заявкой, которая инициировала период занятости прибора, и покидают систему вместе с ней.

Пусть  $\xi_{mps}(t)$  – число заявок в СМО в момент времени  $t \geq 0$ ,  $\xi_{mps}(t) \in \{0, 1, \dots\}$ . Процесс  $\xi_{mps}(t)$  является марковским случайным процессом, у которого при любых значениях  $\lambda_{mps} > 0$  и  $\mu_{mps} > 0$  существует стационарное распределение вероятностей

$$p_n = P\{\xi_{mps}(t) = n\}, \quad n \geq 0,$$

вида

$$p_n = \frac{1}{1 + \rho_{mps}} \left( \frac{\rho_{mps}}{1 + \rho_{mps}} \right)^n, \quad n \geq 0,$$

где  $\rho_{mps} = \lambda_{mps} / \mu_{mps}$ .

Введем случайный процесс  $\{X_{mps}(t), t \geq 0\}$ ,  $X_{mps}(t) \in \{0, 1\}$ , описывающий поведение  $(m, p, s)$ -пути, и его стационарное распределение

$$\pi_{mps}(x_{mps}) = P\{X_{mps}(t) = x_{mps}\}, \quad x_{mps} \in \{0, 1\}.$$

Процесс  $\{X_{mps}(t), t \geq 0\}$  является марковским, а его стационарное распределение связано с распределением МП  $\xi_{mps}(t)$  и имеет вид

$$\begin{aligned} \pi_{mps}(0) &= p_0 = \frac{1}{1 + \rho_{mps}}, \\ \pi_{mps}(1) &= \sum_{n \geq 1} p_n = \frac{\rho_{mps}}{1 + \rho_{mps}}. \end{aligned} \tag{1.3}$$

Стационарное распределение МП  $\{X_{mps}(t), t \geq 0\}$  удовлетворяет уравнениям детального баланса

$$\pi_{mps}(0)\lambda_{mps} = \pi_{mps}(1)\mu_{mps},$$

поэтому процесс  $X_{mps}(t)$  является обратимым марковским процессом (ОМП). Отметим, что построенная модель логического пути не учитывает возможных потерь запросов из-за блокировок и, следовательно,  $\pi_{mps}(1)$  можно рассматривать также как предложенную нагрузку на  $(m, p, s)$ -путь.

Для описания функционирования всех логических путей сети введем составной процесс  $\tilde{X}(t) = (X_{mps}(t))_{m \in M_s, p \in P_s, s \in S}$ , который по построению является ОМП на множестве  $\tilde{X}$  со стационарным распределением

$$\tilde{\pi}(\mathbf{x}) = \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \pi_{mps}(x_{mps}), \quad \mathbf{x} \in \tilde{X},$$

причем

$$\tilde{\pi}(\vec{0}) = \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \frac{1}{1 + \rho_{mps}}.$$

Поведение системы с учетом ограниченной емкости звеньев может быть описано сужением этого процесса на множество  $X$ . Пусть случайный процесс  $X(t)$  является сужением ОМП  $\tilde{X}(t)$  на множество  $X \subseteq \tilde{X}$ .

**Теорема 3.1.** *Равновесное распределение  $\pi(\mathbf{x}) = P\{X(t) = \mathbf{x}\}$  случайного процесса  $\{X(t), t \geq 0\}$  имеет мультипликативный вид*

$$\begin{aligned} \pi(\mathbf{x}) &= \pi(\vec{0}) \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}}, \quad \mathbf{x} \in X, \\ \pi^{-1}(\vec{0}) &= \sum_{\mathbf{x} \in X} \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}}. \end{aligned} \tag{1.4}$$

**Доказательство.** Из основного свойства сужения ОМП следует, что

$$\begin{aligned}
\pi(\mathbf{x}) &= \frac{\tilde{\pi}(\mathbf{x})}{\sum_{\mathbf{x} \in X} \tilde{\pi}(\mathbf{x})} = \frac{\prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \pi_{mps}(x_{mps})}{\sum_{\mathbf{x} \in X} \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \pi_{mps}(x_{mps})} = \\
&= \frac{\prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \frac{\rho_{mps}^{x_{mps}}}{1 + \rho_{mps}}}{\sum_{\mathbf{x} \in X} \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \frac{\rho_{mps}^{x_{mps}}}{1 + \rho_{mps}}} = \frac{\prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}}}{\sum_{\mathbf{x} \in X} \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}}} = \\
&= \pi(\vec{0}) \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}}, \quad \mathbf{x} \in X. \quad \blacksquare
\end{aligned}$$

### 3.1.3. Вероятностные характеристики модели

Зная стационарное распределение вероятностей состояний системы, можно получить выражения для основных ее вероятностных характеристик. Для удобства записи для любого множества  $\Omega \subseteq X$  определим функцию  $G(\Omega)$  соотношением

$$G(\Omega) = \sum_{\mathbf{x} \in \Omega} \prod_{s \in S} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}}. \quad (1.5)$$

Теперь нормирующую константу распределения (1.3) можно представить в виде

$$\pi^{-1}(\vec{0}) = G(X).$$

Выражения для многих важных характеристик сети, которые задаются вероятностью некоторого события, т.е. подмножества  $\Omega$  пространства состояний системы  $X$ , могут быть получены посредством функции  $G(\Omega)$  следующим образом:

$$P\{\mathbf{x} \in \Omega\} = \sum_{\mathbf{x} \in \Omega} \pi(\mathbf{x}) = \frac{G(\Omega)}{G(X)}. \quad (1.6)$$

К таким характеристикам относятся *вероятности блокировки* установления соединений и ряд других характеристик, к рассмотрению которых мы переходим.

Пусть  $B_{mps}$  – множество состояний блокировок  $(m, p, s)$ -пути, т.е. подмножество состояний логических путей сети, в которых запрос пользователя на предоставление  $(m, s)$ -услуги по физическому пути  $p \in P_s$  блокируется по причине нехватки свободных ресурсов звеньев. 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

$$B_{mps} = \{ \mathbf{x} \in X : \exists l \in L_{ps} : y_{ms}^l(\mathbf{x}) = 0, c_l(\mathbf{x}) + b_{ms} > C_l \}. \quad (1.7)$$

Значение вероятности  $B_{mps} = P\{ \mathbf{x} \in B_{mps} \}$  события, заданного соотношением (1.7), можно получить по формуле (1.6). Заметим, что речь идет о блокировках соединений по времени (в отличие от блокировок по вызовам – см. раздел 1.4.5).

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

$$F_{mps} = \{ \mathbf{x} \in X : x_{mps} = 1 \} \quad (1.8)$$

и

$$H_{mps} = \{ \mathbf{x} \in X : x_{mps} = 0, \forall l \in L_{ps} c_l(\mathbf{x}) + b_{ms} \leq C_l \vee y_{ms}^l(\mathbf{x}) = 1 \}. \quad (1.9)$$

Вычислить соответствующие вероятности  $F_{mps} = P\{ \mathbf{x} \in F_{mps} \}$  и  $H_{mps} = P\{ \mathbf{x} \in H_{mps} \}$  вновь можно по формуле (1.6). 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  $X$ , поэтому вероятности этих событий связаны соотношением

$$B_{mps} + F_{mps} + H_{mps} = 1. \quad (1.10)$$

**Упражнение 3.3.** Докажите, что для любой тройки  $(m, p, s)$ ,  $m \in M_s$ ,  $p \in P_s$ ,  $s \in S$ , выполняется соотношение

$$F_{mps} = \rho_{mps} H_{mps}. \quad (1.11)$$

**Решение.** Зафиксируем некоторую тройку  $(\hat{m}, \hat{p}, \hat{s})$ ,  $\hat{m} \in M_s$ ,  $\hat{p} \in P_s$ ,  $\hat{s} \in S$ . По определению множества  $H_{mps}$  соответствующая  $(\hat{m}, \hat{p}, \hat{s})$ -пути компонента для любого  $\mathbf{x} \in H_{\hat{m}\hat{p}\hat{s}}$  равна нулю, т.е.  $x_{\hat{m}\hat{p}\hat{s}} = 0$ , и для каждого  $l \in L_{\hat{p}\hat{s}}$  выполняется соотношение

$$\tilde{c} = \sum_{s \in S \setminus \{\hat{s}\}} \sum_{m \in M_s} b_{ms} y_{ms}^l(\mathbf{x}) + \sum_{m \in M_s \setminus \{\hat{m}\}} b_{m\hat{s}} y_{m\hat{s}}^l(\mathbf{x}) \leq C_l - b_{\hat{m}\hat{s}}.$$

Здесь  $c_l(\mathbf{x}) = \tilde{c}$ , если  $y_{\hat{m}\hat{s}}^l(\mathbf{x}) = 0$ , и  $c_l(\mathbf{x}) = \tilde{c} + b_{\hat{m}\hat{s}}$  в противном случае.

Для любого  $\mathbf{x} \in H_{\hat{m}\hat{p}\hat{s}}$  состояние сети  $\mathbf{x}' = (\mathbf{x} + \mathbf{1}_{\hat{m}\hat{p}\hat{s}})$  принадлежит множеству  $F_{\hat{m}\hat{p}\hat{s}}$  (здесь  $\mathbf{1}_{\hat{m}\hat{p}\hat{s}}$  – вектор,  $(\hat{m}, \hat{p}, \hat{s})$ -я компонента которого равна 1, а остальные – 0). Действительно:  $x'_{\hat{m}\hat{p}\hat{s}} = 1$ , для  $l \in L \setminus L_{\hat{p}\hat{s}}$  выполняется равенство  $c_l(\mathbf{x}') = c_l(\mathbf{x})$ , и для всех  $l \in L_{\hat{p}\hat{s}}$

$$\begin{aligned} c_l(\mathbf{x}') &= \sum_{s \in S \setminus \{\hat{s}\}} \sum_{m \in M_s} b_{ms} y_{ms}^l(\mathbf{x}) + \sum_{m \in M_s \setminus \{\hat{m}\}} b_{m\hat{s}} y_{m\hat{s}}^l(\mathbf{x}) + b_{\hat{m}\hat{s}} = \\ &= \tilde{c} + b_{\hat{m}\hat{s}} \leq C_l. \end{aligned}$$

С другой стороны, для любого  $\mathbf{x}' \in F_{\hat{m}\hat{p}\hat{s}}$  найдется такое состояние  $\mathbf{x} \in H_{\hat{m}\hat{p}\hat{s}}$ , что  $\mathbf{x}' = (\mathbf{x} + \mathbf{1}_{\hat{m}\hat{p}\hat{s}})$ . Таким образом, отображение  $\mathbf{x}' = \varphi_{\hat{m}\hat{p}\hat{s}}(\mathbf{x}) = (\mathbf{x} + \mathbf{1}_{\hat{m}\hat{p}\hat{s}})$  является биекцией из  $H_{\hat{m}\hat{p}\hat{s}}$  в  $F_{\hat{m}\hat{p}\hat{s}}$ .

Переходя к вероятностям, получаем для любого  $\mathbf{x}' \in F_{\hat{m}\hat{p}\hat{s}}$

$$\begin{aligned} \pi(\mathbf{x}') &= G^{-1}(\mathbf{X}) \prod_{s \in \mathbf{S} \setminus \{\hat{s}\}} \prod_{p \in P_s} \prod_{m \in M_s} \rho_{mps}^{x_{mps}} \prod_{p \in P_{\hat{s}} \setminus \{\hat{p}\}} \prod_{m \in M_{\hat{s}}} \rho_{mp\hat{s}}^{x_{mp\hat{s}}} \times \\ &\times \prod_{m \in M_{\hat{s}} \setminus \{\hat{m}\}} \rho_{m\hat{p}\hat{s}}^{x_{m\hat{p}\hat{s}}} \rho_{\hat{m}\hat{p}\hat{s}} = \rho_{\hat{m}\hat{p}\hat{s}} \pi(\mathbf{x}), \end{aligned}$$

где  $\mathbf{x} = (\mathbf{x}' - \mathbf{1}_{\hat{m}\hat{p}\hat{s}}) \in H_{\hat{m}\hat{p}\hat{s}}$ , и, окончательно,

$$\begin{aligned} F_{\hat{m}\hat{p}\hat{s}} &= \sum_{\mathbf{x}' \in F_{\hat{m}\hat{p}\hat{s}}} \pi(\mathbf{x}') = \sum_{\mathbf{x} \in H_{\hat{m}\hat{p}\hat{s}}} \pi(\mathbf{x} + \mathbf{1}_{\hat{m}\hat{p}\hat{s}}) = \\ &= \rho_{\hat{m}\hat{p}\hat{s}} \sum_{\mathbf{x} \in H_{\hat{m}\hat{p}\hat{s}}} \pi(\mathbf{x}) = \rho_{\hat{m}\hat{p}\hat{s}} H_{\hat{m}\hat{p}\hat{s}}. \quad \blacksquare \end{aligned}$$

**Упражнение 3.4.** Докажите, что для любого  $(m, p, s)$ -пути,  $m \in M_s$ ,  $p \in P_s$ ,  $s \in \mathbf{S}$ , верно соотношение

$$F_{mps} = \frac{\rho_{mps}}{1 + \rho_{mps}} (1 - B_{mps}). \quad (1.12)$$

**Решение.** Для доказательства достаточно воспользоваться упражнением 3.3 и соотношением (1.10).  $\blacksquare$

Аналогично модели сети с одноадресными соединениями, полученной формулой можно воспользоваться для вычисления вероятностных характеристик небольших сетей, однако тот факт, что мощность пространства состояний может доходить до  $2^{\sum_{s \in \mathbf{S}} P_s M_s}$ , затрудняет ее применение при больших значениях соответствующих параметров. В связи с этим исследования ведутся в двух направлениях: разработка эффективных точных алгоритмов вычисления нормирующей константы  $G(\mathbf{X})$  для некоторых важных для практики разновидностей сетевых

топологий и поиск методов приближенной оценки вероятностных характеристик сети произвольной конфигурации.

## § 3.2. Модель отдельного звена

### 3.2.1. Построение модели

Рассмотрим сеть мультивещания, в которой для некоторого звена  $l^* \in L$  выполняются соотношения

$$\sum_{s \in S^{l^*}} \sum_{m \in M_s} b_{ms} > C_{l^*};$$

$$\sum_{s \in S^l} \sum_{m \in M_s} b_{ms} \leq C_l, \quad l \in L \setminus \{l^*\}.$$

Таким образом, все звенья, кроме звена  $l^*$ , имеют неограниченные ресурсы для обслуживания запросов пользователей. Такую сеть иногда называют сетью с выделенным звеном. Очевидно, что задача анализа блокировок в сети с выделенным звеном сводится к анализу сети, состоящей из одного звена  $L = \{l^*\}$  и имеющей один источник рассылки, который предоставляет услуги из множества  $M = \bigcup_{s \in S^{l^*}} M_s$ . Для краткости записи

индексы  $l^*$  и  $s$  далее в этом разделе не используются.

Функционирование сети с выделенным звеном будем описывать с помощью многопоточковой мультисервисной системы массового обслуживания. На полнодоступную систему, состоящую из  $C = C_{l^*}$  приборов (единиц емкости звена сети) и не имеющую накопителя, поступают  $M = |M|$  потоков заявок. Будем считать, что все поступающие потоки являются пуассоновскими и независимы в совокупности. Если на момент поступления заявки  $m$ -потока ( $m$ -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  $\mu_m$  и не зависящее ни от длительности обслуживания заявок других потоков, ни от процессов поступления. Все поступившие в течение этого интервала времени  $m$ -заявки принимаются на обслуживания без выделения дополнительных приборов, а по истечении указанного интервала одновременно покидают систему и  $b_m$  приборов освобождаются. Потеря заявки происходит в том случае, если при ее поступлении в системе нет заявок того же потока, а также нет достаточного количества свободных приборов.

Обозначим  $\rho_m = \lambda_m / \mu_m$ , где  $\lambda_1, \dots, \lambda_M$  – интенсивности входящих потоков. Согласно [16] параметры  $\rho_1, \dots, \rho_M$  связаны с интенсивностями потоков запросов пользователей на включение соответствующих логических путей в сети соотношением

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

Положим  $C = \infty$ , в этом случае все поступившие в систему заявки принимаются на обслуживание и потери отсутствуют. Пусть случайный процесс  $\{Y_m(t), t \geq 0\}$ ,  $m = 1, \dots, M$ , находится в состоянии 1, если в момент времени  $t \geq 0$  в системе обслуживается хотя бы одна  $m$ -заявка, и в состоянии 0 в противном случае. Аналогично случайному процессу  $\{X_{mps}(t), t \geq 0\}$ , введенному в §3.1, процесс  $\{Y_m(t), t \geq 0\}$  является ОМП и имеет стационарное распределение

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

Рассмотрим составной случайный процесс  $\{\tilde{Y}(t) = (Y_m(t))_{m \in M}, t \geq 0\}$ , заданный на множестве  $\tilde{Y} = \{0, 1\}^M$ . По построению этот процесс является ОМП на данном множестве и, как следует из формулы (2.2), имеет стационарное распределение



$$\tilde{\pi}(\mathbf{y}) = G^{-1}(\tilde{Y}) \prod_{m=1}^M \rho_m^{y_m}, \quad \mathbf{y} \in \tilde{Y}, \quad (2.3)$$

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

$$G(\Omega) = \sum_{\mathbf{y} \in \Omega} \prod_{m=1}^M \rho_m^{y_m}. \quad (2.4)$$

Из (2.4) следует, что нормирующая константа  $G(\tilde{Y})$  распределения вероятностей процесса  $\{\tilde{Y}(t), t \geq 0\}$  равна

$$G(\tilde{Y}) = \prod_{m=1}^M (1 + \rho_m). \quad (2.5)$$

Процесс  $\{\tilde{Y}(t), t \geq 0\}$  с пространством состояний  $\tilde{Y}$  и распределением вероятностей (2.3) описывает состояние рассматриваемой системы при  $C = \infty$ . Пусть теперь  $C < \infty$  и, следовательно, возможны потери заявок. Будем считать, что потерянные заявки не оказывают влияние на интенсивность породившего их потока, т.е. система функционирует с явными потерями. В этом случае функционирование системы описывает случайный процесс  $\{Y(t), t \geq 0\}$ , являющийся сужением процесса  $\{\tilde{Y}(t), t \geq 0\}$  на множество

$$Y = \{\mathbf{y} \in \tilde{Y} : c(\mathbf{y}) \leq C\}, \quad (2.6)$$

где  $c(\mathbf{y}) = \sum_{m=1}^M b_m y_m$  – число занятых приборов системы в состоянии  $\mathbf{y} \in \tilde{Y}$ .

Как сужение обратимого процесса  $Y(t)$  также обратим, и, следовательно, справедлива следующая теорема.

**Теорема 3.2.** *Стационарное распределение вероятностей состояний процесса  $\{Y(t), t \geq 0\}$  имеет мультипликативный вид*

$$\pi(\mathbf{y}) = G^{-1}(Y) \prod_{m=1}^M \rho_m^{y_m}, \quad \mathbf{y} \in Y, \quad (2.7)$$



принята на обслуживание. Напомним, что условием потери заявки помимо недостаточного числа свободных приборов является отсутствие в системе заявок данного потока. Следовательно, множество потерь  $m$ -заявок имеет вид

$$B_m = \{y \in Y : c(y) + b_m > C, y_m = 0\}. \quad (2.9)$$

Множество таких состояний, что  $m$ -заявка находится в системе, имеет вид

$$F_m = \{y \in Y : y_m = 1\}, \quad (2.10)$$

тогда как множество таких состояний, что  $m$ -заявок в системе нет, но если заявка поступит, то будет принята на обслуживание, принимает вид

$$H_m = \{y \in Y : c(y) + b_m \leq C, y_m = 0\}. \quad (2.11)$$

Физический смысл первого события состоит в том, что соответствующие  $m$ -услуге данные передаются через рассматриваемое звено. Во втором случае  $m$ -услуга через звено не предоставляется, но ресурсов достаточно, чтобы по запросу пользователя инициировать ее предоставление.

Легко видеть, что и здесь для любого  $m = 1, \dots, M$  система множеств  $B_m, F_m, H_m$  является разбиением пространства состояний  $Y$ . Следовательно, вероятности данных событий связаны соотношением

$$B_m + F_m + H_m = 1. \quad (2.12)$$

**Упражнение 3.6.** Докажите соотношения

$$F_m = \rho_m H_m \quad (2.13)$$

и

$$F_m = \frac{\rho_m}{1 + \rho_m} (1 - B_m). \quad (2.14)$$

**Решение.** Соотношения (2.13) и (2.14), устанавливающие дополнительную связь между вероятностями  $B_m, F_m$  и  $H_m$ , доказываются аналогично упражнениям 3.3 и 3.4 из раздела 3.1.3 для всей сети. ■

При анализе отдельного звена сети мультивещания интерес представляет характеристика СВ  $\gamma$ , принимающая значение  $c(\mathbf{y})$ .  $\gamma$  является случайной величиной числа занятых приборов в рассматриваемой системе и соответствует случайному числу занятых единиц емкости звена сети. Если за единицу емкости принять величину одной передаточной единицы, то  $\gamma$  представляет собой СВ ширины полосы пропускания (ШПП), занятой на звене сети при обслуживании установленных через него соединений. Среднее значение занятой ШПП, т.е. среднее число занятых приборов в рассматриваемой модели, можно найти как математическое ожидание  $c^{(1)}$  СВ  $\gamma$ , а именно

$$c^{(1)} = \sum_{\mathbf{y} \in Y} c(\mathbf{y})\pi(\mathbf{y}). \quad (2.15)$$

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

**Упражнение 3.7.** Покажите, что среднее число занятых приборов определяется по формуле

$$c^{(1)} = \sum_{m=1}^M b_m F_m. \quad (2.16)$$

**Решение.**

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

**Упражнение 3.8.** Покажите, что среднее число занятых приборов в системе выражается формулой

$$c^{(1)} = \sum_{m=1}^M b_m \frac{\rho_m}{1 + \rho_m} (1 - B_m). \quad (2.17)$$

**Решение.** Утверждение вытекает из соотношений (1.12) и (2.16). ■

Необходимо заметить, что вычисление введенных вероятностных характеристик непосредственно по представленным формулам представляет собой непростую задачу, поскольку подразумевает перебор всего пространства состояний  $Y$ , имеющего комбинаторный характер и, при решении практических задач, большую размерность. В следующем пункте мы рассмотрим более эффективный метод численного анализа рассматриваемой системы, предложенный в [16].

### 3.2.3. Алгоритм свертки

Известно большое количество методов и алгоритмов, которые могут быть использованы для вычисления вероятностных характеристик рассматриваемого класса моделей. К таким методам относятся модифицированный метод Бузена, эффективные сверточные алгоритмы, метод усеченных сверток, обобщенная рекурсия Кофмана-Робертса<sup>1</sup> и др.

Для вывода алгоритма расчета вероятностных характеристик отдельного звена сети мультимедийного вещания прежде всего необходимо исследовать свойства множества  $Y$  и получить алгоритм для расчета нормирующей константы  $G(Y)$ .

Введем для  $m \in M$  и  $n = 0, \dots, C$  множества

$$Y(m, n) = \{ \mathbf{y}(m) = (y_1, \dots, y_m) : c(\mathbf{y}(m)) = n \}.$$

Доопределим данную систему множеств для значений  $m = 0$  и  $n < 0$  следующим образом:

---

<sup>1</sup> 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. [9, 13]).

$$Y(m, n) = \begin{cases} Y(m, n), & m = 1, \dots, M, n = 1, \dots, C; \\ \{0\}, & m = 0, \dots, M, n = 0; \\ \emptyset, & m = 0, n = 1, \dots, C; \\ \emptyset, & m = 0, \dots, M, n < 0. \end{cases} \quad (2.18)$$

По построению множества  $Y(m, n)$  удовлетворяют соотношениям

$$Y(m, n) \cap Y(m, \tilde{n}) = \emptyset, \quad n \neq \tilde{n}, \quad \text{для любого } m = 0, \dots, M \text{ и } Y = \bigcup_{n=0}^C Y(M, n). \quad \text{Для}$$

всех  $m = 1, \dots, M$  и  $n = 1, \dots, C$  множество  $Y(m, n)$  представимо в виде

$$Y(m, n) = Y(m-1, n) \times \{0\} \cup Y(m-1, n-b_m) \times \{1\}. \quad (2.19)$$

Введем функцию

$$g(m, n) = \sum_{y(m) \in Y(m, n)} \prod_{i=1}^m \rho_i^{y_i}$$

и заметим, что

$$G(Y) = \sum_{n=0}^C g(M, n). \quad (2.20)$$

**Лемма 3.1.** *Функцию  $g(m, n)$  можно вычислить по формуле*

$$g(m, n) = \begin{cases} 0, & m = 0, n = 1, \dots, C; \\ 0, & m = 0, \dots, M, n < 0; \\ 1, & m = 0, \dots, M, n = 0; \\ g(m-1, n) + \rho_m g(m-1, n-b_m), & \begin{matrix} m = 1, \dots, M, \\ n = 1, \dots, C. \end{matrix} \end{cases} \quad (2.21)$$

**Доказательство.** Первые три строки формулы очевидным образом следуют из (2.18). Докажем утверждение четвертой строки. В силу (2.19) имеем:

$$g(m, n) = \sum_{y(m) \in Y(m, n)} \prod_{i=1}^m \rho_i^{y_i} =$$

$$\begin{aligned}
&= \sum_{\mathbf{y}^{(m)} \in Y^{(m-1, n)} \times \{0\}} \prod_{i=1}^m \rho_i^{y_i} + \sum_{\mathbf{y}^{(m)} \in Y^{(m-1, n-b_k)} \times \{1\}} \prod_{i=1}^m \rho_i^{y_i} = \\
&= \sum_{\mathbf{y}^{(m-1)} \in Y^{(m-1, n)}} \prod_{i=1}^m \rho_i^{y_i} + \left( \sum_{\mathbf{y}^{(m-1)} \in Y^{(m-1, n-b_m)}} \prod_{i=1}^m \rho_i^{y_i} \right) \rho_m = \\
&= g(m-1, n) + \rho_m g(m-1, n-b_k). \quad \blacksquare
\end{aligned}$$

Лемма 3.1 и формула (2.20) определяют алгоритм для расчета нормирующей константы  $G(Y)$ . Вывод формул для расчета вероятностных характеристик проведем для услуги с номером  $M$ . Это не ограничивает общности, поскольку всегда можно перенумеровать услуги и после перенумерации провести необходимые вычисления.

**Теорема 3.3.** *Вероятностные характеристики  $B_M$ ,  $F_M$ ,  $H_M$  и  $c^{(1)}$  звена сети мультимещания вычисляются по формулам*

$$B_M = \left( \sum_{n=0}^C g(M, n) \right)^{-1} \sum_{n=C-b_M+1}^C g(M-1, n). \quad (2.22)$$

$$F_M = \left( \sum_{n=0}^C g(M, n) \right)^{-1} \rho_M \sum_{n=0}^{C-b_M} g(M-1, n), \quad (2.23)$$

$$H_M = \left( \sum_{n=0}^C g(M, n) \right)^{-1} \sum_{n=0}^{C-b_M} g(M-1, n), \quad (2.24)$$

$$c^{(1)} = \left( \sum_{n=0}^C g(M, n) \right)^{-1} \sum_{n=1}^C n g(M, n). \quad (2.25)$$

**Доказательство.** Докажем формулу (2.22). Множество блокировок представимо в виде

$$\begin{aligned}
B_M &= \{ \mathbf{y} \in \tilde{Y} : C - b_M + 1 \leq c(\mathbf{y}) \leq C, y_M = 0 \} = \\
&= \{ \mathbf{y}^{(M-1)} : C - b_M + 1 \leq c(\mathbf{y}^{(M-1)}) \leq C \} := \hat{Y}^{(M-1)} \times \{0\},
\end{aligned}$$

отсюда, с использованием введенного обозначения,

$$G(\mathbf{B}_M) = \sum_{\mathbf{y}^{(M-1)} \in \hat{Y}^{(M-1)}} \prod_{i=1}^{M-1} \rho_i^{y_i} \cdot 1 = \sum_{n=C-b_M+1}^C g(M-1, n).$$

Из формул (1.6) и (2.20) следует (2.22). ■

**Упражнение 3.9.** Докажите формулы (2.23)–(2.25) по аналогии с доказательством формулы (2.22).

Формулы (2.22)–(2.25) определяют эффективный алгоритм расчета вероятностных характеристик отдельного звена сети мультивещания.



## **Приложение. КОНЦЕПЦИЯ КАЧЕСТВА ОБСЛУЖИВАНИЯ В СЕТЯХ NGN**

Термин NGN (Next Generation Network – сеть следующего поколения) стал встречаться в специальной литературе начиная с 2000 г., однако четкого определения данного понятия не сформировалось до сих пор. Активно использующие этот термин специалисты сходятся в одном: сети NGN основаны на пакетных технологиях передачи и обеспечивают функциональные возможности «Triple Play» (коммерческой концепции, применяемой поставщиками телекоммуникационных услуг для обозначения совместного предоставления пользователю услуг передачи речи, видео и данных). Разработкой международных стандартов для NGN занимаются Международный союз электросвязи (МСЭ, рекомендации серии Y.2000 «Глобальная информационная структура, аспекты протокола Интернет и сети следующего поколения»), проект TISPAN (Telecoms & Internet converged Services & Protocols for Advanced Networks) в рамках Европейского Института по стандартизации в области телекоммуникаций (European Telecommunications Standards Institute, ETSI), а также организация 3rd Generation Partnership Project (3GPP). TISPAN и 3GPP являются авторами концепции IMS (Internet Protocol-based Multimedia Subsystem) и разрабатывают основанное на подсистеме IMS ядро сети NGN, которое может использоваться для построения как беспроводных, так и фиксированных сетей связи. В России вопросами контроля качества предоставления услуг связи занимается Федеральное агентство по надзору в сфере связи (орган Госсвязьнадзора РФ).

Следует понимать, что NGN – не конкретная сеть, а концепция развития сетей связи в направлении создания универсальной сетевой инфраструктуры, которая позволяла бы переносить любые виды информации и предоставлять пользователям любые услуги, независимо от времени и места расположения. Способом реализации концепции NGN

является создание мультисервисных сетей связи, ориентированных на предоставление самого широкого спектра услуг в рамках единой сетевой структуры. На сегодняшний день концепция NGN находится на этапе исследований и становления, причем чаще формулируются требования к NGN и задачи изучения, а не конкретные решения. Приведем определения NGN в соответствии с документами ведущих международных организаций по стандартизации. Под *сетями связи следующего поколения* институт ETSI понимает *«сеть связи на базе технологии коммутации пакетов, способную обеспечивать предоставление услуг связи и пригодную для использования различных транспортных высокоскоростных технологий, обеспечивающих качество обслуживания, в которой функции, связанные с предоставлением услуг, не зависят от технологии транспортной сети»*. В документах МСЭ-Т это определение дополнено следующим образом: *«...NGN обеспечивает неограниченный доступ пользователей к сетям связи и конкурирующим между собой поставщикам услуг и/или услугам связи, выбираемым самими пользователями. Сеть поддерживает универсальную мобильность, которая обеспечивает постоянное и повсеместное предоставление услуг связи пользователям»*.

Ключевыми особенностями сети NGN являются:

- использование режима коммутации пакетов для передачи данных;
- разделение функций управления на функции, связанные с управлением транспортом, управлением вызовами/сессиями и приложениями/услугами;
- отделение процесса предоставления услуг от процесса транспорта, использование открытых интерфейсов;
- поддержка большого набора услуг, приложений и механизмов, основанных на конструктивных блоках, включая потоковые услуги, услуги в режиме реального и нереального времени, мультимедийные услуги;

- поддержка широкополосных технологий со сквозным («из конца в конец», end-to-end) обеспечением качества обслуживания;
- взаимодействие с существующими сетями через открытые интерфейсы;
- мобильность в обобщенном смысле (generalized mobility);
- неограниченный доступ пользователей к различным поставщикам услуг;
- множество схем идентификации абонента;
- одни и те же характеристики для одинаковых с точки зрения пользователя услуг;
- конвергенция услуг мобильных и фиксированных сетей;
- независимость услуго-ориентированных функций от используемых транспортных технологий;
- поддержка различных технологий для реализации сети доступа и др.

## **§П.1. Современная концепция качества обслуживания в сетях связи**

Начнем с пояснения основных терминов, часть из которых широко используется в рекомендациях и международных документах, посвященных вопросам предоставления телекоммуникационных услуг, а также в специальной литературе.

*Инфокоммуникационная услуга / приложение* (далее – *услуга*) – возможность удовлетворить потребность человека в автоматизированной обработке, хранении или предоставлении по запросу информации различного рода с использованием средств вычислительной техники как на входящем, так и на исходящем конце соединения. Примерами услуги/приложения могут служить электронная почта (e-mail), приложение для удаленного доступа типа Telnet, услуга конференции в реальном времени (веб-чат, мгновенный обмен сообщениями), услуга

индивидуального просмотра видеофильма по запросу (video-on-demand) и др.

*Служба* – совокупность организационно-технических мер, позволяющих предоставить услугу.

*Сеть связи* – совокупность ресурсов (оборудования, программных средств, данных), обеспечивающая реализацию служб.

*Пользователь* – физическое лицо или юридическое лицо, являющееся абонентом оператора связи.

*Оператор связи* – предприятие (юридическое лицо), оказывающее услуги связи и имеющее соответствующую лицензию.

*Поставщик услуг* (service provider) – индивидуальный предприниматель или юридическое лицо, оказывающее инфокоммуникационные услуги и не обладающее собственной инфраструктурой связи.

*Сеть доступа* (access network) – сеть связи, обеспечивающая подключение терминальных устройств пользователя к оконечному узлу транспортной сети связи, входящей в состав мультисервисной сети.

*Транспортная сеть* – совокупность всех ресурсов, выполняющих функции транспортирования, т. е. переноса информации между разнесенными пунктами. Она включает не только системы передачи информации, но и относящиеся к ним средства контроля, оперативного переключения, резервирования, управления.

*Традиционная сеть связи* – специализированная сеть связи, такая как телефонная сеть общего пользования (ТфОП), сеть передачи данных общего пользования (СДОП), сеть кабельного телевидения и т. п., изначально предназначенная для предоставления услуг связи одного вида.

*Качество связи* – степень соответствия предоставляемых услуг связи потребностям и ожиданиям пользователя, заказавшего и оплатившего эту услугу.

*Терминал* (абонентский терминал) – оконечная часть сети доступа, например, аналоговый телефонный аппарат, SIP-телефон (SIP – Session Initiation Protocol – протокол инициирования сеансов связи), персональный компьютер, карманный персональный компьютер и т. п.

Сеть следующего поколения, в соответствии с Рекомендацией Y.2011, можно представить в виде двухуровневой модели (рис. П.1):

- уровень услуг (NGN Service Stratum), где реализуются функции передачи пользовательских данных и функции управления и поддержки ресурсов, необходимых для предоставления услуг;
- транспортный уровень (NGN Transport Stratum), реализующий функции управления и эксплуатационной поддержки транспортных ресурсов для передачи данных между терминальными устройствами.

На рис. П.1 показано также взаимное соответствие уровней сети следующего поколения и модели взаимодействия открытых систем (МВОС).

Одними из наиболее актуальных вопросов при предоставлении услуг связи являются вопросы качества обслуживания. Высокий уровень качества важен как для пользователя, так и для поставщика услуг, при этом принципиальным является гарантированность качества предоставляемых услуг, когда пользователь имеет возможность выбирать требуемый ему уровень обслуживания, а поставщик услуг принимает на себя обязательства этот уровень обеспечить.

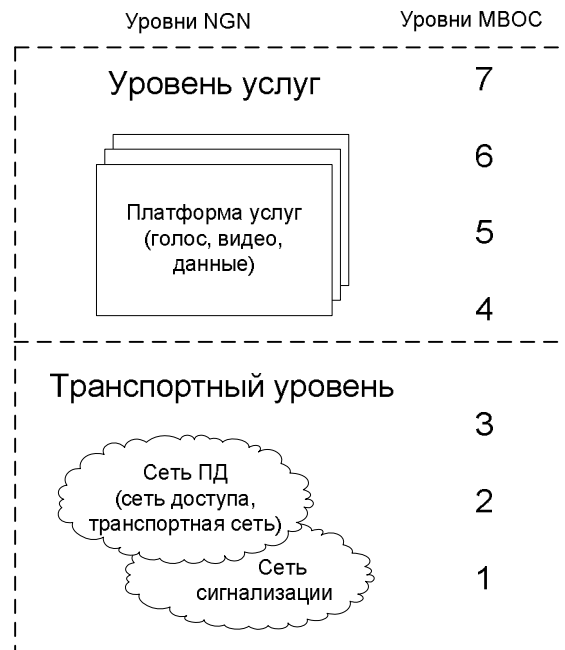


Рис. П.1. Архитектура сети следующего поколения

Оценка качества в сети следующего поколения производится на трех уровнях:

- на уровне пользователя оцениваются показатели субъективного мнения человека, например субъективная оценка качества восприятия отдельного вида информации;
- на уровне услуг оцениваются различные аспекты качества услуги, такие как скорость передачи данных, механизмы кодирования и многое другое;
- на транспортном уровне оценивается качество функционирования сети: задержки, потери, вариация задержки и т. д.

На каждом уровне определены соответствующие параметры оценки качества: показатели качества восприятия (Quality of Experience, QoE) на уровне пользователя, показатели качества обслуживания (Quality of Service, QoS) на уровне услуг и показатели качества функционирования сети (Network Performance, NP) на транспортном уровне. На рис. П.2 показана трехуровневая модель оценки качества и соответствующие параметры оценки.

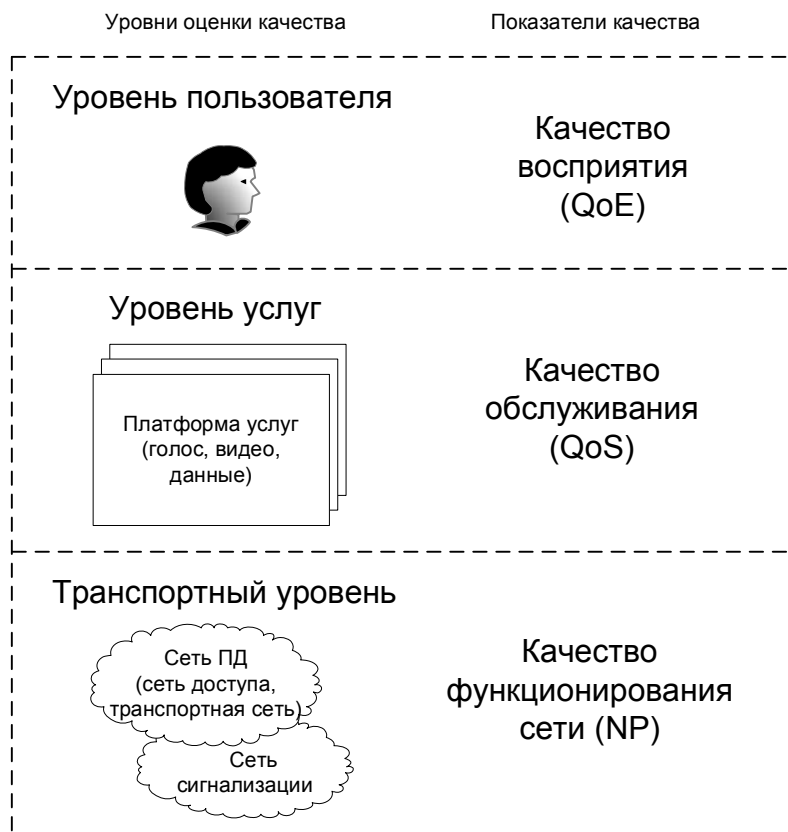


Рис. П.2. Трехуровневая модель оценки качества

*Качество восприятия* определяется как общая приемлемость услуги или приложения с точки зрения конечного пользователя. Влияние на оценку пользователя может оказывать как эффективность функционирования всех элементов системы, включая терминалы, сеть, клиентское оборудование, инфраструктуру сервисов, так и субъективные факторы, такие как ожидание пользователя, связанные с предоставляемой ему услугой, и контекст применения, а также параметры его зрения и слуха. Пользователь интуитивно оценивает качество данной услуги, сравнивая его с качеством подобных услуг других операторов. Качество восприятия с точки зрения пользователя может быть выражено совокупностью параметров. Эти параметры описываются в терминах, понятных как службе, предоставляющей услугу, так и пользователю, и не зависят от структуры сети. Они ориентированы преимущественно на эффект, воспринимаемый пользователем, должны быть гарантированы

пользователю службой и поддаваться объективному измерению в точке доступа к услуге.

*Качество обслуживания* согласно Рекомендации E.800 представляет собой «суммарный эффект показателей качества услуги, который определяет степень удовлетворенности пользователя услуги». Показатели QoS специфицируют характеристики и свойства конкретных приложений, однако требования для различных приложений могут отличаться. Например, для телемедицины точность доставки информации более важна, чем вариация задержки передачи, в то время как для IP-телефонии значение и вариация задержки являются ключевыми параметрами и должны быть минимизированы. Термин «качество обслуживания» также относится к совокупности сетевых технологий (механизмов QoS), целью которых является предоставление поставщику возможности управлять уровнем качества предоставляемых им услуг. С помощью механизмов QoS поставщик может распределить ресурсы сети в зависимости от потребностей той или иной услуги, а также того или иного пользователя, и снизить нагрузку на сеть.

*Качество функционирования сети* согласно Рекомендации I.350 измеряется посредством параметров, которые рассматриваются оператором связи и используются при разработке, конфигурации, эксплуатации и техническом обслуживании сети. Такие показатели определяются независимо от производительности окончного оборудования и действий пользователя, но зависят от используемой сетевой технологии. Для каждой сетевой технологии определяется система уровней качества обслуживания, описываемых с помощью наборов требований, которые носят название классов QoS этой технологии. Разработаны сетевые классы QoS для протоколов IP, ATM (Asynchronous Transfer Mode), Frame Relay и т. д.



На рис. П.3 представлена общая эталонная конфигурация QoE, QoS и NP для сетей NGN, в таблице П.1 приведено краткое описание этих концепций.



Рис. П.3. Общая эталонная конфигурация QoE, QoS и NP для сетей NGN

Концепции QoE, QoS и NP оценивают качество на разных уровнях и с различных точек зрения. Параметры качества восприятия носят в большей степени субъективный характер и зависят от действий и желаний пользователя. Для оператора сети наибольший интерес представляют параметры NP, детализирующие аспекты QoS и представляющие собой проекцию общих требований к функционированию системы на уровень сети и отдельные сетевые технологии. Параметры качества функционирования сети могут быть определены, измерены или вычислены, а также контролируются оператором.

Параметры качества обслуживания определяются с учетом влияния всех компонентов и участков сети, в том числе с учетом влияния сетей различных операторов, поэтому под термином QoS подразумевается прежде всего сквозное качество, или качество «из конца в конец». Параметры качества функционирования сети могут определяться как для всей сети, так и для отдельных ее участков и объектов, например: качество функционирования транспортной сети, абонентского участка сети и т. д.

Таблица П.1. Общая характеристика концепций QoE, QoS и NP

Хар-ка	Уровень NGN	Краткое описание	Примеры показателей
QoE	Уровень пользователя	Общая приемлемость услуги с точки зрения конечного пользователя	Субъективная оценка качества восприятия отдельного вида информации (например, громкость, разборчивость при передаче речи) по 5-балльной шкале MOS (Mean Opinion Score)
QoS	Уровень услуг	Суммарный эффект показателей производительности услуги, который определяет степень удовлетворенности пользователя услугой	Скорость передачи данных, скорость механизмов кодирования, доступность обслуживания
NP	Транспортный уровень	Характеристика работы сети, измеряемая посредством параметров, рассматриваемых оператором и используемых для проектирования, настройки и эксплуатации сети	Односторонняя задержка передачи по сети, коэффициент потери пакетов

Взаимосвязь между показателями QoS и NP может определяться эмпирическим путем, если для этого не существует аналитической методики. Знание взаимосвязи между показателями QoS и NP позволяет, с одной стороны, по измеренным значениям параметров NP предсказать значения параметров QoS или, с другой стороны, по целевым значениям параметров QoS определить требуемые для их поддержания значения параметров NP. Для удовлетворения требований международных стандартов и рекомендаций и для поддержания конкурентоспособности услуг в МСС необходимо учитывать потребности пользователя в услугах с

определенным уровнем качества. В связи с этим при расчете и проектировании мультисервисных сетей связи целесообразно использовать второй подход и при определении показателей NP отталкиваться от нормированных значений показателей QoS.

## **§II.2. Характеристики качества функционирования сети**

### ***II.2.1. Методика измерения параметров NP на уровне IP***

Транспортная сеть NGN строится на основе пакетных технологий передачи информации. Задачей транспортной сети является прозрачная передача информации – как пользовательской, так и управляющей. Основными технологиями построения транспортных сетей NGN являются IP/MPLS (Multiprotocol Label Switching) и ATM, причем технология ATM, как правило, используется на канальном уровне в качестве транспорта для IP. Измерения и параметры качества функционирования на уровне IP позволяют определить эталонные величины для требований к сети, которые не зависят от основных технологий передачи данных и подходят для использования при сквозной оценке качества. Поэтому мы ограничимся рассмотрением характеристик качества функционирования сети на уровне IP.

Существует два основных подхода к измерению качества функционирования сети – активное (active, intrusive) и пассивное (passive, non-intrusive). Активное измерение производится с помощью тестового потока в условиях обычного функционирования сети. Такой тип измерения позволяет детально выделить характеристики NP, например, время односторонней задержки, влияние размера блока данных и т. д., однако активное измерение вносит дополнительную нагрузку на сеть. Пассивное измерение производится с помощью сбора информации в узлах сети (маршрутизаторах уровня IP, коммутаторах ATM и Ethernet) с использованием базового протокола управления сетью (Simple Network

Management Protocol, SNMP) и других технологий. Метод не вносит дополнительной нагрузки на сеть и позволяет производить измерения для каждого устройства или звена сети, однако измерения могут быть ограничены в рамках одного домена или сети в связи с использованием разных протоколов.

Согласно общей модели качества функционирования сети на уровне IP, определенной в Рекомендации Y.1540, основными сетевыми компонентами являются хост, маршрутизатор, хост-источник, хост-получатель и звено. Под *точкой измерения* (Measurement point, MP) понимается граница между хостом и смежным звеном, на которой можно констатировать *эталонные события* (Reference Event, RE) и произвести измерение характеристик передачи. Примерами эталонного события могут служить такие события, как «пакет покидает хост» или «пакет достигает хоста». Передача пакета констатируется, когда пакет проходит точку измерения, при условии что проверка контрольной суммы его заголовка стандартными методами дает положительный результат и значения адресных полей заголовка соответствуют IP-адресам ожидаемых источника и получателя.

Также важными понятиями модели являются сегмент сети (network section) и звено обмена (exchange link). Под *сегментом сети* понимается совокупность хостов и соединяющих их звеньев, которые совместно обеспечивают передачу информации на уровне IP между хостом-источником и хостом-получателем и находятся в пределах одной автономной системы. *Звено обмена* – это звено, соединяющее либо хост-источник или хост-получатель с соседним хостом (например, маршрутизатором), который может принадлежать иной юрисдикции (в этом случае звено обмена называют также звеном доступа); либо два маршрутизатора, принадлежащих к различным сегментам сети. Звено обмена, сегмент сети, хост-источник и хост-получатель являются

базовыми сегментами сети, ограничивающимися точками измерения. *Ансамбль сегментов* (network section ensemble) – любая совокупность связанных между собой сегментов сети и всех звеньев обмена, их соединяющих.

### ***П.2.2. Основные показатели NP транспортной сети***

Показатели качества функционирования сети подразделяются на четыре основные категории.

#### **1. Задержка передачи в базовом сегменте сети.**

К этой категории относятся два показателя NP уровня IP для версии протокола IPv4: задержка передачи (или переноса) IP пакетов (IP packet transfer delay, IPTD) и вариация задержки IP пакетов, или джиттер задержки (IP packet delay variation, IPDV).

#### **2. Ошибки и потери передачи в базовом сегменте сети.**

Основными показателями NP в этой категории являются доля потерянных IP пакетов, или коэффициент потери IP пакетов (IP packet loss ratio, IPLR), и доля искаженных IP пакетов, или коэффициент ошибок IP пакетов (IP packet error ratio, IPER).

#### **3. Готовность базового сегмента сети.**

*Готовность* – свойство объекта быть в состоянии выполнять требуемую функцию при заданных условиях в данный момент времени или в течение заданного интервала времени при условии обеспечения необходимыми внешними ресурсами. Готовность является ключевым показателем, определяющим качество функционирования сети. Основой для определения показателя готовности сети на уровне IP служит параметр IPLR – доля потерянных IP пакетов, или коэффициент потерь IP пакетов.

#### **4. Пропускная способность базового сегмента сети.**

Полезно иметь возможность охарактеризовать качество

функционирования сети с помощью показателей, связанных с понятием пропускной способности. При введении таких показателей необходимо учитывать, что на пропускную способность сети на уровне IP оказывают влияние протоколы вышележащих уровней, например протокол управления передачей (Transmission Control Protocol, TCP). Можно рассматривать такие показатели, характеризующие ширину полосы пропускания, как пропускная способность звена/канала (capacity of a link/path), коэффициент использования звена (link utilization), доступная ширина полосы пропускания звена (available bandwidth of a link), достижимая ШПП канала (achievable bandwidth on a path). Однако работа МСЭ-Т над определением показателей данной категории пока не завершена.

### **§П.3. Требования к качеству обслуживания типовых услуг NGN**

Требования к параметрам качества функционирования сети определяются требованиями к качеству обслуживания при предоставлении той или иной инфокоммуникационной услуги. Услуги мультисервисной сети связи подразделяются на три основные категории по общему характеру требований к QoS:

- услуги передачи речи (например, телефония);
- услуги передачи видео (например, IP-телевидение);
- услуги передачи данных (например, предоставление доступа в Интернет).

Заметим, что речь идет о классификации наиболее востребованных услуг NGN, попадающих в рамки концепции «Triple Play». Можно привести примеры услуг NGN, обладающих специфическими требованиями к QoS, однако мы не выделяем их в отдельные категории в силу их недостаточной распространенности.

В данном параграфе рассматриваются основные требования к качеству, предъявляемые услугами выделенных категорий, и влияние на них различных параметров NR.

### ***П.3.1. Услуги передачи речи***

Предоставление услуг передачи речи в сетях IP явилось первым шагом на пути конвергенции в сфере инфокоммуникаций и развивается на протяжении более чем пяти лет. Работа над стандартами качества при предоставлении таких услуг в целом ряде международных и национальных организаций по стандартизации продвинулась весьма далеко.

Среди услуг по передаче речи выделяют три основные группы:

- телефония;
- голосовые сообщения;
- потоковая речь.

На качество обслуживания при передаче речи значительное влияние оказывает односторонняя задержка передачи по сети. Результатом влияния этого параметра являются два эффекта: возникновение эха и нарушение динамики разговора. Первый эффект становится заметен при значении величины задержки выше нескольких десятков миллисекунд, что делает механизмы, ограничивающие эффект эха, обязательными для оборудования IP-телефонии. Второй – когда величина задержки превышает сотни миллисекунд. Для голосовой связи рекомендована задержка не более 150 мс.

К кратковременным вариациям задержки человеческое ухо малочувствительно. В зависимости от типа кодека не воспринимается вариация задержки в пределах 15–50 мс. На практике для всех голосовых услуг вариация задержки компенсируется посредством использования буфера для сглаживания фазового дрожания.

Поскольку голосовые пакеты не передаются повторно, при их потере или искажении качество восприятия речи на принимающей стороне может заметно ухудшаться. Частые потери голосовых пакетов могут привести к ухудшению разборчивости речи и полной невозможности общения. Искажения при потере пакетов зависят от типов применяемых кодеков. При телефонии хорошего качества допустимый уровень потерь пакетов не должен превышать 1%.

Для услуги передачи голосовых сообщений требования к уровню потери информации такие же, как для телефонии. Ключевое различие между указанными услугами состоит в том, что при передаче голосовых сообщений параметр задержки не оказывает настолько сильного влияния. Значение имеет только задержка между запросом пользователя на прослушивание нового голосового сообщения и тем моментом, когда он услышит начало сообщения. Для услуг голосовой телефонии считается приемлемой величина задержки до нескольких секунд.

Требования к задержке при передаче потоковой речи (например, для услуги радио через Интернет) могут быть ослаблены в связи с односторонней передачей. Однако к потерям предъявляются более жесткие требования, чем для телефонии.

Значения параметров качества при предоставлении услуг передачи речи согласно Приложению I к Рекомендации G.1010 приведены в таблице П.2.

Характеристики задержек и потерь при передаче речи для различных классов обслуживания ETSI TIPHON (Telecommunications and Internet Protocol Harmonization Over Networks) показаны в таблице П.3.



Таблица П.2. Значения показателей качества передачи речи

Приложение	Степень симметрии	Скорость передачи данных, кбит/с	Значения ключевых параметров		
			Односторонняя задержка, мс	Вариация задержки, мс	Потеря пакетов
Голосовая телефония	Двусторонняя	4–64	Предпочтит. < 150, предельная < 400	< 1	< 3%
Голосовые сообщения	Преимущественно односторонняя	4–32	< 1 с для записи, < 2 с для воспроизведения	< 1	< 3%
Потоковая речь высокого качества	Преимущественно односторонняя	16–128	< 10 с	<< 1	< 1%

Таблица П.3. Классы качества для услуг передачи речи ETSI TIPHON

Характеристика	Классы обслуживания			
	наилучший (Gold)	высокий (Silver)	средний (Bronze)	низкий
Сквозная односторонняя задержка, мс	< 150	< 250	< 350	< 450
Вариация задержки, мс	< 10	10–20	20–40	–
Коэффициент потерь пакетов	< 0,5%	0,5–1 %	1–2%	–

### ***П.3.2. Услуги передачи видео***

Предоставление услуги передачи видео на основе IP-сетей – сравнительно молодое, хотя и очень перспективное направление развития

отрасли. На сегодняшний день для передачи видео нет согласованных стандартов, регламентирующих качество предоставления таких услуг.

Среди видеослужб выделяют две основные категории:

- интерактивное видео (например, видеоконференции);
- потоковое видео (например, IPTV).

Интерактивное видео подразумевает двусторонний обмен как видео-, так и аудиоинформацией. В связи с этим требования к характеристикам качества обслуживания и качества функционирования сети для интерактивного видео такие же, как для голосовой телефонии. Из-за односторонней задержки передачи по сети возникают эффекты эха и нарушения динамики. Кроме того, при передаче трафика интерактивного видео возникает дополнительное требование синхронизации видео- и аудиоинформации.

Человеческий глаз нечувствителен к небольшой потере информации, поэтому допустим невысокий уровень потерь, который зависит от особенностей видеокодека и способов защиты от потерь информации. Для предотвращения значительных потерь рекомендуется использовать механизмы, реализующие алгоритмы приоритезации трафика, например DiffServ (Differentiated Services).

Отличительной особенностью потокового видео, например IPTV, является отсутствие диалогового элемента, благодаря чему требования к задержкам могут быть снижены по сравнению с требованиями для интерактивного видео. Вариация задержки незначительна, поскольку компенсируется адаптивными буферами на стороне пользователя. При этом задержка, добавляемая буфером, может достигать 100–500 мс.

Для сохранения хорошего качества изображения требуется низкая величина потери пакетов. Для гарантии эффективного предоставления видеослужб по сети IP величина коэффициента потерь пакетов IP не должна превышать  $10^{-5}$ .

Модель измерения качества видеослужб согласно Рекомендации J.241 показана на рис. П.4. Здесь заданы четыре точки измерений:

- А – кодер видео;
- В – уровень IP на стороне головной станции (IP-трафик);
- С – уровень IP на стороне оборудования пользователя (IP-трафик);
- D – декодер видео.



Рис. П.4. Модель измерения показателей качества при предоставлении услуг передачи видео

В таблице П.4 приведены значения параметров NR при передаче видеoinформации.

Таблица П.4. Значения показателей NR при передаче видео

Приложение	Степень симметрии	Скорость передачи данных, бит/с	Значения ключевых параметров			
			Одностор. задержка, мс	Вариация задержки, мс	Потеря пакетов, %	Доп. параметры
Интерактивное видео	Двустор.	16–384	Предпочт. < 150, допуст. < 400	-	< 1	Рассинхронизация видео и аудио < 80 мс
Потоковое видео	Одностор.	16–384	< 10 с	-	< 1	-

### ***П.3.3. Услуги передачи данных (best effort)***

Традиционно в сетях IP трафик передается по методу «негарантированной доставки» (best effort). Сеть старается обработать поступающий трафик как можно быстрее, но при этом не дается гарантий относительно результата. По методу best effort обслуживаются преимущественно веб-услуги, электронная почта, обмен сообщениями, передача файлов. Также популярны приложения реального времени, связанные с передачей данных, например интерактивные игры.

Для всех видов услуг по передаче данных основным требованием к качеству передачи является гарантия отсутствия потерь информации. Вариация задержки при передаче незначительно влияет на качество, тем не менее определенные ограничения на уровень синхронизации между информационными потоками мультимедийной сессии (например, аудиоинформация при просмотре широковещательной презентации) должны быть установлены. Требования к задержке при передаче отличаются для различных типов приложений. На основе требований к задержкам можно выделить следующие классы услуг по передаче данных:

- Веб-услуги, под которыми понимается поиск и просмотр текстовых компонентов веб-страниц. Остальные компоненты, такие как изображения и видеоклипы, относятся к другим категориям в рамках данного раздела. С точки зрения пользователя основополагающим фактором качества является скорость появления страницы после запроса. Для данной категории услуг задержка менее 10 секунд считается приемлемой.
- Класс объемных данных (файлы). Поскольку при передаче файлов пользователю предоставляется возможность просмотра информации о ходе передачи (скорости передачи, проценте полученной информации от общего размера файла и т. д.), то

- требования к задержке можно ослабить по сравнению с требованиями для веб-страниц.
- Высокоприоритетные услуги по передаче данных, например электронная коммерция. Основные требования пользователя к данной категории услуг – немедленное начало передачи после запроса пользователя. Задержка при передаче в пределах нескольких секунд считается приемлемой.
  - Передача изображений различных форматов, некоторые из которых могут быть терпимы к потерям во время передачи, при передаче других потери недопустимы. Однако, учитывая тот факт, что искажение даже одного бита может вызвать видимое ухудшение качества изображения, принято считать, что в общем случае необходима гарантия отсутствия потерь информации. Требования к задержке при передаче изображений не критичные и сравнимы с требованиями при передаче файлов, поскольку изображение обычно появляется на экране пользователя по мере поступления данных, что иллюстрирует процесс передачи.
  - Интерактивные игры, требования к параметрам качества которых в значительной степени зависят от специфики игры. В общем случае для данной категории задержки при передаче должны быть минимальными и не превышать доли секунды.
  - Приложения для удаленного доступа типа Telnet, обеспечивающие удаленный доступ пользователя к узлам сети. Значения задержек при передаче должны быть минимальны и не превышать доли секунды, для того чтобы разница во времени между набором символов пользователем и их отображением на экране не была ощутима для пользователя.
  - Электронная почта. При передаче сообщений электронной почты различают процесс обмена информацией между пользователем и

локальным почтовым сервером и обмен информацией между серверами. В первом случае задержки при передаче информации не должны превышать нескольких секунд. Во втором случае требования к задержкам минимальны, допустимыми считаются задержки в пределах нескольких минут или даже часов.

- Мгновенный обмен сообщениями, как правило, подразумевает обмен текстовой информацией, однако может также включать обмен аудио- и видеоданными. Несмотря на название данной категории, она не относится к категориям передачи данных в реальном времени, и задержки в пределах нескольких секунд считаются приемлемыми.
- Фоновые приложения. Для данной категории единственным требованием к качеству передачи данных является отсутствие ошибок при передаче. Требования к задержкам минимальны. К данной категории относится передача факсимильных сообщений, задержка при передаче которых не должна превышать 30 секунд, однако допустимы потери информации в некоторых пределах. К фоновым услугам также относится обмен короткими сообщениями, при котором допускается задержка в пределах 10 секунд.

## РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

### Основная

- [1] Башарин Г. П. Лекции по математической теории телетрафика. – М.: Изд-во РУДН, 2007. – 268 с.: ил.
- [2] Бабков В. Ю., Полынцев П. В., Устюжанин В. И. Качество услуг мобильной связи. Оценка, контроль и управление. – М.: Горячая линия-Телеком, 2005. – 160 с.
- [3] Деарт В. Ю. Мультисервисные сети связи. – М.: Инсвязьиздат, 2007. – 166 с.
- [4] Корнышев Ю. Н., Пшеничников А. П., Харкевич А. Д. Теория телетрафика. Учебник для вузов. – М.: Радио и связь, 1996. – 272 с.
- [5] Крылов В. В., Самохвалова С. С. Теория телетрафика и ее приложения. – СПб.: БХВ-Петербург, 2005. – 288 с.
- [6] Кучерявый А. Е., Цуприков А. Л. Сети связи следующего поколения. – М.: ФГУП ЦНИИС, 2006. – 280 с.
- [7] Лагутин В. С., Степанов С. Н. Телетрафик мультисервисных сетей связи. – М.: Радио и связь, 2000. – 320 с.
- [8] Соколов Н. А. Телекоммуникационные сети. Монография. – М.: Альварес Пабблишинг, 2004.
- [9] Телекоммуникационные системы и сети: Уч. пособие. В 3-х т. Том 3. – Мультисервисные сети / Величко В. В. и др. / под ред. проф. В. П. Шувалова – М.: Горячая линия-Телеком, 2005. – 592 с.
- [10] Шнепс-Шнеппе М. А. Системы распределения информации. Методы расчета: Справочное пособие. – М.: Связь, 1979. – 344 с.
- [11] Iversen V. V. Teletraffic Engineering Handbook. – ITU-D, Nov. 2005. – 323 p.
- [12] Kelly F. P. Reversibility and stochastic network. – Chichester: Wiley, 1979. – 630 p.
- [13] Ross K. W. Multiservice loss models for broadband telecommunication networks. – London: Springer-Verlag, 1995. – 343 p.

### **Дополнительная**

- [14] Башарин Г. П., Гайдамака Ю. В., Самуйлов К. Е. Яркина Н. В. Управление качеством и вероятностные модели функционирования сетей связи следующего поколения. Учебное пособие. – М.: Изд-во РУДН, 2008. – 130 с.
- [15] Гайдамака Ю. В., Зарипова Э. Р., Самуйлов К. Е. Модели обслуживания вызовов в сети сотовой подвижной связи: Учебно-метод. пособие. – М.: Изд-во РУДН, 2008. – 72 с.
- [16] Наумов В. А., Самуйлов К. Е., Яркина Н. В. Теория телетрафика мультисервисных сетей. – М.: Изд-во РУДН, 2007. – 192 с.



## СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ

$a_{n,m}$	– интенсивность перехода из состояния $n$ в состояние $m$
$\mathbf{A}$	– матрица интенсивностей переходов МП $X(t)$
$b_k$	– требования к ШПП для заявок $k$ -го поступающего потока
$b_{ms}$	– число единиц емкости звена, требуемое для предоставления услуги $m \in M_s$ в сети мультивещания
$B_1$	– вероятность блокировки нового вызова
$B_2$	– вероятность блокировки хэндовер-вызова
$B_{v,N}(\rho)$	– вероятность блокировки по вызовам в модели Энгсета
$B_{mps}$	– множество блокировок $(m, p, s)$ -пути сети мультивещания
$B_{mps}$	– вероятность блокировки $(m, p, s)$ -пути сети мультивещания
$B_m$	– вероятность блокировки $m$ -услуги на отдельном звене сети мультивещания
$C$	– число радиоканалов в соте ССПС
$F_{mps}$	– вероятность того, что $(m, p, s)$ -путь включен
$F_m$	– вероятность того, что $m$ -услуга на отдельном звене сети мультивещания включена
$H_{mps}$	– вероятность того, что $(m, p, s)$ -путь выключен, но в сети достаточно ресурсов для его включения
$H_m$	– вероятность того, что $m$ -услуга на отдельном звене сети мультивещания выключена, но в сети достаточно ресурсов для ее включения
$G(\Omega)$	– нормирующая константа множества $\Omega$
$c^{(1)}$	– среднее число занятых единиц ШПП на отдельном звене сети мультивещания
$C_l$	– емкость звена сети мультивещания, $l \in L$
$C_{v,N}(\rho)$	– вероятность блокировки по нагрузке в модели Энгсета
$EX$	– среднее значение СВ $X$

$E_n$	– состояние СМО $n$ занятыми линиями
$E_v(\rho)$	– вероятность блокировок для первой модели Эрланга
$E_{v,N}(\rho)$	– вероятность блокировки по времени в модели Энгсета
$J$	– пространство состояний МП $X(t)$
$K$	– число потоков, услуг
$L$	– 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
$(m, p, s)$	– логический путь, $(m, p, s)$ -путь, $m \in M_s$ , $p \in P_s$ , $s \in S$ , в сети мультивещания
$M_s$	– множество услуг, предоставляемых источником $s \in S$ сети мультивещания
$n(t)$	– мгновенная нагрузка в момент $t$
$N$	– число источников в модели Энгсета
$p$	– физический путь в сети мультивещания, $p \in P_s$
$\{p_n, n \in J\}$	– стационарное распределение МП $X(t) \in J$
$P_s$	– множество всех физических путей к источнику $s \in S$ в сети мультивещания
$P_s^l$	– множество физических путей к источнику $s \in S^l$ , проходящих через звено $l \in L$
$q_{n,m}$	– переходные вероятности МП $X(t)$
$\bar{q}$	– средняя длина очереди
$Q$	– стохастическая матрица МП $X(t)$
$r$	– емкость накопителя в СМО с ожиданием
$R$	– общая емкость СМО
$S$	– число источников информации сети мультивещания
$S$	– множество источников информации сети мультивещания
$S^l$	– множество источников информации, предоставляющих услуги через звено $l \in L$ , в сети мультивещания

$\langle S, A \rangle$	– СМО со структурой ресурсов $S$ и алгоритмом $A$ их распределения между входящими потоками заявок
$u(x)$	– функция Хевисайда, $u(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$
$\nu$	– число приборов в СМО
$\mathbf{x}$	– детальное состояние всех логических путей сети мультивещания, $\mathbf{x} \in X$
$x_{mps}$	– состояние логического пути в сети мультивещания
$X$	– пространство состояний логических путей сети
$X(t)$	– случайный процесс, $t \geq 0$
$Y_{\text{обсл}}$	– обслуженная нагрузка
$Y_{\text{предл}}$	– предложенная нагрузка
$Y_{\text{пот}}$	– потерянная нагрузка
$\alpha$	– вероятность занятия единственного канала в модели Энгсета
$\varepsilon$	– интенсивность генерации заявок источником в модели Энгсета
$\{\pi(\mathbf{x}), \mathbf{x} \in X\}$	– стационарное распределение МП $X(t) \in X$
$\lambda$	– интенсивность поступающего потока
$\lambda_1$	– интенсивность ПП новых вызовов
$\lambda_2$	– интенсивность ПП хэндовер-вызовов
$\lambda_k$	– интенсивность поступления заявок $k$ -го потока
$\lambda_{mps}$	– интенсивность ПП запросов на установление $(m, p, s)$ -пути в сети мультивещания, $m \in M_s, p \in P_s, s \in S$
$\Lambda_j$	– суммарная интенсивность предложенного пуассоновского потока $j$ -заявок в модели Энгсета
$\mu$	– параметр экспоненциального обслуживания
$\mu_k$	– параметр экспоненциального обслуживания заявок $k$ -го потока

- $\mu_{mps}$  – параметр экспоненциального обслуживания запросов на установление  $(m, p, s)$ -пути в сети мультивещания,  $m \in M_s, p \in P_s, s \in S$
- $\rho$  – интенсивность поступающей нагрузки
- $\rho_{mps}$  – нагрузка, создаваемая запросами на установление  $(m, p, s)$ -пути в сети мультивещания
- $\omega$  – длительность ожидания заявкой начала обслуживания
- – символ окончания доказательства

## ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- IP-телефония, 88, 95
- Network Performance, 86, 89, 90, 93, 99
- Quality of Experience, 86, 89, 90
- Quality of Service, 49, 50, 86, 88, 94
- Вероятность
  - блокировки, 10, 15, 27, 33, 67, *См также* Вероятность потери
    - по времени, 15, 33, 41, 44, 47, 68
    - по вызовам, 44, 45, 47
    - по вызовам, 68
    - по нагрузке, 44, 45, 46
  - переходная, 14
  - потери, 9, 26, 34, 51, 54, 68, 74, *См. также* Вероятность блокировки
  - предельная, 18
- Вызов, 10, 17, 27, 37, 44, 48, 82
  - новый, 48
  - хэндовер, 48, 50, 51, 52
- Готовность, 93
- Дерево мультивещания, 59, 61, 62
- Загрузка, 24, 26
- Заявка, 9, 10, 20, 26, 34, 35, 43, 51, 64, 71
- Звено обмена, 92
- Интенсивность, 9, 13, 20, 24, 28, 50, 53, 72
- Интенсивность нагрузки, 11, 17, 19, 23, 33, 39, 45, 46
- Источник, 35, 58, 59, 60, 61, 62, 63, 71, 92
- Качество
  - восприятия, 87, 89, 96
  - обслуживания, 81, 82, 88, 89, 94, 95, 96, 98
  - функционирования сети, 88, 89, 91, 93, 98
- Качество обслуживания, 4
- Модель
  - Энгсета, 5, 35, 37, 38, 42
  - Эрланга, 5, 37
    - вторая, 26, 27, 31
    - первая, 9, 15, 21, 28, 34, 35, 37, 44, 47, 51
- Мультивещание, 57, 58, 63
- Нагрузка, 17, 18, 19, 20, 22, 24, 34, 42, 43, 51, 52, 88, 91
  - мгновенная, 19, 21, 22, 23
  - обслуженная, 21, 24, 26, 46, *См. также* Нагрузка принятая
  - поступающая, 11, 21, 46, *См. также* Нагрузка предложенная
  - потерянная, 21, 24, 46
  - предложенная, 21, 22, 34, 44, 46, 47, 66, *См. также* Нагрузка поступающая
  - принятая, 33, 46, *См. также* Нагрузка обслуженная
  - пуассоновская
    - второго рода, 38, 47
    - первого рода, 38, 47

Обслуживание, 10, 20, 21, 26, 34, 38, 48, 52, 53, 65, 71, 72, 75, 76, 85, 88, 90, 96  
Оператор связи, 84, 88, 89, 90  
Передача данных, 82, 85, 86, 90, 91, 94, 100, 102  
    многоадресная, 57  
    одноадресная, 57  
Пользователь, 4, 57, 58, 59, 68, 71, 72, 75, 81, 82, 96, 98, 100  
Порядок занятия, 24  
    случайный, 9, 24  
    упорядоченный, 9, 24  
Поставщик услуги, 82, 83, 84, 85, 88  
Поток, 10, 12, 13, 15, 20, 29, 33, 39, 50, 63, 71, 75, 91, 100  
    входящий, 18, 28  
    заблокированный, 37, 38, *См. также* Поток потерянный  
    обслуженный, 33, *См. также* Поток принятый  
    поступающий, 9, 11, 20, 27, 71, *См. также* Поток предложенный  
    потерянный, 38, *См. также* Поток заблокированный  
    предложенный, 15, 28, 33, 35, 37, 38, *См. также* Поток поступающий  
    принятый, 28, 33, 37, 38, *См. также* Поток обслуженный  
    пуассоновский, 9, 10, 11, 13, 35, 38, 50, 51, 53, 71  
    Энгсета, 38  
Пространство состояний, 11, 14, 28, 37, 42, 51, 54, 59, 63, 67, 69, 70, 73  
Процесс  
    марковский, 5, 12, 18, 28, 65, 72  
    размножения и гибели, 11, 28, 37, 54  
    случайный, 5, 11, 12, 19, 27, 51, 54, 63, 65, 66, 72, 73  
Путь  
    логический, 62, 63, 65  
    физический, 60, 61, 63  
Работа, 21, 22  
Распределение  
    биномиальное, 42, 43  
    усеченное, 41  
Пуассона, 11, 16, 20, 44  
    усеченное, 16  
равновесное, 14, 18, 29, 39, 63, 66  
стационарное, 52, 65, 67, 72, 73  
экспоненциальное, 9, 12, 13, 16, 26, 28, 36, 38, 51, 53, 64, 72  
Энгсета, 39, 41, 42, 44, 46  
Эрланга, 14, 16, 44  
    второе, 27, 31, 32, 33  
    первое, 15, 20, 32, 44  
Сегмент сети, 92  
Сеть  
    доступа, 83, 84, 85  
    мультивещания, 58, 59, 60, 63, 71, 74, 76, 77, 79, 80  
    мультисервисная, 4, 6, 57, 58, 59, 84  
    связи, 84  
    следующего поколения, 4, 5, 81, 82, 85, 86, 89, 91  
    сотовой подвижной связи, 6, 48  
    транспортная, 82, 84, 91

Служба, 84, 87  
Терминал, 84, 85, 87  
Точка измерения, 92  
Услуга, 4, 6, 58, 59, 62, 68, 71, 75, 79, 81, 82, 90, 94, 97, 100  
Формула Эрланга  
    вторая, 33, 34  
    первая, 14, 15  
Хэндовер, 48, 52  
Эталонное событие, 92

## 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 дисциплиной по выбору для студентов, обучающихся по направлениям 010300 «Математика. Компьютерные науки», 010400 «Информационные технологии» или 010500 «Прикладная математика и информатика».

Курс входит в состав модуля «Управление инфокоммуникациями» профиля специализации в бакалавриате. Студенты, выбравшие данный профиль, должны также прослушать следующие дисциплины:

«Основы формальных методов описания бизнес процессов»;

«Основы разработки корпоративных инфокоммуникационных систем»;

«Основы управления инфокоммуникационными компаниями».

Лицам, успешно окончившим бакалавриат по указанным выше направлениям и прослушивавшие перечисленные выше курсы могут быть рекомендованы для обучения в магистратуре по направлению 010400



«Информационные технологии», магистерская программа «Управление инфокоммуникациями».

### ***Цели курса***

- Знакомство с принципами функционирования сетей связи следующего поколения
- Изучение вопросов качества обслуживания в этих сетях
- Освоение методов построения вероятностных моделей для анализа качества обслуживания в терминах теории массового обслуживания и теории марковских процессов.
- Изучение методов анализа и расчета показателей качества обслуживания сетей связи следующего поколения.

### ***Задачи курса***

После успешного прохождения курса слушатели должны

*знать:*

- основные понятия и определения, относящиеся к концепции сетей связи следующего поколения;
- принципы построения сетей связи следующего поколения;
- требования международных стандартов к показателям качества обслуживания – QoS-параметрам;

*уметь:*

- с помощью аппарата теории вероятностей, теории случайных процессов теории массового обслуживания и теории телетрафика строить простые модели узлов и звеньев NGN, а также простые модели сети в целом;

- для построенных моделей составлять и решать системы уравнений равновесия (СУР), получать вероятностные характеристики моделей, связанные с показателями качества обслуживания;
- применять численные методы при анализе полученных характеристик моделей.

## **2. Инновационность курса**

*По содержанию.*

Современная концепция «Сетей связи следующего поколения» (NGN) отражает конвергенцию информационно-телекоммуникационных сетей в единую глобальную сеть. Движение в этом направлении только началось, на сегодняшний день международные стандарты содержат, главным образом, не конкретные решения, а лишь требования к NGN, из которых вытекает множество задач изучения. Одним из направлений исследований является анализ качества обслуживания в NGN. Усложнение конфигураций сетей, по сравнению с телефонными сетями общего пользования, появление новых видов обслуживания, предусматривающих возможность выбора пользователем услуги с заранее заданным уровнем качества, привело к необходимости построения новых моделей для анализа качества обслуживания в сети. В рамках курса слушатели получают необходимый объем 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:* 1 час 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.* Введение в концепцию сетей связи следующего поколения (NGN – Next Generation Networks).

- 1.1. Основные понятия и определения NGN. Архитектура, протоколы и услуги сетей связи следующего поколения.
- 1.2. Вопросы качества обслуживания в NGN. Трехуровневая модель оценки качества. Требования международных стандартов к параметрам качества обслуживания. Требования к качеству обслуживания типовых услуг NGN.
- 1.3. Принципы построения моделей для анализа качества обслуживания.

*Тема 2.* Классические моносервисные модели.

- 2.1.  $M \left| \begin{matrix} M \\ \lambda \\ \mu \end{matrix} \right| c \left| \begin{matrix} 0 \\ \\ \end{matrix} \right.$  - первая модель Эрланга: распределение Эрланга и случай  $c \rightarrow \infty$ . Вычисление  $E_p(c)$ .
- 2.2. Нагрузка и ее характеристики. Методы измерения нагрузки.
- 2.3.  $M \left| \begin{matrix} M \\ \lambda \\ \mu \end{matrix} \right| c \left| \begin{matrix} r \\ \\ \end{matrix} \right.$  - вторая модель Эрланга (с ожиданием и блокировками).
- 2.4.  $M \left| \begin{matrix} M \\ N, \varepsilon \\ \mu \end{matrix} \right| c \left| \begin{matrix} 0 \\ \\ \end{matrix} \right.$  - модель Энгсета: распределение Энгсета при  $N > c$ , потери по времени и по вызовам.

*Тема 3.* Сети сотовой подвижной связи (ССПС). Беспriorитетные и приоритетные модели процесса обслуживания вызовов в одной соте ССПС.

- 3.1. Принципы функционирования ССПС. Эстафетная передача обслуживания (хэндовер).
- 3.2. Модель соты ССПС в виде полнодоступной СМО с потерями: вывод СУР, основные вероятностно-временные характеристики (ВВХ) модели.

3.3. Модель соты ССПС в виде недоступной СМО с потерями: СУР, ВВХ модели.

*Тема 4.* Подход к анализу мультисервисных сетей связи (МСС).

4.1. Модель МСС с одноадресными соединениями: построение пространства состояний, равновесное распределение вероятностей, вероятность блокировок установления соединения.

4.2. Метод расчёта отдельного звена МСС с одноадресными соединениями: алгоритм Кауфмана-Робертса.

4.3. Модель МСС с многоадресными соединениями: пространство состояний, равновесное распределение вероятностей, вероятностные характеристики модели.

4.4. Метод расчёта отдельного звена МСС с многоадресными соединениями: алгоритм свертки.

*Темы семинарских занятий*

*Тема 1.* Основные протоколы сетей связи следующего поколения: H.323, SIP, MGCP, MEGACO/H.248, BICC, Sigtran.

*Тема 2.* Ступенчатые марковские процессы (МП). Конструктивное описание и примеры. Процессы размножения и гибели (ПРГ). Локальный и глобальный баланс. Система уравнений равновесия (СУР). Диаграмма интенсивностей переходов.

*Тема 3.* Первая модель Эрланга. Вычисление и свойства  $E_c(\rho)$ . Применение при проектировании телекоммуникационных систем вероятностей блокировок по Эрлангу.

*Тема 4.* Вторая модель Эрланга: вероятность блокировок, вероятность ожидания и средняя длина очереди.

*Тема 5.* Распределение Эрнгсета: свойства, связь между вероятностями блокировок по времени, по вызовам и по нагрузке.

*Тема 6.* Методы определения момента инициации хэндовера. Процедуры жесткой и мягкой эстафетной передачи. Параметры качества: вероятности блокировок нового вызова, хэндовера, многократного хэндовера в ССПС.

*Тема 7.* Основные алгоритмы (RPB/TRPB, RPM, CBТ) и протоколы (IGMP, BGMP, CBТ, DVMP, MOSPF, PIM) сетей связи с многоадресной доставкой информации.

#### *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 № 1, включающей 2 вопроса по темам № 1, № 2 и № 3 содержания курса.

Примерный перечень вопросов контрольной работы № 1:

1. Назовите основные особенности построения сетей в соответствии с концепцией NGN.
2. Перечислите услуги, которые можно предоставлять пользователям с помощью сети NGN.
3. Назовите основные уровни архитектуры NGN.
4. Назовите основные функциональные элементы сети NGN.
5. Назовите основные типы шлюзов сети NGN.
6. Каково назначение программного коммутатора Softswitch?

7. Назовите уровни, на которых производится оценка качества в сети следующего поколения. На каждом уровне укажите соответствующие параметры оценки качества: Перечислите параметры качества обслуживания QoS.
8. Расскажите об основных показателях качества функционирования транспортной сети. Перечислите шесть классов обслуживания QoS, рекомендованных МСЭ-Т.
9. Перечислите шесть классов обслуживания QoS, рекомендованных МСЭ-Т.
10. Назовите основные протоколы сигнализации для управления соединениями в NGN.
11. Модель Эрланга с явными потерями. Выпишите диаграмму интенсивностей переходов, инфинитезимальную матрицу  $A$ , СУР и условие существования решения СУР. Выведите стационарное распределение вероятностей.
12. Выпишите инфинитезимальную матрицу  $A$  для ПРГ, описывающего функционирование модели Эрланга с явными потерями, и матрицу  $Q$  переходных вероятностей для соответствующей ЦМ.
13. Докажите, что 
$$\frac{\rho}{c} [1 - E_c(\rho)] = \frac{E_c(\rho)}{E_{c-1}(\rho)}.$$
14. Дайте определение нагрузки, назовите ее виды. Перечислите методы измерения нагрузки. Для статистической оценки характеристик нагрузки приведите примеры, когда за единицу времени естественно принять минуты, секунды и миллисекунды.
15. Расскажите о влиянии порядка занятия приборов на общую обслуженную нагрузку и загрузку каждого прибора.

16. Модель Эрланга с ожиданием и блокировками. Выведите СУР и получите ее решение. Перечислите основные вероятностные характеристики второй модели Эрланга.

17. Докажите, что в СМО  $M \left| \begin{matrix} M \\ \lambda \end{matrix} \right| \begin{matrix} M \\ \mu \end{matrix} \left| c \right| r$  при  $r < \infty$  и любом  $\rho < \infty$  средняя длина очереди  $\bar{q}$  имеет вид

$$\bar{q} = p_c \frac{\rho}{c} \frac{1 + r \left(\frac{\rho}{c}\right)^{r+1} - (r+1) \left(\frac{\rho}{c}\right)^r}{\left(1 - \frac{\rho}{c}\right)^2}.$$

18. Модель Энгсета. Выведите СУР и получите ее решение. Перечислите основные вероятностные характеристики модели Энгсета.

19. Для ПРГ выведите систему уравнений локального баланса (СУЛБ) из системы уравнений глобального баланса (СУГБ).

20. Для ПРГ выпишите СУЛБ с помощью диаграммы интенсивностей переходов.

21. Объясните связь между вероятностями блокировок по времени, по вызовам и по нагрузке. Приведите примеры.

22. Расскажите о принципах организации связи в ССПС на примере сети GSM. Опишите основные функциональные элементы этой системы (мобильная станция, подсистема базовых станций, сетевая подсистема).

23. Опишите физическую модель процесса обслуживания вызовов в соте ССПС, дайте определение нового вызова, хэндовера, зоны хэндовера. Определите параметры качества в ССПС: вероятности блокировок нового вызова, хэндовера, многократного хэндовера.



24. Опишите процедуру эстафетной передачи (хэндовера). Как определяется момент инициации хэндовера? Объясните разницу между «жесткой» и «мягкой» эстафетной передачей.
25. Постройте модель соты ССПС в виде полностью доступной СМО с потерями. Получите СУР из условий глобального баланса. Решите СУР, получите стационарное распределение вероятностей и вероятностные характеристики модели.
26. Постройте модель соты ССПС в виде частично доступной СМО с потерями. Выведите и решите СУР, получите равновесное распределение и вероятностные характеристики модели.

Примерный перечень задач контрольной работы № 1:

1. Рассматривается АТС, обслуживающая  $10^4$  абонентов, каждый из которых создает нагрузку 0,04 эрл. Известно, что 10% звонков – междугородные. Рассчитать число каналов для междугородней связи таким образом, чтобы вероятность блокировки вызова на них не превышала 2%. Вычислить вероятность блокировки междугороднего вызова для случая, когда нагрузка в 2,5 раза превысит заданную в условии.
2. Имеется 6 телефонов, соединенных с офисной АТС двумя каналами. Каждый телефон используется 2 раза в час, при этом длительность разговора экспоненциально распределена со средним 3 минуты. Вычислить вероятность блокировки и коэффициент использования канала для СМО с потерями. Вычислить среднее время ожидания и коэффициент использования канала для СМО с ожиданием.
3. В ЛВС имеется 20 серверов, функционирующих в режиме разделения времени, и 600 рабочих станций. С каждой рабочей станции 1 раз в день случайным образом на серверы приходит задача. Длительности решения задач экспоненциально распределены

со средним 12 минут. Вычислить вероятность того, что задаче придется ожидать в очереди, и среднее время ожидания. Вычислить вероятность того, что время ожидания превысит 6 минут (для дисциплин FIFO и RANDOM).

4. Рассматривается случайный поток пакетов данных, поступающих с интенсивностью 6 пакетов в секунду. Длины пакетов имеют геометрическое распределение со средним 1200 бит. Пакеты передаются по двум каналам с быстродействием 4800 бит/сек. каждый. Рассчитать размер накопителя, необходимый для того, чтобы вероятность блокировки передачи пакета не превышала  $10^{-5}$ . Для рассчитанной емкости накопителя вычислить среднее время ожидания и вероятность того, что оно не превысит 1 секунды.

*Итоговый контроль знаний.*

Контроль уровня знаний осуществляется в виде письменной контрольной работы № 2, включающей 2 вопроса по теме № 4 содержания курса.

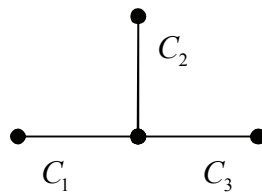
Примерный перечень вопросов контрольной работы № 2:

1. Определите понятие мультисервисной сети связи. Приведите примеры нагрузки, требующей передачи в режиме реального времени, и нагрузки, допускающей задержку.
2. Постройте пространство состояний для математической модели МСС с одноадресными соединениями.
3. Получите стационарные вероятности состояний марковского процесса, описывающего функционирование сети с одноадресными соединениями.
4. Получите в аналитическом виде выражения для вероятностей блокировок запросов в сети с одноадресными соединениями.

5. Постройте пространство состояний для математической модели МСС с многоадресными соединениями.
6. Получите стационарные вероятности состояний марковского процесса, описывающего функционирование сети с многоадресными соединениями.
7. Получите в аналитическом виде выражения для вероятностей блокировок запросов в сети с многоадресными соединениями.
8. Постройте математическую модель отдельного звена сети с многоадресными соединениями.
9. Выпишите алгоритм расчета вероятности блокировки услуги на отдельном звене сети с многоадресными соединениями.

Примерный перечень задач контрольной работы № 2:

1. Для сети с одноадресными соединениями, изображенной на рисунке,

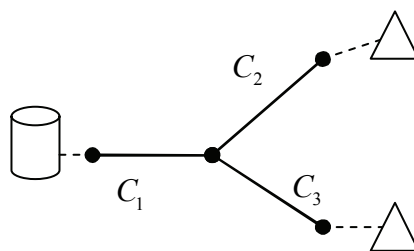


выписать пространство состояний сети и множества блокировок соединений каждого класса при следующих значениях исходных параметров:

$$L = \{1, 2, 3\}, C_1 = 3, C_2 = 2, C_3 = 3;$$

$$K = \{1, 2, 3\}, R_1 = \{1, 2\}, R_2 = \{2, 3\}, R_3 = \{1, 3\}; d_1 = 1, d_2 = 2, d_3 = 1.$$

2. Для сети с многоадресными соединениями, изображенной на рисунке,



при следующих значениях исходных параметров:

$$L = \{1, 2, 3\}, C_1 = 5, C_2 = 5, C_3 = 4;$$

$$S = \{1\}, M_1 = \{1, 2\}, b_{11} = 3, b_{21} = 2$$

1) выписать множества

$P_s = \{1, \dots, P_s\}$  – множество всех физических путей к источнику  $s \in S$ ;

$L_{ps} \subseteq L$  – множество всех звеньев физического пути  $p \in P_s$  к источнику  $s \in S$ ;

$S^l = \{s \in S : P_s^l \neq \emptyset\}$  – множество источников информации, предоставляющих услуги через звено  $l \in L$ ;

$P_s^l = \{p \in P_s : l \in L_{ps}\}$  – множество физических путей к источнику  $s \in S^l$ , проходящих через звено  $l \in L$ ;

2) выписать пространство состояний сети  $X$ ;

3) выписать  $B_{mps}$  – множество состояний блокировок  $(m, p, s)$ -пути,  $m \in M_s, p \in P_s, s \in S$ ;

4) построить дерево мультивещания от каждого из источников.

3. Оценить сверху мощность пространства состояний сети с многоадресными соединениями.

4. Доказать, что для любой тройки  $(m, p, s)$ ,  $m \in M_s, p \in P_s, s \in S$ , выполняется соотношение  $F_{mps} = \rho_{mps} H_{mps}$ .

5. Доказать, что для любого  $(m, p, s)$ -пути,  $m \in M_s, p \in P_s, s \in S$ , верно соотношение  $F_{mps} = \frac{\rho_{mps}}{1 + \rho_{mps}} (1 - B_{mps})$ .

6. Изобразить графически пространство состояний отдельного звена сети с многоадресными соединениями для случая  $M = 3$ .

7. Для отдельного звена сети с многоадресными соединениями

$$C = 3, M = 3, P = 2, b_1 = 1, b_2 = 2, b_3 = 1$$

выписать пространство  $\tilde{Y}$  состояний сети и множества  $B_1, B_2, B_3$  блокировок каждой услуги.

8. Показать, что среднее число занятых приборов для модели отдельного звена сети с многоадресными соединениями

$$c^{(1)} = \sum_{m=1}^M b_m F_m.$$

9. Доказать, что среднее число занятых приборов для модели отдельного звена сети с многоадресными соединениями имеет вид

$$c^{(1)} = \sum_{m=1}^M b_m \frac{\rho_m}{1 + \rho_m} (1 - B_m).$$

10. Оценить вычислительную сложность алгоритма расчёта нормировочной константы  $G(\tilde{Y})$  для отдельного звена сети с многоадресными соединениями.

## *Литература*

### *Обязательная литература*

- [1] Башарин Г. П. Лекции по математической теории телетрафика. – М.: Изд-во РУДН, 2007. – 268 с.: ил.
- [2] Бабков В. Ю., Полынцев П. В., Устюжанин В. И. Качество услуг мобильной связи. Оценка, контроль и управление. – М.: Горячая линия-Телеком, 2005. – 160 с.
- [3] Гольдштейн Б.С., Гольдштейн А.С. SoftSwitch. СПб.: БХВ – Санкт-Петербург. - 2006. - 368 с.: ил.
- [4] Деарт В. Ю. Мультисервисные сети связи. – М.: Инсвязьиздат, 2007. – 166 с.
- [5] Корнышев Ю. Н., Пшеничников А. П., Харкевич А. Д. Теория телетрафика. Учебник для вузов. – М.: Радио и связь, 1996. – 272 с.
- [6] Крылов В. В., Самохвалова С. С. Теория телетрафика и ее приложения. – СПб.: БХВ-Петербург, 2005. – 288 с.
- [7] Кучерявый А. Е., Цуприков А. Л. Сети связи следующего поколения. – М.: ФГУП ЦНИИС, 2006. – 280 с.

- [8] Лагутин В. С., Степанов С. Н. Телетрафик мультисервисных сетей связи. – М.: Радио и связь, 2000. – 320 с.
- [9] Соколов Н. А. Телекоммуникационные сети. Монография. – М.: Альварес Пабблишинг, 2004.
- [10] Телекоммуникационные системы и сети: Уч. пособие. В 3-х т. Том 3. – Мультисервисные сети / Величко В. В. и др. / под ред. проф. В. П. Шувалова– М.: Горячая линия-Телеком, 2005. – 592 с.
- [11] Шнепс-Шнеппе М. А. Системы распределения информации. Методы расчета: Справочное пособие. – М.: Связь, 1979. – 344 с.
- [12] Iversen V. B. Teletraffic engineering and network planning. – ITU-D, May 2006. – 590 p.  
<http://oldwww.com.dtu.dk/education/34340/telenook.pdf>
- [13] Kelly F. P. Reversibility and stochastic network. – Chichester: Wiley, 1979. – 630 p.
- [14] Ross K. W. Multiservice loss models for broadband telecommunication networks. – London: Springer-Verlag, 1995. – 343 p.

*Дополнительная литература и источники Интернет*

- [15] Башарин Г. П., Гайдамака Ю. В., Самуйлов К. Е. Яркина Н. В. Управление качеством и вероятностные модели функционирования сетей связи следующего поколения. Учебное пособие. – М.: Изд-во РУДН, 2008. – 130 с.
- [16] Гайдамака Ю. В., Зарипова Э. Р., Самуйлов К. Е. Модели обслуживания вызовов в сети сотовой подвижной связи: Учебно-метод. пособие. – М.: Изд-во РУДН, 2008. – 72 с.
- [17] Клейнрок Л. Теория массового обслуживания. Пер. с англ. Под ред. В.И. Неймана // М.: Машиностроение, 1979. – 452 с
- [18] Кучерявый А.Е., Цуприков А.Л. Сети связи следующего поколения // М.: ФГУП ЦНИИС, 2006.
- [19] Наумов В. А., Самуйлов К. Е., Яркина Н. В. Теория телетрафика мультисервисных сетей. – М.: Изд-во РУДН, 2007. – 192 с.
- [20] Шварц М. Сети связи: протоколы, моделирование и анализ: в 2-х ч. Пер. с англ. В.И. Неймана // Ч.1: М.: Наука, Гл. ред. физ.-мат. лит., 1992. – 336 с., Ч.2: М.: Наука, Гл. ред. физ.-мат. лит., 1992. –272 с.
- [21] <http://www.itu.int>
- [22] <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.

Второй модуль трудоемкостью 2 кредита составляют:

- теоретический материал, излагаемый в лекциях 11 – 19 календарного плана курса;
- содержание семинарских занятий в течение 18 академических часов.

В конце модуля проводится итоговый контроль знаний.

### Календарный план курса

Виды и содержание учебных занятий				
Неделя	Лекции	Число часов	Лабораторные занятия	Число часов
1	Введение в концепцию сетей связи следующего поколения (NGN – Next Generation Networks). Основные понятия и определения NGN. Архитектура и услуги сетей связи следующего поколения.	2	Основные протоколы сетей связи следующего поколения: H.323, SIP.	2
2	Вопросы качества в NGN. Требования международных стандартов к параметрам качества обслуживания: классы QoS для различных транспортных технологий.	2	Основные протоколы сетей связи следующего поколения: MGCP, MEGACO/H.248.	2
3	Принципы построения моделей для анализа качества обслуживания.	2	Основные протоколы сетей связи следующего поколения: BICC, Sigtran.	2
4	$M M c 0$ - первая модель Эрланга: распределение Эрланга и случай $c \rightarrow \infty$ , вычисление $E_c(\rho)$ .	2	Ступенчатые Марковские процессы (МП). Конструктивное описание и примеры. Процессы размножения и гибели (ПРГ).	2



Виды и содержание учебных занятий				
Неделя	Лекции	Число часов	Лабораторные занятия	Число часов
5	Нагрузка и ее характеристики. Методы измерения нагрузки.	2	Локальный и глобальный баланс. Система уравнений равновесия (СУР).	2
6	$M M c r$ - вторая модель Эрланга (с ожиданием и блокировками).	2	Вероятность блокировок, вероятность ожидания и средняя длина очереди для второго распределения Эрланга.	2
7	$M M c 0$ - модель Энгсета: распределение Энгсета при $N > c$ , потери по времени и по вызовам.	2	Свойства распределения Энгсета.	2
8	Принципы функционирования ССПС. Эстафетная передача обслуживания (хэндовер).	2	Связь между вероятностями блокировок по времени, по вызовам и по нагрузке. Примеры.	2
9	Модель соты ССПС в виде полностью доступной СМО с потерями: вывод СУР, основные вероятностно-временные характеристики (ВВХ) модели.	2	Методы определения момента инициации хэндовера. Процедуры жесткой и мягкой эстафетной передачи.	2

<b>Виды и содержание учебных занятий</b>				
<b>Неде- ля</b>	<b>Лекции</b>	<b>Число часов</b>	<b>Лабораторные занятия</b>	<b>Число часов</b>
<b>10</b>	Модель соты ССПС в виде неполнодоступной СМО с потерями: СУР, ВВХ модели.	<b>2</b>	Параметры качества: вероятности блокировок нового вызова, хэндовера, многократного хэндовера.	<b>2</b>
<b>11</b>	<b>Промежуточный контроль знаний (Контрольная работа № 1)</b>			<b>2</b>
<b>12</b>	Подход к анализу мультисервисных сетей связи (МСС).	<b>2</b>	Алгоритмы и протоколы сетей с многоадресной доставкой информации	<b>2</b>
<b>13</b>	Модель МСС с одноадресными соединениями: построение пространства состояний, равновесное распределение вероятностей, вероятность блокировок установления соединения.	<b>2</b>	Алгоритмы RPB/TRPB и RPM.	<b>2</b>
<b>14</b>	Модель МСС с одноадресными соединениями: построение пространства состояний.	<b>2</b>	Алгоритм СВТ.	<b>2</b>

<b>Виды и содержание учебных занятий</b>				
<b>Неде- ля</b>	<b>Лекции</b>	<b>Число часов</b>	<b>Лабораторные занятия</b>	<b>Число часов</b>
<b>15</b>	Модель МСС с одноадресными соединениями: равновесное распределение вероятностей, вероятность блокировок установления соединения.	<b>2</b>	Протоколы IGMP и BGMP.	<b>2</b>
<b>16</b>	Метод расчёта отдельного звена МСС с одноадресными соединениями: алгоритм Кауфмана-Робертса.	<b>2</b>	Протокол DVMP.	<b>2</b>
<b>17</b>	Модель МСС с многоадресными соединениями: построение пространства состояний.	<b>2</b>	Протоколы CBT и MOSPF.	<b>2</b>
<b>18</b>	Модель МСС с многоадресными соединениями: равновесное распределение вероятностей, вероятностные характеристики модели.	<b>2</b>	Протоколы PIM-DM и PIM-SM.	<b>2</b>
<b>19</b>	Метод расчёта отдельного звена МСС с многоадресными соединениями: алгоритм свертки.	<b>2</b>	Консультации по подготовке к итоговому контролю знаний.	<b>2</b>
<b>20</b>	<b>Итоговый контроль знаний (Контрольная работа № 2)</b>			<b>2</b>

#### 4. Описание системы контроля знаний

##### *Шкала бально-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 начисляются дополнительные баллы: по 1 баллу за каждый балл свыше 60.	86 – 100	5
		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***

1. Начисление баллов за семестр.

Студент посетил 65% занятий, на семинарских занятиях постоянно принимал участие в обсуждениях, три раза решал задачи у доски.

*Набранные баллы: 6 баллов.*

Контрольная работа № 1: студент письменно отвечал на следующие вопросы:

Вопрос 1. Перечислите шесть классов обслуживания QoS, рекомендованных МСЭ-Т.

Ответ полный, без ошибок.

*Набранные баллы: 15 баллов.*

Вопрос 2. Расскажите о влиянии порядка занятия приборов на общую обслуженную нагрузку и загрузку каждого прибора.

Ответ полный, однако, при выводе формулы загрузки  $i$ -го прибора допущена ошибка.

*Набранные баллы: 5 балл.*

Работа над рефератом.

Тема. Технология Sigtran передачи сигнальной информации в сетях связи следующего поколения.

Реферат содержит обзор концепции сетей связи следующего поколения, описана архитектура сети, основное оборудование, услуги, предоставляемые сетью. Для технологии Sigtran описан стек протоколов Sigtran, подробно изложен механизм передачи информации по протоколу транспортного уровня Stream Control Transmission Protocol (SCTP).

При написании реферата студент помимо рекомендованной литературы самостоятельно подобрал и использовал дополнительные источники. Объем реферата составил 25 страниц, реферат содержит иллюстрации и таблицы, 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: 30 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:  $15 + 5 = 20$  баллов

Работа над рефератом:  $30 + 10 = 40$  баллов

Баллы за семестр  $N = 66$  баллов

Общая сумма баллов:  $66 + (66 - 60) = 72$  балла

Для оценки работы в семестре применяется первая строка шкалы балльно-рейтинговой системы, поскольку  $61 \leq N \leq 80$ .

Студент имеет право получить автоматическую итоговую оценку по 5-балльной шкале 4 (*хорошо*) с общей суммой баллов 72 балла. Это соответствует оценке C по шкале ECTS.

По желанию студент может пройти итоговый контроль знаний с целью повышения итоговой оценки.

## 2. Начисление баллов за итоговый контроль знаний.

На контрольной работе № 2 (итоговый контроль знаний) студент письменно отвечал на следующие вопросы:

Вопрос 1. Постройте пространство состояний для математической модели МСС с одноадресными соединениями.

Ответ на вопрос полный, без замечаний.

*Набранные баллы: 10 баллов.*

Вопрос 2. Выпишите алгоритм расчета вероятности блокировки услуги на отдельном звене сети с многоадресными соединениями.

Ответ на вопрос полностью соответствует требованиям, при выводе формул студент не допустил ошибок.

*Набранные баллы: 10 баллов.*

Таким образом, в течение семестра и за итоговый контроль знаний студент набрал следующие баллы.

Баллы за семестр  $N =$  66 баллов

Итоговый контроль знаний  $M =$  10 + 10 = 20 баллов

Общая сумма баллов  $N + M =$  66 + 20 = 86 баллов

Итоговая оценка по 5-балльной шкале: 5 (отлично).

Итоговая оценка шкале ECTS: B.

### ***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.