СОДЕРЖАНИЕ

Тема номера:

"ИНФОКОММУНИКАЦИОННО-УПРАВЛЕНЧЕСКИЕ СЕТИ. РАСЧЕТ И ОПТИМИЗАЦИЯ СИСТЕМ СВЯЗИ"

Башарин Г.П., Штатнов С.В.
Мультисервисная модель обслуживания эластичного трафика с конечным числом источников 4

Самуйлов К.Е., Сопин Э.С., Чукарин А.В.
Оценка характеристик сигнального трафика в сети связи на базе подсистемы IMS 8

Адаму Амину, Гайдамака Ю.В., Самуйлов А.К.
Анализ производительности одной схемы многоканальной передачи потоковых данных в одноранговых сетях 14

Бузюкова И.Л., Бязров А.С., Гайдамака Ю.В.
Имитационная модель узла управления услугами интеллектуальной сети 18

Гудкова И.А., Лузгачев М.В.
Модели разделения ресурсов звена мультисервисной сети с эластичным трафиком 22

Бутурлин И.А., Плаксина О.Н.
К анализу вероятностей блокировок трафика услуг IPTV 25

Коннон А.М.
Приближенные методы расчета характеристик мультисервисной системы с ограниченным числом источников 27

Шнепс-Шнеппе М.А.
Домовая инфокоммуникационная сеть на базе концепции ITU G.hn 30

Мамедов Г.А., Ибрагимов Б.Г.
Метод повышения эффективности функционирования терминального оборудования сетей с интеграцией служб 33

Маликова Е.Е.
Метод повышения пропускной способности систем телеметрии и мониторинга на базе беспроводных сетей 36

Малиновский С.Т., Беленькая М.Н., Спиридонов А.А.
Метод повышения производительности транспортного протокола TCP в глобальных корпоративных сетях передачи данных 39

Подольский Д.В.
К вопросу выбора аналитической модели функционирования звена сети следующего поколения 43

Хатунцев А.Б.
Перспективы и предпосылки внедрения сетей связи следующего поколения в России 46

Попова А.Г., Чан Туан Минь.
Анализ результатов расчета операторской системы 50

Амарян С.Р., Попова А.Г., Степанова И.В.
Направления развития методов описания и построения корпоративных сетей связи 53

T•Comm
Телекоммуникации и транспорт

№7-2010

"T-Comm — Telecommunications and Transport" magazine

Журнал включен в перечень периодических научных изданий, рекомендуемый ВАК Минобразования России для публикации научных работ, отражающих основное научное содержание кандидатских и докторских диссертаций и рекомендован УМО по образованию в области телекоммуникаций для студентов высших учебных заведений.

Учредитель
ООО "Издательский дом Медиа Паблишер"

Главный редактор
В.О. Тихвинский

Издатель
С.С. Дымкова
ds@media-publisher.ru

Редакционная коллегия
А.С. Аджемов, Е.Б. Алексеев, Альберт Вааль, А.А. Гоголь, Юлиус Головачев, В.Л. Горбачев, Ю.А. Громаков, А.И. Демьянов, Б.В. Зверев, Ю.Б. Зубарев, В.Р. Иванов, Юрий Кирхгесснер, Т.А. Кузовкова, В.Н. Лившиц, С.Л. Мишенков, О.Е. Наний, Н.П. Резникова, И.В. Парфенов, Ш.Ж. Сеилов, В.О. Тихвинский, В.В. Фронтов, А.Б. Юрчук

Редакция

Выпускающий редактор
Андрей Волков
va@media-publisher.ru

Редактор
Татьяна Чередниченко

Отдел маркетинга и PR
Ольга Дорошкевич,
директор по маркетингу и PR
ovd@media-publisher.ru

Кристина Маркарова,
менеджер по маркетингу и PR
kristina@media-publisher.ru

Отдел распространения и подписки
info@media-publisher.ru

Предпечатная подготовка
ООО "ИД Медиа Паблишер"

www.media-publisher.ru

Степанова И.В., Покровская М.А.
Проектирование и внедрение информационных центров . . . 56

Попова А.Г., Чан Туан Минь.
Анализ подходов к расчету операторской системы контакт-центра 58

Степанова И.В., Киселев И.В.
Мониторинг и контроль call-центров экстренных служб . . . 61

Степанова И.В., Киселев И.В.
Результаты разработки специализированного программного обеспечения обработки трафика экстренных служб 64

Костин А.Н., Ершова Э.Б.
К вопросу оценки качества обслуживания в сети NGN . . . 66

Данилов А.Н., Кайюмов С.Т.
Прогноз развития сетей связи в Республике Таджикистан 69

Докучаев В.А.
Анализ состояния российских предприятий — производителей телекоммуникационного оборудования 72

Мишенков С.Л., Смирнов Н.И., Караваев Ю.А.
Перспективные требования к сетевой синхронизации и распределению сигналов точного времени в системе связи РФ 77

Алексеев Е.Б., Данилов А.Н., Матвеев В.А., Грош А.О.
Анализ протоколов маршрутизации в сетях NGN 81

Кирик Ю.М., Петренко А.А.
Радиорелейное оборудование с пакетной технологией передачи данных 86

Мельник С.В., Петрова Е.Н., Смирнов Н.И.
Технология мобильной связи четвертого поколения LTE. Основные особенности и перспективы внедрения в России 89

Савенков Ю.Ю.
Оценка взаимного влияния параметров потоков заявок при организации VPN на общем канальном ресурсе 91

Пшеничников А.П., Степанов М.С.
Асимптотические модели оценки показателей качества обслуживания вызовов call-центрами 93

Васькин Ю.А.
Предоставление инфокоммуникационных услуг на сетях связи следующего поколения 97

Рахматулин А.М.
Классификация структур решетчатых оптических сетей . . . 100

Рахматулин А.М.
Основные принципы разработки методов проектирования решетчатых и квазирешетчатых оптических транспортных сетей связи 102

Абаева Б.К.
Вопросы проектирования сетей IPTV 104

Семейкин В.Д., Скупченко А.В.
Моделирование транспортной сети регионального оператора связи с помощью графовых моделей 107

Астахов И.В.
Тенденции и методики развития инфокоммуникационных сетей на ближайшуюи среднесрочную перспективу 110

Птицын Г.А.
Сокращение потерь сообщений в транспортно-распределительных системах 113

Мушили Н.М., Самуйлов К.Е., Сопин Э.С.
Модель функционирования сервера присутствия в сети NGN 116

Ефимушкин А.В.
Численный анализ блокировок оптической сети с маршрутизацией по длине волны 119

Чукарин А.В.
Об одной задаче управленич сетевыми и инфомрационными ресурсами в сетях NGN 123

Савандюков И.М.
Методика расчета эффективности использования сетевых ресурсов в оптической пакетной транспортной сети с волновым уплотнением 126

Савандюков И.М.
Обзор возможных вариантов организации взаимодействия оптимизатора трафика с различными OSS-системами 128

Абаев П.О., Гайдамака Ю.В., Рудикова Е.В.
Численный анализ модели обслуживания сессий в сети GSM/GPRS 130

Абаев П.О., Хатунцев А.Б.
Построение и анализ модели установления соединения по протоколу SIP в сети связи следующего поколения 134

Нсангу М.М., Самуйлов К.Е., Чукарин А.В.
Построение и анализ Марковской модели установления соединений по протоколу SIP с учетом повторных передач 139

Михайлов С.К., Сергеева Т.П.
Оценка длительности задержки телефонных сообщений от числа приоритетов в очереди и распределения времени обслуживания трафика нижних приоритетов . . . 142

Сергеева Т.П., Михайлов С.К.
Надежность сигнализации в МГ/МН-сети телефонной связи на базе оборудования Softswitch 144

Реймер Д.А.
О проблеме перегрузок в мультимедийных IP-ориентированных сетях 148

Баркова И.В., Сергеева Т.П., Тетекин Н.Н.
Методика получения нижних оценок коэффициента готовности тракта в системе взаимосвязанных кольцевых структур SDH 152

Баркова И.В., Слободской Н.В.
Автоматизированынй расчет коэффициента готовности времени задержки для трактов сети SDH с помощью программного комплекса KaHaT 156

Зарипова Э.Р., Леонтьева Е.А.
Анализ модели Call-центра с двумя классами сообщений и тремя группами операторов 160

Ефимушкин В.А., Углов И.В.
Архитектура QoS для конвергентных сетей и особенности ее применения . . . 162

Ефимушкин В.А., Шуваев Я.В.
Построение адресного пространства в сенсорных сетях. Проблемы стандартных подходов и методы их решения 164

Каледин В.В., Козадаева Л.А.
Вопросы использования коротких и длинных номеров, особенности нумерации в IP-сетях 170

Коротовских Д.В.
Влияние электромагнитной совместимости ЛЭП и линий связи на качество телекоммуникационных услуг 175

Темкина Т.А.
Совершенствование принципов организации баз данных компании — оператора связи 178

Кривоногов В.М.
Тенденции развития интерфейсов терминалов сетей связи общего пользования 180

Егоршев Д.Е.
Задачи оптимизации оптоволоконных сетей связи на участке абонентского доступа 182

Шалагинов В.А.
Моделирование сети следующего поколения при оценке качества инфокоммуникационных услуг 184

Шалагинов В.А.
Проблема учета и тарификации количества информации при предоставлении инфокоммуникационных услуг 187

Шашаев И.Ю.
Стохастическая модель минимизации длины соединительных линий телекоммуникационной сети 190

Шашаев И.Ю.
Применение методов многокритериальной оптимизации структурных характеристик припроектировании телекоммуникационных сетей 193

Бухарев И.А.
Особенности функционирвоания сетей беспроводного широкополосного доступа WIMAX с использованием технологии MIMO 195

Полканов Е.И., Ефимушкин И.В.
Современные методы оптимизации бизнес-процессов разработки и внедрения новых инфокоммуникационных услуг 200

Ефимушкина Т.В., Молчанов Д.А., Кучерявый Е.А.
Исследование вероятностно-временных характеристик функционирования соты WiMAX с несколькими режимами модуляции и эластичным трафиком данных 203

Заказ журналов:

- по каталогу "Роспечать" (индекс 80714)
- по каталогу "Интерпочта" (индекс 15241)
- "Деловая пресса" (www.delpress.ru)
- в редакции (info@media-publisher.ru)

Возможен также заказ через региональные альтернативные подписные агентства
http://www.media-publisher.ru/raspr.shtml

Периодичность выхода — шесть номеров в год
Стоимость одного экземпляра 200 руб.

Целевая аудитория по распространению
• Телекоммуникационные компании; • Дистрибьюторы телекоммуникационного оборудования и услуг; • Контент-провайдеры; • Разработчики и производители абонентского оборудования; • Предприятия и организации нефтегазового комплекса; • Энергетические компании; • Автотранспортные предприятия; • Крупные организации с собственным автомобильным автопарком; • Компании, занимающиеся железнодорожными, воздушными и морскими перевозками; • Логистические и экспедиционные компании; • Провайдеры охранно-поисковых услуг; • Геодезические и картографические организации; • Государственные ведомства и организации; • Строительные компании; • Профильные учебные заведения

Тираж 5000 экз. + Интернет-версия

Адрес редакции
111024, Россия, Москва,
ул. Авиамоторная, д. 8, корп. 1, офис 514
e-mail: ds@media-publisher.ru
Тел.: (495) 957-77-43

Журнал зарегистрирован Федеральной службой по надзору за соблюдением законодательства в сфере массовых коммуникаций и охране культурного наследия. Свидетельство о регистрации: ПИ № ФС77-27364

Мнения авторов не всегда совпадают с точкой зрения редакции. За содержание рекламных материалов редакция ответственности не несет

Материалы, опубликованные в журнале — собственность ООО "ИД Медиа Паблишер". Перепечатка, цитирование, дублирование на сайтах допускаются только с разрешения издателя
All artocles and illustrations are copyryght. All rights reserved. No reproduction is permitted in whole or part without the express consent of Media Publisher Joint-Stock Company

Вниманию авторов!
Для начисления авторского гонорара необходиомо указать ваши ФИО, почтовый адрес (с индексом), паспортные данные (серия, номер, кем и когда выдан), ИНН, номер свидетельства пенсионного страхования, дату и место рождения, номер телефона.
Плата с аспирантов за публикацию рукописи не взимается

© ООО "ИД Медиа Паблишер", 2010

www.media-publisher.ru

Мультисервисная модель обслуживания эластичного трафика с конечным числом источников

Быстро развивающиеся сети 3G во всем мире, а с 2007 г. и в России, ставят перед теорией телетрафика новые задачи. Развитие мобильного высокоскоростного доступа в Интернет, а также растущая популярность интерактивных услуг приводит к тому, что все большую долю в общем объеме составляет эластичный трафик. Это вызывает необходимость решения проблем, связанных с управлением доступом при учете эластичности трафика.

Поскольку многие операторы могут предоставлять часть услуг лишь ограниченному числу абонентов, то это также необходимо учитывать при анализе, применяя модели энгсетовского типа. В данной работе представлена K-сервисная модель обслуживания эластичного трафика с конечным числом источников. Приведен эффективный рекуррентный алгоритм расчета основных ВВХ.

Башарин Г.П.,
д.т.н., профессор кафедры систем телекоммуникаций РУДН,
gbasharin@sci.pfu.edu.ru

Штатнов С.В.,
магистрант кафедры систем телекоммуникаций РУДН,
svshtatnov@gmail.com

Введение

В конце XX в. появились мультисервисные системы связи, в которых целесообразно рассматривать два типа заявок: эластичные (передача файлов, эл. почты и др.) и неэластичные (передача речи, видео и др.), также называемые потоковыми [2, Гл.1].

Потоковые l-заявки имеет фиксированную ширину и ееобслуживание может начаться только при наличии необходимого ей числа b_1 свободных каналов, которые она займет на всевремя обслуживания. Эластичная k-заявка не имеет фиксированной ширины и при поступлении займет b_k^{max} каналов, если они свободны, но может удовлетвориться и меньшим их числом b_k, $b_k^{min} \le b_k \le b_k^{max}$ как при поступлении, так и в процессе обслуживания, уступив их позднее поступившей заявке. При этом ееинтенсивность обслуживания соответственно уменьшится, а средняя длительность — возрастет. Существуют и смешанные системы, в которых одновременно обслуживаются потоковые и эластичные заявки. Все это вызывает необходимость решения новых инженерных и теоретических проблем, связанных с управлением доступом [3].

В настоящей статье рассматривается одна из наиболее простых теоретических проблем, которая может быть полезна в реальных практических разработках.

Описание модели

Рассмотрим мультисервисную модель обслуживания пуассоновской нагрузки типа Энгсет-2 [1, Гл.3.1] с эластичным трафиком.

$$\underset{\mathbf{N},\mathbf{e},\mathbf{b}^{\max}}{\mathbf{M}} \left| \underset{\mathbf{м}^{\max}}{\mathbf{M}} \right| \begin{matrix} < C, elast > \\ V \end{matrix} \Big| \, 0. \qquad (1)$$

Схема функционирования данной модели представлена на рис.1.

Предположим, что к широкополосной цифровой линии (ШЦЛ) емкостью C базовых цифровых каналов (БЦК), организован доступ с явными потерями для K информационных потоков от K групп пользователей. К k-услуге имеют доступ $N_k < \infty$ абонентов, каждый из которых может одновременно пользоваться только одной и только этой услугой. Каждый из N_k абонентов создает независимый пуассоновский поток заявок на k-услугу с интенсивностью $\varepsilon_k, 0 < \varepsilon_k < \infty, k = \overline{1,K}$ [1, Гл.3]. Для краткости назовем такую нагрузку мультисервисной пуассоновской типа 2.2 и обозначим МПН-2.2.

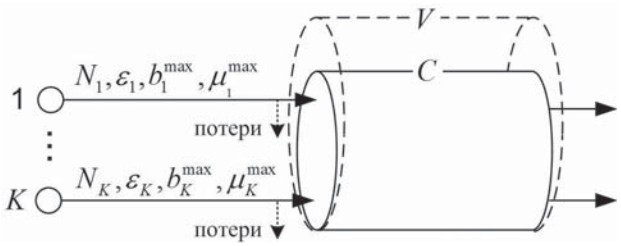


Рис. 1. Схема модели $\underset{\mathbf{N},\mathbf{e},\mathbf{b}^{\max}}{\mathbf{M}} \left| \underset{\mathbf{м}^{\max}}{\mathbf{M}} \right| \begin{matrix} < C, elast > \\ V \end{matrix} \Big| \, 0$

Состояние системы может быть записано в виде вектора $\mathbf{n} = (n_1,...,n_k)$, где n_k — число k-заявок в системе $n_k \le N_k$, $k = \overline{1,K}$. При поступлении каждая k-заявка требует b_k^{max} каналов, $k = \overline{1,K}$. Введем величину v — число "виртуально" занятых в системе каналов, когда все заявки принимаются на обслуживание с максимальными требованиями:

$$\mathrm{v} := \mathrm{v}(\mathbf{n}) = \sum_{k=1}^{K} n_k b_k^{\max}, \; 0 \le \mathrm{v} \le V, \qquad (2)$$

где V — виртуальная емкость системы [4],[5].

Если при поступлении k-заявки в системе нет свободных b_k^{max} каналов, заявка может быть принята с меньшими требованиями

$$b_k := b_k(\mathbf{n}) = b_k^{\max} \gamma(\mathbf{n}), \qquad (3)$$

где

$$\gamma := \gamma(\mathbf{n}) = \begin{cases} 1, & 0 \le \mathrm{v} \le C \\ \dfrac{C}{\mathrm{v}}, & C < \mathrm{v} \le V \end{cases}, \quad \gamma^{\min} \le \gamma(\mathbf{n}) \le 1 \qquad (4)$$

— мгновенная степень сжатия ШПП в состоянии $\mathbf{n}$, а величина $\gamma^{\min} := \dfrac{C}{V}$ одинакова для всех типов заявок и определяет максимальную степень сжатия полосы. При этом выделенная на обслуживание других заявок ШПП также сжимается. Сжатие требований для k-заявки возможно вплоть до значения

$$b_k^{\min} := b_k^{\max} \gamma^{\min}, \quad k = \overline{1,K}. \qquad (5)$$

Если при поступлении k-заявки свободных каналов нет, а дальнейшее сжатие невозможно, то заявка получает отказ и теряется, не оказывая влияния на интенсивность поступлений породившего ее пуассоновского потока.

Будем считать время занятия k-заявкой b_k каналов распределенным экспоненциально с параметром μ_k, $k = \overline{1,K}$. Примем, что для эластичного трафика среднее время μ_k^{-1} занятия заявкой каналов обратно пропорционально числу b_k требуемых БЦК, т.е. выполняется условие

$$\frac{b_k}{\mu_k} = \frac{b_k^{\max}}{\mu_k^{\max}}. \qquad (6)$$

Из (3) и (6) следует, что мгновенная интенсивность обслуживания

$$\mu_k := \mu_k(\mathbf{n}) = \mu_k^{\max}\gamma(\mathbf{n}),\ k = \overline{1,K}. \qquad (7)$$

Обслуживание заявок происходит в соответствии со стратегией FCFS. По завершении обслуживания k-заявка одновременно освобождает все занятые ею b_k БЦК.

Пусть $\tilde{\rho}_k^{\max}$ — интенсивность нагрузки от одного свободного источника при обслуживании с максимальными требованиями, а $\tilde{\rho}_k$ — мгновенная нагрузка:

$$\tilde{\rho}_k^{\max} := \frac{\varepsilon_k}{\mu_k^{\max}}, \qquad (8)$$

$$\tilde{\rho}_k = \tilde{\rho}_k(\mathbf{n}) := \frac{\varepsilon_k}{\mu_k(\mathbf{n})} = \frac{\tilde{\rho}_k^{\max}}{\gamma(\mathbf{n})}. \qquad (9)$$

Пространство S возможных состояний системы и подпространства S_k приема и $\overline{S}_k$ блокировки k-заявок имеют вид

$$\begin{aligned} &S := \{\mathbf{n} : 0 \le \mathrm{v} \le V\}, \\ &S_k := \{\mathbf{n} : 0 \le \mathrm{v} \le V - b_k^{\max}\}, \\ &\overline{S}_k = S \setminus S_k = \{\mathbf{n} : \mathrm{v} > V - b_k^{\max}\}. \end{aligned} \qquad (10)$$

Для иллюстрации, рассмотрим следующий пример.

Пример. Пусть на обслуживание принимаются эластичные заявки двух типов ($K = 2$). Параметры модели приведены в табл. 1.

Таблица 1

Исходные данные для примера

K	$\mathbf{N}$	$\boldsymbol{\varepsilon}$	$\mathbf{b}^{\max}$	$\boldsymbol{\mu}^{\max}$	C	V
2	$(3,2)$	$(\varepsilon_1, \varepsilon_2)$	$(1,2)$	(μ_1', μ_2')	4	6

Здесь максимальная степень сжатия ШПП $\gamma^{\min} = \frac{C}{V} = \frac{2}{3}$, а число занятых виртуальных каналов $\mathrm{v} = \sum_{k=1}^{2} n_k \cdot b_k^{\max} = n_1 + 2n_2$.

Пространство S возможных состояний системы и подпространства $S_1(S_2)$ приема и $\overline{S}_1(\overline{S}_2)$ блокировки 1-заявок (2-заявок) для данного примера имеют следующий вид:

$$\begin{aligned} &S := \{(n_1, n_2) : 0 \le \mathrm{v} \le 6\}; \\ &S_1 := \{(n_1, n_2) : 0 \le \mathrm{v} \le 5\},\ \overline{S}_1 = \{(n_1, n_2) : \mathrm{v} > 5\}; \\ &S_2 := \{(n_1, n_2) : 0 \le \mathrm{v} \le 4\},\ \overline{S}_1 = \{(n_1, n_2) : \mathrm{v} > 4\}. \end{aligned}$$

Диаграмма переходов для примера представлена на рис. 2.

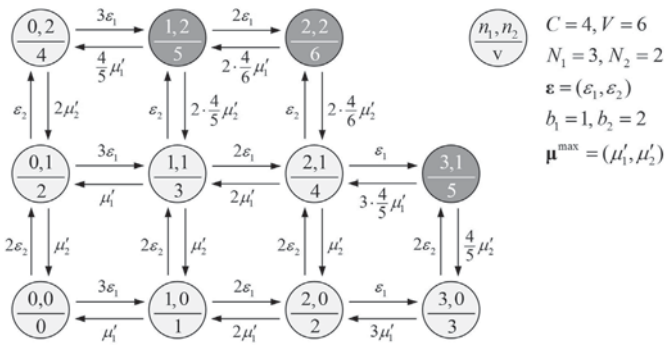


Рис. 2. Диаграмма переходов для частного случая

Построение СтМП и СУГБ

Функционирование данной системы описывается K-мерным СтМП $X(t) = (X_1(t),...,X_k(t))^T$ с пространством состояний S, где $X_k(t)$ — число k-заявок в системе в момент времени $t \ge 0$. Обозначим равновесное распределение процесса $X(t)$ как

$$p(\mathbf{n}) := P\{\mathbf{X} = \mathbf{n}\} = P\{X_k = n_k, k = \overline{1,K}\}. \qquad (11)$$

Используя диаграмму интенсивностей переходов (рис. 3), запишем СУГБ в виде:

$$\begin{aligned} &p(\mathbf{n})\left(\sum_{k=1}^{K}(N_k - n_k)\varepsilon_k I(\mathbf{n} \in S_k) + \sum_{k=1}^{K} n_k \mu_k(\mathbf{n})\right) = \\ &= \sum_{k=1}^{K} p(\mathbf{n} - \mathbf{e}_k)(N_k - n_k + 1)\varepsilon_k u(n_k) + \\ &+ \sum_{k=1}^{K} p(\mathbf{n} + \mathbf{e}_k)(n_k + 1)\mu_k(\mathbf{n} + \mathbf{e}_k) I(\mathbf{n} \in S_k), \quad \mathbf{n} \in S \end{aligned} \qquad (12)$$

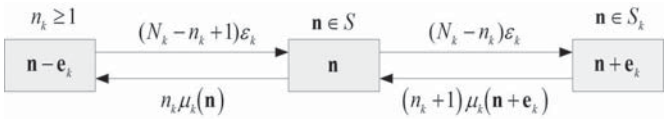


Рис. 3. Диаграмма интенсивностей переходов процесса $X(t)$, $k = \overline{1,K}$

Пусть π_k — вероятность того, что вновь поступившая k-заявка застанет систему в макросостоянии $\overline{S}_k$ и будет заблокирована. Тогда

$$\pi_k := \sum_{\mathbf{n} \in \overline{S}_k} p(\mathbf{n}),\ k \in K. \qquad (13)$$

Рассмотрим в примере замкнутый контур (1,1)-(2,1)-(2,2)-(1,2)-(1,1). Интенсивности переходов между состояниями этого контура не удовлетворяют критерию Колмогорова [1, Прил. В]:

$$2\varepsilon_1 \cdot \varepsilon_2 \cdot \frac{4}{3}\mu_1' \cdot \frac{8}{5}\mu_2' \ne \varepsilon_2 \cdot 2\varepsilon_1 \cdot \frac{4}{3}\mu_2' \cdot 2\mu_1.$$

Это означает, что для процесса $X(t)$ не выполняется свойство обратимости, а, следовательно, и мультипликативности. Поэтому точный расчет ВВХ может быть выполнен средствами имитационного моделирования или с помощью решения СУГБ (12) одним из численных методов, что является достаточно ресурсоемкой задачей. Однако для приближенного расчета основных ВВХ системы можно применить следующий рекуррентный алгоритм.

Рекуррентный алгоритм вычисления макровероятностей

При вычислении ВВХ модели будем использовать не значения вероятностей $p(\mathbf{n})$, а значения макровероятностей $q(v)$ пребывания СтМП $X(t)$ во множестве состояний $S(v)$:

$$S(\mathrm{v}) := \left\{ \mathbf{n} \in S : \sum_{k=1}^{K} n_k b_k^{\max} = \mathrm{v} \right\}, \bigcup_{\mathrm{v}=0}^{V} S(\mathrm{v}) = S. \quad (14)$$

$$q(\mathrm{v}) := P\{\mathbf{n} \in S(\mathrm{v})\} = \sum_{\mathbf{n} \in S(\mathrm{v})} p(\mathbf{n}) \quad (15)$$

— равновесная вероятность того, что в виртуальной системе занято v БЦК.

Утверждение. Макровероятность $q(v)$ удовлетворяет следующему приближенному рекуррентному соотношению:

$$\mathrm{v}q(\mathrm{v}) \approx \sum_{k=1}^{K} \left[\left(N_k + 1 - \frac{\mathrm{v}}{K b_k^{\max}} \right) \frac{\tilde{\rho}_k^{\max}}{\gamma(\mathrm{v})} b_k^{\max} q\left(\mathrm{v} - b_k^{\max}\right) \right]. \quad (16)$$

Для доказательства используем СУЧБ (17).

$$p(\mathbf{n}) n_i \mu_i(\mathbf{n}) = p(\mathbf{n} - \mathbf{e}_i)(N_i - n_i + 1)\varepsilon_i I(n_i > 0), \quad \mathbf{n} \in S \quad (17)$$

Выражая $p(\mathbf{n})$ и учитывая (15), получим

$$q(\mathrm{v}) = \sum_{\mathbf{n} \in S(\mathrm{v})} p(\mathbf{n}) = \sum_{\mathbf{n} \in S(\mathrm{v})} \left[(N_i - n_i + 1) \frac{\tilde{\rho}_i}{n_i} p(\mathbf{n} - \mathbf{e}_i) I(n_i > 0) \right] =$$

$$= \sum_{\mathbf{n} \in S(\mathrm{v})} \left[\left((N_i + 1) \frac{\sum_{k=1}^{K} n_k b_k^{\max}}{\mathrm{v}} - n_i \frac{\mathrm{v}}{\mathrm{v}} \sum_{k=1}^{K} \frac{1}{K} \right) \frac{\tilde{\rho}_i}{n_i} p(\mathbf{n} - \mathbf{e}_i) I(n_i > 0) \right] =$$

$$= \frac{1}{\mathrm{v}} \sum_{\mathbf{n} \in S(\mathrm{v})} \left[\sum_{k=1}^{K} \left((N_i + 1) b_k^{\max} n_k - n_i \frac{v}{K} \right) \frac{\tilde{\rho}_i}{n_i} p(\mathbf{n} - \mathbf{e}_i) I(n_i > 0) \right], \; n_i \geq 1, i = \overline{1, K}.$$

Пологая $i = k$, $n_k \geq 1$, получим

$$\mathrm{v}q(\mathrm{v}) = \sum_{\mathbf{n} \in S(\mathrm{v})} \left[\sum_{k=1}^{K} \left((N_k + 1) b_k^{\max} \cancel{n_k} - \cancel{n_k} \frac{\mathrm{v}}{K} \right) \frac{\tilde{\rho}_i}{\cancel{n_k}} p(\mathbf{n} - \mathbf{e}_k) \right] =$$

$$= \sum_{\mathbf{n} \in S(\mathrm{v})} \left[\sum_{k=1}^{K} \left((N_k + 1) - \frac{\mathrm{v}}{K b_k^{\max}} \right) \tilde{\rho}_k b_k^{\max} p(\mathbf{n} - \mathbf{e}_k) \right] =$$

$$= \sum_{k=1}^{K} \left[\left((N_k + 1) - \frac{\mathrm{v}}{K b_k^{\max}} \right) \tilde{\rho}_k b_k^{\max} \sum_{\mathbf{n} \in S(\mathrm{v})} p(\mathbf{n} - \mathbf{e}_k) \right] = \sum_{k=1}^{K} \left[\left(N_k + 1 - \frac{\mathrm{v}}{K b_k^{\max}} \right) \tilde{\rho}_k b_k^{\max} q\left(\mathrm{v} - b_k^{\max}\right) \right].$$

Используя (9), получаем (16).

Алгоритм.

1. Вводим параметры V, N_k, $b_k^{\max}$, $\tilde{\rho}_k^{\max}$, $k = \overline{1, K}$.
2. Полагаем $q'(0) := 1$, $q'(v) := 0$ при $v < 0$, где $q'(v)$ — ненормированные значения вероятностей $q(v)$.
3. Рекуррентно по $v = \overline{1, V}$ вычисляем

$$q'(\mathrm{v}) := \frac{1}{\mathrm{v}} \sum_{k=1}^{K} \left[\left(N_k + 1 - \frac{v}{K b_k^{\max}} \right) \frac{\tilde{\rho}_k^{\max}}{\gamma(\mathrm{v})} b_k^{\max} q'\left(\mathrm{v} - b_k^{\max}\right) \right].$$

4. Вычисляем нормирующую константу

$$G := \sum_{\mathrm{v}=0}^{V} q'(\mathrm{v}).$$

5. Переходим от ненормированного к нормированному распределению вероятностей:

$$q(\mathrm{v}) := \frac{q'(\mathrm{v})}{G}, \; \mathrm{v} = \overline{0, V}.$$

Полученные значения $q(v)$, $v = \overline{0, V}$ позволяют рассчитать основные ВВХ системы:

— вероятность блокировки k-заявки

$$\pi_k := \sum_{\mathrm{v} = V - b_k^{\max} + 1}^{V} q(\mathrm{v}); \quad (18)$$

— среднее число занятых каналов в виртуальной системе

$$UTIL^{\mathrm{v}} = \sum_{\mathrm{v}=1}^{V} \mathrm{v}q(\mathrm{v}); \quad (19)$$

— среднее число занятых каналов в реальной системе

$$UTIL = \sum_{\mathrm{v}=1}^{C} \mathrm{v}q(\mathrm{v}) + \sum_{\mathrm{v}=C+1}^{V} Cq(\mathrm{v}). \quad (20)$$

Численный анализ

Базовые параметры системы, используемые при численном анализе модели приведены в таблице 2 [4].

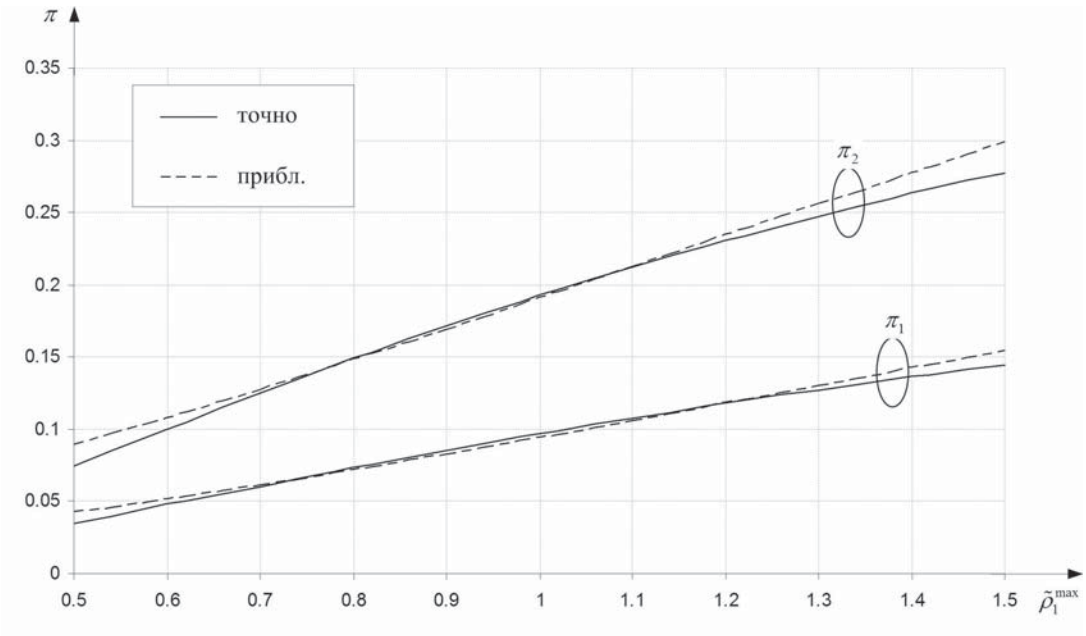


Рис. 4. Зависимость вероятностей потерь от нагрузки $\tilde{\rho}_1^{\max}$ при постоянной нагрузке $\tilde{\rho}_2^{\max}$

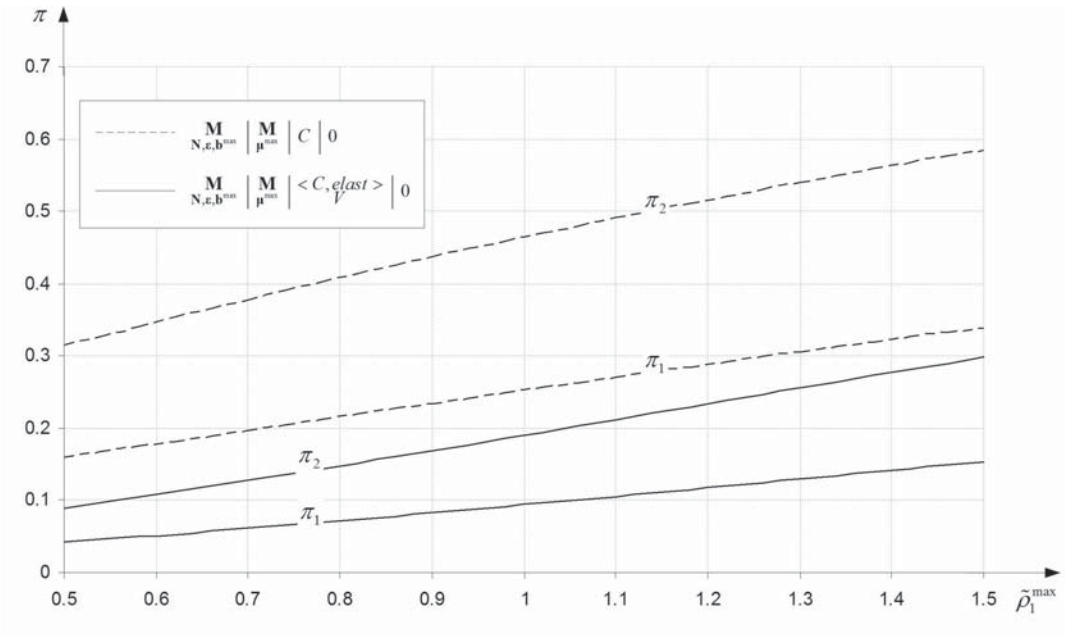


Рис. 5. Сравнение вероятностей потерь для моделей со сжатием и без

Таблица 2

Базовые значения параметров системы.

K	$\mathbf{N}$	$\tilde{\boldsymbol{\rho}}^{max}$	$\mathbf{b}^{max}$	$\boldsymbol{\mu}^{max}$	C	V
2	(20, 20)	(0.5, 0.5)	(1, 2)	(2, 2)	20	30

Исследуем зависимость вероятностей потерь π_1 и π_2 от нагрузки $\tilde{\rho}_1^{max}$ при постоянной нагрузке $\tilde{\rho}_2^{max} = 0{,}5$. Результаты, полученные при точном решении СУГБ (12), сравним с результатами, полученными приближенным алгоритмом. Как видно из рис. 4, в выбранном диапазоне нагрузок приближенный алгоритм дает достаточно точное решение.

На рис. 5 представлена зависимость вероятностей потерь от нагрузки $\tilde{\rho}_1^{max}$ для модели $\underset{\mathbf{N},\boldsymbol{\varepsilon},\mathbf{b}^{max}}{\mathbf{M}} \left| \underset{\boldsymbol{\mu}^{max}}{\mathbf{M}} \right| C \left| 0 \right.$ с неэластичными заявками и модели $\underset{\mathbf{N},\boldsymbol{\varepsilon},\mathbf{b}^{max}}{\mathbf{M}} \left| \underset{\boldsymbol{\mu}^{max}}{\mathbf{M}} \right| \underset{V}{<C, elast>} \left| 0 \right.$ с эластичными заявками при одинаковых структурных и нагрузочных параметрах (табл. 2). На диаграмме хорошо видно, что применение алгоритма со сжатием дает значительное сокращение вероятности потери.

Заключение

Таким образом, в работе построена модель обслуживания МПН-2.2 с эластичным трафиком. Представлен эффективный рекуррентный алгоритм приближенного расчета макрохарактеристик системы. Было исследовано влияние эластичности трафика на основные ВВХ. На основании численных примеров можно сделать вывод о том, что применение алгоритма доступа со сжатием дает возможность значительно сократить потери.

Литература

1. **Башарин Г.П.** Лекции по математической теории телетрафика / изд. 2-е. — М.: Изд-во РУДН, 2007; изд. 3-е, 2009.

2. **Наумов В.А., Самуйлов К.Е., Яркина Н.В.** Теория телетрафика мультисервисных сетей. — М.: Изд-во РУДН, 2007. — 191 с.

3. **Меликов А.З., Пономаренко Л.А., Паладюк В.В.** Телетрафик: Модели, методы, оптимизация. — Киев: ИПК Политехника, 2007. — 256 с.

4. **M. Glabowski, M. Stasiak.** An approximate model of the full-availability group with multi-rate traffic and a finite source population, Proc. 12th GI-ITG Conference on Measuring, Modeling and Evolution of Computer and Communication Systems (MMB) and 3rd Polish-German Teletraffic Symposium (PGTS), Dresden, Germany, P.195-204, Sept. 12-15, 2004. — P. 195 203.

5. **V. Vassilakis, I. Moscholios and M. Logothetis.** Call-level Performance Modelling of Elastic and Adaptive Service-classes with Finite Population, IEICE Trans. Commun., Vol. E91-B, №1, January 2008. — P. 151-163.

Оценка характеристик сигнального трафика в сети связи на базе подсистемы IMS

Сигнальный трафик в мультисервисной сети, построенной на базе подсистемы IMS, генерируется протоколом SIP при установлении, управлении и разъединении сеансов связи, а также протоколом Diameter при выполнении процедур авторизации, аутентификации и учета. До настоящего времени вопрос оценки объема сигнальной нагрузки и времени установления соединения в подсистеме IMS остается мало исследованным. В докладе, опираясь на исходные данные, полученные специалистами British Telecom [1], и на методы, разработанные в [3, 4], сделаны предварительные оценки важнейших характеристик сигнального трафика — величины сигнальной нагрузки между функциональными узлами подсистемы IMS и среднего времени установления соединения.

Самуйлов К.Е.,
зав. кафедрой систем телекоммуникаций РУДН,
ksam@sci.pfu.edu.ru

Сопин Э.С.,
магистрант кафедры систем телекоммуникаций РУДН,
sopin-eduard@yandex.ru

Чукарин А.В.,
доцент кафедры систем телекоммуникаций РУДН,
chukarin@yandex.ru

Постановка задачи исследований

IMS (IP Multimedia Subsystem) является решением, используемым при построении мультисервисных сетей, в которых услуги реализуются на основе протокола инициирования сеансов связи SIP (Session Initiation Protocol) [6]. Изначально концепция IMS разрабатывалась консорциумом 3GPP (3rd Generation Partnership Project) для сетей мобильной связи третьего поколения 3G, однако позднее она была предложена в качестве основы и для сетей фиксированной связи следующего поколения (NGN, Next Generation Network). В ддокладе рассматриваются функциональные элементы IMS, применяемые как в мобильных, так и в фиксированных сетях, и поэтому их специфика не влияет на общность излагаемых далее результатов.

Аналогично тому, как для сетей ТфОП, GSM и ИСС — система ОКС 7 является основным стеком протоколов сигнализации, для подсистемы IMS таким протоколом является SIP, который используется для установления, управления и разъединения сеансов связи. Для процедур авторизации, аутентификации и учета в IMS используется также протокол Diameter.

Дадим краткую характеристику подсистемы IMS в объеме, достаточном для дальнейшего изложения результатов. На рис. 1 показан набор функциональных объектов IMS, взаимодействующих при предоставлении услуг абонентам [4, 5, 6]. Предполагается, что абонент A осуществляет вызов абонента B, причем оба абонента находятся в своих домашних сетях (Home Network). Согласно стандарту сеть абонента A в таком случае называется Origination Network, а сеть абонента B — Terminating Network.

Основными функциональными элементами IMS являются модуль, реализующий функцию управления сеансами связи CSCF (Call Session Control Function) и сервер домашних абонентов HSS (Home Subscriber Server), фактически являющийся базой пользовательских данных.

В состав модуля CSCF включены три функции — функция обслуживания (S-CSCF, Serving CSCF), функция запроса (Interrogating CSCF, I-CSCF) и функция прокси-сервера (P-CSCF, Proxy CSCF). Функция S-CSCF предназначена для регистрации абонентов в базе данных HSS, а также осуществляет контроль и управление сеансами связи. Функция I-CSCF реализует доступ в домашнюю сеть и, в случае необходимости, может выполнять функции сетевого экрана, скрывая внутреннюю топологию сети. Функция P-CSCF обеспечивает первую точку контакта для терминала IMS, а также может выполнять шифрование и сжатие сообщений. Подчеркнем еще раз, что определенные выше объекты являются функциональными, т.е. для подсистемы IMS определена только функциональная архитектура, а физическая реализация данных функций зависит от решения конкретного производителя. Важно, что в общем случае при управлении сеансами связи функциональные узлы IMS (далее, для краткости — узлы IMS) обмениваются сигнальными сообщениями и создают сигнальный трафик, который должен обслуживаться с заданными параметрами качества, которые могут определяться требованиями стандартов, а также соглашениями об уровне обслуживания (SLA,

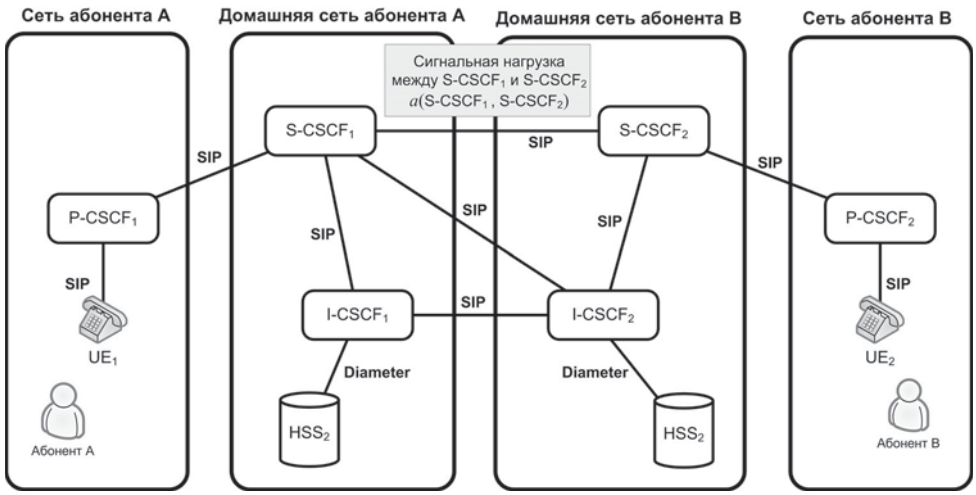


Рис. 1. Функциональные объекты IMS

Service Level Agreement). Узлы CSCF обмениваются между собой сообщениями протокола SIP, а для обращения к серверу HSS используется протокол Diameter. Терминалы абонентов (UE, User Equipment) соединяются посредством узлов P-CSCF тех сетей, в которых они на данный момент находятся и через эти функции взаимодействуют с другими узлами IMS, находящимися в их домашних сетях. При регистрации абонента в некоторой сети, функция P-CSCF этой сети определяет домашнюю сеть абонента и соответствующую функцию S-CSCF, в которой он обслуживается.

Введем теперь определения основных характеристик сигнального трафика, являющихся предметом исследования в данной работе. Под сигнальной нагрузкой $a(x, y)$ будем понимать нагрузку, создаваемую сообщениями протоколов SIP и Diameter между узлами x и y подсистемы IMS. Например, на рис. 1 показана величина $a(\text{S-CSCF}_1, \text{S-CSCF}_2)$ сигнальной нагрузки между узлами S-CSCF_1 и S-CSCF_2. Напомним, что величина a нагрузки определяется как $\alpha = \lambda T$, где λ — интенсивность потока сообщений, T — среднее время передачи сообщения.

Время установления соединения является важнейшим параметром качества обслуживания и определяется с момента, когда терминальное оборудование вызывающего абонента передало всю информацию, необходимую для установления соединения, до момента, когда это оборудование получило сигнал о состоянии терминального оборудования вызываемого абонента. На рис. 2 изображен фрагмент диаграммы одной из процедур, на которой показана случайная величина (СВ) Δ_{IMS} времени установления соединения в подсистеме IMS [4]. В рассматриваемой процедуре время установления соединения определяется, начиная с момента передачи терминалом абонента А к узлу P-CSCF сообщения INVITE до момента принятия терминалом абонента А сообщения 180 Ringing, отправленного терминалом абонента В.

Покажем на примере фрагмента процедуры установления соединения, отмеченном заштрихованными областями на рис. 2, как определяется СВ Δ_{IMS}. На рис. 3 детально показаны составные части СВ Δ_{IMS}, причем задержка, соответствующая той части процедуры, которая не приведена на рис. 3, обозначена как СВ $\tilde{\Delta}$.

Из рис. 3 следует, что формула для расчета СВ вид

$$\Delta_{IMS} = \Delta_{IP} + \Delta_{P-CSCF} + \Delta_{IP} + \Delta_{S-CSCF} + \\ + \Delta_{IP} + \tilde{\Delta} + \Delta_{IP} + \Delta_{S-CSCF} + \Delta_{IP} + \Delta_{P-CSCF} + \Delta_{IP}, \quad (1)$$

где $\Delta_{\text{P-CSCF}}$ и $\Delta_{\text{S-CSCF}}$ — времена обработки сообщений в узлах P-CSCF и S-CSCF соответственно, Δ_{IP} — задержка при передаче сообщения по IP-сети. Составляющие СВ $\tilde{\Delta}$, определяются аналогично и не включены в формулу (1) для краткости изложения. Заметим, что составляющие СВ Δ_{IMS} в формуле (1) для упрощения записи не снабжены индексами, отражающими их принадлежность тому или иному сегменту сети.

В подсистеме IMS стандартом определен набор сигнальных процедур для различных вариантов установления соединения. В разделе 2 доклада на основании результатов, представленных в [2], показано, что статистически на время установления соединения существенное влияние оказывают восемь основных процедур, диаграммы которых не приводятся, а расчет СВ Δ_{IMS} для каждой процедуры проводится по формуле аналогичной формуле (1).

Таким образом, мы определили основные понятия, необходимые для постановки задачи исследования, которая состоит в:

— разработке модели сигнального трафика подсистемы IMS, создаваемого сообщениями процедур протоколов SIP и Diameter (см. раздел 2);

— оценке значения величины сигнальной нагрузки $a(x, y)$ между

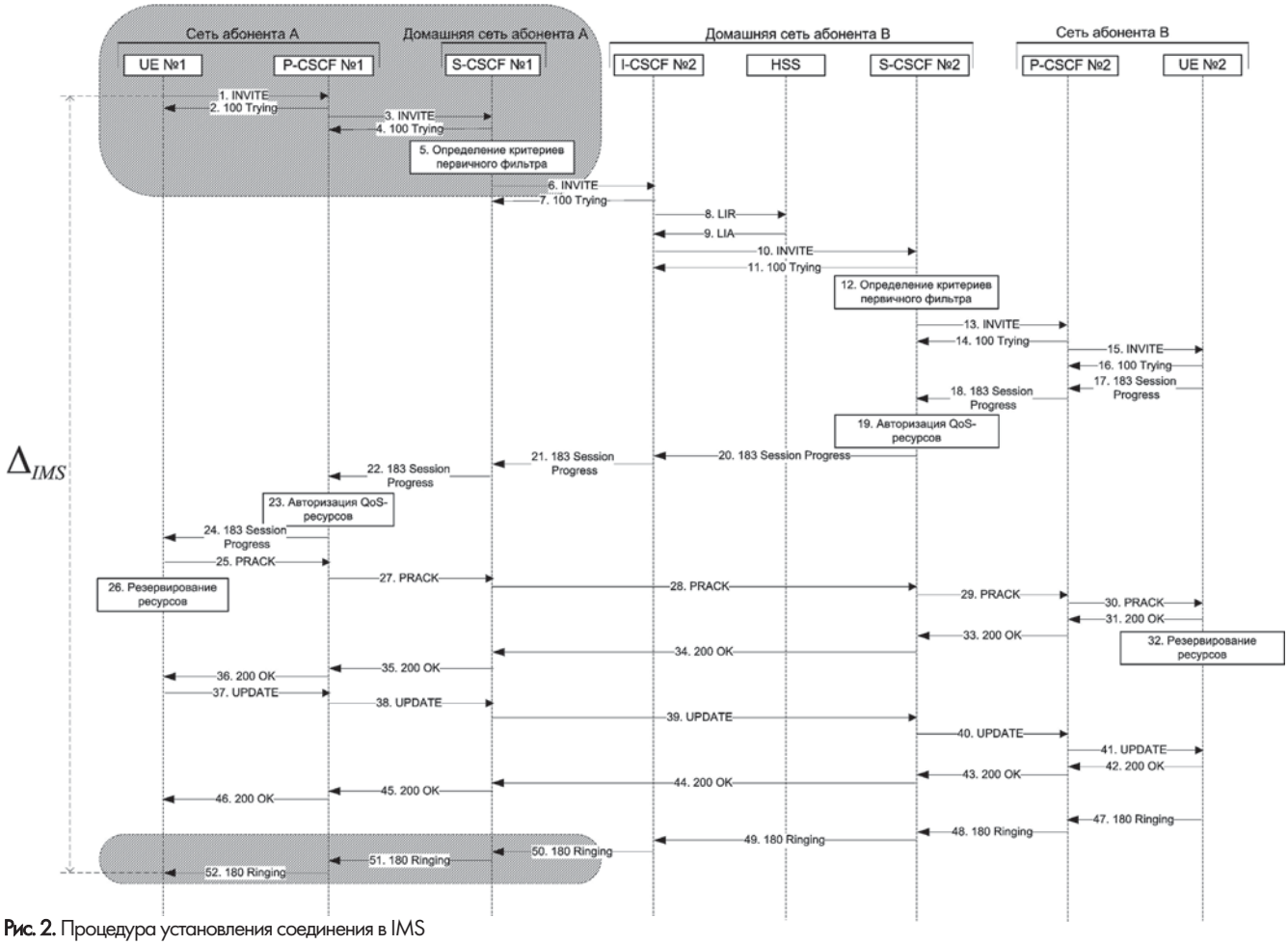


Рис. 2. Процедура установления соединения в IMS

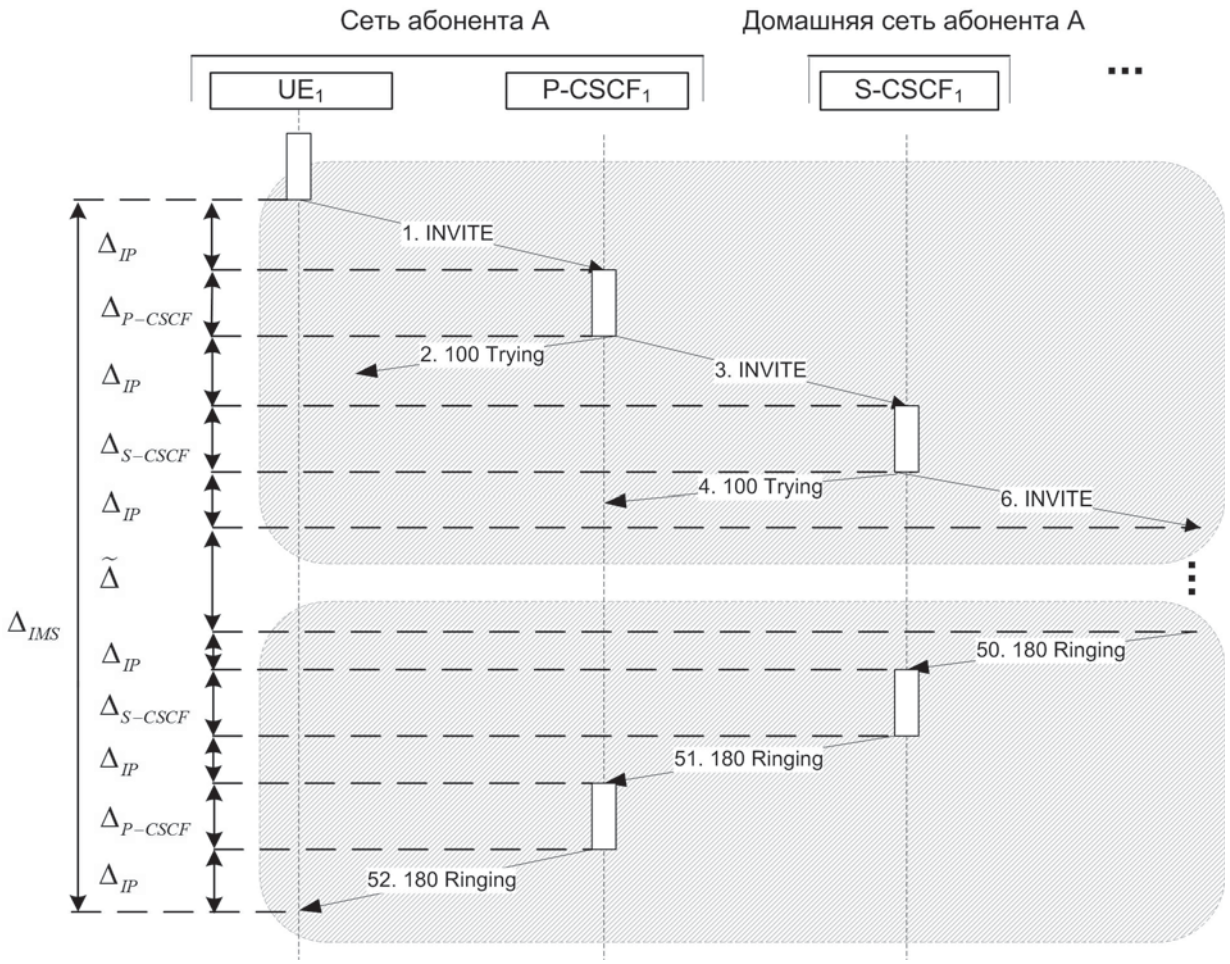


Рис. 3. Фрагмент процедуры установления соединения

узлами *x* и *y* подсистемы IMS (см. раздел 3);

— оценке среднего значения Δ_{IMS} времени установления соединения (см. раздел 4).

Модель сигнального трафика подсистемы IMS

Известно [7], что модель сигнального трафика включает: набор сигнальных процедур; значения объема передаваемой сигнальной информации, генерируемой сигнальными процедурами; процентное соотношение процедур в общем объеме трафика, порождаемого вызовами абонентов; состояния вызовов и их процентное соотношение.

В таблицах 1 и 2 приведены результаты построения модели сигнального трафика подсистемы IMS. При построении модели в расчет брались процедуры, описанные в версии 5 (Release 5) стандартов 3G консорциума

Таблица 1

№	Описание процедуры	Доля в общем трафике	Средняя длина сообщения (байт)	Среднее число сообщений	Средний объем сигнальной информации (байт)
1	Установление соединения, пользователь зарегистрирован	0,087	750	42	31500
2	Установление соединения, пользователь не зарегистрирован	0,01	750	44	33000
3	Регистрация	0,006	753	16	12048
4	Перерегистрация	0,051	855	6	5130
5	Отмена регистрации, инициированная сетью	0,001	798,33	6	4790
6	Отмена регистрации, инициированная пользователем	0,006	832	8	6656
7	Подписка на информацию присутствия	0,003	891	14	12474
8	Уведомление об изменении информации присутствия	0,836	816	4	3264

3GPP [4] (установление соединения, регистрация, перерегистрация, отмена регистрации), а также процедуры, реализующие новые услуги, описанные в версии 6 (Release 6) [5] (подписка на информацию присутствия и уведомление о статусе присутствия).

В таблице 1 приводятся данные по процентному соотношению процедур в общем объеме трафика, а также информация по объемам сигнальной информации. Расчеты долей производились на основе данных о структуре сигнального трафика подсистемы IMS сети оператора British Telecom [2].

Каждая процедура из табл. 1 с определенной вероятностью может находиться в некотором состоянии. Для процедур установления соединения (1 и 2) в соответствии с [8] определены состояния, перечисленные в табл. 2, а для всех остальных процедур определено единственное состояние — успешное выполнение с вероятностью 1.

Расчет длины сообщений протокола SIP основан на статистических данных [8], в то время как длины сообщений Diameter вычислены приблизительно, исходя из числа пар AVP (Attribute Value Pair) в каждом сообщении и средней длины поля AVP. При расчетах также учитываются длины заголовков протоколов нижних уровней, которые добавляются к сообщениям SIP при передаче по IP-сети. Отметим, что длины сообщений рассчитывались без учета их сжатия. Данные о длинах сигнальных сообщений представлены в табл. 3.

Построенная модель трафика содержит все данные, необходимые для оценки сигнальной нагрузки, создаваемой сообщениями соответствующих протоколов между узлами подсистемы IMS. Расчеты, результаты которых приведены в разделе 3, проведены с использованием методов, разработанных в [1].

Для расчета среднего времени установления соединения, где в качестве исходных данных также используется модель сигнального трафика, была построена математическая модель в виде открытой неоднородной сети массового обслуживания [9]. При разработке модели были использованы и частично модифицированы методы, разработанные в [1]. В данной работе мы не приводим весьма объемное и рутинное описание и анализ модели, отсылая заинтересованного читателя к упомянутым выше статьям. По мнению авторов, интерес представляют результаты вычислений и их обсуждения, представленные в разделе 4.

Таблица 2

№	Состояние процедуры (1) или (2)	Вероятность состояния
1	Успешное соединение	0,65
2	Абонент не отвечает	0,18
3	Абонент занят/отказ	0,17

Таблица 3

Сообщение SIP	Длина (байт)	Сообщение SIP	Длина (байт)	Сообщение Diameter	Длина (байт)
INVITE	976	ACK	676	UAR	618
100 Trying	496	BYE	556	UAA	728
183 Session Progress	956	REGISTER	856	SAR	728
PRACK	496	401 Unauthorized	726	SAA	798
200 OK	1036	SUBSCRIBE	946	LIR	548
UPDATE	496	NOTIFY	596	LIA	728
180 Ringing	496	486 Busy	496	MAR	728
				MAA	798

Оценка величины сигнальной нагрузки

Для расчета сигнальной нагрузки для пары $(x; y)$, где x и y — узлы IMS, обменивающиеся сигнальными сообщениями, необходимо определить средний объем сигнальной информации, передаваемой от x к y ($L(x,y)$) и от y к x ($L(y,x)$). Для расчета данных величин для каждого состояния каждой процедуры определяется общий объем передаваемой сигнальной информации на этом участке в обоих направлениях, а затем средний объем сигнальной информации с учетом приведенных долей вызовов и вероятностей состояний. Тогда сигнальная нагрузка вычисляется по формуле

$$a(x;y) = (L(y,x) + L(x,y))\frac{\Lambda(x;y)}{V}, \qquad (2)$$

где $\Lambda(x; y)$ — суммарная интенсивность выполнения сигнальных процедур, а V — скорость передачи данных.

На рис. 4 представлен результат вычислений сигнальной нагрузки между узлами подсистемы IMS для случая скорости передачи 2 Мбит/с. Видно, что наиболее загруженными при такой архитектуре сети являются звенья сети между P-CSCF и S-CSCF. Расчеты проводились при суммарной интенсивности выполнения всех процедур, соответствующей 15 выз/сек, при этом без ограничения общности, предполагается, что число вызовов в направлении от P-CSCF1 к P-CSCF2 и обратно совпадают.

Приведенный пример показывает, что сигнальные процедуры оказывают существенное влияние на объем трафика в IMS. Так, например, нагрузка порядка 0,1 Эрл (между P-CSCF и S-CSCF) является весьма значительной, учитывая, что рассматривается канал с шириной полосы пропускания 2 Мбит/с. Для сравнения, такая нагрузка соответствует нагрузке шестнадцати стандартных звеньев ОКС7 (64 Кбит/с) в нормальных условиях функционирования сети (т.е. при нагрузке на звено 0,2 Эрл).

Проведенные расчеты подтверждают, что сигнальный трафик IMS создает ощутимую нагрузку на IP-сеть и должен учитываться при расчетах сетей следующего поколения или сетей 3G. Разработанный метод может использоваться для расчета требуемой ширины полосы пропускания при резервировании ресурсов, необходимом для обеспечения требуемого качества обслуживания.

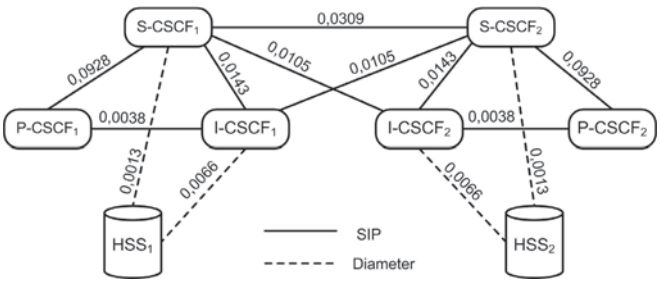


Рис. 4. Сигнальная нагрузка между узлами IMS

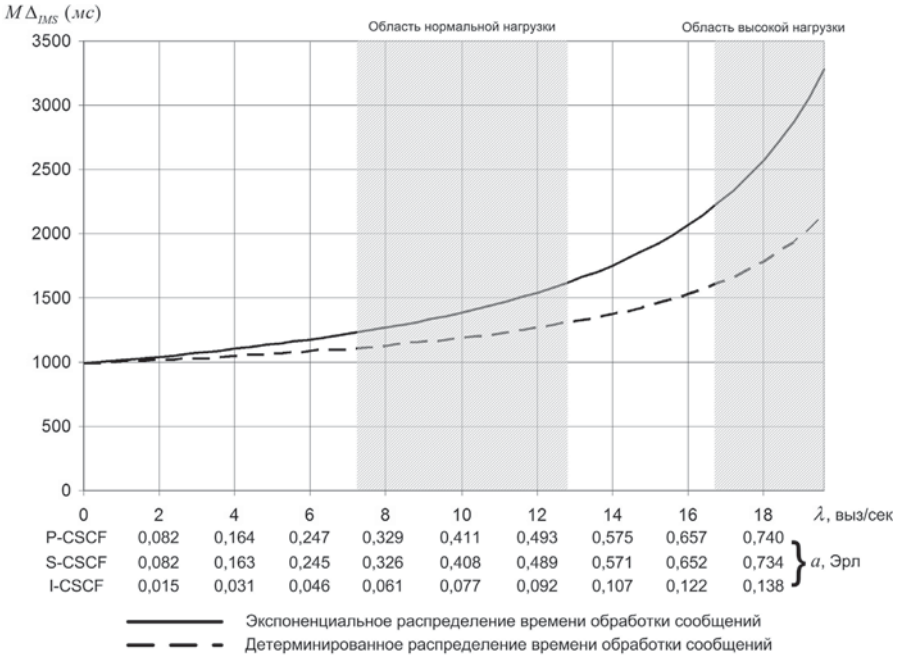


Рис. 5. Среднее время установления соединения в IMS

Оценка среднего времени установления соединения

Итак, время установления соединения согласно формуле (1) складывается из времен передачи сообщений по сети и времен обработки сообщений на узлах. В предположении, что все узлы функционируют независимо, и независимо передаются сообщения по сети, можно оценить среднее время установления соединения по формуле

$$M\Delta_{IMS} = M\Delta_{P-CSCF} + M\Delta_{S-CSCF} + M\widetilde{\Delta} + M\Delta_{S-CSCF} + \\ + M\Delta_{P-CSCF} + 3RTT. \quad (3)$$

В наших расчетах для простоты мы предполагаем, что время передачи по сети постоянно, и не зависит от маршрутизации сообщений, т.е. $M\Delta_{IP} = \frac{1}{2}RTT$. Таким образом, для вычисления $M\Delta_{IMS}$ необходимо вычислить величины средних времен обработки на узлах.

В [1] была разработана модель процедуры установления соединения по протоколу SIP в виде открытой сети массового обслуживания (СеМО) с заявками различных классов, которая позволяла оценить времена сериализации сообщений. Для расчета среднего времени установления соединения в IMS этот метод был модифицирован, при этом учитывались задержки в узлах сети, вызванные загрузкой в связи с обработкой сообщений процедур всех типов (табл. 1). При построении СеМО было сделано предположение, что вероятность ошибки или потери сообщения равна нулю, т.е. все сообщения всегда передаются и обрабатываются успешно, не передаются повторно. Считалось, что поток вызовов каждой процедуры является пуассоновским, время обработки с учетом их сериализации не зависит от типа сообщения, а все процедуры завершаются успешно. Узлы СеМО моделируют функциональные узлы IMS, заявкам же соответствуют выполняемые в подсистеме процедуры различных типов. Вызовы от абонентов поступают на узлы P-CSCF$_1$ и P-CSCF$_2$, и, таким образом, на СеМО поступают 16 потоков заявок (8 типов процедур в двух направлениях). Каждая заявка помимо расположения на определенном узле характеризуется отношением к некоторому классу, который используется для маршрутизации заявок по СеМО в соответствии с моделируемыми процедурами. Например, заявка класса 1 на узле 1 соответствует процедуре установления соединения (пользователь зарегистрирован), в которой на данный момент обрабатывается сообщение INVITE на узле P-CSCF1.

Полученная модель позволяет оценить величины средних времен обработки на узлах как средние времена пребывания заявок в узлах СеМО. Предполагая, что на каждый узел СеМО поступает пуассоновский поток заявок, для вычисления среднего времени пребывания необходимо знать интенсивность поступающего потока заявок на каждый узел. Такие интенсивности могут быть получены из системы уравнений равновесия, однако заметим, что для нашего случая (без повторных передач сообщений) решение тривиально и интенсивность потока в j-узле i-процедуры равна $\gamma_{ij} = n_{ij}\Lambda_i$, где n_{ij} — число сообщений, обрабатываемых j-узлом на одну i-процедуру, Λ_i — интенсивность j-потока заявок, поступающих на СеМО. Тогда интенсивность потока в узле $\lambda_j = \sum_{i=1}^{16}\lambda_{ij}$. Теперь, считая, что узлы функционируют независимо друг от друга, можно рассчитать среднее время выполнения каждой процедуры суммированием средних времен пребывания каждого сообщения этой процедуры на узлах, через которые проходят сообщения, а также добавив суммарное время передачи сообщений, связанных с передачей по IP-сети. Ниже приводятся численные результаты, полученные с помощью описанного выше метода, где предполагается, что время передачи сообщений по сети равно 10 мс.

Рассмотрим два случая — для экспоненциального и детерминированного распределения времени обслуживания в узлах СеМО. Среднее время обработки сообщения положим одинаковым (20 мс) для всех функциональных узлов (без учета ожидания обработки) и, без ограничения общности, будем считать, что интенсивности потоков вызовов, поступающих на P-CSCF$_1$ и P-CSCF$_2$, одинаковы. Графики зависимости среднего времени установления соединения от суммарной интенсивности поступления вызовов всех процедур представлены на рис. 5.

Из графика на рис. 5 видно, что среднее времени установления соединения меньше при детерминированном распределении времени обслуживания сообщений, поскольку в этом случае время ожидания в очереди на обслуживание на функциональных узлах в два

раза меньше по сравнению с экспоненциальным распределением. В области реальных значений нагрузок (0,3-0,5 Эрл) среднее время установления соединения находится в диапазоне 1-1,5 секунд. Учитывая, что рассматривались задержки только для протоколов SIP и Diameter, задержки в установлении соединения получились ощутимыми — напомним, что для телефонии МСЭ-Т определяет требования к задержкам при установлении соединения (до 3 сек на местные и до 5 сек на междугородные вызовы). Таким образом, вклад подсистемы IMS в задержки при передаче сообщений является значимым и должен учитываться при планировании сетей NGN.

Литература

1. **К.Е. Самуйлов, М.В. Лузгачев, О.Н. Плаксина.** Разработка вероятностной модели для анализа показателей качества протокола инициирования сеансов связи // Вестник Российского университета дружбы народов. Серия Математика. Информатика. Физика, 2007. — № 3-4. — С.53-63.

2. **V.S. Abhayawardhana, R. Babbage.** A traffic model for the IP Multimedia Subsystem (IMS) // 2007, IEEE Magazine. — P.783-787.

3. **Bobrikov N., Chukarin A., Luzgachev M.** Signaling and Intelligent Large-scale Network Dimensioning and Planning // The 13th IEEE Mediterranean Electrotechnical Conference, Melecon 2006, Benalmadena, Spain. — 2006. — P.725-728.

4. 3GPP TS 24.228 v5.13.0 (2005-06) 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Signalling flows for the IP Multimedia Call Control based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (Release 5).

5. 3GPP TS 29.328 v6.6.0 (2005-06) 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; IP Multimedia (IM) Subsystem Sh Interface; Signalling flows and message contents (Release 6).

6. Яновский Г.Г. IP Multimedia Subsystem: принципы, стандарты и архитектура // М.: Вестник связи, 2006. — №3 — С.71-76.

7. **Самуйлов К.Е.** Методы анализа и расчета сетей ОКС 7. — М.: Изд-во РУДН, 2002. — 292 с.

8. **Андрианов Г.А.; Самуйлов К.Е.; Гайдамака Ю.В.** Анализ модели трафика ОКС 7 по результатам обработки статистики измерений //Вестник связи, 2007. — №11. — С. 17-23

9. **Вишневский В.М.** Теоретические основы проектирования компьютерных сетей. — М.: Техносфера, 2003. — 512 с.

Анализ производительности одной схемы многоканальной передачи потоковых данных в одноранговых сетях

Современные динамические peer-to-peer (P2P) системы передачи потокового видео насчитывают сотни тысяч пользователей и предлагают сотни, а в ближайшем будущем — тысячи и более динамических каналов. В работе проведен анализ схемы многоканальной передачи потоковых данных, в которой учтено, что пользователи меняют каналы значительно чаще, чем открывают и закрывают видео-приложение.

Адаму Амину,
аспирант кафедры систем телекоммуникаций РУДН,
aminu@mail.ru

Гайдамака Ю.В.,
доцент кафедры систем телекоммуникаций РУДН,
ygaidamaka@mail.ru

Самуйлов А.К.,
бакалавр кафедры систем телекоммуникаций РУДН,
asam1988@gmail.com

Введение

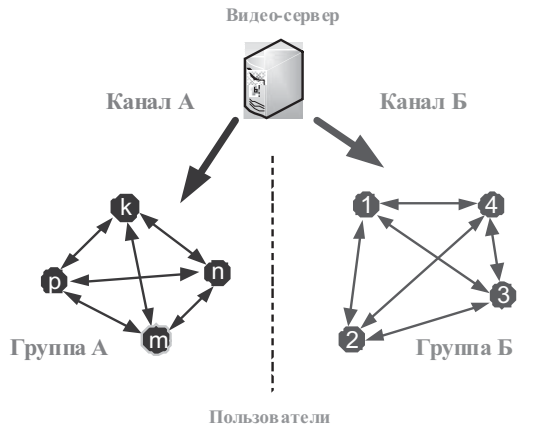
Технология P2P — это технология построения распределенной сети, где каждый узел может одновременно выступать и в роли клиента (получателя информации), и в роли сервера (поставщика информации). В P2P-сети (ее еще называют одноранговой, пиринговой или децентрализованной) в качестве клиента каждый из узлов может посылать запросы на предоставление каких-либо ресурсов другим узлам в пределах этой сети и получать их. Как сервер, каждый узел должен обрабатывать запросы от других узлов сети, отсылать то, что было запрошено, а также выполнять некоторые вспомогательные и административные функции. По сравнению с системами типа клиент-сервер система P2P потенциально имеет значительно большую масштабируемость и стабильность предоставления услуги из-за отсутствия точек критических отказов. Кроме того, она накладывается на уже существующие сети, то есть не требует создания дополнительной сетевой структуры.

Одним из приложений P2P-технологии является P2P-телевидение, которое предназначено для перераспределения видео потоков в P2P-сетях в режиме реального времени. Распределенные видео потоки могут иметь разные источники, но обычно они являются телевизионными каналами со всего мира. В системах P2P-телевидения каждый клиент, скачивая видео поток из сети, одновременно рассылает ("отдает") этот поток другим клиентам, занимая таким образом всю доступную полосу пропускания, снижая нагрузку на источник информации и зависимость каждого клиента от источника. Большинство крупномасштабных динамических P2P систем передачи потокового видео, запущенных в последние годы (напр., CoolStreaming, PPLive, PPStream, UUSee, Sopcast), основаны на едином принципе передачи информации, получившем название "модель изолированного канала" (ISO, isolated-channel model) [1]. Все пользователи, смотрящие один и тот же канал, образуют группу, внутри которой они обмениваются данными. Таким образом, пользователи внутри группы могут получать пакеты видео данных не только от сервера — источника трансляции, но и от остальных пользователей группы. При этом, как правило, поступающие к пользователям видео потоки запаздывают на некоторое время по сравнению с сервером — источником.

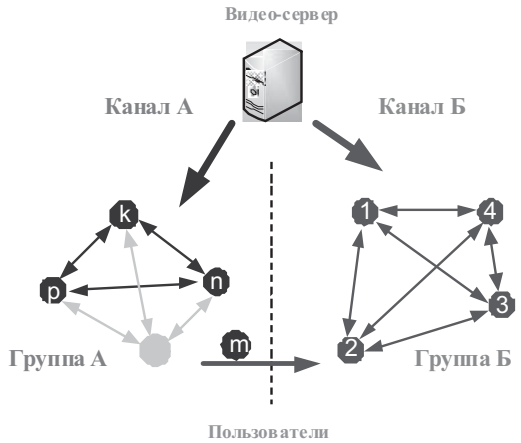
В работе для схемы "модель изолированного канала" построена аналитическая модель в виде замкнутой экспоненциальной сети массового обслуживания (CeMO), с помощью которой определены и исследованы некоторые показатели производительности схемы для разных по популярности каналов.

Функциональная модель процесса распределения видео потоков в P2P-сети

Рассмотрим P2P-сеть, которую образуют пользователи и видеосерверы, транслирующие телевизионные каналы. Для каждого пользователя определена скорость, с которой он может загружать данные из сети (скорость загрузки), и скорость, с которой он может отправлять данные в сеть (скорость отдачи), для видео-сервера определена только скорость отдачи. Для каждого канала определены требования к минимальной скорости загрузки пакетов видео данных пользователем. Все пользователи, смотрящие один и тот же канал, образуют одну группу: например, для канала А это группа А, состоящая из пользователей k, m, n, p, для канала Б группа Б, включающая пользователей 1-4 (рис. 1а).



а)



б)

Рис. 1. Фрагмент сети P2P-телевидения

Пользователь, просматривающий канал А, может получать пакеты видео данных от сервера — источника трансляции, а также от своих соседей по группе А. Когда пользователь (рис. 1б) переключается с канала А на канал Б, он прекращает отправлять видео данные пользователям группы А. В этом случае его соседи должны найти другие источники данных, чтобы поддерживать беспрерывное воспроизведение видео. В свою очередь, в группе Б вновь присоединившийся пользователь должен найти соседей, которые имеют возможность (достаточную скорость отдачи) отправлять ему видео данные. Заметим, что при переключении канала пользователем на единицу уменьшается число пользователей группы А и увеличивается число пользователей группы Б. Состояние сети может измениться при открытии/закрытии пользователем видео-приложения (например, при выключении персонального компьютера, с которого осуществлялся доступ в сеть P2P-телевидения), что приводит к увеличению/уменьшению на единицу общего числа пользователей в P2P-сети. Очевидно, что пользователи меняют каналы значительно чаще, чем открывают и закрывают видео-приложение. Если же в качестве устройств, подключенных к P2P-сети, рассматривать не только компьютеры, но и устройства домашней электроники с жестким диском и мощным процессором (например, цифровой видеомагнитофон или телевизионную приставку), постоянно подключенные к электропитанию и сети Интернет, то интенсивность смены пользователей будет настолько мала, что ей можно пренебречь. Будем рассматривать сеть, в которой пользователи подключены к P2P-сети постоянно, т.е. общее число пользователей сети не меняется, пользователи могут только переключаться с канала на канал.

Модель поведения пользователей в P2P-сети в виде сети массового обслуживания

Процесс переключения каналов пользователем сети P2P-телевидения может быть описан с помощью замкнутой СеМО $Q=\langle M,\Theta,N;\ \mu_j,c_j=\infty,r_j=0,j=\overline{1,M};\ FCFS\rangle$, состоящей из M узлов, каждый из которых соответствует каналу P2P-телевидения [2,3]. В сети циркулирует постоянное число N заявок, соответствующих пользователям сети P2P телевидения.

Узел j СеМО представляют собой систему массового обслуживания типа Infinite Server $\mathrm{M}|\mathrm{M}|\infty$, $j=\overline{1,M}$. Длительность обслуживания заявок в узле имеет экспоненциальное распределение с параметром μ_j, что соответствует среднему времени $1/\mu_j$ просмотра пользователем j-го канала, $j=\overline{1,M}$.

Процесс переключения каналов пользователем сети P2P телевидения рассматривается как процесс блуждания некоторой заявки (пользователя) между узлами СеМО (каналами). Пусть t_k — случайный момент k-го ($k=0, 1,...$) окончания просмотра канала пользователем (и переключения на следующий канал), и пусть $v_k=v(t_k+0)$ — номер канала, на который переключился пользователь, $v_k=\overline{1,M}$. Тогда процесс блуждания заявки описывает однородная цепь Маркова (ЦМ) $\{v_k,\ k\geq 0\}$, вложенная по моментам поступления заявки в узлы сети. Также обозначим $\theta_{ij}=P\{v_{k+1}=j\,|\,v_k=i\}$ вероятность перехода из состояния i в состояние j за один шаг, $i,j=\overline{1,M}$, тогда матрица $\Theta=(\theta_{ij}),\ \ i,j=\overline{1,M}$ — матрица переходных вероятностей ЦМ $\{v_k,\ k\geq 0\}$.

В силу неразложимости матрицы Θ все состояния рассматриваемой вложенной конечной ЦМ сообщаются и, следовательно, все они положительно-возвратны. Поэтому ранг системы уравнений равновесия

$$x_j=\sum_{i=1}^{M}x_i\theta_{ij},\ j=\overline{1,M},$$

или в матричном виде $\mathbf{x}^T\Theta=\mathbf{x}^T$, равен $M-1$, и она имеет положительное решение $\mathbf{x}^T=(x_1,\ldots,x_M)>\mathbf{0}^T$, определяемое с точностью до произвольной постоянной.

Если нормировать единицей это решение:

$$p_j=\frac{x_j}{\sum_{i=1}^{M}x_i},\ i=\overline{1,M},$$

то мы получим стационарное распределение вероятностей $\mathbf{p}^T=(p_1,\ldots,p_M)>\mathbf{0}^T$ для данной ЦМ. Стационарные вероятности p_j можно интерпретировать как число попаданий в состояние j на большом интервале времени, $j=\overline{1,M}$.

Если же выделить некоторый узел i^* и нормировать решение $x^T=(x_1,..., x_M)$ так, чтобы

$$h_j=\frac{p_j}{p_{i^*}},\ j\neq i^*,\ h_{i^*}=1.$$

то h_j можно интерпретировать как среднее число посещений некоторой заявкой узла j за период длительности τ_{i^*}, начавшийся с момента поступления этой заявки в узел i^* и до момента ее возвращения в этот узел. Поэтому в замкнутых сетях величина

$$d_j=\frac{h_j}{\mu_j} \tag{1}$$

имеет смысл среднего значения суммарной трудоемкости, предъявляемой каждой заявкой узлу j, $j=\overline{1,M}$, в указанном интервале времени.

Пусть $X_j(t)$ — число заявок в узле j в момент t, $t\geq 0$. Тогда M-мерный ступенчатый марковский процесс $X(t)=(X_1(t),..., X_M(t)$, $t\geq 0$, описывающий функционирование замкнутой сети при сделанных ранее предположениях, будет регулярным и неприводимым над конечным пространством состояний

$$\mathbb{X}(M,N)=\left\{\mathbf{n}:\ n_j=0,1,\ldots,N;\ \sum_{j=1}^{M}n_j=N\right\}. \tag{2}$$

При этом $|\mathbb{D}(M,N)|=D(M,N)$ — число способов, которыми можно разместить N неразличимых частиц (заявок) в M ячейках (узлах), которое легко получить из комбинаторных соображений:

$$D(M,N)=\binom{M+N-1}{M-1},\qquad M\geq 1, N\geq 0.$$

По теореме Гордона-Ньюэла [4] для замкнутой экспоненциальной СеМО независимо от начального распределения существует равновесное распределение $\{p(\mathbf{n}),\ \mathbf{n}\in\mathbb{D}(M,N)\}$ ступенчатого процесса $X(t)$, $t\geq 0$, которое имеет мультипликативный вид:

$$p(\mathbf{n})=G^{-1}(M,N)\prod_{j=1}^{M}\frac{d_j^{n_j}}{\beta_j(n_j)},\ \mathbf{n}\in\mathbb{X}(M,N), \tag{3}$$

где d_j определяются формулой (1), а нормирующая константа $G(M, N)$ находится из нормировочного условия для распределения (3):

$$G(M,N)=\sum_{\mathbf{n}\in\mathbb{X}(M,N)}\prod_{j=1}^{M}\frac{d_j^{n_j}}{\beta_j(n_j)},\qquad M\geq 1,\quad N\geq 0. \tag{4}$$

Согласно [5] для замкнутой однородной СеМО с узлами типа Infinite Server стационарное распределение вероятностей СП $X(t)$, $t\geq 0$, имеет вид

$$p(\mathbf{n})=N!\prod_{j=1}^{M}\frac{\rho_j^{n_j}}{n_j!},\ \mathbf{n}\in\mathbb{X}(M,N), \tag{5}$$

где $\rho_j = \frac{d_j}{\sum_{i=1}^{M} d_i}$ — относительная популярность канала $j, j = \overline{1, M}$, $\sum_{j=1}^{M} \rho_j = 1$.

Формула (5) является частным случаем формулы (3). Действительно, из (2) $\sum_{j=1}^{M} n_j = N \ \forall \mathbf{n} \in \mathbb{Z}(M, N)$. Кроме того, для узла типа Infinite Server $\beta_j(n_j) = n_j!$, откуда нормировочная константа (4) имеет вид $G(M, N) = N!$.

Очевидным следствием формулы (5) является биномиальный вид маргинальных распределений числа n_j заявок по узлам СеМО:

$$P(n_j) = \frac{N!}{n_j!(N-n_j)!} \rho_j^{n_j} \left(1-\rho_j\right)^{N-n_j}, \ j = \overline{1, M}. \quad (6)$$

Основные показатели качества функционирования одноранговой сети

Для того, чтобы пользователь сети P2P-телевидения мог смотреть выбранный канал с заданным уровнем качества (минимальными требованиями являются отсутствие пауз в воспроизведении, синхронизация звука и изображения и т.п.), нужно, чтобы пользователь загружал необходимые для просмотра видео пакеты со скоростью, удовлетворяющей требованиям просматриваемого канала. Таким образом, скорость загрузки пользователя должна быть не меньше требуемой для просматриваемого канала. Как правило, скорость отдачи пользователя ниже скорости загрузки, поэтому проблемы при просмотре могут возникать, если в какой-то момент пользователь не сможет найти источники (видео-сервер, транслирующий канал, или другого пользователя группы), которые отдавали бы ему пакеты видео данных с требуемой скоростью. Заданный уровень качества услуги P2P-телевидения при модели изолированного канала легче обеспечить для пользователей, просматривающих более популярные каналы, поскольку качество телевизионного канала зависит от того, сколько пользователей в данный момент смотрят этот канал: чем многочисленнее группа, тем быстрее каждый ее пользователь найдет источники загрузки и получит необходимые для просмотра видео пакеты.

Будем говорить, что какой-либо канал находится в состоянии всеобщей передачи, если каждый пользователь из группы этого канала получает видео данные со скоростью, удовлетворяющей требованиям просматриваемого канала. Если же все пользователи сети P2P-телевидения получают видео данные со скоростями, удовлетворяющими требованиям просматриваемых каналов, то система P2P-телевидения находится в состоянии всеобщей передачи. Поскольку ресурсы сети ограничены, состояние всеобщей передачи для какого-либо канала или для системы в целом возможно лишь в некоторые периоды времени. Вероятности нахождения канала или всей системы в состоянии всеобщей передачи являются характеристиками качества функционирования одноранговой сети.

Определим $\sigma_j(n_j)$ — индекс ресурсов канала *j* следующим образом [1]:

$$\sigma_j(n_j) = \frac{b_j}{d_j(n_j)}, \ j = \overline{1, M},$$

где n_j — случайное число пользователей группы канала *j*; b_j — суммарная скорость отдачи пакетов видео данных сервером — источником и другими пользователями группы канала *j*, доступная пользователю, просматривающему указанный канал; $d_j(n_j)$ — суммарная скорость загрузки пакетов видео данных, необходимая для того, чтобы обеспечить состояние всеобщей передачи для канала *j* при заданном числе n_j пользователей группы канала *j*.

Тогда вероятность PU_j состояния всеобщей передачи для канала *j* определяется формулой

$$PU_j = P\left\{\sigma_j(n_j) \geq 1\right\}, \ j = \overline{1, M},$$

а вероятность *PS* нахождения системы в состоянии всеобщей передачи определяется формулой

$$PS = P\left\{\sigma_j(n_j) \geq 1, \ j = \overline{1, M}\right\}.$$

Введем следующие обозначения:

N — постоянное число пользователей сети P2P-телевидения;

N_j — множество пользователей, просматривающих канал $j, j = \overline{1, M}, \left|N_j\right| = n_j \leq N$;

v_j — скорость отдачи пакетов данных канала *j* видео-сервером;

u_i — скорость отдачи потока пользователем $i, i = \overline{1, N}$;

r_j — требования к минимальной скорости загрузки пакетов видео данных канала *j* пользователем.

С учетом введенных обозначений уточним определение вероятности нахождения системы в состоянии всеобщей передачи:

$$PS = P\left\{v_j + \sum_{i \in \mathbb{Z}_j} u_i \geq n_j r_j, \ j = \overline{1, M}\right\}.$$

Предположим, что в сети имеется два типа пользователей: N^l пользователей с низкой скоростью u^l отдачи и N^h пользователей с высокой скоростью u^h отдачи видео потока, $N^l + N^h = N$. Обозначим n_j^l и n_j^h случайное число пользователей группы канала *j*, имеющих соответственно низкую и высокую скорости отдачи: $n_j^l + n_j^h = n_j, 0 \leq n_j^l \leq N^l, 0 \leq n_j^h \leq N^h$.

Тогда вероятность PU_j состояния всеобщей передачи для канала *j* может быть представлена в виде

$$PU_j = P\left\{v_j + \sum_{i \in \mathbb{N}_j} u_i \geq n_j r_j\right\} = P\left\{v_j + u^h n_j^h + u^l n_j^l \geq (n_j^h + n_j^l) r_j\right\} =$$

$$= \sum_{x=0}^{N^h} \sum_{y=0}^{N^l} P\left\{v_j + u^h x + u^l y \geq (x+y) r_j \mid n_j^h = x, n_j^l = y\right\} \cdot P\left\{n_j^h = x\right\} \cdot P\left\{n_j^l = y\right\} =$$

$$= \sum_{x=0}^{N^h} \sum_{y=0}^{N^l} I(v_j + u^h x + u^l y \geq (x+y) r_j) \cdot P\left\{n_j^h = x\right\} P\left\{n_j^l = y\right\}, \ j = \overline{1, M}, \quad (7)$$

где $I(a)$ — индикаторная функция:

$$I(a) = \begin{cases} 1, & a \text{ - истинно;} \\ 0 & \text{в противном случае.} \end{cases}$$

Вероятности $P\left\{n_j^h = x\right\}$ и $P\left\{n_j^l = y\right\}$ согласно формуле (6) имеют вид

$$P\left\{n_j^h = x\right\} = \binom{N^h}{x} \rho_j^x \left(1-\rho_j\right)^{N^h - x}, \ j = \overline{1, M}, \quad (8)$$

$$P\left\{n_j^l = y\right\} = \binom{N^l}{y} \rho_j^y \left(1-\rho_j\right)^{N^l - y}, \ j = \overline{1, M}. \quad (9)$$

Пример численного анализа

Целью численного анализа является сравнение вероятностей PU_j состояния всеобщей передачи для разных по популярности каналов. Расчет проводился по формулам (7-9) с использованием программного средства Mathcad.

Рассмотрен фрагмент сети P2P-телевидения, состоящий из *N* = 1800 пользователей и видео-сервера, транслирующего *M* = 20 каналов [1]. Все каналы имеют одинаковые требования к минимальной скорости загрузки пакетов видео данных пользователями:

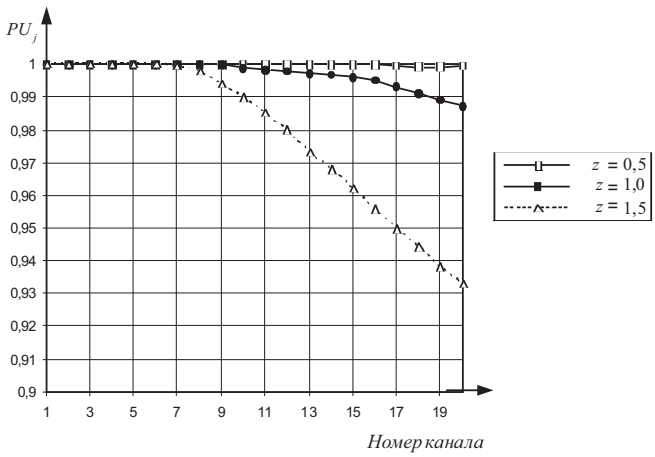


Рис. 2. Вероятности PU_j состояния всеобщей передачи канала в зависимости от популярности канала при различных значениях параметра Ципфа

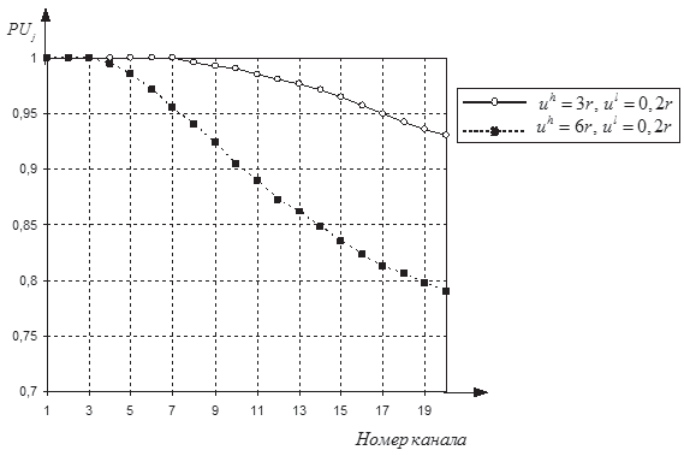


Рис. 3. Вероятности PU_j состояния всеобщей передачи канала в зависимости от популярности канала при различных значениях высокой скорости отдачи пользователей

$r_j = r,\ j = \overline{1, M}.$ Пользователи сети имеют различные скорости отдачи: часть пользователей (их доля f) низкую скорость отдачи u^l, остальные ($1-f$) — высокую скорость отдачи u^h, $0 \le f \le 1$.

Величины популярности каналов распределены по закону Ципфа, т.е.

$$\rho_j = \frac{1}{j^z \sum_{i=1}^{M} \frac{1}{i^z}},\ j = \overline{1, M},$$

где z — параметр распределения Ципфа. Таким образом, каналы упорядочены в порядке убывания популярности.

На рис. 2 для трех значений параметра распределения Ципфа $z \in \{0,5;\ 1;\ 1,5\}$ приведены значения вероятностей PU_j состояния всеобщей передачи для канала $j,\ j = \overline{1, 20}.$ Расчеты проводились в предположении о равном количестве пользователей с низкой и с высокой скоростями отдачи, т.е. $f = 0,5$, $N^l = N^h = 900$. Значения скоростей отдачи выбраны равными $u^l = 0,2r$, $u^h = 3r$.

Из рис. 2 видно, что при уменьшении популярности канала вероятность состояния всеобщей передачи для канала падает. Для первых пяти по популярности каналов вероятности состояния всеобщей передачи близки к единице при любом значении параметра z. Для менее популярных каналов вероятности состояния всеобщей передачи для разных значений z различаются: так, для наименее популярного канала (канал №20) при $z = 0,5$ вероятность $PU_{20} = 0,999$, а при $z = 1,5$ вероятность $PU_{20} = 0,933$. Менее популярные каналы сильнее страдают при больших значениях параметра z, когда среднее число пользователей, просматривающих соответствующий канал, становится маленьким. Низкие значения вероятностей состояния всеобщей передачи в этом случае объясняются тем, что из-за малого количества пользователей в группе часто не получается эффективно организовать сеть и моментально переключить потоки в случае, когда кто-то из пользователей прекращает передачу данных.

На рис. 3 приведены значения вероятностей PU_j состояния всеобщей передачи канала $j,\ j = \overline{1, 20}$ для двух различных значений высокой скорости отдачи пользователей: $u^h = 3r$, $u^h = 6r$. Значение низкой скорости отдачи для обоих случаев выбрано равным $u^l = 0,2r$. В каждом случае доля f пользователей, имеющих низкую скорость отдачи, подобрана так, чтобы средняя скорость отдачи $\overline{u} = u^l f + u^h (1-f)$ в обоих случаях была одинаковой. Расчеты проводились для значения параметра распределения Ципфа $z = 1,5$.

Из рис. 3 видно, что вероятности состояния всеобщей передачи канала для случая $u^h = 6r$ ниже, чем для случая $u^h = 3r$. Действительно, при одинаковой средней скорости $\overline{u}$ отдачи доля (и, соответственно, число) пользователей, имеющих высокую скорости отдачи, меньше для случая $u^h = 6r$, чем для случая $u^h = 3r$. А чем меньше размер группы пользователей канала, тем труднее каждому ее пользователю найти источники, которые обеспечили бы требуемую для просматриваемого канала скорость загрузки.

В заключение авторы выражают благодарность заведующему кафедрой систем телекоммуникаций РУДН д.т.н., проф. К.Е. Самуйлову за советы в подборе литературы и полезные дискуссии в процессе исследований.

Литература

1. **Wu D., Liu Y., Ross K.W.** Queuing Network Models for Multi-Channel Live Streaming Systems // IEEE Infocom, Rio de Janeiro, April 2009. — P. 73-81.

2. **Вишневский В.М.** Теоретические основы проектирования компьютерных сетей. — М.: Техносфера, 2003. — 512 с.

3. **Башарин Г.П.** Лекции по математической теории телетрафика: Учеб. Пособие. Изд. 3-е, испр. и доп. — М.: РУДН, 2009. — 342 с.

4. **Harrison J.M., Lemoine A.J.** A note on networks of infinite-server queues. Journal of Applied Probability, 1981, 18, №2. — P. 561-567.

5. **Gordon W.J., Newell G.F.** Closed queueing systems with exponential servers. Operations Research, 1967, 15, №2. — P. 254-265.

Имитационная модель узла управления услугами интеллектуальной сети

Методика расчета задержек, возникающих в процессе установления соединения для услуг интеллектуальной сети связи (ИСС), предложенная в предыдущих работах авторов, предусматривает оценку времени обработки сообщений в узлах ИСС. В докладе построена имитационная модель процесса обработки сообщения в узле управления услугами SCP (Signalling Control Point) в виде открытой сети массового обслуживания с экспоненциальным и детерминированным временем обслуживания в узлах, которая позволяет оценить некоторые показатели функционирования ИСС.

Бузюкова И.Л.,
аспирант кафедры сетей связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ),
irbuz@mail.ru

Бязров А.С.,
бакалавр кафедры систем телекоммуникаций РУДН,
alanbyazrov@gmail.com

Гайдамака Ю.В.,
доцент кафедры систем телекоммуникаций РУДН,
ygaidamaka@mail.ru

Введение

Одним из основных параметров качества обслуживания в интеллектуальной сети связи (ИСС) является задержка "из-конца-в-конец", возникающая в процессе установления соединения. В предыдущих работах авторов [1,2] была предложена методика расчета задержек для основных услуг, реализованных на ИСС России. Для оценки задержек при предоставлении услуг интеллектуальной сети согласно этой методике необходимо уметь рассчитывать время обработки сигнальных сообщений в узлах ИСС. В работе построена модель узла управления услугами (SCP Signaling Control Point), представляющего собой программно-управляемую базу данных, доступ к которой реализован при помощи протоколов верхнего уровня, таких как TCAP (Transaction Capabilities Applications Part — подсистема управления возможностями транзакций) и INAP (Intelligent Network Application Part прикладная подсистема интеллектуальной сети). Функциональная архитектура узла SCP определяется тем, как подсистемы MTP (Message Transfer Part подсистема передачи сообщений), SCCP (Signalling Connection Control Part подсистема управления соединениями сигнализации), TCAP и INAP системы сигнализации по общему каналу №7 (ОКС 7) делят между собой ресурсы сети процессоры узла SCP. В работе рассмотрена модель узла SCP с распределенной архитектурой [3].

Функциональная модель узла SCP

Распределенная архитектура узла управления услугами SCP (рис. 1) предусматривает один центральный модуль с центральным процессором для обслуживания функциональных процессов подсистем TCAP и INAP, а также связанных периферийных модулей, в каждом из которых обрабатываются сообщения, входящие/выходящие по закрепленным за этим модулем звеньям сигнализации. Внутри каждого периферийного модуля функциональные процессы подсистем MTP2 обслуживаются одним процессором, подсистем MTP3 и SCCP другим процессором.

Процессоры на рис. 1 схематически обозначены белыми прямоугольниками. Внутренняя (по отношению к SCP) сеть передачи данных позволяет передавать сообщения между центральным и периферийными модулями. Сигнальное сообщение, поступившее в узел SCP по звену сигнализации, попадает в один из периферийных модулей, в котором обрабатывается в соответствии с процедурами и функциями подсистемы MTP2, а после этого в соответствии с процедурами и функциями подсистем MTP3 и SCCP. Затем сообщения со всех периферийных модулей по внутренней сети передачи данных поступают в центральный модуль, где в порядке поступления становятся в единую очередь для обработки подсистемами TCAP и INAP. После обработки центральным процессором сообщение поступает в соответствующий исходящему звену сигнализации периферийный модуль, где обрабатывается подсистемами SCCP — MTP3 — MTP2 для передачи в пункт назначения [4,5].

Для анализа задержки обработки сигнальных сообщений в узле SCP воспользуемся методом декомпозиции и агрегации. Суть метода состоит в разбиении сложной системы на блоки и последующем детальном анализе отдельных блоков системы. Такими блоками для узла SCP являются процессы обработки сообщения одним процессором. На каждую подсистему узла SCP поступают потоки сообщений, относящихся к большому числу соединений, поэтому потоки

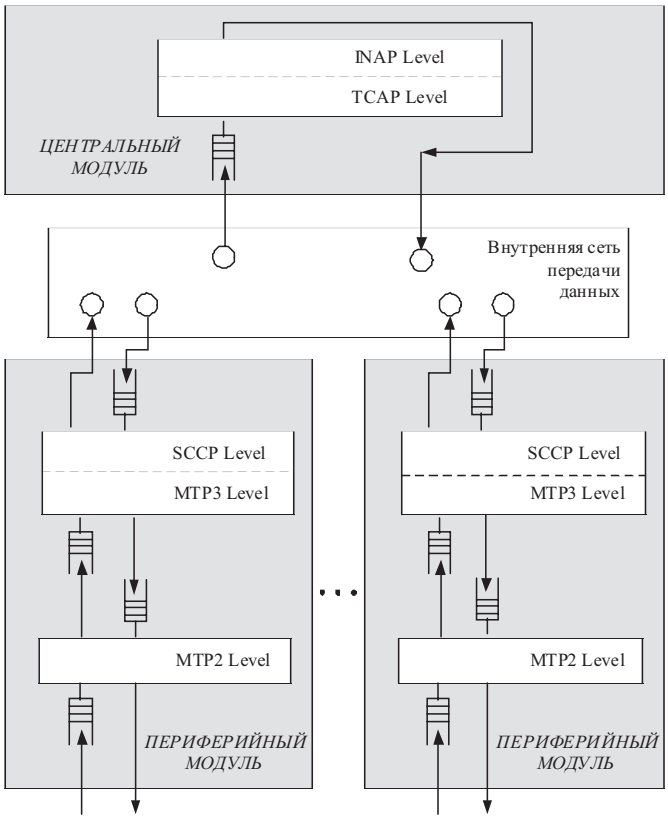


Рис. 1. Функциональная модель узла SCP с распределенной архитектурой

сообщений, принадлежащие к индивидуальным соединениям, не отличаются от соответствующего агрегированного потока. Агрегированный процесс поступления сообщений в каждую подсистему аппроксимируется пуассоновским процессом [6,7,8]. Применение метода декомпозиции и агрегации позволяет после вычисления времени обработки сообщения в каждом процессоре узла SCP получить общую величину времени обработки сообщения в узле SCP.

Модель узла SCP в виде сети массового обслуживания

Процесс обработки сигнального сообщения в узле SCP с распределенной архитектурой может быть описан с помощью открытой сети массового обслуживания (СеМО), схематично изображенной на рис. 2. Сеть, в которую поступают заявки из внешнего источника, состоит из $(2n + 2)$-х узлов. Заявки в рассматриваемой СеМО соответствуют сигнальным единицам, обрабатываемым узлом управления услугами. Поток заявок из внешнего источника является суперпозицией пуассоновских потоков (по числу периферийных модулей) интенсивности λ/n каждый. Будем считать все потоки, поступающие в узлы сети, пуассоновскими. Это предположение верно только для случая экспоненциального распределения времени обслуживания заявки в узлах СеМО, однако, такое предположение значительно упрощает анализ СеМО. В этом случае все узлы являются взаимно независимыми и для каждого в отдельности могут быть получены в аналитическом виде формулы для расчета задержек в этом узле. Оценка точности построенной с учетом этого предположения аналитической модели сделана ниже с помощью имитационного моделирования.

В СеМО имеются узлы четырех типов: узлы типа 1 — узлы i ($i = 1,\ldots, n$), узлы типа 2 узлы j ($j = n + 1,\ldots,2n$), узел типа 3 — узел $(2n + 1)$, узел типа 4 — узел $(2n + 2)$.

Обслуживание заявки в узле типа 1 соответствует обработке сообщения подсистемой MTP2 одного из периферийных модулей и моделируется с помощью однолинейной системы массового обслуживания (СМО) $M_2/G_2/1/\infty/f_1^0$ с относительным приоритетом.

Обслуживание заявки в узле типа 2 СеМО соответствует обработке сообщения подсистемами MTP3 и SCCP того же периферийного модуля и моделируется с помощью однолинейной СМО $M_2/G_2/1/\infty/$.

Передача сообщения по внутренней сети передачи данных между центральным и периферийными модулями соответствует обслуживанию заявки в узле типа 3, которое моделируется с помощью СМО типа $M/G/\infty$ обслуживание без ожидания Infinite Server (IS).

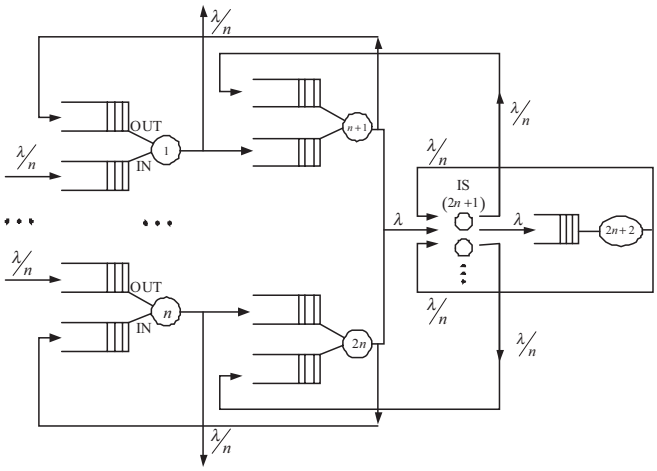


Рис. 2. Открытая СеМО для узла SCP

Обслуживание заявки в узле типа 4 соответствует обработке сигнального сообщения подсистемами TCAP и INAP центрального модуля и моделируется с помощью однолинейной СМО $M/G/1/\infty$.

Так как интенсивность поступления сигнальных сообщений одинакова для каждого из периферийных модулей, будем считать интенсивности входящих потоков для периферийных модулей равными. Поскольку процессоры одного и того же уровня для каждого из периферийных модулей имеют одинаковую производительность, будем считать средние времена обслуживания заявки узлами одного типа равными.

Маршрутная матрица $\Theta = (\theta_{i,j}),\ i, j = \overline{1,\ 2n+2}$, для рассматриваемой модели сети имеет вид, показанный в табл. 1. Здесь узел 0 соответствует внешней среде. Заметим, что матрица Θ — стохастическая (сумма по строке равна 1).

Имитационная модель узла SCP

Имитационная модель, представленная в работе, выполнена с помощью общецелевой системы имитационного моделирования GPSS World Student Version (General Purpose Simulation System,

Таблица 1

Матрица переходов Θ

	0	1	...	n-1	n	n+1	n+2	...	2n-1	2n	2n+1	2n+2
0	0	$1/n$	…	$1/n$	$1/n$	0	0	…	0	0	0	0
1	1/2	0	…	0	0	1/2	0	…	0	0	0	0
2	1/2	0	…	0	0	0	1/2	…	0	0	0	0
...	…	0	…	0	0	0	0	…	0	0	0	0
n-1	1/2	0	…	0	0	0	0	…	1/2	0	0	0
n	1/2	0	…	0	0	0	0	…	0	1/2	0	0
n+1	0	1/2	…	0	0	0	0	…	0	0	1/2	0
n+2	0	0	…	0	0	0	0	…	0	0	1/2	0
...	…	…	…	…	…	…	…	…	…	…	…	…
2n-1	0	0	…	1/2	0	0	0	…	0	0	1/2	0
2n	0	0	…	0	1/2	0	0	…	0	0	1/2	0
2n+1	0	0	…	0	0	$1/2n$	$1/2n$	…	$1/2n$	$1/2n$	0	1/2
2n+2	0	0	…	0	0	0	0	…	0	0	1	0

GPSS). Имитационная модель позволяет оценить среднее время пребывания заявки в СеМО и сравнить эту оценку с результатами для аналитической модели. Исходными данными для моделирования являются интенсивность поступающего в СеМО извне потока заявок и распределение случайного времени обслуживания заявки для каждого узла СеМО. Интенсивность входящего потока соответствует интенсивности потока сигнальных сообщений, поступающих в узел управления услугами в часы наибольшей нагрузки (ЧНН). Случайные длительности обслуживания заявки в узлах СеМО представляют собой длительности обработки сигнального сообщения соответствующими процессорами узла SCP. Расчеты проведены для исходных данных из [2,8]:

b_1 = 0,01875 с — среднее время обработки сообщения подсистемой MTP2;

b_2 = 0,00650 с — среднее время обработки сообщения подсистемами MTP3/SCCP;

b_3 = 0,00100 с — среднее время передачи сообщения по внутренней сети передачи данных;

b_4 = 0,00750 с — среднее время обработки сообщения подсистемами TCAP/INAP.

Среднее время пребывания заявки в СеМО было вычислено для 100 тыс. заявок, прошедших через имитационную модель. Время работы программы составило менее 10 секунд, в зависимости от интенсивности поступающего на узел потока за это время было проведено моделирование от 13 минут (при λ = 130 вызовов/с) до 27 часов (при λ = 1 вызов/с) работы узла управления услугами.

Результаты численного анализа

Целью численного анализа является оценка случайного времени обработки сигнального сообщения в узле SCP для случаев экспоненциального и детерминированного распределений времени обработки сообщения процессорами узла SCP. Эта характеристика соответствует времени пребывания заявки в СеМО.

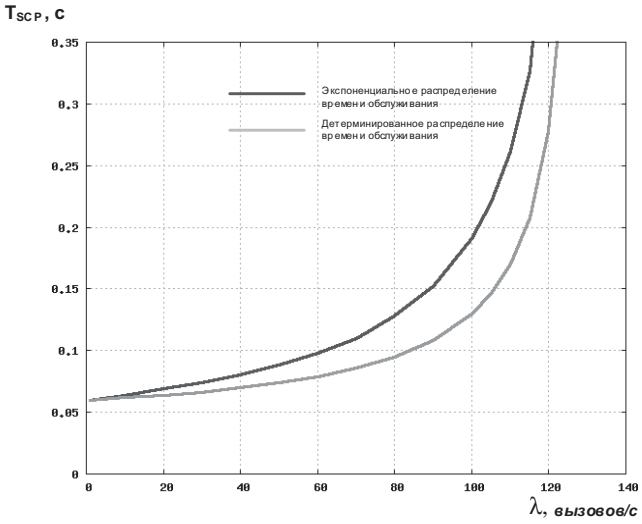


Рис. 3. Зависимость среднего времени обработки сообщения в узле SCP от интенсивности входящего потока

Расчет проводился для интенсивности поступления вызовов λ от 0 до 130 вызовов/с. Диапазон изменения значений λ выбран таким, чтобы выполнялись условия существования стационарного режима для отдельных узлов СеМО.

На рис. 3 приведен график зависимости среднего времени T_{SCP} обработки сигнального сообщения в узле управления услугами от интенсивности входящего потока для обоих рассмотренных случаев распределения времени обслуживания заявки в узлах СеМО.

Таблица 2

Результаты имитационного моделирования

λ, вызовов/с	$T_{SCP}^{эксп}$, с	$T_{SCP}^{дет}$, с	Абсолютная погрешность α	Относительная погрешность δ
1	0,060	0,060	0	0
10	0,064	0,062	0,032258	0,002
20	0,069	0,064	0,078125	0,005
30	0,074	0,066	0,121212	0,008
40	0,080	0,070	0,142857	0,010
50	0,088	0,074	0,189189	0,014
60	0,098	0,079	0,240506	0,019
70	0,110	0,086	0,279070	0,024
80	0,128	0,095	0,347368	0,033
90	0,152	0,108	0,407407	0,044
100	0,191	0,130	0,469231	0,061
110	0,262	0,170	0,541176	0,092
120	0,448	0,278	0,611511	0,170
130	1,753	1,211	0,447564	0,542

Из рис. 3 видно, что время $T_{SCP}^{эксп}$ обработки сигнального сообщения в узле SCP при экспоненциальном распределении длительностей обслуживания в узлах СеМО служит оценкой сверху для времени $T_{SCP}^{дет}$ обработки сигнального сообщения в узле SCP при детерминированном распределении длительностей обслуживания. О точности этой оценки можно судить по значениям абсолютной α и относительной δ погрешностей, значения которых приведены в табл. 2. Здесь

$$\alpha = \left|T_{SCP}^{эксп} - T_{SCP}^{дет}\right|, \delta = \frac{\left|T_{SCP}^{эксп} - T_{SCP}^{дет}\right|}{\left|T_{SCP}^{дет}\right|}.$$

Из таблицы 2 видно, что при значениях нагрузки до 100 вызовов/с относительная погрешность не превышает 5%, что приемлемо для инженерных расчетов. Таким образом, предложенная в предыдущих работах авторов [1,2] методика расчета задержек для основных услуг, реализованных на ИСС России, основанная на анализе модели открытой экспоненциальной СеМО, может быть рекомендована для расчета этой характеристики. Заметим, что для значения λ = 15 000 вызовов/час (или 4,167 вызовов/с) интенсивности потока сигнальных сообщений, которое соответствует реальному значению интенсивности суммарного потока сигнальных сообщений при предоставлении шести основных услуг набора возможностей CS1, реализованных на интеллектуальной сети РФ, в ЧНН, среднее время $T_{SCP}^{эксп}$ обработки сообщения в узле SCP составляет 62 мс.

Заключение

Результаты вычислений позволяют оценить временные задержки сообщений в узле управления услугами интеллектуальной сети. При нагрузке сети до 100 вызовов/с среднее значение задержки в узле SCP составляет от 60 мс до 200 мс. Полученный результат соответствует требованиям международных стандартов: время обработки сообщения в узле SCP с вероятностью 99,999% не должно превышать 300 мс [9].

Литература

1. **Бязров А.С., Гайдамака Ю.В.** Имитационная модель процесса установления соединения для одной услуги интеллектуальной сети // Сб. "Современные телекоммуникации и математическая теория телетрафика" // Тр. XLV Всеросс. научной конф. факультета физ.-матем. наук РУДН, апрель 2009. — М.: Изд-во РУДН, 2009. — С. 160-162.

2. **Buzyukova Irina, Gaidamaka Yulia, Yanovsky Gennady.** Estimation of GoS parameters in Intelligent Network // Proc. of 9th Int.Conf. NEW2AN 2009, September 15-18, 2009. — Saint-Petersburg, 2009. — P.143-153.

3. EURESCOM: Project P308, Methods and specifications for tools to dimension Intelligent Networks — Dimensioning and Technical Planning for Intelligent Networks. Del. №3, Vol.2, March 1996. — 130 p.

4. ITU-T: White Book, Recommendation Q.706: Signalling System №7 — Message Transfer Part Signalling Performance. Geneva, March 1993. — 36 p.

5. ITU-T: White Book, Recommendation Q.1600. Signalling System № 7 — Interaction between ISUP and INAP. Geneva, December 2000. — 45 p.

6. **Bafutto M., Kuehn P., Willmann G.** Capacity and performance analysis of signalling networks in multivendor environments. IEEE Journal on selected areas in communications. Vol.12. №3. April 1994. — P.490-500.

7. **Willman G., Kuehn P.J.** Performance Modelling of Signalling System. IEEE Communications Magazine Vol.7, №28, 1990. — P.44-56.

8. **Самуйлов К.Е., Филюшин Ю.И.** Оценка среднего значения времени установления соединения для услуг интеллектуальной сети. Электросвязь, №6, 1996. — С. 14-16.

9. ITU-T Recommendation Q.1218. Intelligent network. Interface recommendation for Intelligent Network CS-1. ITU, October 1995. — 32 p.

Модели разделения ресурсов звена мультисервисной сети с эластичным трафиком

Рассматриваются различные стратегии разделения ресурсов звена мультисервисной сети, основанные на модели эластичного трафика с потерями. Предложено несколько постановок задачи разделения пропускной способности между виртуальными частными сетями. Отдельно проанализирован случай без требований к пропускной способности, т. е. когда вероятности блокировок запросов на передачу блоков данных равны нулю.

Гудкова И.А.,
ассистент кафедры систем телекоммуникаций РУДН,
igudkova@gmail.com

Лузгачев М.В.,
аспирант кафедры систем телекоммуникаций РУДН,
mluzgachev@mail.ru

Введение

Виртуальные частные сети (Virtual Private Network, VPN) являются широко применяемым решением для организации передачи данных по сетям общего пользования [4], они могут использоваться для передачи трафика различного типа. Существуют две основные категории трафика: потоковый [1] и эластичный [2]. Ниже рассматриваются сети, в которых передается только эластичный трафик [2,3,6]. С точки зрения планирования VPN важной задачей является разделение ресурсов сети, в частности пропускной способности звеньев, — необходимо оптимально разделить пропускную способность между VPN, критериями оптимальности могут быть доход, вероятности блокировок, время передачи данных и др. Задача разделения пропускной способности для потокового трафика рассматривалась в [4,5], причем были рассмотрены как одноадресные, так и многоадресные соединения [1]. В [4] исследована модель с потерями, критерием оптимальности является доход. Предложен алгоритм для звена с одноадресными и многоадресными соединениями. В [5] предложен метод решения задачи оптимального планирования VPN (маршрутизация и разделение пропускной способности).

Рассмотрим задачу разделения пропускной способности сети между VPN. Пусть имеется сеть из L звеньев, C^l — пропускная способность l-звена сети. Предположим также, что в сети построены VPN, а C_n^l — пропускная способность l-звена VPN. В VPN передается эластичный трафик разного типа. Необходимо оптимально разделить пропускные способности звеньев между VPN, критериями оптимальности могут служить вероятности блокировок, величина времени передачи блоков, величина обслуженной нагрузки. Необходимо найти такие C_n^l, $\sum_{n=1}^{N} C_n^l = C^l$, что целевая функция принимает максимальное (минимальное) значение, $n=\overline{1,N}$, $l=\overline{1,L}$. В предположении, что пропускные способности всех кроме одного звеньев сети бесконечны, данная постановка сводится к задаче разделения пропускной способности отдельного звена. Предложенные нагрузки определяются исходя из маршрутизации на сети. Далее в первом разделе представлена модель отдельного звена мультисервисной сети с эластичным трафиком, во втором — модели разделения пропускной способности.

Анализ вероятностно-временных характеристик звена с эластичным трафиком

Модели с эластичным трафиком, в первую очередь, характеризуются дисциплиной обслуживания типа разделения процессора (Processor Sharing, PS) [2]. Входящий поток либо зависит от состояния системы [6], либо не зависит он него [3,6]. Управление доступом блоков эластичных данных к звену может осуществляться как при помощи введения различных порогов на максимальное число блоков всех типов в звене [6], так и посредством задания требований блоков к емкости звена. В работе [6] предлагается в качестве требований использовать максимально возможное число единиц емкости, которую может занять один блок. Минимальные требования, необходимые для обслуживания эластичного трафика, например, использовались в модели [3], но ее авторы, тем не менее, не предложили эффективного алгоритма расчета характеристик модели.

Рассматривается звено сети емкостью единиц, по которому передаются блоки эластичных данных K типов. Запросы на передачу k-блоков образуют пуассоновский поток интенсивности λ_k, а длина k-блока является случайной величиной, распределенной экспоненциально со средним θ_k. Блок k-типа характеризуется минимальным требованием b_k к емкости звена, $k=\overline{1,K}$, причем $1 \le b_1 < b_2 < ... < b_k \le C$. Обслуживание блоков происходит по дисциплине справедливого разделения процессора (Egalitarian Processor Sharing, EPS) [2].

Пусть состояние модели описывает вектор $(n_1,...,n_k)$, где n_k — число k-блоков на обслуживании, $k=\overline{1,K}$. Обозначим $n=\sum_{k=1}^{K} n_k$ суммарное число блоков в состоянии $(n_1,...,n_k)$, тогда пространство всех возможных ее состояний имеет вид

$$S=\left\{(n_1,\ldots,n_K):\quad n_k=0,1,\ldots,\ k=\overline{1,K}:\quad n\cdot \max_{k=\overline{1,K}:\ n_k>0}(b_k)\le C\right\} \tag{1}$$

Нетрудно убедиться, что функционирование модели описывается марковским процессом с мультипликативным стационарным распределением вероятностей состояний

$$p(n_1,\ldots,n_K)=p(0,\ldots,0)\frac{n!}{C^n}\prod_{k=1}^{K}\frac{\rho_k^{n_k}}{n_k!},\quad (n_1,\ldots,n_K)\in S. \tag{2}$$

Основными вероятностно-временными характеристиками модели являются вероятность блокировки запроса на передачу k-блока, определяемая по формуле $B_k=\mathrm{P}\{(n_1,\ldots,n_K)\in S_k\}$, где S_k — множество состояний блокировок, и среднее время T_k передачи k-блока, вычисляемое по формуле Литтла через среднее число k-блоков в системе. Для эффективного вычисления B_k и T_k применим метод, основанный на разбиении пространства состояний модели по двум параметрам: максимальному из требований к емкости звена блоков на обслуживании и общему числу передаваемых блоков. Характеристики выражаются через соответствующие ненормированные вероятности $q(k, n)$ того, что по звену передаются блоков с максимальным из требований b_k:

$$B_k=\beta_k(C)=\left(1+\sum_{i=1}^{K}\sum_{n=1}^{N_i}q(i,n)\right)^{-1}\times \\ \times\left(\sum_{i=1}^{k-1}\sum_{n=N_k}^{N_i}q(i,n)+\sum_{i=k}^{K}q(i,N_k)\right),\quad k=\overline{1,K}, \tag{3}$$

$$T_k = \tau_k(C) = \left(1 + \sum_{i=1}^{K}\sum_{n=1}^{N_i} q(i,n)\right)^{-1} \frac{\theta_k}{C(1-B_k)} \times$$
$$\times \left(1 + \sum_{i=1}^{k}\sum_{n=2}^{N_k} nq(i,n) + \sum_{i=k+1}^{K}\sum_{n=2}^{N_i} nq(i,n)\right), \quad k = \overline{1,K}, \tag{4}$$

где $q(k, n)$ определяется из рекуррентного соотношения

$$q(k,n) = \begin{cases} \dfrac{\rho_k}{C}, & n = 1, \\ \dfrac{1}{C}\left(\rho_k \sum_{i=1}^{k-1} q(i,n-1) + q(k,n-1)\sum_{i=1}^{k}\rho_i\right), & n = 2,\ldots,N_k, \end{cases}$$

$N_k = \lfloor C/b_k \rfloor$, $\rho_k = \lambda_k \theta_k$. Заметим, что вероятность блокировки (3) и среднее время (4) для k-блока представляют собой соответственно функции $\beta_k(C)$ и $\tau_k(C)$ от параметра C, $k = \overline{1,K}$. При решении задачи оптимального разделения емкости звена между имеющими к нему доступ виртуальными частными сетями, значение переменной будет равняться пропускной способности звена в конкретной VPN.

Модели разделения ресурсов звена

Для виртуальных частных сетей с эластичным трафиком могут быть использованы различные стратегии разделения ресурсов. Каждой стратегии разделения ресурсов соответствует задача оптимального разделения ресурсов звена, но перед тем как перейти к данным задачам, рассмотрим модель звена виртуальной частной сети.

Пропускная способность звена равна C единиц. Пусть C_n — пропускная способность VPN_n, $\sum_{n=1}^{N} C_n = C$. Обозначим θ_k среднюю длину блока данных, b_k — требование к минимальной пропускной способности. Запросы на передачу k-блоков в VPN_n образуют пуассоновский поток интенсивности λ_n^k. В случае нехватки пропускной способности происходит блокировка запроса на передачу блока данных с вероятностью $\beta_n^k(C_n)$. Обозначим $\tau_n^k(C_n)$ среднее время передачи k-блока в VPN_n, $k = \overline{1,K}$, $n = \overline{1,N}$.

Перейдем к рассмотрению задач разделения пропускной способности. Задача имеет вид

$$\min \sum_{n=1}^{N}\sum_{k=1}^{K} \tau_n^k(C_n),$$
$$\sum_{n=1}^{N} C_n = C,$$
$$0 \le C_n \le C, \quad n = \overline{1,N}. \tag{5}$$

В данном примере целевая функция — сумма средних времен передачи блоков данных всех типов во всех VPN. Это лишь одна из возможных целевых функций (стратегий). В таблице также приведены другие возможные стратегии.

Для простоты в рассмотренных моделях не используются веса, однако они могут быть введены. Также могут использоваться ограничения на вероятности блокировок и средние времена передачи блоков данных. Отметим, что для модели без ограничений ($b_k = 0$, $k = \overline{1,K}$.) могут использоваться только первые две стратегии разделения, так как вероятности блокировок равны нулю.

На рис. 1 приведены результаты расчета примера для первых двух стратегий. Рассматривается модель трафика без ограничений. Пропускная способность делится между двумя VPN, в сети передаются блоки данных только одного типа. Для модели без ограничений среднее время передачи блока k-типа может быть рассчитано по формуле

$$\tau_n^k(C) = \frac{\theta_k}{C - \sum_{k=1}^{K} \lambda_n^k \theta_k}, \quad k = \overline{1,K}, \quad n = \overline{1,N}. \tag{6}$$

На рис. 1 а) изображена зависимость суммы средних времен от (стратегия 1), на рис. 1 б) — зависимость максимального времени передачи блока от (стратегия 2), а на рис. 1 в) — средние времена в VPN1 и VPN2 по отдельности. На графиках видно и аналитически может быть доказано, что для такой задачи решения для обеих стратегий совпадают. Также видно и может быть показано, что при оптимальном разделении ресурсов средние времена передачи блоков в двух VPN равны.

Стратегии разделения пропускной способности

№	Название стратегии	Целевая функция
1	Минимизация суммарного среднего времени передачи блоков данных	$\min \sum_{n=1}^{N}\sum_{k=1}^{K} \tau_n^k(C_n)$
2	Минимизация максимального среднего времени передачи блоков данных	$\min \max_{k=\overline{1,K},\, n=\overline{1,N}} \tau_n^k(C_n)$
3	Минимизация средней величины вероятностей блокировок	$\min \dfrac{\sum_{n=1}^{N}\sum_{k=1}^{K} \lambda_n^k \beta_n^k(C_n)}{\sum_{n=1}^{N}\sum_{k=1}^{K} \lambda_n^k}$
4	Минимизация максимальной величины вероятности блокировки	$\min \max_{k=\overline{1,K},\, n=\overline{1,N}} \beta_n^k(C_n)$
5	Максимизация обслуженной нагрузки	$\max \sum_{n=1}^{N}\sum_{k=1}^{K} \lambda_n^k \theta_k \left(1 - \beta_n^k(C_n)\right)$

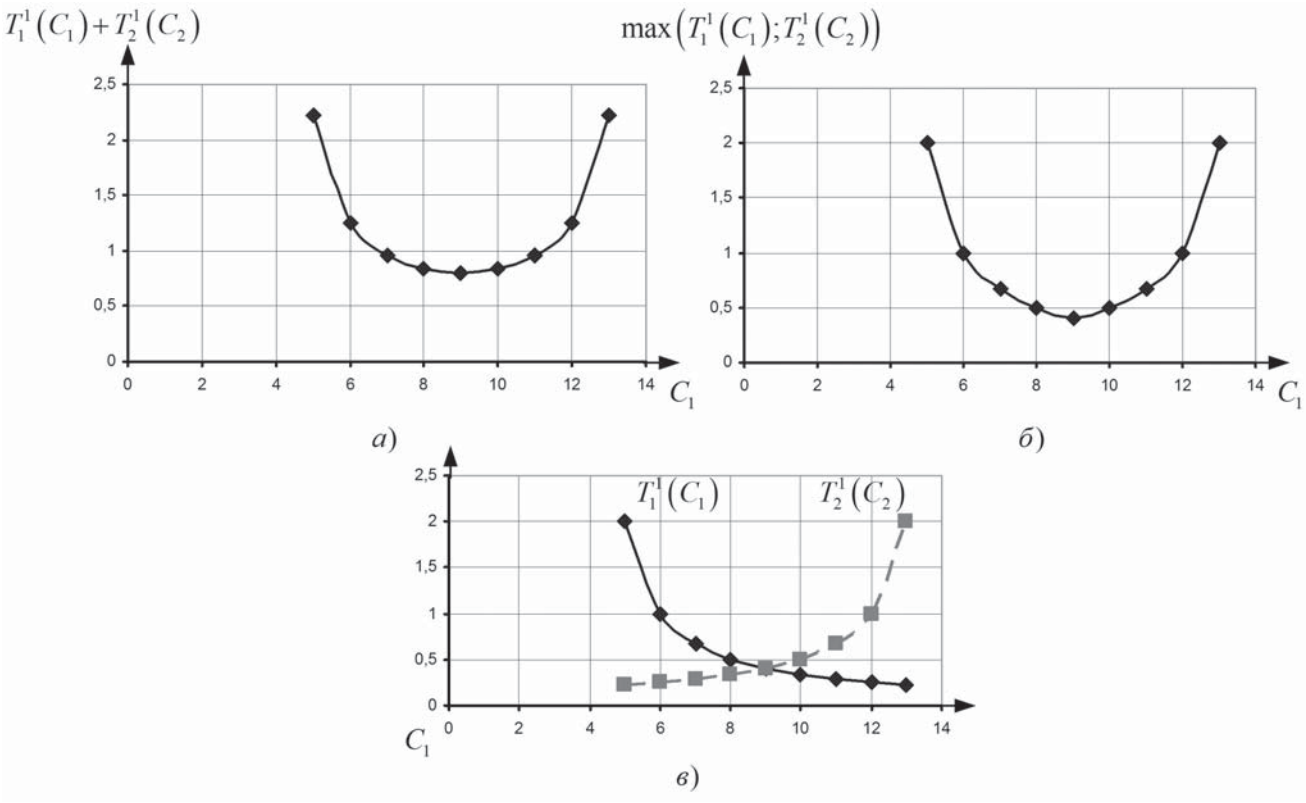


Рис. 1. а) Зависимость суммы средних времен от C_1; б) Зависимость максимального среднего времени от C_1; в) Зависимость средних времен от C_1

Заключение

В дальнейших исследованиях планируется рассмотрение стратегий разделения ресурсов, основанных на различных моделях эластичного трафика, как в одной VPN, так и в разных. Также в задаче разделения будут добавлены ограничения на параметры качества (границы для вероятностей блокировок и времени передачи блока данных). Еще одним важным открытым вопросом является решение задачи разделения ресурсов мультисервисной сети с эластичным трафиком.

Литература

1. **Башарин Г.П., Самуйлов К.Е., Яркина Н.В., Гудкова И.А.** Новый этап развития математической теории телетрафика // Автоматика и телемеханика, 2009. — № 12. — С. 16-28.

2. **Яшков С.Ф.** Математические вопросы теории систем обслуживания с разделением процессора // Итоги науки и техники. Серия "Теория вероятностей. Математическая статистика. Теоретическая кибернетики", 1990. — № 29. — С. 3-82.

3. **Boxma O.J., Gabor A.F., N??ez-Queija R., Tan H.P.** Performance analysis of admission control for integrated services with minimum rate guarantees // Proc. of the 2nd Conference on Next Generation Internet Design and Engineering, Val?ncia, Spain, 3-5 April 2006. — P. 41-47.

4. **Luzgachev M., Samouylov K.** On the resource allocation problem for a multiservice network link with unicast and multicast connections // Proc. of the International Conference on Ultra Modern Telecommunications, St.-Petersburg, 12-14 October 2009.

5. **Mitra D., Morrison J.A., Ramakrishnan K.G.** Virtual private networks: joint resource allocation and routing design // Proc. of the 18th Annual Joint Conference of the IEEE Computer and Communications Societies, New York, USA, 21-25 March 1999. — Vol. 2. — P. 480-490.

6. **Vassilakis V., Moscholios I., Logothetis M.** Call-Level performance modelling of elastic and adaptive service-classes with finite population // IEICE Transactions on Communications. — 2008. — Vol. 91-B, №1. — P.151-163.

К анализу вероятностей блокировок трафика услуг IPTV

Рассматривается модель звена сети мультивещания в виде системы массового обслуживания с прозрачными заявками. Исследуется приближенный метод расчета вероятности блокировки вызова услуги для двух сценариев дисциплин обслуживания.

Бутурлин И.А.,
магистр 1-го года обучения кафедры системы телекоммуникаций РУДН,
ivan_buturlin@mail.ru

Плаксина О.Н.,
ассистент кафедры системы телекоммуникаций РУДН,
oplaksina@sci.pfu.edu.ru

Введение

Услуга мультивещания занимает фиксированную емкость звена сети в независимости от числа обслуживаемых пользователей, что позволяет значительно экономить сетевые ресурсы. Мультивещание (multicasting) является одной из составляющих пакетной услуги TriplePlay [1], включающей передачу данных, речи и видео, а одним из наиболее актуальных примеров использования мультивещания является передача трафика вещательного телевидения по IP-сетям. Кроме того, мультивещание применяется в ряде других приложений, в том числе, при организации видеоконференций, дистанционного обучения, рассылке корпоративной информации, сетевых играх и пр.

Рассматриваются два сценария предоставления услуги мультивещания. В первом случае (дисциплина $П_1$) принципиальную роль играет пользователь, активизирующий услугу, и, который определяет начало и окончание сессии мультивещания. Остальные пользователи могут присоединяться к активизированной услуге, однако, завершение сессий всех пользователей происходит одновременно в момент завершения сессии пользователем-инициатором услуги. Такой сценарий может применяться в сетевых играх или видеоконференциях. Разработке соответствующих математических моделей посвящены, например, работы [1-4]. Во втором случае (дисциплина $П_2$) пользователи могут присоединяться к услуге мультивещания и отключаться в любой момент времени. Этот сценарий соответствует просмотру пользователями канала вещательного телевидения. Модель и приближенный метод анализа ее характеристик разработаны в [5].

Задача состоит в расчете вероятности блокировки запроса пользователя на предоставление услуги мультивещания. Метод расчета вероятностей блокировок, применяемый в [5,6] для дисциплины $П_2$, расширен для дисциплины $П_1$.

Модель звена сети мультивещания

Рассматривается звено сети мультивещания емкостью C единиц, по которому однородным пользователям предоставляются услуги из множества $M = \{1,...,M\}$, причем m-услуга требует b_m единиц емкости звена. Каждая услуга может находиться в состоянии "1" — включена, если она предоставляется хотя бы одному из пользователей, или в состоянии "0" — выключена, в противном случае.

Функционирование звена описывается в терминах системы массового обслуживания $M|M|C|0|П$ с "прозрачными заявками" [4], на которую поступает M пуассоновских потоков заявок с интенсивностями $\lambda_1,..., \lambda_M$. Исследуются две дисциплины "прозрачного" обслуживания. В первом случае (дисциплина $П_1$) период занятости b_m приборов m-заявками определяет первая из поступивших заявок, а в момент окончания ее обслуживания систему одновременно покидают все m-заявки, поступившие за время ее обслуживания, незамедлительно освобождая занятые приборы. Длительность обслуживания заявки-инициатора периода занятости распределена по экспоненциальному закону. При дисциплине $П_2$ m-заявки, обслуживаясь одновременно на b_m приборах, покидают систему независимо друг от друга, а их длительности обслуживания имеют экспоненциальное распределение со средним значением μ_m^{-1}.

Положим $C = \infty$, в этом случае все поступившие в систему заявки принимаются на обслуживание и потери отсутствуют. Пусть случайный процесс $\{Y_m(t),\ t \ge 0\},\ m \in \mathbf{M}$, находится в состоянии "1", если в момент времени е $t \ge 0$ в системе обслуживается хотя бы одна m-заявка, и в состоянии "0" в противном случае. Можно показать, что процесс $Y_m(t)$ является обратимым марковским процессом и имеет стационарное распределение

$$\pi_m(y_m) = \mathrm{P}\{Y_m(t) = y_m\} = \begin{cases} \dfrac{\rho_m^{\,y_m}}{1+\rho_m} & \text{при дисциплине } П_1, \\ u(y_m) + (-1)^{y_m} e^{-\rho_m} & \text{при дисциплине } П_2, \end{cases} \tag{1}$$

где $y_m \in \{0,1\},\ \rho_m = \dfrac{\lambda_m}{\mu_m}$.

Рассмотрим составной случайный процесс $\left\{\tilde{Y}(t) = \left(Y_m(t)\right)_{m\in\mathbf{M}},\ t \ge 0\right\}$, заданный на множестве $\tilde{\mathbf{Y}} = \{0,1\}^M$. Пусть теперь $C < \infty$ и, следовательно, возможны потери заявок. В этом случае функционирование системы описывает случайный процесс $\{Y(t),\ t \ge 0\}$, являющийся сужением процесса $\{\tilde{Y}(t),\ t \ge 0\}$ на множество $\mathbf{Y} = \{y \in \tilde{\mathbf{Y}} : c(y) \le C\}$, где $c(y) = \sum_{m=1}^{M} c_m(y_m) = \sum_{m=1}^{M} b_m y_m$ — число занятых приборов системы в состоянии $y \in \mathbf{Y}$. Заметим, что $c_m(y_m)$ является дискретной случайной величиной с распределением $\{\pi_m(0),\ \pi_m(1)\}$ и, следовательно, с производящей функцией (ПФ)

$$\psi_m(z) = \pi_m(0) + \pi_m(1) z^{b_m},\ m \in \mathbf{M}. \tag{2}$$

Условием потери m-заявки помимо недостаточного числа свободных приборов является отсутствие в системе заявок данного типа. Следовательно, множество потерь m-заявок имеет вид

$$\mathbf{L}_m = \{y \in \mathbf{Y} : c(y) + b_m > C,\ y_m = 0\}. \tag{3}$$

Пусть τ_n^m — n-й момент поступления m-заявки. Обозначим $Y_n^m = Y(\tau_n^m - 0)$ и будем исследовать вероятность потерь m-заявок по вызовам, т.е.

$$L_m = \lim_{n\to\infty} P\left\{Y_n^m \in \mathbf{L}_m\right\},\ m \in \mathbf{M}. \tag{4}$$

В работах [5,6] представлен метод расчета вероятностей L_m для дисциплины $П_2$. Этот метод, уточненный и модифицированный для дисциплины $П_1$, изложен в следующем разделе.

Метод расчета вероятностей блокировок запроса пользователя

Расчет значения вероятностей L_m предлагается производить по формуле

$$L_m = P_m^0 B_m, \tag{5}$$

где P_m^0 — вероятность того, что *m*-услуга находится в состоянии "0", а B_m— вероятность того, что на звене нет свободных b_m единиц емкости. Для нахождения вероятности P_m^0 строится цепь Маркова (ЦМ), диаграмма состояний которой показана на рис. 1.

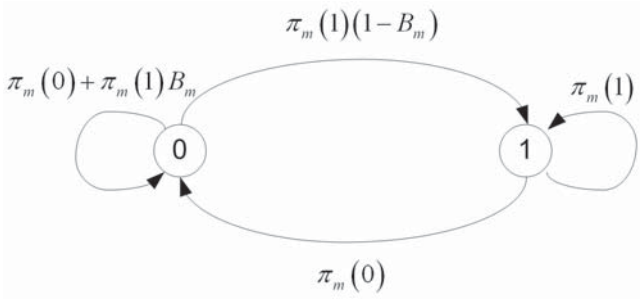


Рис. 1. Диаграмма состояний цепи Маркова

Из диаграммы видно, что для перехода из состояния "0" в состояние "1" необходимо поступление заявки, что происходит с вероятностью $\pi_m(1)$, и с вероятностью $1-B_m$ заявка не должна быть заблокирована. Если с вероятностью $\pi_m(0)$ заявка не поступила, или с вероятностью $\pi_m(1)B_m$ была заблокирована, то система остается в состоянии "0". Для перехода из состояния "1" в состояние "0" заявка должна быть обработана и система должна освободиться с вероятностью $\pi_m(0)$. Для того, чтобы ЦМ оставалась в состоянии "1", система должна быть не пуста с вероятностью $\pi_m(1)$. Очевидно, что вероятность P_m^0 рассчитывается по формуле

$$P_m^0 = \frac{\pi_m(0)}{1-\pi_m(1)B_m}. \tag{6}$$

Для расчета вероятности B_m используется мультисервисная модель Энгсета с множеством M источников и C приборами, где источники соответствуют услугам мультивещания. В этой модели вероятность B_m соответствует вероятности блокировки *m*-источника по вызовам.

Обозначим $P_{-m}(j)$ стационарное распределение числа занятых приборов в модели Энгсета с $M/\{m\}$ источниками и неограниченным числом приборов ($C=\infty$). Введем также производящую функцию $\varphi_{-m}(z)$ для распределения $\{p_{-m}(j), j\geq 0\}$. Так как $P_{-m}(j)$ является распределением числа занятых приборов, то ПФ $\varphi_{-m}(z)$ может быть найдена с помощью ПФ $\{\psi_m(z), m\in \mathbf{M}\}$ по формуле

$$\varphi_{-m}(z) = \sum_{j=0}^{\infty} p_{-m}(j) z^j = \prod_{\substack{i=1,\dots,M \\ i\neq m}} \left(\pi_i(0) + \pi_i(1) z^{b_i}\right) \tag{7}$$

Теперь получаем очевидное соотношение распределений $\{p_{-m}(j), j\geq 0\}$ и $\{\pi_m(y_m), m\in \mathbf{M}\}$:

$$p_{-m}(j) = \frac{1}{j!}\varphi_{-m}^{(j)}(0), \tag{8}$$

где $\varphi_{-m}^{(j)}(0)$ — *j*-я производная ПФ $\varphi_{-m}(z)$ в точке 0.

Известно (см., например, [7]), что в таком случае вероятность блокировки *m*-источника по вызовам в модели Энгсета равна вероятности блокировки того же источника по времени в модели с $M/\{m\}$ источниками и рассчитывается по формуле

$$B_m = \frac{\sum_{j=C-b_m+1}^{C} p_{-m}(j)}{\sum_{j=0}^{C} p_{-m}(j)} = \frac{\sum_{j=C-b_m+1}^{C} \frac{1}{j!}\varphi_{-m}^{(j)}(0)}{\sum_{j=0}^{C} \frac{1}{j!}\varphi_{-m}^{(j)}(0)}. \tag{9}$$

Теперь, для расчета искомых вероятностей L_m достаточно подставить в (5) значения P_m^0 и B_m, вычисляемые по формулам (6) и (9) соответственно. Отметим, что в [8] показана инвариантность модели Энгсета относительно распределения длительности обслуживания. Таким образом, полученные выше результаты могут быть использованы для случая функции распределения длительности обслуживания общего вида, при этом вид распределения (1) не изменится.

Заключение

Рассмотрен приближенный метод расчета вероятностей блокировок запросов пользователей на услуги мультивещания, предложенный в [5,6] для частного случая модели звена сети мультивещания с обслуживанием $П_2$. Метод был уточнен и применен к модели с дисциплиной обслуживания $П_1$. Проведена оценка точности метода с помощью имитационного моделирования.

Литература

1. **Башарин Г.П., Самуйлов К.Е., Яркина Н.В., Гудкова И.А.** Новый этап развития математической теории телетрафика // Автоматика и телемеханика, 2009. — №12. — С. 16-28.

2. **Наумов В.А., Самуйлов К.Е., Яркина Н.В.** Теория телетрафика мультисервисных сетей. Монография. — М.: Изд-во РУДН, 2007. —191 с.

3. **Konstantin E. Samouylov, Olga N. Plaksina.** Approximating blocking probabilities for multiservice network link with unicast and multicast connections // International IEEE Conference EUROCON 2009 Russia, Saint Petersburg. — May 18-23, 2009. — P.1814-1817.

4. **Рыков В.В.** Сети обслуживания прозрачных требований // Автоматика и телемеханика, 2001. — №5. — С.147-158.

5. **Karvo J., Virtamo J., Aalto S. Martikainen O.** Blocking of dynamic multicast connections in a single link, Proceedings of the IFIP TC6/WG6.2 Fourth International Conference on Broadband Communications: The future of telecommunications. — April 01-03, 1998. — P.473-484.

6. **Lu Y., Kuipers F.A., Janic M., and Mieghem P.V.** E2E Blocking Probability of IPTV and P2PTV // in Proc. Networking, 2008. — P. 445-456.

7. **Башарин Г.П.** Лекции по математической теории телетрафика. Учебное пособие. Изд. 3 дораб. и доп. — М.: Изд-во РУДН, 2009. — 342 с.

8. **Cohen J.W.** The generalized Engset formula // Phillips Telecommun. Rev. 18, 1957. — P.158-170.

Приближенные методы расчета характеристик мультисервисной системы с ограниченным числом источников

Рассматриваются два приближенных метода, которые дают возможность вычислить стационарные вероятности для процесса, описывающего числа занятых каналов в системе с ограниченным числом источников, и провести оценку нужных вероятностно-временных характеристик. Проводится сравнительный анализ результатов.

Коннон М.А.,
ассистент кафедры системы телекоммуникаций РУДН,
mkonnon@gmail.com

Введение

В нынешней беспроводной и мобильной индустрии новые тенденции предлагают переход от первично голос-ориентированных услуг к передаче данных и мультимедийных приложений. Конвергенция сетей различных типов породила множество классов сетевого трафика. Эти классы различаются своими характеристиками, объемом необходимых сетевых ресурсов, а также требованиями к качеству обслуживания (QoS). Основным отличительным фактором между этими классами трафика является чувствительность к задержкам. Классы не чувствительные к задержкам, в основном, используются для передачи трафика не в реальном времени (например такие интернет-приложения, как WWW, электронная почта и FTP). Трафик такого рода имеет характеристики, которые сложно описывать традиционным пуассоновским потоком. В [1] и [2] построена модель системы с ограниченным числом источников для описания такого трафика и предложен приближенный алгоритм расчета стационарных вероятностей. В данной работе выводится новая приближенная рекуррентная формула, на основе рассуждений источника [3, гл.3].

Для удобства будем использовать индекс (1) сверху для модели с бесконечным числом источников (мультисервисная модель Эрланга), а (2) — для модели с ограниченным числом источником (мультисервисная модель Энгсета).

Данная статья организована следующим образом. В разделе 1 исследована модель системы с бесконечным числом источников. В раздел 2 описывается первое приближение расчета характеристик системы с ограниченным числом источников. Раздел 3 посвящен выводу второго приближенного метода. В разделе 4 проводится сравнительный анализ на основе численного примера.

1.Модель трафика с бесконечным числом источников

$$\left(\begin{array}{c|c|c|c} \vec{M} & \vec{M} & V<\infty & 0 \\ \vec{\lambda},\vec{b} & \vec{\mu} & & \end{array}\right)$$

Рассмотрим многоскоростную мультисервисную модель с бесконечным количеством источников и пуассоновскими потоками. Данную модель принято называть мультисервисной моделью Эрланга с явными потерями [3, гл. 2]. Пусть система поддерживает K различных типов услуг, т.е. возможно обслуживание заявок на передачу k типов данных (k-заявки), $k=\overline{1,K}$ Вся ширина полосы пропускания системы делится на фиксированное число V единиц канального ресурса (ЕКР, Кбит/с).

Пусть k-заявка при поступлении одновременно занимает b_k ЕКР, $k=\overline{1,K}$. Здесь b_k — целые числа, причем $b_k \in \{1,2,\ldots,V\}$ и $1 \le b_1 \le b_2 \le \ldots \le b_K$. Потоки поступления k-заявок пуассоновские, соответственно имеют постоянные интенсивности λ_k, $k=\overline{1,K}$, и независимы в совокупности. Время занятия k-заявкой b_k ЕКР распределено по экспоненциальному закону с параметром μ_k, $k=\overline{1,K}$.

Определим $\vec{n}^{(1)}=\left(n_1^{(1)},\ldots,n_K^{(1)}\right)^T$, где $n_k^{(1)}$ — число k-заявок в некоторый момент времени $t>0$ и $\vec{b}^T=(b_1,\ldots,b_K)$. Пространство $\Omega^{(1)}$ всех возможных состояний системы имеет вид :

$$\Omega^{(1)}:=\left\{\vec{n}^{(1)}: n_k^{(1)}=0,1,\ldots,\left\lfloor\frac{V}{b_k}\right\rfloor,\ k=\overline{1,K},\ \ 0\le\vec{b}^T\vec{n}^{(1)}\le V\right\}. \quad (1.1)$$

Процесс функционирования данной системы описывается K-мерным СтМП $\vec{X}^{(1)}(t)=\left(X_1^{(1)}(t),\ldots,X_K^{(1)}(t)\right)^T, t\ge 0,$ с пространством состояний $\Omega^{(1)}$, где $X_k^{(1)}(t)$ — число k-заявок в системе в момент времени $t\ge 0$ (рис 1).

Обозначим через:

$p\left(\vec{n}^{(1)}\right):=P\left\{\vec{X}^{(1)}=\vec{n}^{(1)}\right\}=P\left\{\left(X_k^{(1)}\right)=\left(n_k^{(1)}\right)k=\overline{1,K}\right\}$ — равновесную вероятность того, что на передаче находятся $n_k^{(1)}$ k-заявок, $\vec{n}^{(1)}\in\Omega^{(1)}, k=\overline{1,K},$ а $q^{(1)}(\mathrm{v})=\sum_{\vec{n}^{(1)}\in\Omega^{(1)}(\mathrm{v})} p\left(\vec{n}^{(1)}\right)$ — макро вероятность пребывания СтМП $\vec{X}^{(1)}(t)$ во множестве состояний $\Omega^{(1)}(\mathrm{v})$, где $\Omega^{(1)}(\mathrm{v}):=\left\{\vec{n}^{(1)}\in\Omega^{(1)}:\vec{b}^T\vec{n}^{(1)}=\mathrm{v}\right\}$, v — число занятых ЕКР, $\mathrm{v}=\overline{0,V}$.

На рис. 1 видно, что

$$n_k^{(1)}p\left(\vec{n}^{(1)}\right)=\rho_k p\left(\vec{n}^{(1)}-\vec{e}_k\right)I\left(n_k^{(1)}>0\right), \quad (1.2)$$

где $\rho_k=\lambda_k/\mu_k$

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

Суммируя по состояниям $\vec{n}^{(1)}\in\Omega^{(1)}(\mathrm{v})$, левая и правая части уравнения (1.2) согласно [3, гл. 2] будут принимать соответственно следующий вид :

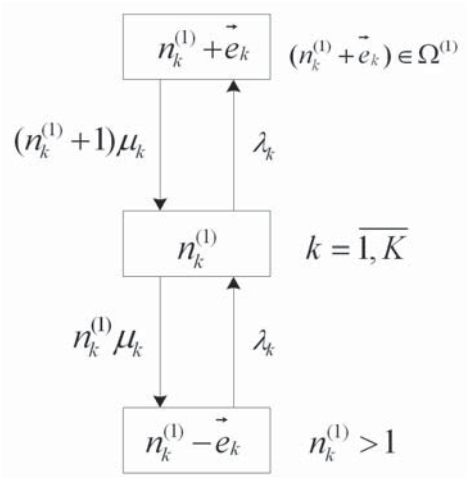


Рис. 1. Диаграмма интенсивностей переходов для процесса $\vec{X}^{(1)}(t)$

$$\sum_{\vec{n}^{(1)}\in\Omega^{(1)}(\mathrm{v})} n_k^{(1)} p\left(\vec{n}^{(1)}\right) = q^{(1)}(v) E(X_k^{(1)} \mid \mathrm{v});$$
$$\sum_{\vec{n}^{(1)}\in\Omega^{(1)}(\mathrm{v})} \rho_k p\left(\vec{n}^{(1)} - \vec{e}_k\right) = \rho_k q^{(1)}(\mathrm{v} - b_k). \quad (1.3)$$

Таким образом:

$$E(X_k^{(1)} \mid \mathrm{v}) = \begin{cases} \rho_k q^{(1)}(\mathrm{v} - b_k)/q(\mathrm{v}), & 1 \le \mathrm{v} \le V \\ 0, \textit{в противном случае}. \end{cases}; \quad (1.4)$$

Здесь $E(X_k^{(1)} \mid \mathrm{v})$ — среднее число k-заявок в системе, когда занято ровно v ЕКР.

Предполагая, что $E(X_k^{(1)} \mid \mathrm{v}) \equiv n_k^{(1)}(\mathrm{v})$ и суммируя по k, получим из (1.3) следующую рекуррентную формулу:

$$\sum_{k=1}^{K} b_k \rho_k q^{(1)}(\mathrm{v} - b_k) = q^{(1)}(\mathrm{v}) \sum_{k=1}^{K} b_k n_k^{(1)}(\mathrm{v}) = \mathrm{v} q^{(1)}(\mathrm{v}). \quad (1.5)$$

2. Первый приближенный алгоритм для модели трафика с ограниченным числом источников $\begin{matrix}\vec{M}\\ N,\vec{\varepsilon},\vec{b}\end{matrix}\left|\begin{matrix}\vec{M}\\ \vec{\mu}\end{matrix}\right| V<\infty \left| 0\right.$

Рассмотрим многоскоростную мультисервисную модель с ограниченным числом источников, в которой имеется V единиц канального ресурса, и N_k абонентов, $N_k < \infty$, каждый из которых может одновременно пользоваться одной и только этой услугой. Каждый из N_k абонентов создает независимый пуассоновский поток сообщений на k-услугу с интенсивностью ε_k, $0 < \varepsilon_k < \infty$ и параметрами b_k, μ_k, $k = \overline{1,K}$. Данная модель является одной из мультисервисных моделей Энгсета с явными потерями [3, гл. 3].

Определим $\vec{n}^{(2)} = \left(n_1^{(2)}, \ldots, n_K^{(2)}\right)^T$, где $n_k^{(2)}$ — число k-заявок в некоторый момент времени $t > 0$ и $\vec{b}^T = (b_1, \ldots, b_K)$. Тогда $n_k^{(2)} \le N_k$, $k = \overline{1,K}$. Пространство всех возможных состояний системы имеет вид:

$$\Omega^{(2)} = \left\{ \vec{n}^{(2)} : n_k^{(2)} \le N_k, k = \overline{1,K}, 0 \le \vec{b}^T \vec{n}^{(2)} \le V \right\} \quad (2.1)$$

Обозначим через $\tilde{\rho}_k$ интенсивность k-нагрузки от одного абонента без учета требуемого числа ЕКР:

$$\tilde{\rho}_k = \frac{\varepsilon_k}{\mu_k}, k = \overline{1,K}.$$

Определим для K-мерного СтМП $\vec{X}^{(2)}(t) = \left(X_1^{(2)}(t), \ldots, X_K^{(2)}(t)\right)^T$:

$$p\left(\vec{n}^{(2)}\right) := P\left\{\vec{X}^{(2)} = \vec{n}^{(2)}\right\} = P\left\{X_k^{(2)} = n_k^{(2)}, k = \overline{1,K}\right\},$$

где $X_k^{(2)}(t)$ — число k-заявок в системе в момент времени $t \ge 0$, $\vec{n}^{(2)} \in \Omega^{(2)}$, $k = 1, K$, и $q^{(2)}(\mathrm{v}) = \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} p\left(\vec{n}^{(2)}\right)$,

где $\Omega^{(2)}(\mathrm{v}) := \left\{\vec{n}^{(2)} \in \Omega^{(2)} : \vec{b}^T \vec{n}^{(2)} = \mathrm{v}\right\}$, v — число занятых ЕКР, $\mathrm{v} = \overline{0,V}$.

Имеет место следующее уравнение для k-услуги:

$$n_k^{(2)} p\left(\vec{n}^{(2)}\right) = \tilde{\rho}_k p\left(\vec{n}^{(2)} - \vec{e}_k\right)(N_k - n_k^{(2)} + 1) I\left(n_k^{(2)} > 0\right). \quad (2.2)$$

Суммируя по состояниям $\vec{n}^{(2)} \in \Omega^{(2)}(\mathrm{v})$, левая и правая части уравнения (2.2) согласно [2, 3, гл. 3] будут принимать соответственно следующий вид:

$$\sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} n_k^{(2)} p\left(\vec{n}^{(2)}\right) = q^{(2)}(\mathrm{v}) E(X_k^{(2)} \mid \mathrm{v});$$
$$\sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} (N_k - n_k^{(2)} + 1)\tilde{\rho}_k p\left(\vec{n}^{(2)} - \vec{e}_k\right) = (N_k - E(X_k^{(2)} \mid \mathrm{v} - b_k))\tilde{\rho}_k q^{(2)}(\mathrm{v} - b_k). \quad (2.3)$$

Здесь $E(X_k^{(2)} \mid \mathrm{v})$ — среднее число k-заявок в системе Энгсета, когда занято ровно v ЕКР. Предполагая, что $E(X_k^{(2)} \mid \mathrm{v}) \equiv n_k^{(2)}(\mathrm{v})$ и суммируя по $\vec{n}^{(2)} \in \Omega^{(2)}(\mathrm{v})$ из (2.3) получим следующую рекуррентную формулу:

$$\sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} (N_k - E(X_k^{(2)} \mid \mathrm{v} - b_k))\tilde{\rho}_k q^{(2)}(\mathrm{v} - b_k) = q^{(2)}(\mathrm{v}) \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} b_k n_k^{(2)}(\mathrm{v}) = \mathrm{v} q^{(2)}(\mathrm{v}). \quad (2.4)$$

Приведем теперь алгоритм расчета $q^{(2)}(\mathrm{v})$, $\mathrm{v} = \overline{0,V}$.

Шаг 1. Вычисляем $q^{(1)}(\mathrm{v})$, $\mathrm{v} = \overline{1,V}$ по формуле (1.5), приняв $\rho_k = N_k \tilde{\rho}_k$, $k = \overline{1,K}$.

Шаг 2. Вычисляем среднее количество k-заявок $E(X_k^{(1)} \mid \mathrm{v})$ по формуле (1.4).

Шаг 3. Полагая $E(X_k^{(2)} \mid \mathrm{v}) \equiv E(X_k^{(1)} \mid \mathrm{v})$, $k = \overline{1,K}$, $\mathrm{v} = \overline{1,V}$ вычисляем макровероятности $q^{(2)}(\mathrm{v})$, $\mathrm{v} = \overline{0,V}$, используя рекуррентную формулу (2.4).

3. Второй приближенный алгоритм для модели трафика с ограниченным числом источников

Рассмотрим многоскоростную мультисервисную модель Энгсета, описанную в разделе 2.

Суммируя левую и правую части уравнения (2.2) по состояниям $\vec{n}^{(2)} \in \Omega^{(2)}(\mathrm{v})$, получим:

$$q^{(2)}(\mathrm{v}) = \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} p(\vec{n}^{(2)}) = \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} (N_k - n_k^{(2)} + 1)\frac{\tilde{\rho}_k}{n_k^{(2)}} p(\vec{n}^{(2)} - \vec{e}_k) =$$
$$= \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} \left((N_k + 1)\frac{\sum_{j=1}^{K} b_j n_j^{(2)}}{\mathrm{v}} - n_k^{(2)} \right) \frac{\tilde{\rho}_k}{n_k^{(2)}} p(\vec{n}^{(2)} - \vec{e}_k) =$$
$$= \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} \left(\frac{\sum_{j=1}^{K} b_j n_j^{(2)}}{\mathrm{v}} (N_k + 1) - n_k^{(2)} \frac{\mathrm{v}}{\mathrm{v}} \sum_{j=1}^{K} \frac{1}{K} \right) \frac{\tilde{\rho}_k}{n_k^{(2)}} p(\vec{n}^{(2)} - \vec{e}_k) =$$
$$= \frac{1}{\mathrm{v}} \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} \sum_{j=1}^{K} \left(b_j n_j^{(2)} (N_k + 1) - n_k^{(2)} \frac{\mathrm{v}}{K} \right) \frac{\tilde{\rho}_k}{n_k^{(2)}} p(\vec{n}^{(2)} - \vec{e}_k).$$

Пусть $n_j^{(2)} \ge 1$ и $k = j$, тогда

$$\mathrm{v} q^{(2)}(\mathrm{v}) = \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} \sum_{j=1}^{K} \left(b_j n_j^{(2)} (N_j + 1) - n_j^{(2)} \frac{\mathrm{v}}{K} \right) \frac{\tilde{\rho}_j}{n_j^{(2)}} p(\vec{n}^{(2)} - \vec{e}_j) =$$
$$= \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} \sum_{j=1}^{K} \left(b_j (N_j + 1) - \frac{\mathrm{v}}{K} \right) \tilde{\rho}_j p(\vec{n}^{(2)} - \vec{e}_j) =$$
$$= \sum_{j=1}^{K} \left(b_j (N_j + 1) - \frac{\mathrm{v}}{K} \right) \tilde{\rho}_j \sum_{\vec{n}^{(2)}\in\Omega^{(2)}(\mathrm{v})} p(\vec{n}^{(2)} - \vec{e}_j).$$

$$\mathrm{v} q^{(2)}(\mathrm{v}) = \sum_{j=1}^{K} \left(b_j (N_j + 1) - \frac{\mathrm{v}}{K} \right) \tilde{\rho}_j q^{(2)}(\mathrm{v} - b_j), k = \overline{1,\mathrm{K}}, \mathrm{v} = \overline{1,\mathrm{V}}. \quad (3.1)$$

4. Оценка приближенных алгоритмов

Одной из важных характеристик рассматриваемой модели Энгсета является вероятность блокировки π_k, $\overline{1,K}$ для k-заявок, которая определяется по следующей формуле:

$$\pi_k^{(2)} = \sum_{\mathrm{v}=V-b_k+1}^{V} q^{(2)}(\mathrm{v}), k = \overline{1,\mathrm{K}}. \quad (4.1)$$

Проведем сравнительный анализ результатов, полученных предложенными алгоритмами и численным методом для системы со структурными параметрами, представленными в таблице.

Исходные данные для сравнительно анализа на численном примере ($K = 3$, $V = 6$)

Параметр \ k	1	2	3
N_k	5	4	3
b_k	1	2	3
$\tilde{\rho}_k$	0,1	0,05	0,025 – 0,009

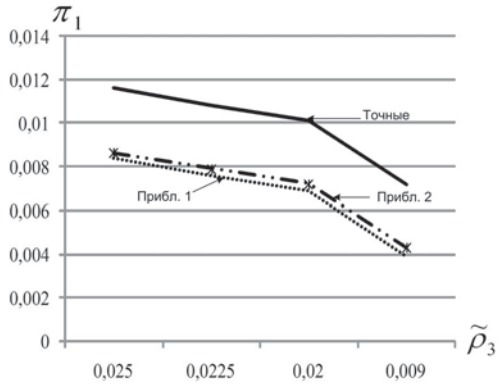


Рис. 2. Зависимость вероятности π_1 от $\tilde{\rho}_3$

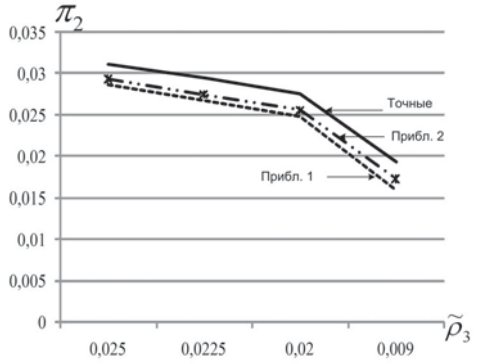


Рис. 3. зависимость вероятности π_2 от $\tilde{\rho}_3$

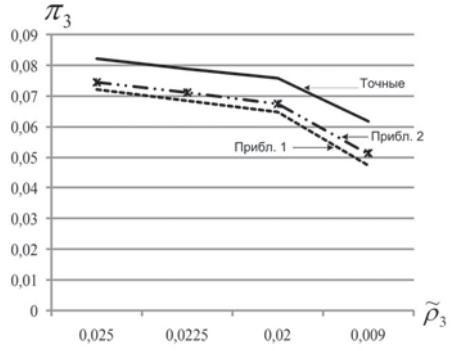


Рис. 4. Зависимость вероятности π_3 от $\tilde{\rho}_3$

В приближениях 1 и 2 были использованы соответственно формулы (2.4) и (3.1).

Результаты численного анализа показывают, что в определенном диапазоне выбранных нагрузочных параметров с минимальной погрешностью $2 \cdot 10^{-3}$ вероятности блокировки, полученные предложенными алгоритмами, совпадают с точными результатами. Заметно, что разница между вероятностями, полученными предложенными алгоритмами, не существенна, при этом приближение 2 дает лучшие результаты.

В целом приближение 1 всегда дает не плохие результаты, когда как точные расчеты и результаты приближения 2 иногда сильно расходятся.

Заключение

В данной статье предложены 2 приближенных алгоритма расчета характеристик систем с ограниченным числом источников. В результате сравнительного анализа установлено, что приближение 1 может быть всегда использовано для определения стационарных вероятностей системы. Но, для оценки вероятностно-временных характеристик в определенном диапазоне нагрузочных параметров, или для решения задач нахождения оптимальных параметров при заданных значениях QoS, лучше использовать приближение 2, так как оно более эффективно, легко реализуемо, и может быть применимо в сетях с большими параметрами.

Литература

1. **Glabowski M., Stasiak M.** An approximate model of the full- availability group with multi-rate traffic and a ?nite source population // Proc. 12th GI-ITG Conference on Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB) / 3rd Polish-German Teletraffic Symposium (PGTS). — Dresden, 2004. — P.195-204.

2. **Stamatelos G. M., Hayes J. F.** Admission-control techniques with application to broadband networks // Computer Communications, 1994. — Vol. 17, №9. — P. 663-673.

3. **Башарин Г. П.** Лекции по математической теории телетрафика. Изд. 3-е, перераб. и доп. — М.: Изд-во РУДН, 2009. — 342 с.

4. **Vassilakis V. G. , Moscholios I. D., Logothetis M. D.** Call-Level Performance Modelling of Elastic and Adaptive Service-Classes with Finite Population // IEICE Transactions on Communications, 2008. — V.E91-B. — №1. — P.151-163.

5. **Stasiak M.** An approximate model of a switching network carrying mixture of different multichannel traffic streams // IEEE Transactions on Communications, 1993. — V.41. — №6. — P. 836-840.

Домовая инфокоммуникационная сеть на базе концепции ITU G.hn

Следуя концепции G.hn, разработанной МСЭ, изложены принципы построения умного дома SmartHouse. Описаны протоколы концепции G.hn. В качестве примера разработки домовых устройств приведены изделия для учета потребления воды по протоколу meter-bus по проводам и по радио.

Шнепс-Шнеппе М.А.,
Генеральный директор ЗАО "ЦКБ-Абаванет",
д.т.н., профессор, академик МА,
sneps@abavanet.ru

1. Телефонная линия как основа домовых сетей

12 декабря 2008 г. в Женеве произошло событие, примечательное для мирового инфокоммуникационного сообщества: Международный Союз Электросвязи одобрил архитектуру нового поколения домовых сетей G.hn (сокращение от слов "next generation" home network), а 9 октября 2009 г. принял базовую Рекомендацию G.9960 [1]. Это итог работы группы ведущих мировых телефонных операторов, производителей средств связи и автоматизации дома. Суть нового стандарта G.9960 состоит в том, что основной физической средой домовых сетей признается телефонная линия (витая пара). Наряду с этим, допускается коаксиальный или оптический кабель, силовая линия, а также радиоканал — как резервный вариант.

Основным принципом модуляции признается частотная модуляция с ортогональными поднесущими OFDM (Orthogonal Frequency-Division Multiplexing). Сам алгоритм OFDM и связанный с ним метод быстрого преобразования Фурье (FFT, Fast Fourier transform) разработаны еще 1960-х гг., но только сейчас стали доступны дешевые сигнальные процессоры в 10 и более миллионов операций в секунду, которые должны быть в каждом домовом устройстве, выходящем на домовую сеть. Развитие микроэлектроники вернуло в жизнь и другое полузабытое достижение — исправление ошибок передачи сигнала при помощи кодов LDPC (Low-Density-Parity-Check).

Надеясь на окончание многолетнего непримиримого соперничества разных подходов к построению домовых сетей, безоговорочную поддержку новому стандарту выразили многие компании, например, телефонная компания AT&T и Best Buy (крупнейший торговец бытовой электроникой в США). Поддержали и организации по автоматизации дома, которые в настоящее время вынуждены пользоваться множеством нестыкующихся технологий: Continental Automated Buildings Association (CABA), CEPCA, HomePNA и Universal Powerline Association.

Привлекательность нового подхода состоит в том, что он позволяет создавать домашнюю сеть или сеть малого офиса с использованием существующей телефонной проводки. История создания стандартов G.hn берет начало с 1996 г., когда ведущими производителями телекоммуникационного оборудования был образован альянс, получивший название Home Phoneline Networking Alliance, и в 1998 г. появился стандарт передачи данных по телефонным линиям, названный HPNA. Самая современная на данный момент версия — HPNA 3.1 была разработана в компании Coppergate Communications (Израиль) и одобрена в ITU как рекомендация G.9954 (в январе 2007 г.). Допускается скорость передачи до 240 Мбит/с. Используемый частотный диапазон 4-10 МГц. Максимальное число пользователей — до 64 (но с ростом их числа скорость обмена падает). Одну из 64 позиций занимает домовой шлюз RG (выполняя роль ведущего устройства, master). Остальные пользователи выполняют роль ведомых устройств (slaves). Это решение оптимально для передачи голоса, данных и видеоизображений на расстоянии до 600 метров.

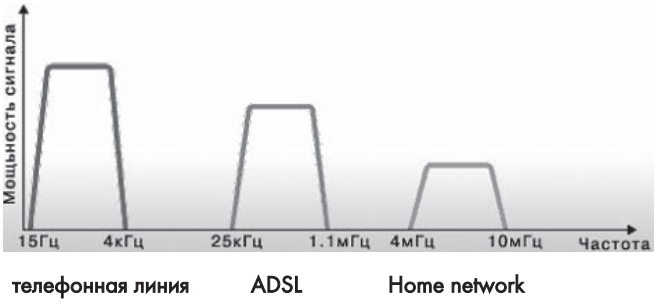


Рис. 1. Разделение полосы частот в домовой сети G.hn

2. Основы технологии G.hn

В настоящее время широко внедрены услуги triple play с использованием ADSL модемов, что сохраняется и в сетях G.hn. Передача данных по технологии ADSL реализуется через обычную аналоговую телефонную линию при помощи модема ADSL и мультиплексора доступа DSLAM, находящегося на той же АТС, что и абонентский комплект. Чтобы не мешать использованию телефонной сети по ее-прямому назначению (для передачи голоса в полосе 0...4 кГц), в ADSL нижняя граница диапазона частот находится на уровне 25 кГц. Верхняя же граница, исходя из требований к скорости передачи данных и возможностей телефонного кабеля, составляет 1,1 МГц (или 2,2 МГц — по технологии ADSL2+). Эта полоса делится на две части: частоты от 26 кГц до 138 кГц отведены восходящему потоку данных (от дома к АТС), а частоты от 138 кГц до 1,1 МГц — нисходящему. Такое частотное разделение позволяет по одной и той же линии одновременно как разговаривать по телефону, так и производить двухсторонний обмен данными.

Существующие домовые устройства, которые имеют стык Ethernet, не прийдется менять, так как в сети G.hn уровень передачи данных (Data Link Layer) работает по технологии Ethernet, т.е. данные от приложений поступают в виде фреймов IEEE 802.3 Ethernet, а на

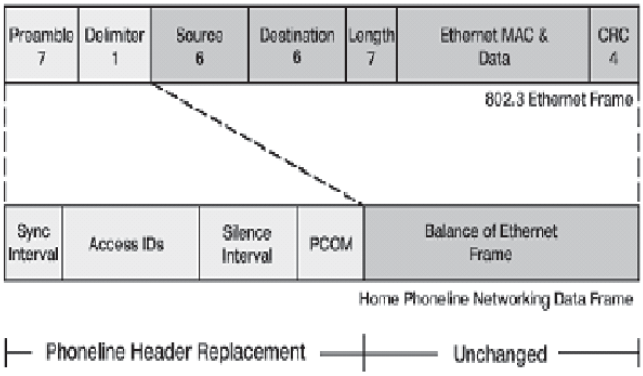


Рис. 2. Работа сети G.hn (нижняя часть рисунка) основана на использовании технологии Ethernet (верхняя часть рисунка)

физическом уровне PHY — восемь первых октетов убираются и заменяются на PHY-заголовок (рис. 2). На приемном конце фреймы меняются в обратном порядке.

3. Стек протоколов G.hn

Стек протоколов G.hn (рис. 3) предполагается реализовать в чипсетах G.hn. Стек содержит два уровня (уровень передачи данных и физический уровень) и шесть подуровней. В системе G.hn управление средой передачи MAC (Media Access Control) не зависит от физической среды передачи (отсюда обозначение MII — Media Independent Interface) и основано на временном делении канала согласно архитектуре TDMA. И что важно, в последней версии G.hn не применяется метод IEEE 802.3 CSMA/CD (т.е. отсутствуют коллизии и установлением соединений между устройствами занимается домовой шлюз).

На физическом уровне G.hn PHY используется модуляция OFDM и метод быстрого преобразования Фурье FFT. На каждой поднесущей частоте сигналы кодируются методом квадратурной амплитудной модуляции QAM (вплоть до 12-битовой сложности). Физический уровень G.hn PHY завершается интерфейсом MDI (Media Dependent Interface), зависящим от среды передачи (силовая линия, телефонная линия или коаксиальный кабель). Это единственное место в стеке протоколов, где учитываются особенности среды передачи. Остальные пять подуровней стека протоколов не зависят от среды передачи, а согласование работы сети в разных средах обеспечивает домовой шлюз (Residential Gateway).

За качество связи QoS отвечает канальный уровень. Введены восемь приоритетов. При установлении очередного соединения домовой шлюз выбирает необходимое качество QoS, и так как в среде передачи отсутствуют конфликты (не применяется метод CSMA/CD), то требуемое качество удается обеспечить.

4. Как строить умный дом

SmartHouse — это термин, которым обозначают европейскую концепцию умного дома (синонимы: интеллектуальный дом, цифровой дом и т.п). Термин введен нормирующим документом по благоустройству жилых домов [2], который в 2005 г. обнародовал Европейский комитет по стандартизации в электротехнике CENELEC.

В архитектуре умного дома (рис. 4) выделяются три подсети: инфокоммуникационная (телефонная связь, интернет, телевидение, т.е. нак называемый набор услуг triple play), развлекательная и автоматизации дома. Они пользуются общей инфраструктурой передачи информации внутри квартиры (дома). Через сеть доступа (Communication Networks) домовой шлюз RG общается с другими домами, сетями операторов и сервис-провайдерами и отвечает за доведение сервисов до пользователя. Слева на рисунке перечислены сервис-провайдеры (идем сверху вниз), которые борются за клиента: коммунальные услуги (utilities), телекоммуникационные услуги, правительственные, общественные, корпоративные (например, работа на дому), развлекательные, вещание.

Разрабатывая требования к умному дому, CENELEC следовала принципу консенсуса и считала допустимыми любые информационные технологии, что порождает чрезвычайное многообразие протоколов работы домовых устройств и увеличивает стоимость домовых сетей. Этот недостаток устраняет концепция ITU G.hn (рис. 5). Домовой шлюз RG берет на себя управление доступом к G.hn-шине, а домовые устройства выходят на эту шину через устройства G.hn node, в которых реализуется стек протоколов, что на рис. 3.

5. Интерфейс Meter-Bus в домовой сети. Пример реализации

Ранее нами [3] опубликованы вопросы разработки программного обеспечения домовых шлюзов, а в настоящей статье расматриваются особенности разработки оборудования домовых шлюзов и систем автоматизации дома в целом. И наша цель — показать, что производство домовых устройств вполне по силам отечественным производителям.

Рисунок 6 демонстрирует текущее представление об умном доме. Ключевым элементом является домовой (квартирный) контроллер. К нему поступает информация от множества контролируемых домовых устройств. А далее по телефону (фиксированному или мобильному), через интернет или через диспетчера она передается клиенту. Для сбора информации испольуется множество домовых устройств иностранного производства, которые работают по разным интерфейсам (BACnet, LonWorks, KNX, CAN и др.).

Для подключения домовых устройств нами предлагается использовать интерфейс Meter-Bus (M-Bus), что, на наш взгляд, позволит производить дешевые домовые устройства. M-bus разработан для удаленного считывания показателей счетчиков бытовых энергоресурсов, например потребления газа или воды в доме, применим так-

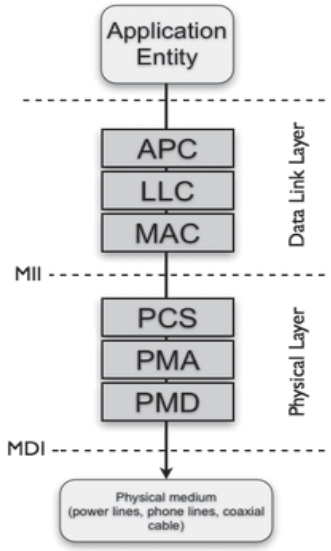


Рис. 3. Стек протоколов G.hn

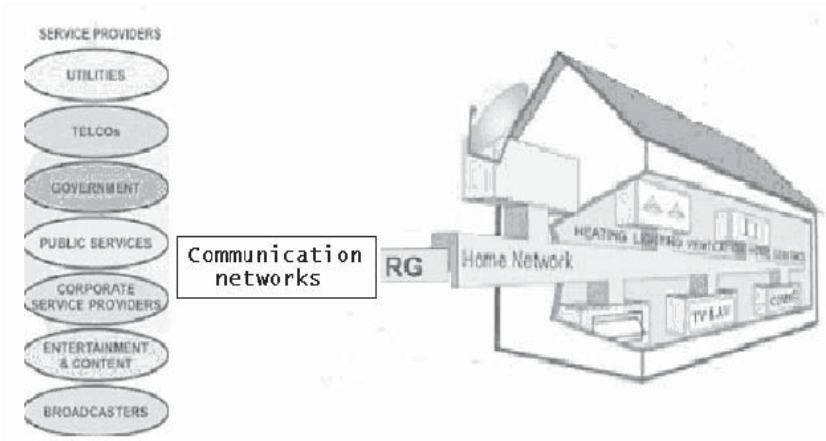


Рис. 4. Общая архитектура SmartHouse и ключевая роль домового шлюза RG

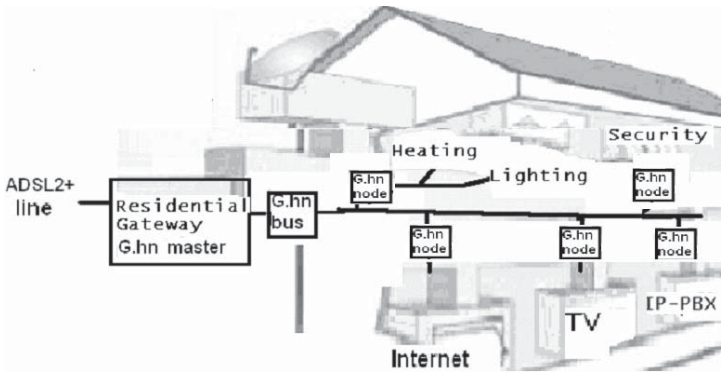


Рис. 5. Домовая сеть по концепции ITU G.hn

Рис. 6. Существующее представление об умном доме

Рис. 7. Две схемы учета потребления воды

же в системах безопасности, управления освещением или теплом. Применение M-Bus упрощается наличием интегральной микросхемы транссивер TSS721 (Texas Instruments), а внедрение в России облегчает наличие ГОСТа РЕН 1434-3-2006 "Теплосчетчики" (на базе M-Bus).

Разработан прототип домового шлюза (Web-сервер) для считывания показаний счетчиков воды (доступен по адресу: http://mbus.linkstore.ru) и реализованы две схемы учета потребления воды (рис. 7):

1) считывание показаний по проводам (левая часть на рисунке). Обратим внимание на роль домового счетчика с клапаном. В случае аварии клапан перекрывает утечку воды. Этот же клапан может помочь в борьбе с хищением воды: надо сравнивать суммарные показания квартирных счетчиков и домового счетчика.

2) считывание показаний и передача по радио на частоте 868 МГц и по m-bus. У счетчика в квартире размещен радиопередатчик (с автономным питанием от батарейки), а в коридоре размещен приемопередатчик, который далее общается с конвертером по M-bus.

6. Заключение. Что производить

Домовая сеть по стандартам G.hn и система автоматизации дома в целом состоит из множества устройств, но самой важной и дорогой частью является программное обеспечение (ПО), что включает в себе ПО общедомового компьютера, ПО квартирного контроллера и ПО домовых устройств. Сложность задачи неизбежно требует разработки многих версий ПО, поэтому, планируя разработку и производство домовых устройств, важно иметь собственное ПО и службу сопровождения ПО для эксплуатации систем автоматизации дома.

В качестве подразделов промышленной продукции в области автоматизации дома можно указать:

- Учет энергоресурсов (вода холодная и горячая), отопление, газ, освещение,
- Управление климатом (термоизоляция, вентиляция и кондиционирование),
- Охранно пожарная сигнализация,
- Система контроля доступа в помещения, контроль протечек воды, утечек газа. Система видео наблюдения,
- Медицинские приборы для слежения за больными,
- Механизация здания (открытие/закрытие ворот, шлагбаумов, электроподогрев ступеней),
- Сети связи по стандартам G.hn (в том числе телефон и локальная сеть здания, квартирная сеть, домовая сеть),
- Программное обеспечение системы автоматизации дома.

Как показывает опыт, разработку и производство домовых устройств можно планировать немедленно, так как уже в 2010 году на мировом рынке появятся чипсеты и первые изделия по стандартам ITU G.hn.

Литература

1. ITU-T Recommendation G.9960. Next generation wireline based home networking transceivers — Foundation, 2009.

2. SmartHouse. Code of Practice. CWA 50487, CENELEC, 2005.

3. Schneps-Schneppe M. ITU G.hn concept and home automation//International Conference NEW2AN/ruSMART Proc, LCNS 5764 (Springer), 2009.— P.1-7.

Метод повышения эффективности функционирования терминального оборудования сетей с интеграцией служб

Приведены результаты исследования принципов построения и алгоритмов функционирования многофункциональных абонентских и сетевых терминалов, работающих в сетях общего пользования, на основе которых предложены модели обслуживания неоднородного трафика. Получены аналитические выражения, позволяющие оценить вероятностно-времнные характеристики систем передачи речевого и неречевого трафика.

Мамедов Г.А.,
д.т.н., профессор, зав. кафедрой "Автоматика и управления" Азербайджанского технического университета

Ибрагимов Б.Г.,
д.т.н., профессор кафедры "Многоканальные телекоммуникационные системы" Азербайджанского технического университета, i.bayram@mail.ru

Введение

Развитие цифровых интегральных узкополосных и широкополосных мультисервисных сетей связи на базе современных технологий требует создания и интеграции сетей общего пользования. Решение этой сложной задачи неразрывно связано с повышением уровня технического обеспечения телекоммуникационных систем (ТС) многофункциональными терминальными комплексами (ТК), обеспечивающими интеграцию процессов совместного обслуживания неоднородного трафика.

В настоящее время одним из основных стратегических направлений развития и технической политики МСЭ-Т является интеграция ТС и их технических средств для передачи неоднородного трафика по каналам единой сети связи, где главным звеном являются многофункциональные ТК и терминальные элементы управления, представляющие собой распределенную систему управления на базе TMN (Telecommunication Management Network). Кроме того, внедрение современных информационных, компьютерных и сетевых технологий в системах телекоммуникации с использованием служб DSP (Digital Signal Processing), IP (Internet Protocol), ATM (Asynchronous Transfer Mode) открывают новые возможности для интеграции сетей телекоммуникаций общего пользования, объединения и обслуживания с помощью терминальных комплексов нескольких видов информации[1,2] . Поэтому решение проблемы совместного обслуживания неоднородного трафика с помощью многофункционального терминального оборудования приобретает большую актуальность в телекоммуникационных системах.

Анализ показывает, что во многих работах подробно исследованы модели информационных потоков при совместном обслуживании речевых и неречевых сообщений [2,3,4]. Однако, в полном объеме не рассмотрены проблемы создания технических средств, обеспечивающих передачу и обслуживание неоднородного трафика в зависимости от распределения пропускной способности терминального оборудования и каналов связи с учетом коэффициента сжатия трафика.

Постановка задачи. Для решения данной задачи рассматриваются вопросы создания многофункциональных терминальных комплексов на основе блочно-модульных систем, базирующихся на современных технологиях, и исследуются методы оценки интеграции неоднородного трафика при совместном обслуживании в системах, где имеется большой объем информации.

Установлено, что основные характеристики эффективного функционирования многофункциональных терминальных комплексов на сетевом уровне определяются вероятностно-временными характеристиками (ВВХ) системы передачи пакетов $Q_{ввх}$ и производительностью системы $W_{пс}$, в частности, суммарной пропускной способностью терминального оборудования при совместном обслуживании неоднородного трафика [2].

Формализация задачи интеграции процессов совместного обслуживания информационных потоков, где возможное их количество-n, реализуемых с помощью многофункционального ТК, могут быть представлены следующими целевыми функциями:

$$W_{пс} = Arg \max_i \left[\sup_i \left(C_{i\,.\max} \right) \right], \qquad Q_{ввх} = Arg \min_i \left[\inf_i \left(T_{i\,сз} \right) \right], \; i = \overline{1,n} \tag{1}$$

при следующих ограничениях

$$N_{б.i} \le N_{б.i.доп.}\,;\; C_{i.a} \le C_{i.a..доп.};\quad K_{i.сж} \ge K_{i.сж.доп.}\,,\; i = \overline{1,n}\,,$$

где $C_{i.max}$ — пропускная способность многофункциональных терминальных комплексов при передаче *i*-го трафика; $T_{i.сз}$ — среднее время задержки при передаче *i*-го трафика; $N_{б.доп.}$, $C_{а.доп}$, $K_{i.сж.доп.}$ — векторы требуемых показателей, соответственно, емкости буферных накопителей (БН) терминала при передаче *i*-го трафика, стоимости аппаратных и программных средств системы при передаче *i*-го трафика, коэффициенты сжатия данных *i*-го трафика.

В работах [5,6] рассмотрены проблемы создания терминальных комплексов, реализуемых цифровой принцип передачи информации с использованием телекоммуникационных технологий, объединенных в единую систему для приема и передачи неоднородного трафика. В перспективе современные телекоммуникационные сети общего пользования будут строиться на базе многофункциональных терминальных комплексов, которые состоят из терминального оборудования различных технологий: PDH&SDH, IP, ATM, NGN (Next Generation Network), служащих для передачи неоднородного трафика в соответствии с проводимой МСЭ-Т технической политикой глобализации и персонализации связи. Под неоднородным трафиком подразумеваются речевые сигналы, видео информация и неречевые сообщения документального обмена.

Создание алгоритмов функционирования многофункциональных абонентских и сетевых терминалов. Терминалы для обмена неречевого трафика с документированием текста, данных, факсимиле, являются абонентскими терминалами телеграфного типа, имеющими программно-управляемые анализирующие и синтезирующие устройства, модемы, интерфейсы ввода-вывода и терминальные

адаптеры, которые позволяют предоставлять любым абонентам возможность передачи информации со скоростью от 0,05 Кбит/с до 2Мбит/с на базе типовых каналов цифровой иерархии с использованием современных технологий [5]. Абонентские терминалы документального типа, работающие в режиме "текстофакс", при обработке документальной информации смешанного типа, определяются временными характеристиками документальной электросвязи, которые можно описать в следующей форме:

$$T_{общ} = \left[T_{пер}, K_{сж}, N_б, T_{i\,сл}, T_{пр} \right], \qquad (2)$$

где $T_{общ}$ — общее время обработки информации смешанного типа; $T_{пер}$, $T_{пр}$ — время передачи и приема бланка; $T_{i.сл}$ — время выполнения служебных операций; $N_б$ — емкость буферных накопителей (БН).

Для устранения избыточности и сокращения объема передаваемой информации смешанного типа, в частности, факсимильной и графической, использовано кодирование длины серии с предварительным описанием на базе модифицированного кода Хаффмэна и алгоритма кодирования Шеннона-Фано [4].

Терминалы для реализации телефонных служб и передачи речевых сообщений в сетях с коммутацией пакетов на основе протоколов X.25 и TCP/IP, гарантирующих приоритезацию голосовых сообщений, получили название абонентских терминалов телефонного типа, имеющих связные процессоры, порты ввода-вывода, шлюзы, аналого-цифровые преобразователи (АЦП и ЦАП), модули микрофона и динамика с аудио платой, работающие в реальном масштабе времени. Для расширения функциональных возможностей и услуг абонентские терминалы речевого трафика снабжены сетевым стыком, терминальным элементом управления, модулем сборки-разборки пакетов. Терминалы речевого трафика определяются множеством показателей каждой из подсистем, т.е.

$$W_р = W[F_b, T_{сз}(\lambda_р); N_{бр}; Q_{MOS}; N_п],$$

где $T_{сз}(\lambda_р)$ — средняя взвешенная задержка передачи речевых сообщений с интенсивностью $\lambda_р$; $N_{бр}$ — емкость БН; $N_п$ — количество потерь пакетов; Q_{MOS} — коэффициент, учитывающий качество передачи речевых сообщений $Q_{MOS} \leq (5,\ldots,2,5)$.

Алгоритмы передачи и приема речевых сообщений рассмотрены в [5,6] и отмечены, что основными блоками абонентского терминала являются модули АЦП и ЦАП, состоящие из Ku временных отсчетов, каждый из которых кодируется кодовыми комбинациями. При этом время вывода цифровой информации $T_{вв}(L_п, V_{AT})$ из АЦП терминала определяется следующим образом:

$$T_{вв}(L_п, V_{AT}) = (L_п - r_k)\cdot(K_и \cdot V_{AT})^{-1}, \quad V_{AT} = 2F_b \cdot K_И, \qquad (3)$$

где K_u — длина информационного кода, отображающего отсчеты; F_b — верхняя граничная частота спектра речевого сигнала, которая для речи и факсимильного трафика составляет 4 кГц; r_k — проверочные кодовые комбинации.

На основании временных отсчетов и алгоритма передачи речевого трафика скорость абонентского терминала определяется следующим выражением:

$$V_{AT} = K_и \cdot (\Delta t)^{-1} \qquad (4)$$

Для передачи речевых сообщений на базе технологии пакетного преобразования, необходимо использовать равномерное квантование с шагом Δt, который определяется как:

$$\Delta t_k = L_n \cdot (V_{AT})^{-1}, \quad L_п = L_u + L_c, \qquad (5)$$

где L_u, L_c — соответственно длина информационного и служебного пакета.

Анализ показывает, что речевые сообщения описываются двумя состояниями: "0" -пауза речи и "1" активное состояние. Длительности активного диалога речи t_a и паузы $t_п$ выражаются следующим образом:

$$B(t) = \sum_{r=1}^{n} t_{ar} + \sum_{k=1}^{e} t_{пк}, \; r = \overline{1,h}, \quad k = \overline{1,e}. \qquad (6)$$

Из последнего следует, что обработка речевых сообщений сопровождается большой избыточностью из-за передачи паузы речи, которая остро нуждается в сокращении избыточности при передачи трафика. Сжатия речевых сообщений осуществляются с помощью кодеров-декодеров с применением АДИКМ по рекомендации МСЭ-Т G.729 на основе заполнения пауз посредством шлюза абонентского терминала. Последние в значительной степени определяют пропускную способность абонентских терминалов и качество передачи речи.

Терминалы видео трафика характеризуются следующими параметрами: средней задержкой $T_{сзв}$ передачи; средним количеством потерь информации $N_п$; скоростью обработки видео потоков — V_b. Видео терминалы описываются множеством показателей каждой из подсистем, т.е.

$$W_b = W[T_{сзв}(\lambda_b), Q_{MOS}, N_п, V_b, N_{бb}].$$

В связи с увеличением объемов обрабатываемого трафика и интеграцией методов передачи (синхронной и асинхронной) и коммутации (каналов, пакетов и ячеек), а также внедрением гибридных технологий VoFR, VoIP и VoATM открываются новые возможности создания многофункциональных ТК для передачи неоднородного трафика большого объема.

На рис.1 представлена архитектура ИТМ для передачи речи (РС), данных (ПД) и видео (ВИ) трафика.

Реализация данной задачи на базе абонентских терминалов рассмотрена в [5,6]. В них подробно изложены принципы построения и алгоритмы функционирования многофункционального терминального комплекса. Для передачи неоднородного трафика информационный поток после АЦП и кодера поступает в модем, при помощи которого неоднородный трафик представляется в одинаковой цифровой форме. После сжатия цифровых сигналов информационный поток поступает на модуль сборки пакетов с использованием протокола RTP (Real Time Protocol). Последний за счет модемного сигнала различает сигналы речи, факсов, данных и передает управляющий сигнал на БН мультиплексора. Мультиплексор на основе приоритетности поступающих потоков трафика с помощью виртуальных каналов связи передает сообщение необходимому адресату.

Приоритет имеют пакеты речевых сообщений, сигналы управления и взаимодействия, а также видео и пакеты данных, если не определены жесткие приоритетности трафика. Аналогично работает приемная часть терминальных комплексов, которая выполняет обратное преобразование.

Для оценки эффективности многофункциональных терминальных комплексов при совместном обслуживании неоднородного трафика, поступающего в поле ИТМ, необходимо создание модели интеграции обработки информационных потоков.

Интеграция и модели обслуживания неоднородного трафика. Важным блоком управления в многофункциональных терминальных комплексах является мультиплексор интегрального типа — интегральный терминальный мультиплексор (ИТМ) для объединения неоднородного трафика при совместном обслуживании с применени-

Рис. 1. Структурная схема и архитектура ИТМ

ем эффективных алгоритмов кодирования сообщений. На рис. 2 изображена модель интеграции неоднородного трафика при совместном обслуживании многофункциональным терминальным комплексов.

Из описания схемных решений видно, что ИТМ представляет собой устройство управления RTP, БН терминал и канал связи. В БН поступают потоки трафика от различных источников сообщений, работающих по "FIFO" и "LIFO", и на приоритетной основе обслуживаются. При этом емкости БН $N_{i.б}$, обеспечивающие отправление всех пакетов, находящихся в очереди, в рамках допустимой средней задержки $T_{i.сз.доп.}$, определяются как:

$$N_{i.б} = (C_{i.k} - V_{i.т})T_{i.сз.доп.} + \Delta N_{i.б}, \quad C_{i.k} \geq V_{i.т}, \; i = \overline{1,n}, \qquad (7)$$

где $C_{i.k}$ — скорость передачи *i*-го трафика по каналам связи; $V_{i.т}$ — скорость работы абонентского терминала при обработке *i*-го трафика; $\Delta N_{i.б}$ — максимальный всплеск при передаче *i*-го потока пакетов; $T_{i.сз.доп.}$ — допустимое среднее время задержки при передаче *i*-го трафика, $T_{i.сз.доп.} \leq (150,....,300)мс$.

Предположим, необходим переход с режима обработки неречевых сообщений на режим обработки речевых сообщений, который осуществляется по протоколу RTP на базе терминального элемента управления, распознающего, какой вид сообщения — текст, речь, видео, факс содержит данный трафик для сжатия и передачи информации в реальном масштабе времени.

Из алгоритма передачи и приема неоднородного трафика следует, что среднее время задержки складывается: из задержки на модуле кодера, АЦП и ЦАП, модеме; сжатия; сборки-разборки пакета; передачи по каналам и приеме информации. В общем виде эта величина может быть определена следующим образом:

$$T_{i.сз} = \frac{1}{K_{i.сж} \cdot N_{i.m}} \sum_{j=1}^{N_{i.m}} T_{i,j.сз} \quad , \quad T_{j.сж} = t_{j.ож} + t_{j\,пер.},$$

$$j = 1,2,...N, \qquad (8)$$

где $N_{i.m}$ — общее количество блочно-модульных систем абонентского терминала при передаче *i*-го трафика, $i = \overline{1,n}$; $T_{i.сз}$ — средневзвешенная временная задержка, так как суммирование ведется по всем модулям функциональных схем, по которым проходит трафик; $t_{i.ож}$, $t_{i.пер}$ — среднее время ожидания и передачи неоднородного трафика, который проходит *j*-ю блочно-модульную систему абонентского терминала. В частности, для передачи речевого трафика необходимо соблюдать, чтобы $T_{i.сз} \leq (150 \div 300)$ мс [2,5].

Заключение

Результаты исследования показали, что имеется принципиальная возможность для построения многофункциональных ТК неоднородного трафика на основе блочно-модульных систем с использованием DSP-технологий, которые являются интегральным элементом создания широкополосных мультисервисных сетей, способствующих возможности гибкого предоставления ресурсов как терминалов, так и каналов.

Литература

1. **Деарт В.Ю.** Мультисервисные сети связи. Транспортные сети и сети доступа. — М.: Инсвязьиздат, 2008. — 168 с.

2. **Шварц М.** Сети связи: протоколы, моделирование и анализ. В 2-х т.ч. II/Пер. с англ. — М.: Наука, 1992. — 272 с.

3. **Вишневский В.М.** Теоретические основы проектирования компьютерных сетей. М.: Техносфера, 2003. — 512 с.

4. **Лагутин В.С.** Анализ эффективности совместного обслуживания новых информационныхпотоков на ГТС большой емкости // Электросвязь, 1999. —№3. — С.28-30.

5. **Суздалев А.В., Чутреев О.С.** Передача данных в локальных сетях связи, М.: Радио и связь, 1987. —168 с.

6. **Ибрагимов Б.Г.** Терминальные комплексы на базе современных технологий // Вестник связи, 1999. — №10. — С.44.

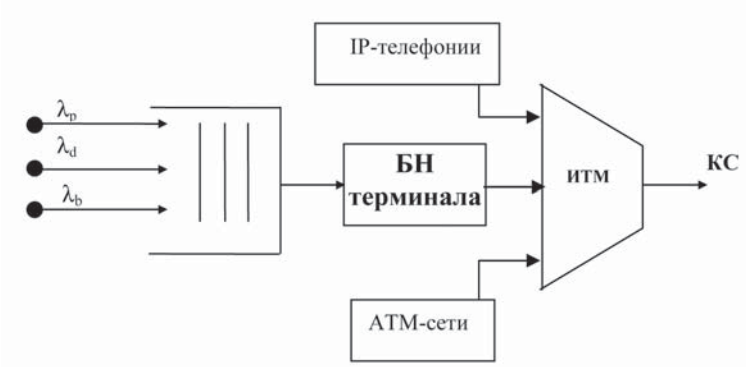


Рис. 2. Модель обслуживания и интеграции неоднородного трафика

Метод повышения пропускной способности систем телеметрии и мониторинга на базе беспроводных сетей

Рассматривается метод повышения пропускной способности систем сбора и обработки телеметрической информации, для которых характерен скачкообразный характер нагрузки и ее долговременная зависимость. Это является причиной появления в них самоподобного трафика, требующего дополнительной пропускной способности на случай критических ситуаций. В обычном состоянии этот ресурс можно использовать для других пользователей с более низким приоритетом.

Маликова Е.Е.,
старший преподаватель каф. АЭС МТУСИ

Введение

Повсеместное проникновение общедоступных мультисервисных беспроводных сетей, а также создание сенсорных сетей (СС) [1-3] позволило создавать на их базе различные общедоступные и ведомственные системы мониторинга и телеметрии. Такие системы мониторинга могут применяться для наблюдения за транспортными средствами, для охраны дач, квартир, контроля отопительного и технологического оборудования, пропадания напряжения 220 В, передачи данных о работе теплосетей, бойлерных и др.

В настоящее время быстрыми темпами развивается мониторинг экологической ситуации. В городах или на территории вокруг аварийно опасных промышленных объектов (атомных станций, химических предприятий и т.д.) устанавливаются станции экологического мониторинга, в которых смонтированы радиационные датчики и датчики химически опасных веществ. Эти датчики для передачи информации могут использовать каналы передачи данных сети подвижной сотовой связи (СПСС) стандарта GSM и CDMA , широкополосные беспроводные сети типа Wi-Fi и WiMAX, а в труднодоступных местах — спутниковые каналы.

Стремительно развивается служба экстренной помощи при автомобильных авариях. Известно, что значительное количество смертельных исходов связано с неоказанием своевременной медицинской помощи в течение "золотого часа" после аварии. Изменить эту ситуацию можно, установив на автомобилях датчики, которые будут посылать SMS с определением местоположения в службу скорой помощи, что в разы может увеличить нагрузку служб передачи данных мобильных сетей.

Для предотвращения утечки газа в домах устанавливаются датчики обнаружения газа, которые также могут осуществлять передачу данных через беспроводные сети. При организации контроля зданий, находящихся в аварийном состоянии, они оборудуются видеокамерами и тензометрическими датчиками, прикрепленными в наиболее опасных местах: при критическом нарастании скорости деформации стен и конструкций через службы передачи данных мобильных сетей передается сигнал тревоги. Кроме того, с помощью службы SMS мобильных сетей возможно дистанционное управление различным оборудованием и осуществление видеонаблюдения за результатами управления.

Приоритетные направления современных разработок предполагают создание разнообразных сенсорных биочипов. Сенсорные сети находят применение в военной области, управлении кризисными и чрезвычайными ситуациями и в борьбе с терроризмом. В зависимости от типа сенсоров СС могут быть развернуты на земле, в воздухе, над и под водой, в зданиях, непосредственно на теле человека. Подобные применения СС в медицине открывают колоссальные перспективы для фармацевтических разработок и синтеза разнообразных биологических препаратов. При этом перспективные сенсорные сети могут заменить существенную часть существующих технологических сетей, они обеспечат новое функциональное наполнение сетей связи и преобразуют их во всепроникающие сети, на которых могут быть развернуты новые системы мониторинга и телеметрии.

Особенностью систем мониторинга и телеметрии является наличие разнородного по своей природе трафика. В обычном состоянии они периодически передают данные с малой интенсивностью и небольшими скоростями, а в случае возникновения критических ситуаций включаются широкополосные системы передачи большого числа параметров телеметрии, или видеоинформации, использующие максимальные скорости. Для систем телеметрии характерен как скачкообразный характер нагрузки, так и ее долговременная зависимость, возникающая из-за передачи информации по заданному расписанию. Интеграция в сетях коммутации пакетов различного по своей природе трафика, а также особенности используемых в них технологий передачи данных, повышенная в случае критических ситуаций активность абонентов, являются некоторыми из причин проявления самоподобного характера сетевого трафика, или иначе, его фрактальных свойств [4,5].

Из-за свойств самоподобия трафика сети, на которых развернуты системы мониторинга, вынуждены резервировать большую пропускную способность на случай возникновения критических ситуаций, поэтому в нормальной обстановке целесообразно использовать этот зарезервированный ресурс для передачи информации других пользователей. При этом дополнительный трафик должен иметь более низкий приоритет по сравнению с основным. Пакетные сети имеют развитые механизмы обеспечения качества обслуживания, однако появление самоподобного трафика требует специального резервирования ресурсов.

Особенности моделей трафика для мультисервисных сетей

Последнее десятилетие ознаменовалось существенными результатами исследований случайных процессов, протекающих в современных мультисервисных сетях. Выявлено широкое распространение [5,6] явления самоподобия при описании статистических характеристик трафика этих систем, в частности, для описания свойств трафика реального времени (видео сервисов и пакетной телефонии). Также проведено множество как экспериментальных, так и теоретических исследований в области самоподобных свойств мультисервисного трафика [4].

В [6] проанализированы методы классической теории телетрафика, разработанной для задач проектирования телефонных сетей [7], основанной на рассмотрении во многих случаях стационарных процессов пуассоновского типа в ЧНН. Показано, что при исследовании абонентской нагрузки уже в ISDN потоки требований на предоставление сеанса связи не являются пуассоновскими, причем ЧНН для разных услуг чаще всего не совпадают. Фундаментальный недостаток пуассоновских моделей состоит в том, что они сильно недооценивают образование пачечной нагрузки, при которой как по-

токи пакетов, так и потоки требований не являются пуассоновскими, длительности сеансов описываются распределениями логонормального, а не показательного типа, т.е. имеют "тяжелые хвосты".

Одним из важных статистических свойств нагрузки, создаваемой телеметрическими приложениями, является наличие зависимости между отдельными ее частями, которая не исчезает при изменении шкалы времени в сторону укрупнения, как это происходит с традиционными случайными процессами, используемыми для описания нагрузки в теории телетрафика [8]. Так, например, для пуассоновского процесса соответствующая корреляция всегда равна нулю, а для марковского процесса с конечным числом состояний — экспоненциально убывает к нулю. При этом поступления пакетов, а также требования на установление сеансов связи коррелированны.

Расчеты, основанные на традиционных представлениях о том, что мультиплексирование большого числа независимых случайных потоков приводит к пуассоновскому процессу, явились причиной грубых ошибок при проектировании коммутаторов ATM первого поколения. Когда такие коммутаторы с небольшими накопителями были пущены в эксплуатацию, потери ячеек оказались недопустимо большими, что заставило конструкторов внести необходимые исправления [6].

Основной характеристикой самоподобного трафика является параметр Херста — H [5], являющийся мерой самоподобия, или статистической инерцией процесса. Его значения лежат в пределах 0,5< H < 1. Значение 0,5 соответствует Марковскому процессу, а 1 — случаю жесткой долгосрочной зависимости. Свойства самоподобия могут возникать в различных системах связи. Например, в каналах ОКС N7 СПСС, передающих наряду с сообщениями об установлении соединения по протоколу ISUP, сообщения о роуминге абонентов протоколов MAP и TCAP и тексты SMS, в мультисервисных сетях при передаче трафика реального времени. Наиболее наглядно самоподобный характер трафика проявляется в широко развивающихся в последнее время беспроводных системах передачи данных, основанных на различных технологиях радиодоступа. Так, в [1] приведены результаты экспериментального исследования самоподобия GPRS — трафика. Оценка производилась на основе вейвлет — анализа, показано, что трафик имеет сложную мультифрактальную структуру. Показатели Херста лежат в пределах 0,789 — 0,988. Наибольшее значение показателя Херста наблюдается во входящем трафике IP, TCP, UDP.

Математическая модель

В системах мониторинга и телеметрии, развернутых на беспроводных сетях передачи данных, возникает задача обслуживания мультисервисного трафика с приоритетами, причем высокоприоритетный трафик должен иметь гарантированное качество обслуживания, а низкоприоритетный — оптимальное.

Рассмотрим математическую модель обслуживания участком беспроводной сети требований на передачу данных абонентами системам телеметрии и мониторинга. В обычном состоянии это регулярно поступающие сообщения одинаковой длины, однако в случае возникновения критических ситуаций нагрузка резко возрастает, ее объемы в сотни раз превышают среднюю нагрузку в системе. Предполагается, что для передачи данных телеметрии всегда зарезервирована часть ресурсов, которая в обычное время может обслуживать требования других абонентов, не связанных с этими системами, например, для выхода в Интернет, ICQ (популярная программа для общения с помощью мгновенных текстовых сообщений в сети Интернет) и т.д. При этом потоки данных от этих пользователей имеют более низкий приоритет в обслуживании.

Анализ современных технологий в сетях передачи информации показывает, что необходимо рассматривать вектор нагрузки, состоящий из нескольких компонент, каждая из которых должна отражать свойства трафика некоторых групп абонентов с одинаковыми интегральными характеристиками [8]. Поскольку предполагается, что деление на компоненты осуществлено таким образом, что каждая из них описывает трафик принципиально различного происхождения: например одна — голосовой трафик, вторая — трафик данных, третья — трафик видео в режиме реального времени и т.п., то можно считать эти компоненты статистически независимыми. В связи с этим становится понятным, что обслуживаемый трафик должен рассматриваться как сумма нескольких компонент [9]:

$$X_t^i = \lambda_i t + \sqrt{a_i \lambda_i} Z_t^{(i)}, \qquad (1)$$

где X_t^i — i-ая компонента процесса X в момент времени t; λ_i — интенсивность i-го потока сообщений; a_i характеризует дисперсию этого потока сообщений в единицу времени; $Z_t^{(i)}$ — фрактальное броуновское движение со значением параметра Херста Hi (Для описания компонент целесообразно использовать диффузионное приближение).

Таким образом, предметом рассмотрения становится смесь нескольких процессов с различными значениями параметра Херста. Рассматривалась двухкомпонентная смесь, что соответствует ситуации, когда присутствуют компоненты с наибольшим и наименьшим из возможных значений параметра Херста. Данное упрощение дает процесс, который по своим статистическим свойствам хуже или лучше исходного.

Таким образом, трафик исследуемой сети представляется как смесь двух самоподобных процессов:

$$Y_t = (1-c)Z_t^{(1)} + cZ_t^{(2)}. \qquad (2)$$

где c — показывает долю абонентов одного типа в общем трафике, Y_t задает величину отклонения от среднего значения числа поступивших пакетов до момента t, за вычетом обработанных.

Каждая из компонент является дробным броуновским движением:

$$B_H(t) = \frac{1}{\Gamma(H+\frac{1}{2})} \int_{-\infty}^{t} h(t,\tau)dB(\tau), \qquad (3)$$

где H — параметр Херста (параметр самоподобия); $dB(\tau)$ — приращение Винеровского процесса; $\Gamma(H+\frac{1}{2})$ — гамма функция, а функция $h(t, \tau)$ задается следующим образом:

$$h(t,\tau) = \begin{cases} (t-\tau)^{H-\frac{1}{2}}, 0 \le \tau \le t \\ (t-\tau)^{H-\frac{1}{2}} - (-\tau)^{H-\frac{1}{2}}, \tau < 0 \end{cases}. \qquad (4)$$

Таким образом, общий трафик задается параметрами: c, H_1, H_2. При моделировании смеси самоподобных процессов и вычислении вероятности потерь пакетов при пиковых нагрузках возникают явления, вероятность возникновения которых может быть менее 10^{-5}. Применение стандартных методов типа Монте-Карло привело бы к необходимости проведения значительного числа экспериментов, что на самом деле является практически невозможным. Для моделирования столь редких событий в данной работе применен метод случайной замены меры, который состоит в том, что интересующее нас событие рассматривается относительно другой вероятностной меры, относительно которой оно имеет большую вероятность [10]. Так как компоненты дробного броуновского движения считаются независимыми, то можно использовать случайную замену меры для каждой компоненты процесса отдельно, это позволяет применить стандартные методы.

Процесс моделирования производился до момента переполнения буфера или же до момента T, которое выбрано таким образом, что вероятностью наступления интересующего нас события после этого момента можно пренебречь [9]. Для обоих процессов строил-

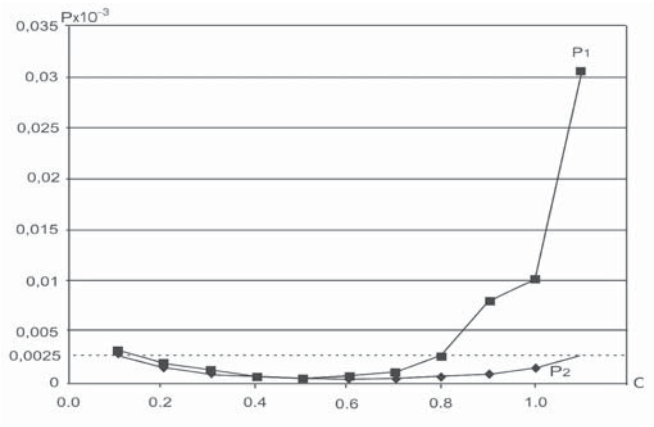


Рис. 1. Результаты моделирования при H1=0,5, H2 =0,8. Длина буфера A=3. Сеть обслуживает трафик с вероятностью потерь 0,0025

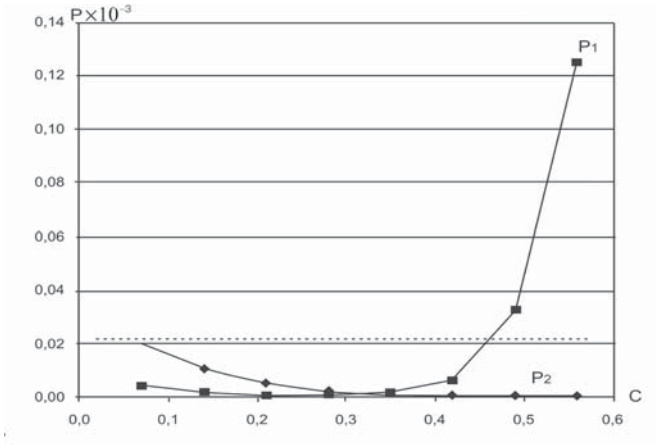


Рис. 2. Результаты моделирования при H1=0,5, H2 =0,8. Длина буфера A=5. Сеть обслуживает дополнительный трафик абонентов с вероятностью потерь 0,022 · 10^{-3}

ся массив точек интегрирования на положительной части интервала, здесь задавался постоянный шаг интегрирования, такой чтобы вычисление производилось с заданной точностью. На отрицательной части интервала точки интегрирования различны, в силу различий параметров Херста для двух потоков. Общее число точек отсчета взято 200, что при *H*, не превосходящих 0,9, обеспечивает погрешность, связанную с отбрасыванием бесконечного промежутка интегрирования, не более 10^{-6}.

Во всех опытах изменялся параметр с. Первые результаты моделируют смесь двух потоков (рис.1): первый — основной трафик мониторинга, а второй — дополнительный трафик. P_2 — вероятность переполнения буфера в том случае, когда у обеих независимых компонент входного потока H = 0,5; P_1 — вероятность переполнения буфера в том случае, когда у второй компоненты параметр Херста равен 0,8. Предполагается, что сеть обслуживает дополнительный трафик с вероятностью потерь 0,0025. На графике видно, что вероятность потерь пакетов превышает пороговое значение, при *c* > 0,7.

Затем увеличивалось значение буфера A до 5, что соответствовало вероятности потерь при обслуживании дополнительного трафика равной 0,000022 (рис. 2). При этом значение с изменялось от 0 до 0,7, так как область устойчивости сети только сужалась. Понятно, что при таких значениях вероятности, без использования метода замены меры, пришлось бы делать огромное количество опытов.

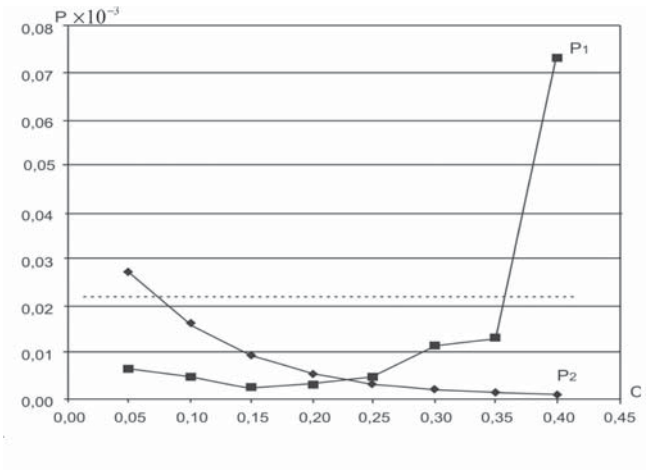


Рис. 3. Результаты моделирования при H1=0.5, H2 =0.9. Длина буфера A=5. Сеть обслуживает дополнительный трафик абонентов с вероятностью потерь 0,022 · 10^{-3}

Далее увеличен параметр Херста у трафика мониторинга (рис.3), что соответствует тому, что он оказывает большее влияние и в результате при тех же параметрах граница устойчивости сети отодвинулась и стала около 0,35.

Данные исследования показывают, что сеть может устойчиво функционировать при оптимальном соотношении между интенсивностями потоков с различными значениями параметра Херста, при этом увеличение доли процесса с меньшим значением параметра увеличивает устойчивость системы.

Выводы

1. Полученные результаты показывают, что вероятность переполнения буфера очень чувствительна к параметру Херста, особенно при его больших значениях.

2. Поведение вероятности переполнения буфера позволяет рассматривать величину параметра Херста как один из параметров, на основе которого нужно строить политику обеспечения качества обслуживания вызовов в беспроводных сетях.

Литература

1. **Кучерявый А.Е. , Цуприков А.Л.** Сети связи следующего поколения. М.: ФГУП ЦНИИС, 2006. — 280 с.

2. **Кучерявый А.Е., Кучерявый Е.А.** От e- России к u- России. Тенденции развития электросвязи. Электросвязь, 2005. — №5.

3. **Майская В.** Беспроводные сенсорные сети. Электроника: Наука, Технология, Бизнес, 2005. —№2.

4. **Шелухин О.И.,Осин А.В. Смольский С.М.** Самоподобие и фракталы. Телеком-муникационные приложения. / Под. ред. О. И. Шелухина. — М.: Физматлит, 2008. — 368 с.

5. **Кроновер Р.** Фракталы и хаос в динамических системах. — М.: Техносфера, 2006. — 448 с.

6. **Нейман В.И.** Новое направление в теории телетрафика // "Электросвязь", 1998. — № 7. — С. 27-30.

7. **Корнышев Ю.Н., Пшеничников А.П., Харкевич А.Д.** Теория телетрафика. Учебник для вузов. — М.: Радио и связь, 1996. — 272 с.

8. **Цитович И.И.** Устойчивые модели трафика мультисервисных сетей // Труды РНТОРЭС им. А.С. Попова, посвященной Дню радио. Выпуск LX-2. М., 2005. — Т. 2. — С. 271-273.

9. **Бубнов Ю.А., Цитович И.И.** О вычислении несобственного интеграла медленно убывающей функции. Труды 59 научной сессии РНТОРЭС им. А.С. Попова, посвященной Дню радио. М., 2004. — Т. 2. — С.179.

10. **Маликова Е.Е.** Задача оценки параметров качества сетей сбора и обработки телеметрической информации // Труды РНТОРЭС им. А.С. Попова, посвященной Дню радио. Выпуск LXIV. М., 2009. — С. 363-365.

Метод повышения производительности транспортного протокола TCP в глобальных корпоративных сетях передачи данных

Рассмотрены основные причины задержки в сетях передачи данных, вызванные протоколом TCP. Приведен пример, подробно описывающий функционирование и недостатки TCP в глобальных корпоративных сетях передачи данных. Предложен метод повышения производительности КСПД за счет снижения задержки TCP, не требующий повсеместной модернизации TCP.

Малиновский С.Т.,
Профессор кафедры МСиУС, МТУСИ

Беленькая М.Н.,
Доцент кафедры МСиУС, МТУСИ

Спиридонов А.А.,
Аспирант кафедры МСиУС, МТУСИ

В настоящее время протокол TCP (Transmission Control Protocol) повсеместно используется, для приложений, работающих в корпоративных сетях передачи данных (КСПД) и Интернете. TCP оперирует на транспортном уровне модели OSI (Open Systems Interconnection) и выполняет следующие основные функции [1]:

• Устанавливает соединение между процессами приложений на хостах, которые обмениваются между собой данными.

• Выполняет гарантированную доставку данных между процессами приложений, взаимодействующих хостов.

• Производит поиск доступной полосы пропускания и осуществляет контроль загрузки сети и, в зависимости от этих параметров, изменяет скорость передачи данных внутри TCP-соединения.

Последняя функция показывает, что протокол TCP оказывает существенное влияние на скорость передачи полезных данных (данных приложений) по сети, а следовательно и на задержку при работе с централизованными сетевыми приложениями. Этот протокол был разработан в 1981 году. Сегодняшние требования пользователей и программного обеспечения к сетям передачи данных значительно отличаются от тех, которые были в 80х годах. За последние годы проводилось большое количество исследований и разработок для его модернизации. Однако, не смотря на это, TCP подвергся лишь незначительным изменениям [4]. Сложность состоит в том, что большое число систем работают, используя данную версию протокола и изменить их все, применив новые протоколы или алгоритмы, модернизирующие TCP, не представляется возможным. Так в документации Microsoft TechNet для одной из наиболее популярных на данный момент серверных операционных систем Windows Server 2003 указано, что протокол TCP реализован на базе спецификации RFC793 и двух спецификаций с алгоритмами повышения производительности TCP — RFC2018 и RFC2581 [5].

• RFC793 Transmission Control Protocol — спецификация протокола TCP от 1981 года.

• RFC2018 TCP Selective Acknowledgment Options — алгоритм повышения производительности TCP — выборочное подтверждение о доставке.

• RFC2581 TCP Congestion Control — алгоритмы повышении производительности TCP — алгоритмы поиска доступной полосы пропускания и контроля загрузки сети TCP Slow Start, TCP Congestion Avoidance, TCP Fast Retransmit, TCP Fast Recovery.

Существуют условия, в которых протокол TCP плохо справляется со своими задачами и часто является причиной замедлений в сетях передачи данных. Виной тому является специфика: протокола TCP, алгоритмов поиска доступной полосы пропускания, алгоритмов контроля загрузки [4].

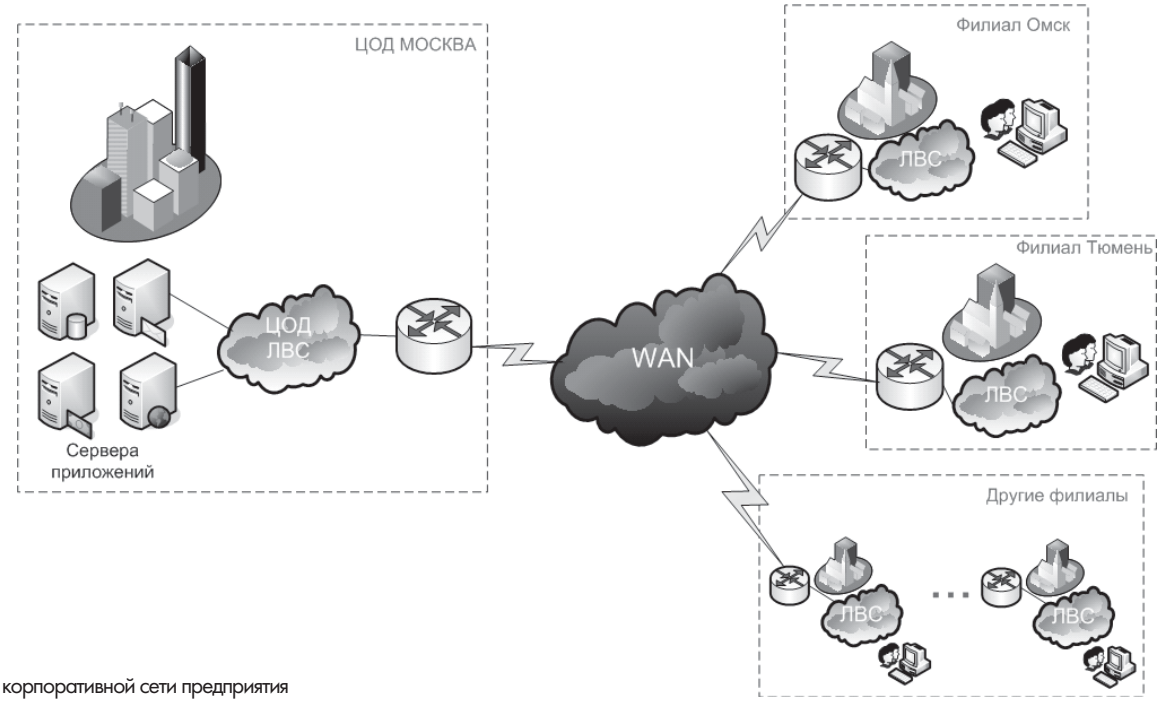


Рис. 1. схема корпоративной сети предприятия

Рассмотрим группу корпоративных территориально-распределенных глобальных сетей (рис. 1) и покажем причины возникновения задержек для пользователей сети. Большое число пользователей работают с сетевыми бизнес-приложениями реального времени и другими централизованными ресурсами через глобальную сеть передачи данных предприятия. Производительность их труда напрямую зависит от того, с каким качеством сеть предоставляет ресурсы центра обработки данных (ЦОД). Для эффективной работы, пользователь не должен испытывать трудностей при работе с приложениями, система должна быстро реагировать на его действия и запросы. У пользователя должно быть минимальное время ожидания на ответ системы. Но на больших расстояниях протокол TCP достаточно медленно производит набор скорости передачи данных для установленного соединения, это негативно отражается на производительности приложений, которые используют не продолжительные TCP-соединения, а так же длительные TCP-соединения, если возникают потери данных и происходит повторная передача. В сети находится множество TCP-соединений (TCP-сессий). В каждом из них работают алгоритмы поиска доступной полосы пропускания и контроля загрузки канала связи. При этом каждое соединения пытается занять свою максимально возможную часть пропускной способности канала. TCP-сессии негативно влияют друг на друга, создают коллизии, вызывают потери данных из-за перегрузки. В современных высоконадежных сетях передачи данных это основная причина, из-за которой теряются пакеты. Скорость передачи данных по каждому отдельному соединению снижается, что в условиях глобальных сетей приводит к замедлением. В связи с этим увеличивается задержка при работе с сетевыми приложениями, что в свою очередь влияет на производительность труда многочисленных территориально удаленных пользователей сети предприятия [1-4].

Рассмотрим величину задержки на примере участка КСПД от пользователя до сервера. Пользователь из Омска работает с централизованным сетевым приложением, установленным на сервере в ЦОД предприятия в Москве (рис. 2). От ПК пользователя поступает запрос на загрузку объекта приложения размером в 50 Kb. Предположим размер сегмента (область полезных данных TCP) равен 500 байт (по умолчанию 512), в этом случае сервер разобьет объект на сегменты размером в 500 байт и начнет передачу. Специфика работы протокола такова, что первым ходом сервер отправит один сегмент, а далее будет дожидаться подтверждения его доставки. После подтверждения начнется второй ход, за который передатчик отправит два сегмента, дождется подтверждения о доставке. И так с каждым шагом число сегментов за один ход будет увеличиваться вдвое. В случае потери, число переданных сегментов за ход сократиться вдвое и в дальнейшем будет возрастать линейно. Этот метод известен так же под названием "Скользящее окно TCP". Если не произойдет потеря или искажение данных во время передачи, протоколу TCP потребуется семь ходов (см. таблицу), чтобы передать запрошенный объект приложения. За каждый ход будет отправлена порция пакетов в одну сторону и передано подтверждение о доставке в другую сторону. В локальной сети задержка на выполнение одного хода мала, около 1 мс. В глобальных сетях задержка может достигать 300 мс, а при использовании спутниковых каналов связи может быть и >500 мс. На примере задержка при передаче данных из Москвы в Омск или из Омска в Москву составляет 100 мс. Следовательно, на один ход требуется 200 мс. А на передачу 50 Кб данных (на 7 ходов) потребуется 1,4 секунды, причем независимо от полосы пропускания канала. По сравнению с локальной сетью задержка возрастает на два порядка.Теперь представим ситуацию, когда пользователь не один и имеет место значительное число TCP-соединений между серверами в Москве и пользователями в Омске. При этом наступает момент, когда данных больше, чем сеть может передать. Заполняются буферы памяти на сетевых устройствах. При переполнении буфера, на сетевом устройстве теряется часть данных. Передающий хост, не отправляет следующую порцию сегментов пока не получит подтверждение о доставке. Как только истекает

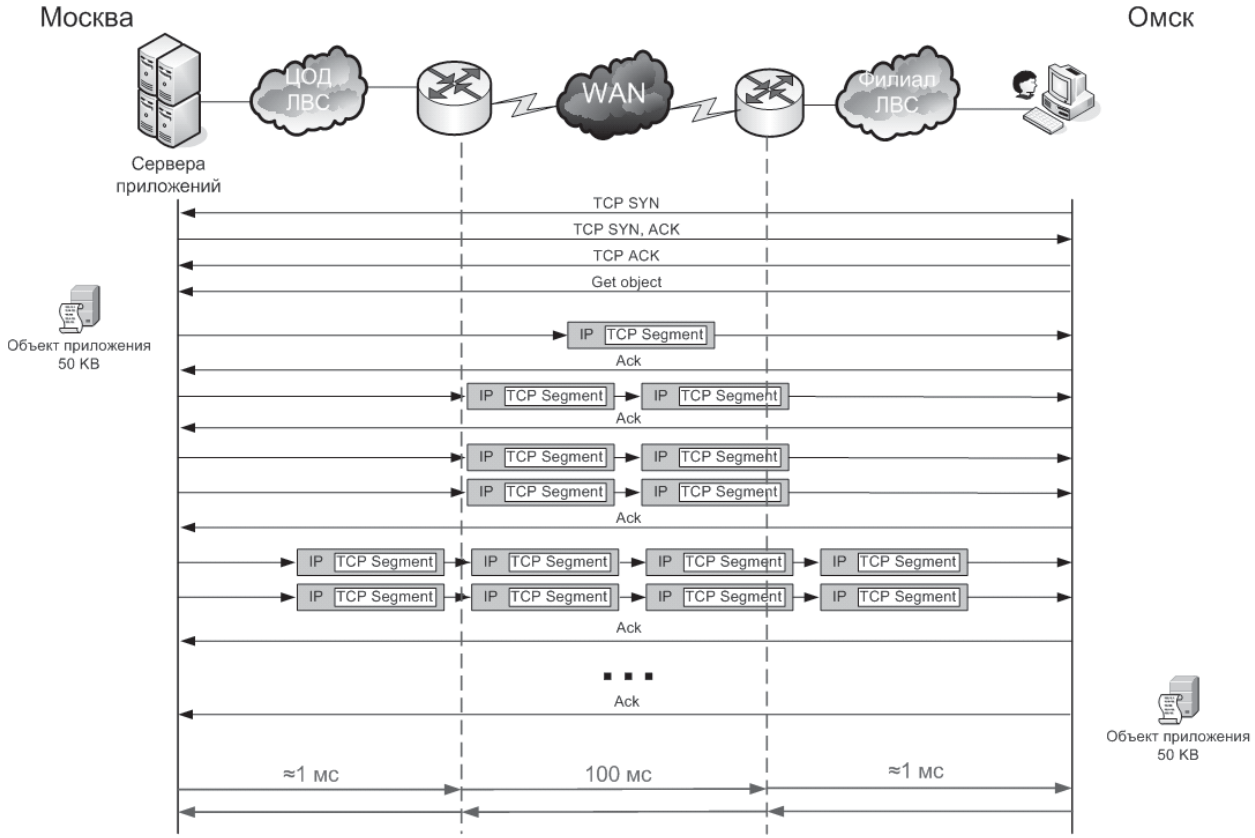


Рис. 2. Пример функционирования TCP

Характеристики ходов TCP-соединения

Номер хода	Число переданных сегментов за ход	Число переданных байтов за ход	Оставшееся число байтов	Задержка, сек
1	1	500	49500	0,2
2	2	1000	48500	0,4
3	4	2000	46500	0,6
4	8	4000	42500	0,8
5	16	8000	34500	1
6	32	16000	18500	1,2
7	64	32000	0	1,4

таймер ожидания подтверждения доставки, считается, что пакеты потеряны. Число передаваемых сообщений за один ход уменьшается вдвое и в дальнейшем возрастает линейно, а в некоторых ситуациях увеличение скользящего окна TCP начинается с начала. В результате еще больше увеличивается задержка, что негативно сказывается на работе удаленных пользователей с централизованными приложениями.

Следовательно, возникает задача повышения производительности сети, за счет снижения задержки в работе протокола TCP. Предлагается снизить задержку TCP за счет следующего метода: идея метода состоит в том, чтобы все многочисленные соединения между пользователями и серверами приложений ЦОДа терминировались локально, а на участке глобальной сети функционировало лишь одно служебное TCP-соединение с наилучшими характеристиками и внутри его передавались пользовательские TCP-сессии. В этом случае исключаются взаимные влияния между TCP-соединениями на участке глобальной сети, они происходят только на участках локальных сетей. Кроме того, исключается алгоритм скользящего окна на участке глобальной сети. Для реализации этого метода необходимо установить аппаратно-программные комплексы повышения производительности TCP (АПКПП TCP) согласно схеме на рис. 3. Алгоритм функционирования АПКПП TCP (рис. 4) состоит в том, чтобы все многочисленные соединения между пользователями и серверами приложений ЦОДа терминировались локально на сервера АПКПП TCP, а на участке глобальной сети устанавливалось одно служебное TCP-соединение. На устройствах АПКПП TCP формируется очередь из IP пакетов с TCP сегментами полезных данных приложений. Пакеты из очереди поступают в сокет TCP на АПКПП TCP и инкапсулируются в TCP поле данных (рис. 5), далее отправляются в рамках соединения между АПКПП TCP. Для исключения бесконечного нарастания очереди, контроль загрузки канала связи осуществляется АПКПП TCP за счет изменения скользящего окна на локальном участке сети. Алгоритмы поиска доступной полосы пропускания и контроля загрузки сети работают на участках высокоскоростных локальных сетей, в глобальной сети соединение установлено независимо от локальных участков и продолжает находиться в состоянии с наилучшими характеристиками все время. Получение наилучших характеристик осуществляется стандартными алгоритмами TCP однократно в момент первоначальной установки соединения между АПКПП TCP, ограничением является лишь полоса пропускания канала связи. Функционирование метода подробно изображено на рис. 4.

При использовании предложенного метода для передачи объекта приложения в 50 Кб из Москвы в Омск понадобился бы всего лишь один ход, то есть 200 мс, что в 7 раз быстрее обычного функционирования TCP.

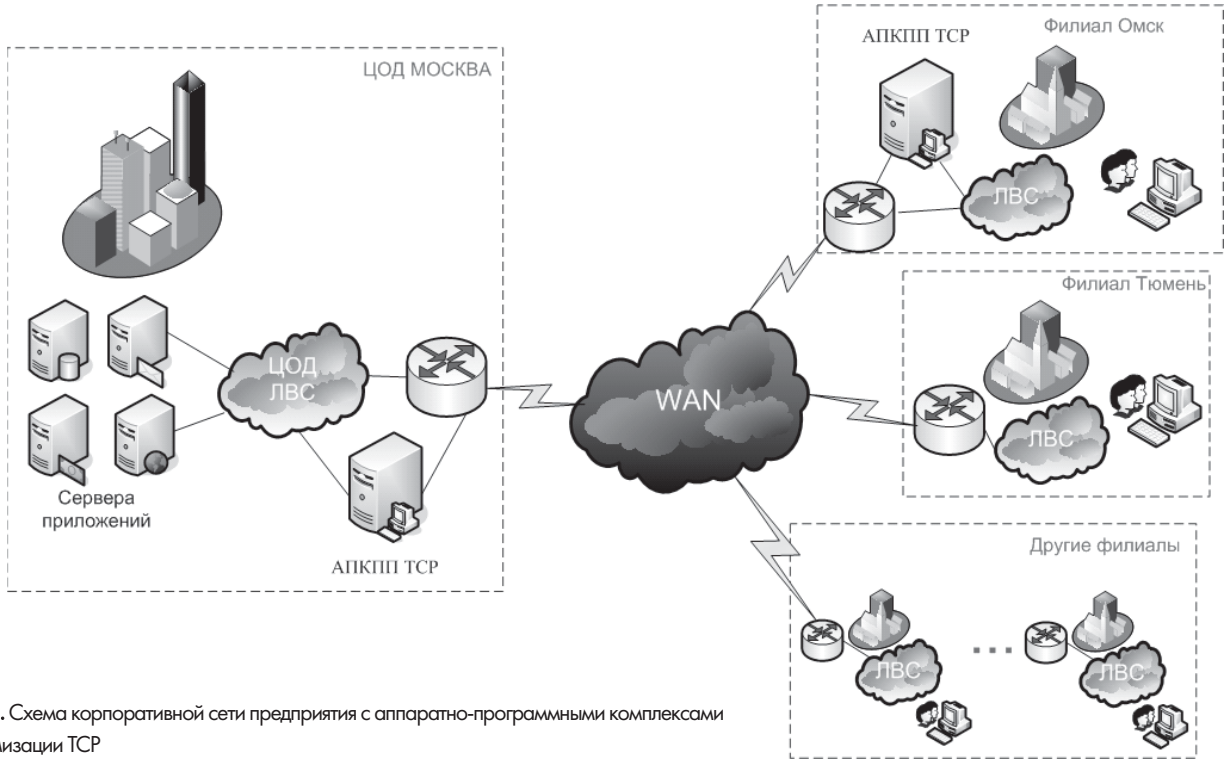


Рис. 3. Схема корпоративной сети предприятия с аппаратно-программными комплексами оптимизации TCP

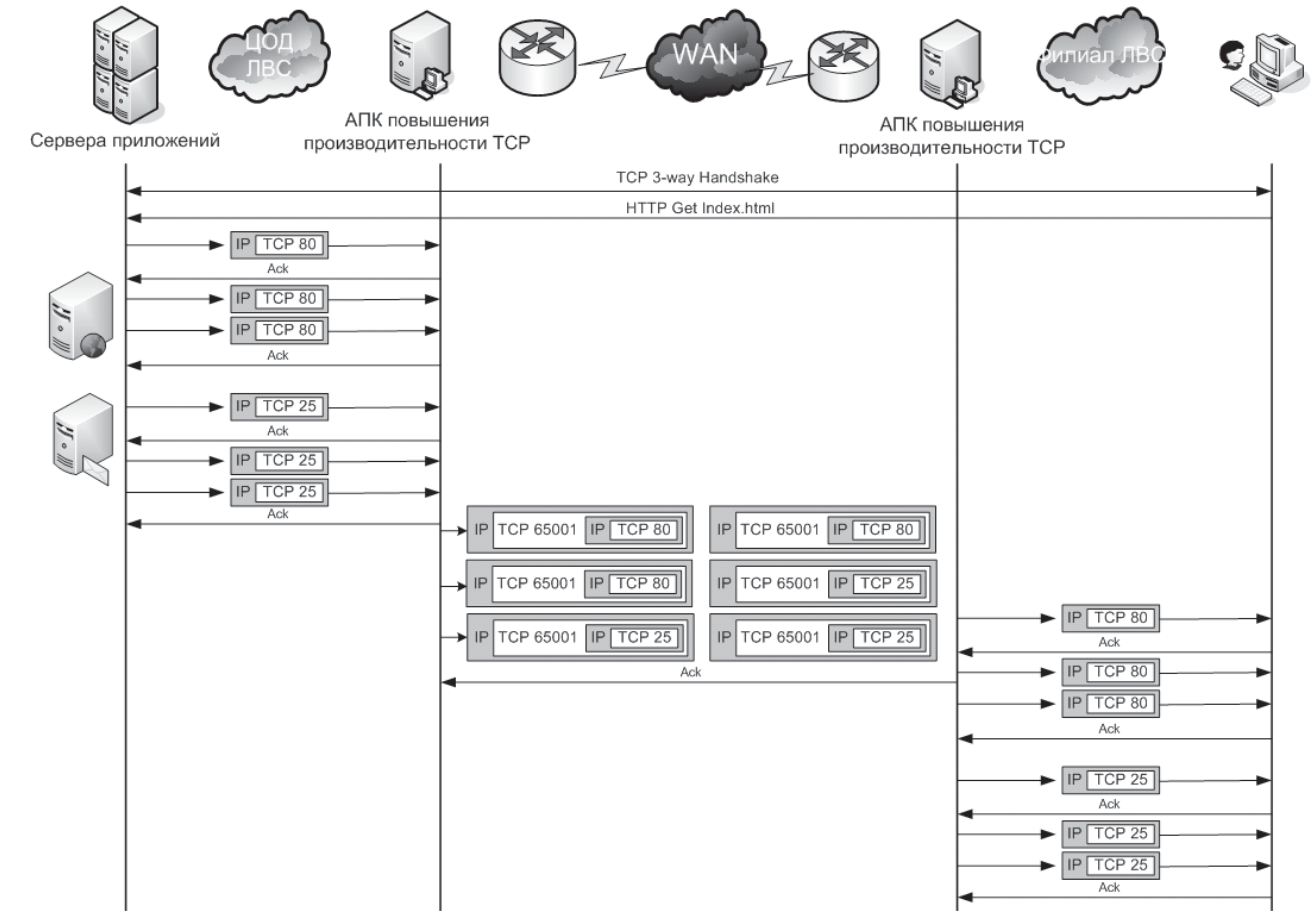


Рис. 4. Функционирование TCP с предложенным алгоритмом

Рис. 5. Схема инкапсуляции

Заключение

За счет предложенного метода исключается длительный процесс увеличения скользящего окна и взаимные влияния TCP-соединений на участке глобальной сети. Эти процессы происходят на участках высокоскоростной локальной сети, что многократно сокращает задержку. Так же увеличивается скорость передачи полезных данных приложений по глобальной сети, так как достигается более полное использование каналов связи. В топологии сети звезда, в которой используются каналы связи точка-точка исключается потеря данных, вызванная повышенной загрузкой канала, так как в канале связи будет функционировать только одно TCP-соединение. В полносвязных топологиях сети, потери значительно сократятся за счет существенного сокращения TCP-сессий. Их число будет ? числа узлов в полносвязной топологии. При использовании данного метода отпадает необходимость модернизировать TCP для всех систем, все модернизации производятся на серверах АПКПП TCP.

Литература

1. RFC793 — Transmission control protocol, DARPA Internet Program, protocol specification, Information Sciences Institute University of Southern California, 1981.

2. RFC2581 — TCP Congestion Control, M. Allman NASA Glenn/Sterling Software, V. Paxson ACIRI / ICSI, W. Stevens Consultant, 1999.

3. Congestion Avoidance and Control, Van Jacobson Lawrense Berkeley Laboratory, Michael J. Karels University of California at Berkeley.

4. Application Acceleration and WAN Optimization Fundamentals — Cisco Press, Ted Grevers Jr.; Joel Christner.

5. Microsoft Technet — http://technet.microsoft.com.

К вопросу выбора аналитической модели функционирования звена сети следующего поколения

Рассматривается математическая модель звена сети следующего поколения с учетом мультисервисного характера трафика. Приводится выражение для стационарного распределения вероятностей состояний исследуемой системы, которое позволяет получить ее основные вероятностно-временные характеристики.

Подольский Д.В.,
аспирант кафедры АЭС МТУСИ

Особенностями мультисервисных сетей следующего поколения является наличие не одного вида, как в традиционных сетях, а некоторого конечного множества видов трафика, создаваемого пользователями. Это множество видов трафика образует профиль трафика.

Учитывая эту особенность сетей следующего поколения, важнейшим вопросом при их проектировании, обслуживании и эксплуатации является решение задачи по обеспечению определенного заданного качества обслуживания для различных видов трафика, обслуживаемых сетью. Принимается, что если трафик от различных по своему характеру источников неразличим сетью и обслуживаются единообразно, то в этом случае качество обслуживания, которое обеспечивает сеть для любого вида трафика, определяется как максимально доступное в момент обслуживания (Best Effort) [1]. В этом случае можно говорить о некоторой обобщенной величине показателя качества обслуживания, которая не учитывает разнородность источников нагрузки, а следовательно, и главное свойство мультисервисной сети. Мультисервисная сеть должна обеспечивать ресурсами каждый тип пользователей в соответствии с предъявляемыми запросами этого типа пользователя. В этом выражается понятие качества обслуживания пользователей для мультисервисной сети.

В общем случае любая мультисервисная сеть может быть охарактеризована собственным профилем трафика и соответствующими видами его обслуживания. Параметр, характеризующий качество обслуживания в сетях NGN, таким образом, может быть выражен в виде функциональной зависимости от профиля трафика:

$$\overline{QoS} = f(\text{профиль_трафика}). \qquad (1)$$

Для обеспечения заданных показателей качества обслуживания для различных видов трафика в мультисервисной сети должны присутствовать механизмы, которые определяют, к какому виду можно отнести поступающий от источника трафик и обеспечивают качество обслуживания в соответствии с видом трафика.

Таким образом, в мультисервисной сети для предоставления сервиса с гарантированным уровнем качеством обслуживания должны быть решены следующие две основные задачи. Это, в первую очередь, задача по сортировке поступающего трафика таким образом, чтобы было возможно обрабатывать одни виды трафика отлично от других. Вторая задача выражается в предоставлении сетевых ресурсов трафику в соответствии с его видом, который определяет требования к качеству обслуживания [2].

Мультисервисный трафик в сети следующего поколения можно условно разделить на три базовых класса: потоковый трафик, эластичный трафик, требующий гарантированной пропускной способности и эластичный трафик, не требующий гарантированной пропускной способности.

Для обслуживания различных классов трафика с различными требованиями к сетевым ресурсам в сетях следующего поколения применяются механизмы управления уровнем качества обслуживания. Наибольшее распространение получил механизм дифференцированного обслуживания DiffServ, согласно которому виды обслуживания трафика делятся в соответствии с классами трафика:

- Expedited Forwarding (EF) — трафик, не допускающий задержки в обслуживании, т.е. требующий немедленной передачи пакетов.
- Assured Forwarding (AF) — трафик, при обслуживании которого требуется гарантированное качество доставки пакетов.
- Best Effort (BE) — трафик, при обслуживании которого приемлемым является максимально доступное на данный момент на сети качество обслуживания.

Для оценки ресурсов, требуемых для поддержания заданного уровня качества обслуживания в сети следующего поколения, сначала рассмотрим звено сети. Функциональная структура звена приведена на рис.1 [2]. Под звеном мы будем понимать, с одной стороны, узел доступа к сети, а с другой стороны — направление передачи к другому (одному) узлу. Маркированный трафик поступает на узел, который анализирует заголовки пакетов и классифицирует их в соответствии с алгоритмом агрегированного резервирования ресурсов [3]. Согласно этому алгоритму пакеты одного класса обслуживаются по единой PHB (per Hop Behavior) политике в соответствующих очередях. Распределение ресурсов звена (пропускной спо-

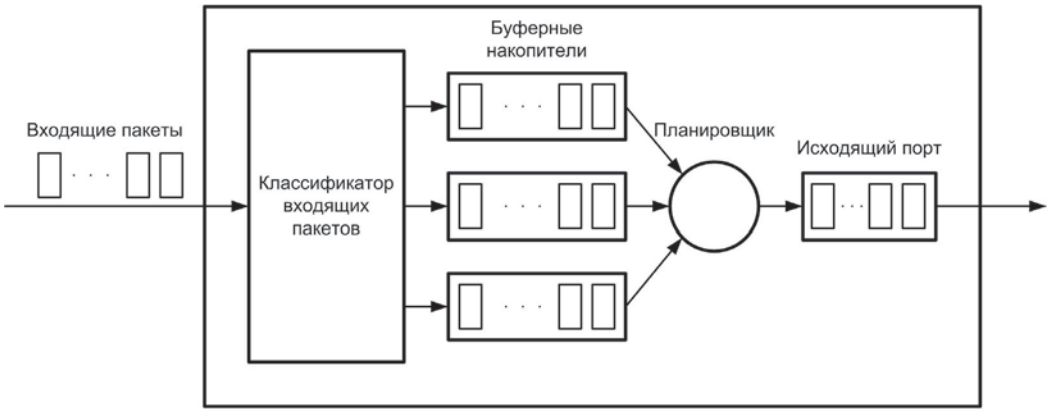


Рис. 1. Функциональная структура звена сети следующего поколения

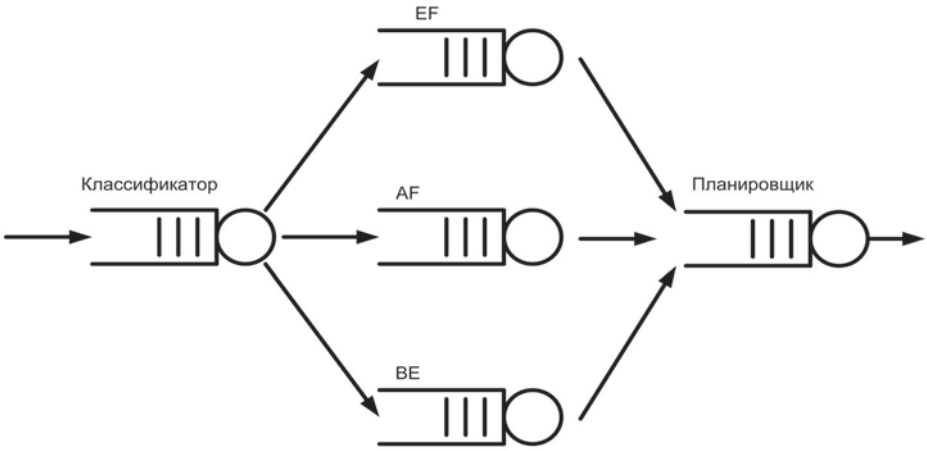


Рис. 2. Модель звена сети следующего поколения

собности) в исходящем направлении между различными классами трафика осуществляется согласно механизму резервирования ресурсов планировщика.

Поскольку рассматриваемый объект является функционально комплексной системой, то для ее описания в качестве инструмента для построения аналитической модели функционирования воспользуемся теорией сетей массового обслуживания (СеМО) [4]. На рис. 2 приведена модель звена сети следующего поколения, представленная в терминах теории сетей массового обслуживания.

Как видно из рис. 2, модель включает классификатор, который сортирует поступающие пакеты по очередям в соответствии с классом пакета, очередей, в которых хранятся заявки соответствующего класса (AF, EF, BF), и планировщика, распределяющего ресурсы исходящего канала между поступающими пакетами различных классов в соответствии с заданной дисциплиной обслуживания.

Для данного исследования примем допущение о том, что время поступления пакетов и время обслуживания пакетов подчиняются экспоненциальному распределению.

Поступающие на узел пакеты вначале попадают в классификатор с буферной памятью емкости $R_{кл}$, где определяется класс трафика пакета. Классификатор распределяет пакеты по очередям в зависимости от класса трафика. Каждая очередь для класса трафика (EF, AF, BF) имеет буфер, с емкостями R_{EF}, R_{AF} и R_{BE} соответственно. Из очереди пакеты попадают на вход планировщика. Планировщик анализирует доступные ресурсы в системе и принимает решение о порядке обслуживания пакетов. Планировщик также обладает собственной памятью емкостью $R_{пл}$.

Исследуемая система может быть представлена с помощью неоднородной открытой сети массового обслуживания, состоящей из $N = 5$ узлов с дисциплинами обслуживания FIFO (First In First Out) и PS (Processor Sharing), с заданным множеством классов сообщений и маршрутной матрицей. В рассматриваемой сети массового обслуживания узлами с дисциплиной обслуживания FIFO моделируются: работа классификатора, в том числе ожидание пакетов в очередях на обслуживание EF, AF и BF. Узел, моделирующий планировщик, реализует дисциплину обслуживания PS.

Указанные параметры системы позволяют отнести ее к классу BCMP-сетей (название класса сетей образовано от первых букв фамилий его авторов Baskett F., Chandy K., Muntz R., Palacios F.), стационарные вероятности состояний которых имеют мультипликативную форму [4]. Следовательно, для построения модели и расчета показателей качества обслуживания можно воспользоваться методами расчета BCMP-сетей.

Маршрутная матрица или матрица переходов $P = [P_{i,r,j,s}]$ определяет марковскую цепь, в которой состояния отмечаются парами (i,r), где i и j определяют номер узла, а r и s — класс заявки. В рассматриваемой системе существует три класса заявок. Так как после поступления в систему заявка не меняет своего класса, то при $r \neq s$ $P_{i,r,j,s} = 0$, следовательно, можно записать $P_{i,r,j,s} = P_{i,r,j,r}$.

Для заявки класса r примем, что $\{e_{ir}, 1 \le i \le N\}$ является решением системы уравнений:

$$\sum_{1 \le i \le N} e_{ir} P_{i,r;j,r} + P_{jr} = e_{jr}, \qquad (2)$$

где e_{ir} — определяет относительную интенсивность потока сообщений класса r, проходящего через i-ый узел сети, а P_{jr} — вероятность того, что входящие из внешнего источника сообщение поступит в j-ый узел сети и станет сообщением класса r [5].

Если сеть открытая, то процесс поступления заявок не зависит от состояния системы и оно определяется исключительно числом заявок в каждом из узлов сети, т.е. каждое состояние описывается вектором $S = (n_1, n_2, ..., n_N)$, где n_i — число заявок в i-ом узле, а N — число узлов в сети. Тогда стационарное распределение вероятностей состояний сети определяется выражением [6]:

$$P(S) = Cd(S)h_1(n_1)h_2(n_2)...h_N(n_N), \qquad (3)$$

где C — нормировочный коэффициент, $d(S)$ — функция полного числа заявок в системе, а $h_i(n_i)$ — функции, определяющие тип обслуживания в i-ом узле.

Для узлов с дисциплиной обслуживания FIFO:

$$h_i(n_i) = \left(\sum_{r \in R} e_{ir}\right)^{n_i} \prod_{j=1}^{n_i} \frac{1}{\mu_i(j)}. \qquad (4)$$

Для узлов с дисциплиной обслуживания PS:

$$h_i(n_i) = \left(\sum_{r \in R} \frac{e_{ir}}{\mu_{ir}}\right)^{n_i}. \qquad (5)$$

Вычисление нормирующего коэффициента требует суммирования выражения для стационарных вероятностей по всем возможным состояниям. Поскольку в нашем случае сеть открытая, то любое число заявок возможно в любом ее узле. Следовательно:

$$C^{-1} = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} ... \sum_{n_N=0}^{\infty} \left(\prod_{i=1}^{N} \lambda^{n_i} h_i(n_i)\right) \qquad (6)$$

или

$$C^{-1} = \left(\sum_{n_1=0}^{\infty} \lambda^{n_1} h_1(n_1) \right) \left(\sum_{n_2=0}^{\infty} \lambda^{n_2} h_2(n_2) \right) \ldots \left(\sum_{n_N=0}^{\infty} \lambda^{n_N} h_N(n_N) \right), \quad (7)$$

где

$$\sum_{n_i=0}^{\infty} \lambda^{n_i} h_i(n_i) = \left(1 - \sum_{r \in R_i} \lambda \frac{e_{ir}}{\mu_i} \right)^{-1} \quad (8)$$

для узлов с дисциплиной обслуживания FIFO и

$$\sum_{n_i=0}^{\infty} \lambda^{n_i} h_i(n_i) = \left(1 - \sum_{r \in R_i} \lambda \frac{e_{ir}}{\mu_{ir}} \right)^{-1} \quad (9)$$

для узлов с дисциплиной обслуживания PS.

Таким образом, стационарное распределение вероятностей состояний сети только от средних значений длительности обслуживания в узлах сети, а не от их функции распределения. Используя выражение для стационарного распределения вероятностей состояний сети, можно получить основные вероятностно-временные характеристики исследуемой системы (звена сети), такие как: маргинальное распределение числа сообщений в каждом из компонентов системы, среднее время обслуживания различных классов пакетов и пропускную способность для различных классов пакетов.

Литература

1. **И.Г. Бакланов.** NGN: принципы построения / под ред. Ю.Н. Чернышева. — М.: Эко-Трендз, 2008. — 400 с.: илл.

2. **Kun I. Park.** QoS in Packet Networks. Springer, 2005. — 245 p.

3. **Z-L. Zhang, Z. Duan, and Y. T. Hou.** Fundamental trade-offs in aggregate packet scheduling. In Proc. of the 9th International Conference on Network Protocols ICNP, page 129. IEEE Computer Society, 2001.

4. **Г. П. Башарин, А.Л. Толмачев.** Теория сетей массового обслуживания и ее приложения к анализу информационно-вычислительных систем, Итоги науки и техн. Сер. Теор. вероятн. Мат. стат. Теор. кибернет., 21, ВИНИТИ, М., 1983. — P.3-119.

5. **В.М. Вишневский.** Теоретические основы проектирования компьютерных сетей. — М.: Техносфера, 2003. — 512 с.

6. **Baskett F., Chandy K.M., Muntz R.R., Palacios F.G.** Open, Closed and Mixed Networks of Queues with Different Classes of Customers//J. Assoc. Comput. Mach, 1975. — Vol. 22, №2. — P. 248-260.

Перспективы и предпосылки внедрения сетей связи следующего поколения в России

Статья посвящена анализу существующих условий перехода на сети связи следующего поколения в Российской Федерации. Рассматриваются как положительные, так и отрицательные аспекты, проанализированы статистические данные, экономические, организационно-правовые и технические факторы, сделан прогноз дальнейшего пути развития сетей связи общего пользования и их интеграции с сетями связи следующего поколения.

Хатунцев А.Б.,
аспирант кафедры АЭС МТУСИ

Введение

Под сетями связи нового поколения подразумевается спектр решений и технологий (подходы, предлагаемые международными организациями, концепции построения от производителей оборудования, руководящие технические документы различных системных групп и т.п.). Основными их элементами являются гибкий коммутатор (Softswitch), мультимедийная подсистема на базе протокола IP (IP Multimedia Subsystem — IMS), различные шлюзы — медиа, транковые, сигнальные и т.д., устройства мультисервисного доступа, транспортные технологии (многопротокольная коммутация по меткам — MPLS и др.) и пр.

В рамках поставленной задачи анализа, условия перехода к сетям следующего поколения следует разделить на три глобальных фактора:

- Экономические факторы;
- Организационно-правовые факторы;
- Технические факторы.

Экономические факторы

Экономические факторы при модернизации сети связи, бесспорно, являются основными для любого оператора связи. С одной стороны, затраты на проведение глобальной модернизации сети довольно велики, а окупаемость является процессом медленным и нацеленным на перспективу. С другой стороны, ввод в эксплуатацию новых средств связи и уход от коммутации каналов позволит оптимизировать дальнейшие затраты и более экономично использовать канальные ресурсы. Основной идеей перехода на сети связи следующего поколения является поэтапная модернизация ныне существующих телекоммуникационных сетей. В результате этого, на сегодняшнем этапе внедрения новых сетевых технологий в РФ, наблюдается так называемый переходный период — совместное функционирование сетей старого и нового поколения, что вносит трудности обеспечения качества предоставляемых пользователям услуг связи.

Нельзя умалять и вклад всей отрасли связи в экономику государства. На современный этапе развития цивилизации стал четко ощущаться переход от индустриального общества к информационному, предполагающему новые формы социальной и экономической деятельности, которые базируются на массовом использовании информационных и телекоммуникационных технологий [1].

Отрасль телекоммуникаций относится к числу инфраструктурообразующих секторов экономики и играет ключевую роль в жизнедеятельности государства, организации и управления производством, а также является основой для создания информационного общества, развития сетевой экономики и производства виртуальной продукции. На отрасль телекоммуникаций ложится и большая социальная нагрузка, поскольку основной целью еесоциально-экономического развития на долгосрочную перспективу является повы-

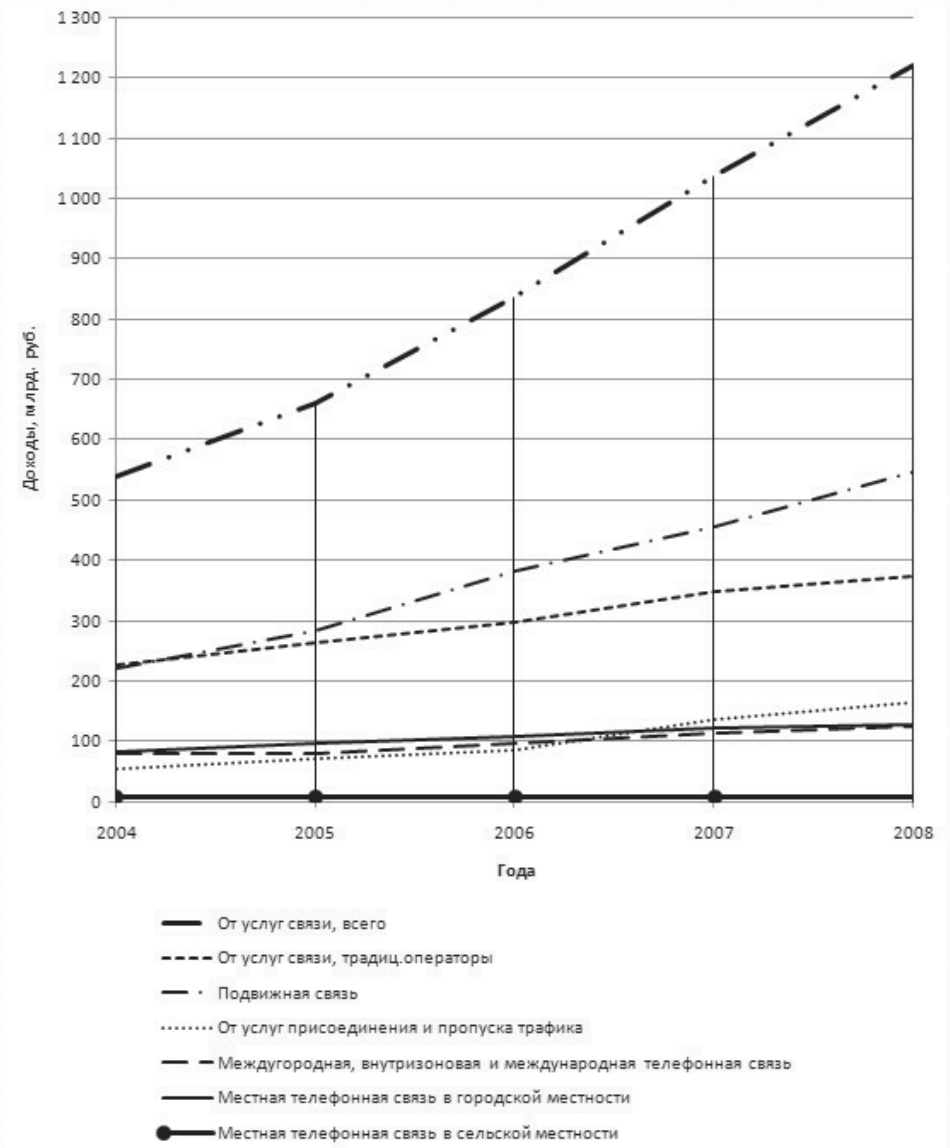


Рис. 1. Рост доходов операторов связи

шение уровня жизни населения и обеспечение доступности услуг связи. Стратегическое значение отрасли телекоммуникаций состоит в повышении эффективности государственного управления и в решении вопросов обеспечения обороноспособности страны и безопасности государства [2].

Вклад отрасли телекоммуникаций в экономику России характеризуется доходами от услуг связи, формирующими значительную часть валового внутреннего продукта (ВВП) страны. К III кв. 2009 г. доходы от услуг связи выросли на 5,5% по сравнению с аналогичным периодом 2008 г. Доходы от услуг связи традиционных операторов возросли на 4,9%. Основные показатели доходности отрасли за 2004-2008 гг. представлены на рис. 1. Все статистические данные основываются на официальной аналитике Министерства связи и массовых коммуникаций РФ [3].

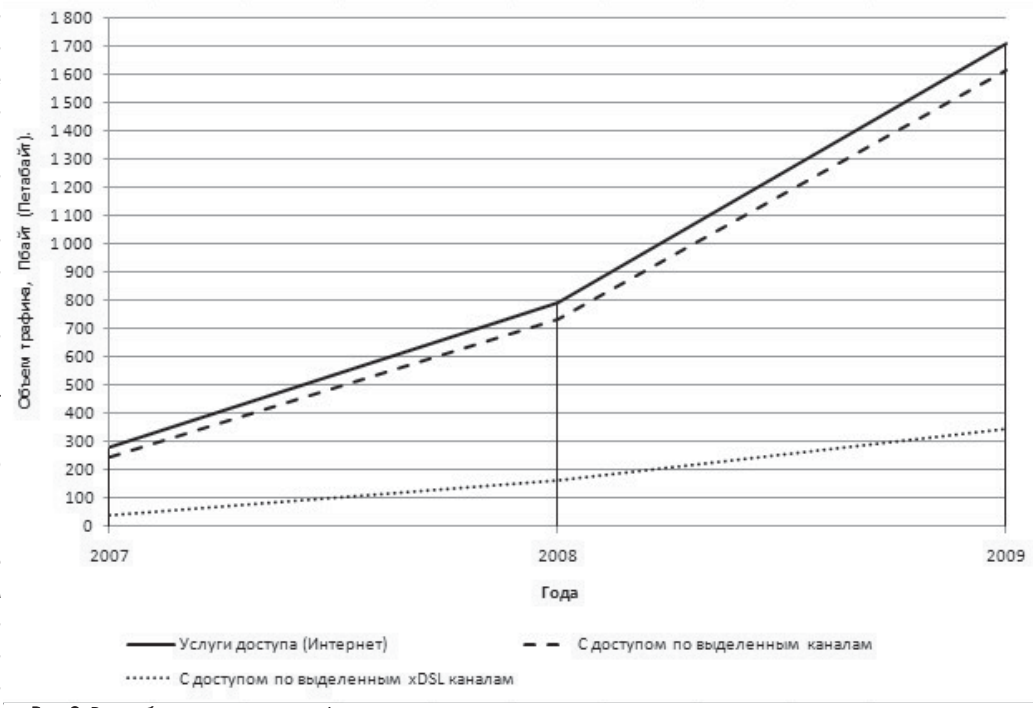


Рис. 3. Рост объема пакетного трафика

Исходя из рис.1, за анализируемый период наблюдается стабильный рост прибыли во всех основных сферах телекоммуникационных услуг.

На рис. 2 представлен анализ Gartner Group [4] общемировой динамики доходов операторов фиксированной и мобильной связи в основных сегментах рынка, в котором четко прослеживается существенное падение удельной составляющей доходов операторов фиксированной телефонной связи и реальный рост доходов операторов мобильной связи. Данное исследование представляет интерес с точки зрения перспектив российских операторов связи, поскольку эти же тенденции неизбежно через определенный промежуток времени проявятся и в России.

В настоящее же время в России при аналогичном мировым тенденциям росте доходов операторов мобильной связи все ещенаблюдается рост доходов операторов фиксированной связи, хотя уже и замедляющимися темпами (рис. 1).

Наряду с услугой телефонии растут также объемы передаваемых данных и обмена трафиком. Так, объем услуги доступа в Интернет к концу III кв. 2009 г. увеличился в 3,5 раза по сравнению с аналогичным периодом 2008 г. и составил 1 707 350 841 765 Мб. Показатели использования услуги доступа в Интернет представлены на рисунке 3.

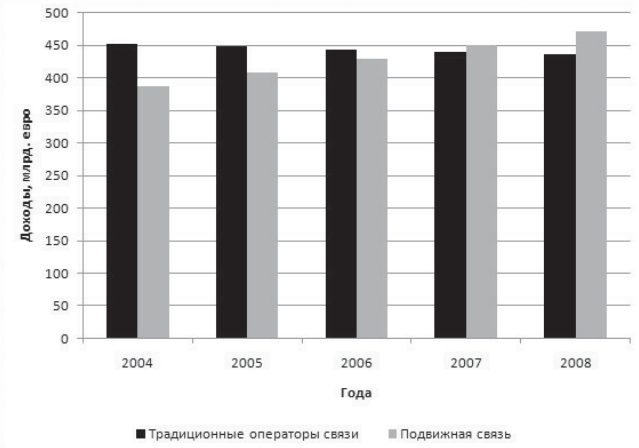


Рис. 2. Изменение пропорций в доходах мировых операторов связи

Как видно из рисунка 3, объемы пакетного трафика за последние три года выросли более чем в 6 раз. С ростом количества абонентов интернет-провайдеров и увеличения домашних пользователей xDSL закономерно уменьшилось использование сети ТфОП для передачи пакетных данных и выхода в Интернет. В связи с этим стала менее востребована услуга доступа в Интернет через пункты коллективного доступа (ПКД).

Одновременно с ростом доходов операторов связи и увеличением внедрения услуг передачи данных и VoIP, наблюдается закономерный спад в объеме классических абонентских устройств [5]. Так к концу III кв. 2009 г. количество квартирных абонентских устройств на местной телефонной сети уменьшилось почти на 193 тысячи по сравнению с аналогичным периодом прошлого года. В то же время выросло количество абонентских устройств операторов подвижной связи — за тот же период более чем на 23 миллиона. Показатели, характеризующие данный процесс, приведены в таблице.

Все вышеуказанные показатели говорят о постепенной и уверенной конвергенции сетей передачи данных и ТфОП, с уменьшением доли присутствия последней. И, судя по всему, такая тенденция сохранится и в будущем.

Организационно-правовые и технические факторы

Что касается организационно-правовых аспектов, то нельзя не признать наличие существенных проблем в осуществлении лицензирования и стандартизации нового оборудования и протоколов. Многие аспекты сетей нового поколения в Российской Федерации до сих пор не нормированы, что тормозит внедрение оборудования и, следовательно, развитие отрасли связи [6]. Но, несмотря на такие, казалось бы, существенные минусы, более поздний по сравнению с европейскими странами переход на сети связи следующего поколения привнес и ряд плюсов. А именно, большой набор обновленных и откорректированных рекомендаций ведущих мировых ис-

Количество абонентских устройств на местной телефонной сети и в сетях подвижной связи

Наименование показателя	Все операторы		
	3-й квартал 2009 года (тыс.единиц)	В % к соотв. периоду пред. года	Прирост за 3-й квартал 2009 года (тыс.единиц)
Абонентские устройства, включая абонентские устройства учрежд. станций	45 388	99,9	-172,0
Абонентские устройства ГТС и СТС	43 193	98,7	-660,9
Основные абонентские устройства ГТС и СТС	42 400	98,9	-593,4
Квартирные абонентские устройства ГТС и СТС	33 817	99,6	-192,4
Основные квартирные абонентские устройства ГТС и СТС	33 760	99,6	-185,7
Абонентские устройства ГТС и СТС, подсоединенные к УПАТС	2 195	131,4	488,8
Абонентские устройства подвижной связи	222 837,9	117,5	23 171,8

следовательских институтов связи, отлаженное оборудование, обширный опыт построения аналогичных сетей связи и созданная инфраструктура специализированных тренинг-центров. Все это позволяет избежать большинства ошибок и проблем, всегда имеющих место при внедрении нового оборудования.

Рассмотрим технические аспекты перехода к сетям связи следующего поколения. Изначально специфической особенностью цифровизации российской ТфОП стало планирование структуры сети на всех уровнях ее иерархии. Опыт развитых и большинства развивающихся стран свидетельствует, что в процессе цифровизации ТфОП существенно сократилось количество коммутационных станций и уровней иерархии, для оптимального построения сети доступа стали использоваться выносные концентраторы. На сегодняшний день развитие инфокоммуникационных услуг осуществляется, в основном, в рамках компьютерной сети Интернет, доступ к услугам которой происходит через традиционные сети связи. Объемы рынка услуг VoIP, и в частности SIP-телефонии, неуклонно растут. В то же время в ряде случаев услуги Интернет, ввиду ограниченных возможностей еетранспортной инфраструктуры, не отвечают современным требованиям, предъявляемым к услугам информационного общества. Тормозящим фактором развития сетей нового поколения является наличие у операторов обширного объема оборудования предыдущих поколений, которое не успело себя окупить и, помимо этого, приносит стабильный доход. Поэтому операторы не спешат отказаться от его использования и осуществлять капиталовложения в новые средства связи.

Отрицательным фактором является и большая территория страны, что влечет за собой сложности внедрения и оптимизации сетей нового поколения. Следствием этого, как было указано выше, является период совместного функционирования технологий коммутации каналов и коммутации пакетов, поэтапное замещение оборудования предыдущих поколений и преодоление трудностей отладки и взаимодействия таких конвергентных сетей.

С другой стороны, скорость изменения требований современного абонента не позволяет оператору остановиться на месте. Если это происходит, то конкуренты немедленно завладевают его наиболее доходными клиентами, вернуть которых, как показывает статистика, нелегко. Высокая межоператорская конкуренция влечет за собой модернизацию сетей связи, так как операторам необходимо искать способы удовлетворения требований клиентов, вводить новые услуги, активно их продвигать и развивать сеть. Важно заметить, что скорость появления новых и совершенствования используемых технологий сегодня чрезвычайно высока. Операторы связи зачастую не только не успевают внедрять новые технологии, но, в ряде случаев, даже тщательно их анализировать. На сегодняшний день большинство абонентов созрели для ухода от узкого спектра услуг сетей ТфОП в сторону повышения своей мобильности, меньшей зависимости от определенного пользовательского терминала, унификации услуг и их регулярно расширяемого диапазона. Иными словами, прозрачность услуг и доступность каждой из них независимо от дислокации пользователя и используемого им терминала является ведущим фактором, благоприятствующим переходу к пакетным сетям следующего поколения.

Особенности перехода к сетям связи следующего поколения

Одной из главных задач для операторов является не только модернизация сети и расширение спектра предоставляемых услуг, но и обеспечение их высокого качества, что подразумевает отлаженный контроль сети, всех ееузлов и обеспечение эффективной работы сети сигнализации на всех участках. Центральную роль в задачах развития и модернизации сетей связи, безусловно, играют производители телекоммуникационного оборудования. Их задача состоит в создании необходимых оператору технических средств на основе конкурентоспособных решений и самых современных технологий [7]. Скорость, с которой производитель вынужден работать для того, чтобы идти в ногу с новыми технологиями, к сожалению, влияет на качество предлагаемых решений. Поэтому в переходный для российский сетей связи период главной задачей является обеспечить эффективное функционирование конвергентных сетей связи без каких-либо ухудшений в качестве обслуживания для клиентов по сравнению с классическими сетями ТфОП и коммутацией каналов.

Следствием этого является тот факт, что сети следующего поколения NGN формируются на основе телефонной сети общего пользования. В настоящее время существуют концепции построения сетей NGN, технологии сопряжения сетей ТфОП с IP-сетями. И задача реализации надежного и производительного оборудования в таких сетях на сегодняшний день является главной.

На начальном этапе внедрения сетей связи следующего поколения исходят из того, что оператору известны требования потенциальных абонентов, заданы ограничения на показатели качества обслуживания всех видов трафика, качества передачи информации и надежности связи. Но выбор архитектуры построения NGN нетривиален. Идеология консорциума IPCC (International Packet Communication Consortium) предполагает использование набора сетей с обеспечением их взаимодействия, что идеально подходит именно для одновременной работы разных сетей. Применение гибкого коммутатора Softswitch в качестве устройства управления позволяет одинаково легко работать с любыми протоколами сигнализации, а значит, и с любым оборудованием, которое эксплуати-

руется в сети. Таким образом, набор сетей операторов, на которых используются абсолютно разные технологии, превращается в единую сеть общего пользования. Сами стыки IP-ТфОП реализуются с использованием различных шлюзов, обеспечивающих единство сети и возможность передачи информации любого вида, что и характеризует их как центральный элемент конвергентных сетей [8].

Кроме этого, сегодня задача конвергенции двух фиксированных сетей (с коммутацией каналов и коммутацией пакетов) дополнилась беспроводными технологиями. В традиционной телефонии мобильная связь, как видно из статистических данных, приведенных на рис. 1-3, постепенно вытесняет проводную. Одновременно с этим в сегменте пакетной сети уверенно лидируют технологии Wi-Fi и WiMAX, последняя из которых в 2009 году совершила успешный рывок к популяризации благодаря брэнду "Yota" [9]. В результате этого термин "конвергенция сетей и услуг связи" сменился на "конвергенция фиксированной и мобильной связи" (FMC — Fixed Mobile Convergence), что означает полную прозрачность всех видов существующих сетей связи для конечного пользователя. Данный подход отражен в концепции IMS (начиная с седьмого релиза рабочих групп 3GPP и TISPAN), которая подразумевает единую сеть, основанную на протоколе SIP (Session Initiation Protocol). Исходя из схожести подходов к реализации сетей NGN и 3G/4G, можно утверждать, что в будущем сети NGN в значительной мере будут опираться на мобильный доступ, используя проводные средства на уровне базовой сети (Core Network).

Выводы

Таким образом, можно с уверенностью сказать, что, не смотря на ряд отрицательных факторов, все основные условия перехода к сетям связи следующего поколения в РФ созданы в полном объеме. В первую очередь к таким глобальным переменам операторов подталкивают сами пользователи, которые, насытившись существующим спектром услуг, стали требовать их унификации и расширения возможностей доступа к ним с любого терминала и из любой точки страны. Но, поскольку этот подход требует больших капиталовложений со стороны операторов, был разработан поэтапный план конвергенции различных технологий. В данный момент отечественные сети находятся на переходном этапе к сетям связи следующего поколения. Существуют и успешно функционируют различные виды сетей связи, разные типы применяющихся технологий транспорта и доступа. Главной задачей на сегодняшний день является их эффективное сопряжение, что позволит наряду с расширением спектра услуг обеспечивать качество их предоставления как минимум без каких-либо заметных ухудшений с точки зрения норм ТфОП.

Литература

1. Концептуальные положения по построению мультисервисных сетей на ВСС России. — Минсвязь РФ, 25.01.2002.

2. **Масленников И.О.** Разработка концепции инфокоммуникационного регулирования, содействующей интеграции связи, вещания и информатизации. — www.rezonance.com.

3. www.minsvyaz.ru.

4. www.gartner.com.

5. **Мальцев А.** Новые стратегии "Связьинвеста" и "Ростелекома". — НП "Телеком Форум", сентябрь 2009.

6. **Брусиловский С.А.** Гармонизация инновационного развития ЕСЭ РФ. — Труды 8-ой Международной конференции "Инфокоммуникационные технологии и услуги. NGN в России. Контент", www.rezonance.com.

7. **Гургенидзе А.** Новый виток в развитии NGN-сетей России. - Connect! Мир Связи, №5, 2009.

8. **Денисьева О.М., Коротовских А.В.** Особенности построения и эксплуатации телефонных станций на базе NGN-решений. — Вестник связи, №10, 2009.

9. **Непомнящий А.** WiMAX-2009: год "взрывного роста". — Вестник связи, №12, 2009.

Анализ результатов расчета операторской системы

Традиционные Call -центры ориентированы на прием и обработку вызовов, поступающих от телефонной сети общего пользования. Операторские системы контакт-центров являются мощным инструментом по обслуживанию поступающего трафика. Для расчета операторских систем, ориентированных на прием и обработку вызовов телефонии, использованыметоды теории телетрафика.

Попова А.Г.,
кандидат технических наук, профессор кафедры АЭС МТУСИ,

Чан Туан Минь,
аспирант МТУСИ

Формирование операторских групп в рамках контакт-центров выгодно отличается от существующих вариантов организации структуры call-центров:

- возможна реализация выбора свободного оператора внутри группы в соответствии со специально задаваемыми критериями, что обеспечивает более равномерную загрузку операторов и позволяет сократить число супервизоров;
- сокращаются общие арендуемые площади;
- может быть использована современная система контроля за параметрами и показателями работы центра (в том числе, с использованием подсистемы обработки данных статистики и наблюдения за работой операторов).

Рассмотрим вариант перехода к обобщенной операторской системе контакт-центра от существующей информационной структуры, которая представлена совокупностью небольших групп операторов, включенных в системы коммутации на правах групп серийного включения. Предположим, что всего имеются три группы емкостью $V_1 = 3$, $V_2 = 7$, $V_3 = 11$ операторов.

Предположим, что в операторской системе сохраняется дисциплина обслуживания абонентов с отказами. Поэтому в качестве базовой расчетной формулы вероятности потери вызовов используем первую формулу Эрланга. Важной характеристикой функционирования операторской системы является ее пропускная способность, то есть интенсивность нагрузки, пропущенной с заданным качеством. Зафиксируем допустимую вероятность потерь по вызовам в операторской системе на уровне Ротк доп =10%. Распределим Ротк доп между первой и второй фазами обслуживания:

- в пучке каналов доступа, принимая в нашем случае число каналов на этом участке Vдоступа = 30 каналов;
- в конкретной группе операторов, на которую направляется вызов.

При проведении расчетов интенсивность нагрузки на входе Yвх изменялась в диапазоне Yвх = 7…29 Эрл. Пропускную способность контакт-центра найдем, проведя расчет зависимости вероятности потери по вызовам на уровне доступа $Ротк_1$ и вероятности потери по вызовам в i-ой группе операторов $Ротк_{2,i}$ от интенсивности нагрузки на входе контакт — центра Yвх. Используем первую формулу Эрланга для расчета уровня доступа $Ротк_1 = Ev_{доступа}(Yвх)$.

Предположим, что система управления будет распределять пропущенную нагрузку Yпроп = Yвх(1- $Ротк_1$) между группами пропорционально числу операторов

$$Yi = Yпроп(Vi/Vобщее). \quad (1)$$

Используем первую формулу Эрланга для расчета операторской системы, определяя величину $Ротк_{2,i} = Evi(Yi)$. Найдем вероятность потери вызова, направленного на i — ую группу операторов как

$$Ротк\ общ\ i = 1 - (1 - Ротк_1)(1 - Ротк_{2,i}). \quad (2)$$

Общая вероятность потерь определялась как

$$Ротк = Роткобщ_1(V_1/Vобщее) + Роткобщ_2(V_2/Vобщее) + Роткобщ_3(V_3/Vобщее). \quad (3)$$

Результаты расчета представлены на рис.1-2. Наибольшее влияние на величину Ротк оказывает первая, самая малочисленная группа операторов. Требуемая величина Ротк < Ротк доп достигается при значениях Yвх < 13,7 Эрланг.

Положительными сторонами рассмотренного варианта являются: сохранение принципов взаимодействия с пользователями (территориальное закрепление пользователей за определенной группой операторов центра); сохранение сложившейся структуры трафика. Отрицательными сторонами рассмотренного варианта являются: возможность отказов на уровне доступа, которое может быть устранено выбором емкости пучка доступа, заведомо превышающей потребности; использование дисциплины обслуживания с отказами, хотя имеется принципиальная возможность перехода к обслуживанию с ожиданием. При использовании доступа по IP-сети необходимо решать вопрос о выборе скорости IP-канала.

Создавая собственный операторский центр контакт-центра следует предварительно просчитывать потенциальные доходы от его эксплуатации. Вполне вероятно, что можно обойтись меньшими затратами, оставив качество обслуживания абонентов на должном уровне. Например, нецелесообразно направлять все входящие в центр вызовы на автоматизированную систему, в то время как большинство операторов свободны от обслуживания вызовов и потенциально готовы к обслуживанию клиентов. Использование очередей ожидания, автоматизированных подсказок и многоуровневых меню повышает производительность работы операторского центра, но лучше при любой возможности подключать к разговору с клиентами живых операторов, а автоматизированные системы будут на 100% использоваться в пиковые часы.

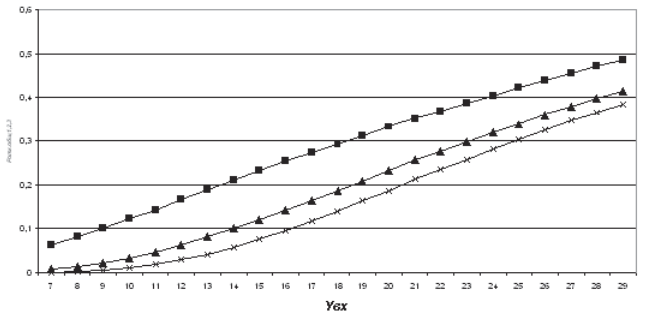


Рис.1. Зависимость вероятности потери вызовов в группе операторов контакт-центра от интенсивности входящей нагрузки

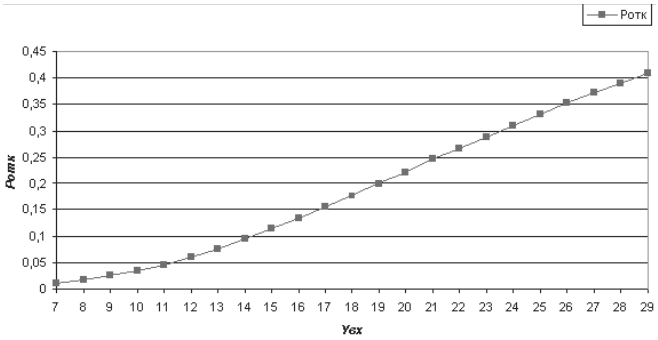


Рис. 2. Зависимость общей вероятности потери вызовов в контакт-центре от интенсивности входящей нагрузки

Рассмотрим принцип построения и возможности системы интерактивного голосового меню IVR на примере системы коммутации каналов MEDIO фирмы StromTelecom сети фиксированной связи. При помощи системы IVR абонент может:

- самостоятельно выбирать и получать различную информацию, например, о задолженности по абонентской плате или по остатку на счете (номер телефона может определяться автоматически, или вводиться абонента в тоновом режиме);
- активизировать роуминг и речевую почту;
- подключаться к внешним базам данных;
- использовать дополнительные сервисы (например, обработка карт "экспресс оплаты", определение собственного номера, автодозвон на занятого абонента).

Оборудование рассматриваемой системы IVR рассчитано на непрерывную и круглосуточную работу без постоянного присутствия персонала технического обслуживания и обеспечивает следующие показатели производительности и качества работы:

- средняя нагрузка на соединительную линию составляет 0,8 Эрл;
- производительность системы составляет до 700000 вызовов в час наибольшей нагрузки;
- вероятность потери из-за блокировок соединительных путей не превышает 0,001;
- потери при обслуживании вызовов не должны превышать 0,002;
- средняя наработка на отказ аппаратных средств хранения данных IVR составляет не менее 10000 часов;
- среднее время восстановления работоспособного состояния оборудования IVR не более 15 минут.

Рассматриваемая система IVR содержит:

- от 8 до 200 интерфейсов E1 с сигнализацией взаимодействия ОКС №7, а также ISUP-R, INAP-R, CAMEL;
- от 250 до 6000 внутренних каналов к модулям детекторов тонального дополнительного набора и генераторов голосовых сообщений.

Произведем расчет для двух вариантов работы информационно-справочной системы (иначе — Call-центра): без использования интерактивного голосового меню (IVR); с использованием IVR.

Рассмотрим вариант обслуживания по системе с отказами без использования IVR. Предварительно заданы такие параметры:

V — число каналов доступа к Call-центру, определяемое в данном конкретном случае как V = n x 30, где n — число потоков E1 на входе Call-центра, предположим, что n = 4;

Mmax — максимальное число операторов информационного системы, определяемое техническими возможностями выбранного оборудования по организации рабочих мест.

Будем характеризовать качество обслуживания пользователей величиной допустимой вероятности потерь по вызовам P, которая состоит из двух составляющих $P = 1 - (1 - P_{вход})(1 - P_{опер})$, где $P_{вход}$ — допустимые потери по вызовам из-за занятости всех каналов на входе центра; $P_{опер}$ — допустимые потери по вызовам из-за отсутствия в момент поступления вызова свободного и доступного оператора.

Если величины $P_{вход}$ и $P_{опер}$ будут составлять несколько процентов, то можно использовать равенство $P = P_{вход} + P_{опер}$. Рассмотрим три варианта построения информационно-справочной системы (Call-центра), которые должны удовлетворять заданным условиям:

1) P = 1% $P_{вход}$ = 0.5% $P_{опер.}$ = 0.5%;
2) P = 5% $P_{вход}$ = 0.5% $P_{опер.}$ = 4.5%;
3) P = 10% $P_{вход}$ = 0.5% $P_{опер.}$ = 9.5%.

Для расчетов используем первую формулу Эрланга:

$$Y_{пост.\,доп.} = Ev(P_{вход}), \quad (4)$$

где $Y_{пост.\,доп.}$ — допустимая поступающая нагрузка на входе системы в эрлангах.

Допустимую нагрузку, поступающую к операторам информационно-справочной системы, рассчитаем по следующей формуле:

$$Y_{call\text{-}центр\,доп.} = Y_{пост.\,доп} \times (1 - P_{вход}). \quad (5)$$

При расчетах предположим, что все операторы объединены в одну группу. Тогда необходимое число рабочих мест операторов можно определить по первой формуле Эрланга как:

$$V_{опер.} = E(Y_{call\text{-}центр\,доп.}; P_{опер.}) \quad (6)$$

Рассчитаем число операторов для каждого из заданных вариантов.

Для первого случая имеем $P_{опер.}$ = 0,5%; $Y_{пост.\,доп.} = Ev(P_{вход})$ = 99,39 Эрл; $Y_{call\text{-}центр\ доп.} = Y_{пост.\ доп.} * (1 - P_{вход})$ = 44,5 Эрл; $V_{опер.} = E(Y_{call\text{-}центр\,доп.}; P_{опер.})$ = 121 операторов.

Для второго случая имеем $P_{опер.}$ = 4,5%; $Y_{пост.\,доп.} = Ev(P_{доп.})$ = 99,39 Эрл; $Y_{call\text{-}центр\,доп.}$ = 44,5 Эрл; $V_{опер.} = E(Y_{call\text{-}центр\,доп.}; P_{опер.})$ = 106 операторов.

Для третьего случая имеем $P_{опер.}$ = 9,5%; $Y_{пост.\,доп.} = Ev(P_{доп.})$ = 99,39 Эрл; $Y_{call\text{-}центр\,доп.}$ = 44,5 Эрл; $V_{опер.} = E(Y_{call\text{-}центр\,доп.}; P_{опер.})$ = 96 операторов.

Из результатов вычислений видно, что наименее затратным является третий вариант построения информационно-справочной системы (Call-центра) на 96 операторов.

Проведем расчет для варианта обслуживания с ожиданием (использование IVR), предполагая, что имеем одинаковую систему доступа по четырем потокам E1. Можно предположить, что имеет место равенство

$$W_{опер.} + W_{ivr} = 1, \quad (7)$$

где $W_{опер.}$ — доля вызовов, направляемых на облуживание к операторам системы; W_{ivr} — доля вызовов, направляемых в IVR.

Интенсивность нагрузки на операторов с учетом того, что часть вызовов будет направляться на IVR, определим, как

$$Y^*_{call\text{-}центр\,доп.} = Y_{call\text{-}центр\,доп.} \times W_{опер.} \quad (8)$$

Соотношение между $W_{опер.}$ и W_{ivr} может меняться в широких пределах. Общее число одновременно поступающих в информационно-справочную систему вызовов ограничено величиной V = n x 30 = 4 x 30 = 120 вызовов.

Для того, чтобы можно было оценить степень влияния перехода на обслуживание с ожиданием на экономичность построения ин-

формационно-справочной системы, будем исходить из следующих требований $P_{вход} = 0,5\%$; $P_{опер} = 9,5\%$. Указанные значения $P_{вход}$ и $P_{опер}$ можно считать достаточно мягкими.

Предположим, что имеет место соотношение $W_{опер.} = 0,1$ и $W_{ivr} = 0,9$. Тогда получим $Y^*_{call\text{-}центр\ доп} = 9,94$ эрланг и $V_{доп.опер.IVR} = 13$ операторов.

Предположим, что имеет место соотношение $W_{опер.} = 0,5$ и $W_{ivr} = 0,5$. Тогда получим $Y^*_{call\text{-}центр\ доп} = 49,5$ эрланг и $V_{доп.опер.IVR} = 55$ операторов.

Предположим, что имеет место соотношение $W_{опер.} = 0,7$ и $W_{ivr} = 0,3$. Тогда получим $Y^*_{call\text{-}центр\ доп} = 69,3$ эрланг и $V_{доп.опер.IVR} = 79$ операторов.

Таким образом, проведенный расчет позволяет сделать предварительный вывод о том, что переход к обслуживанию с ожиданием позволяет значительно сократить необходимое число рабочих мест операторов. А именно в нашем случае при $W_{опер.} = 0,1$ и $W_{ivr} = 0,9$ — в 7,3 раза и при $W_{опер.} = 0,7$ и $W_{ivr} = 0.3$ — на 20%.

Литература

1. **Смит Ш.** Контакт-центры: виртуальная реальность// Технологии и средства связи. — 2004. — №6. — С22-26.

2. **Масленников И.О., Лыкошин А.С.** Тенденции развития рынка решений для организации центров обслуживания вызовов и контакт-центров// Технологии и средства связи, Каталог 2005. — С.76-83.

3. **Брэд Кливленд.** 12 особенностей работы Call-центров, о которых должно знать руководство// Сети и системы связи, 2007. — №4. — С.52-54.

4. Техническая документация по оборудованию MEDIO. — М., 2006. — 250 с.

Направления развития методов описания и построения корпоративных сетей связи

Рассматривается оценка влияние фрактальных свойств трафика на среднее время пребывания требований в системе. Приведены аналитические решения для определения среднего времени пребывания требования в системе, полученные Норосом для систем массового обслуживания вида fBM/D/1, M/M/1, M/D/1. Зависимости даны в системе от параметра Херста.

Амарян С.Р., Попова А.Г., Степанова И.В.,
МТУСИ

Анализ тенденций изменения характера трафика

Современные корпоративные сети связи, в общем случае, представляют собой сложные, многоуровневые и многовариантные структуры, объединяющие в себе сервисы пакетной телефонии, передачи данных, видеоконференцсвязи. Для каждого уровня построения корпоративной сети связи актуально использование специальных методов исследования, позволяющих производить сравнение и обоснованный выбор вариантов организации связи. Характерные особенности современных информационных систем — клиент-серверная архитектура, тяготение к централизации на уровне центров обработки данных, географическая распределенность клиентских компонент. По мере интеграции корпоративных информационных систем сети с пакетной коммутацией приобретают роль транспортной инфраструктуры для различных приложений.

Организация доступа корпоративного клиента к ресурсам сети предполагает выбор варианта интеграции потоков трафика разной природы (речь, данные, видео) и регулирования их интенсивности. В зависимости от используемой технологии интеграция потоков трафика может выполняться в специальном коммутаторе доступа или внутри конвергентной УПАТС, имеющей распределенную структуру и содержащей различные среды передачи [1].

Для сетей с коммутацией пакетов характерны проявления самоподобного (фрактального) характера трафика передачи данных, подразумевающие повторяемость распределения нагрузки во времени в различных масштабах, которую называют масштабной инвариантностью. Под самоподобием принято понимать комплекс статистических свойств трафика, включающий масштабную инвариантность по времени, долговременную зависимость (медленно убывающую корреляцию), "весомые" распределения с высокой вероятностью пиковых значений, степенные законы убывания статистик второго порядка.

В общем случае случайный процесс $x(t)$ считается самоподобным, если для любого действительного a ($a > 0$) справедливо следующее: математическое ожидание $M[x(t)] = a^{-H} M[x(at)]$; дисперсия $D[x(t)] = a^{-2H} D[x(at)]$.

Фрактальные свойства, существующие и в традиционных сетях общего пользования, особенно ярко проявляются в современных корпоративных вычислительных сетях. На такое поведение трафика оказывает влияние как характер источников нагрузки (приложений), так и особенности сетевых протоколов.

Основанием для изучения самоподобных свойств потоков данных в сетях передачи информации послужили работы [2,3]. Так, объектом исследования в работе [4] является сеть Ethernet с предоставлением услуг B-ISDN. Результатами работы явились следующие выводы:

- наличие самоподобия в сети приводит к увеличению пачечности трафика и не соответствует модели Пуассона;
- обычно теоретические представления модели трафика в локальной сети не учитывают свойство самоподобия;
- поступление пакетов на обслуживание показывает очень большой разброс (дисперсию) между моментами поступления (синдром бесконечной дисперсии);
- параметр Херста позволяет удовлетворительно оценить уровень самоподобия обслуживаемых потоков;
- нет стандартных формализованных моделей для оценки влияния самоподобия на качественные характеристики системы связи.

Как показано в работах Leland W.E., Неймана В.И., Ромашковой О.Н., негативное влияние самоподобного трафика трудно обнаружить с помощью обычных средств контроля трафика и методов оценки его параметров. Оно начинает сказываться, если доля трафика передачи данных близка к половине общей интенсивности трафика. Влияние трафика передачи данных в обычных телефонных сетях не приводит к существенному ухудшению качества связи, поскольку доля примеси обычно не достигает указанных значений. Наличие автокорреляции во времени у сетевого трафика негативно влияет на пропускную способность сетей связи и эффективность ее использования.

Использование классических моделей теории телетрафика признается целесообразным в тех случаях, когда интенсивность трафика данных существенно меньше трафика телефонии. Кроме того, допустимо использование моделей теории телетрафика при проведении сравнительных оценок различных вариантов организации телекоммуникационных систем при условии последующей проверки адекватности классических моделей функционированию конкретной системы [5, 6, 7, 8].

Оценка влияния фрактальных свойств трафика на среднее время пребывания требований в системе

Обозначим через величину λ_t интенсивность поступления требований в систему массового обслуживания с ожиданием в промежутке времени (0, t), то есть $\lambda(t) = \frac{\alpha(t)}{t}$, где $\alpha(t)$ — функция времени, отображающая число входящих требований.

Определим величину T_t как время, проведенное одним требованием в системе, усредненное по всем требованиям, находящимся в системе в течение времени (0, t). Величину T_t определяет соотношение $T_t = \frac{\gamma(t)}{\alpha(t)}$, где $\gamma(t)$ — общее время в системе по всем требованиям вплоть до момента t.

Определим величину $\overline{N_t}$ как среднее число требований в системе в промежутке времени (0, t). Эта величина может быть определена как отношение $\gamma(t)$ к длине t промежутка времени наблюдения $\overline{N_t} = \frac{\gamma(t)}{t}$. Используя обозначения для $\gamma(t)$ и для T_t можем записать, что $\overline{N_t} = \lambda_t T_t$.

Предположим, что рассматриваемая система массового обслуживания такова, что при $t \to \infty$ существуют следующие пределы

$\lambda = \lim_{t\to\infty} \lambda_t$ и $T = \lim_{t\to\infty} T_t$, где λ — интенсивность поступления требований в систему и T — среднее время пребывания требования в системе. Если указанные выше два предела существуют, то существует также предел $\overline{N} = \lim_{t\to\infty} \overline{N}_t$, который определяет среднее число требований в системе и называется формулой Дж. Д. К. Литтла:

$$\overline{N} = \lambda T. \quad (1)$$

Таким образом, среднее число требований в системе равно произведению интенсивности поступления требований в систему на среднее время пребывания требования в системе. Соотношение (1) справедливо для любых систем массового обслуживания с ожиданием и не зависит от частных ограничений, накладываемых на вид распределения входящего потока или на вид распределения времени обслуживания. Представим среднее время пребывания требования в системе в нормированном виде как

$$\mu T = \frac{\lambda\mu}{\lambda} T = \frac{\lambda}{\rho} T = \frac{\overline{N}}{\rho}, \quad (2)$$

где $\rho = \lambda / \mu$ — интенсивность нагрузки на систему.

Классическими моделями информационных потоков, используемыми в теории телетрафика, являются следующие модели: *M* — простейший поток вызовов; E_r — поток Эрланга порядка *r*; Г — гамма распределение (иначе — модель потока Эрланга для дробных r). Для представления свойства самоподобных потоков используются модели со следующими распределениями [9, 10]: логарифмически — нормальное; Вейбулла (*W*); Парето (*P*).

Предположим, входной поток системы представляет собой фрактальное броуновское движение. Можно сослаться на результаты исследований Нороса для системы массового обслуживания вида fBM/D/1, которая характеризуется наличием одного сервера с детерминированным временем обслуживания $\chi = 1/\mu$. Аналитические решения для определения среднего времени пребывания требования в системе вида fBM/D/1 были получены Норосом. Среднее время пребывания требования в системе вида fBM/D/1, будучи представлено в нормированном виде, выглядит как [6]:

$$\mu T = \frac{\rho^{(H-1/2)/(1-H)}}{(1-\rho)^{H/(1-H)}}, \quad (3)$$

где *H* — параметр самоподобия Херста.

Для сравнения представим результаты оценки μT для системы массового обслуживания M/M/1 с ожиданием, простейшим потоком требований на обслуживание и экспоненциальным временем обслуживания [10]:

$$\mu T = \frac{1}{(1-\rho)}. \quad (4)$$

Аналогичные результаты для оценки μT для системы массового обслуживания M/D/1 с ожиданием, простейшим потоком требований на обслуживание и постоянным временем обслуживания приведены в [2]:

$$\mu T = \frac{1}{(1-\rho)}\left(1-\frac{\rho}{2}\right). \quad (5)$$

В таблице и на рис. 1 представлены результаты расчета μT по формулам (2) - (4) в зависимости от величины ρ.

Расчеты проводились для фиксированных значений параметра самоподобия Херста.

1. Значение H = 0,724 характеризует локальную вычислительную сеть на 25 станций, для реализации которой использовался канал технологии FastEthernet 100Мбит/с.

Результаты расчетов среднего времени пребывания требования в системе μT в зависимости от величины ρ для различных систем массового обслуживания

Величина $\rho = \lambda / \mu$	Результаты расчета μT				
	Система массового обслуживания вида fBM/D/1			Система массового обслуживания вида M/M/1, H = 0,5	Система массового обслуживания вида M/D/1
	H = 0,724	H = 0,65	H = 0,55		
0,3	0,959	1,158	1,353	1,429	1,214
0,4	1,815	1,744	1,686	1,667	1,333
0,5	3,510	2,692	2,159	2,000	1,500
0,6	7,309	4,406	2,895	2,500	1,750
0,7	17,615	8,032	4,186	3,300	2,166
0,8	56,869	18,061	6,975	5,000	3,000

2. Значение H = 0,65 зафиксировано для локальной сети со смешанным трафиком.

3. Значение H = 0,55 зарегистрировано для телефонной сети связи для часа наибольшей нагрузки.

Зависимости на рис. 1 подтверждают выводы, полученные в работах, о том, что для фрактальных процессов имеет место ухудшение характеристик качества обслуживания по сравнению с любыми другими распределениями потока на входе системы.

Проблема состоит в том, что стремление повысить величину ρ до значений ρ = 0,8,....0,9 приводит, с одной стороны, к резкому росту среднего времени пребывания требований в системе, а с другой стороны, влечет за собой невозможность использования формул Эрланга, описывающих СМО вида M/M/1.

Ранее считалось, что система M/M/1 дает самую завышенную оценку для среднего времени пребывания требования в системе. Но анализ зависимостей рис. 1 показывает, что при использовании системы fBM/D/1 увеличение значения H в диапазоне от 0,55 до 0,724 приводит к существенному росту μT в зависимости от ρ в диапазоне значений $\rho \geq 0,5$.

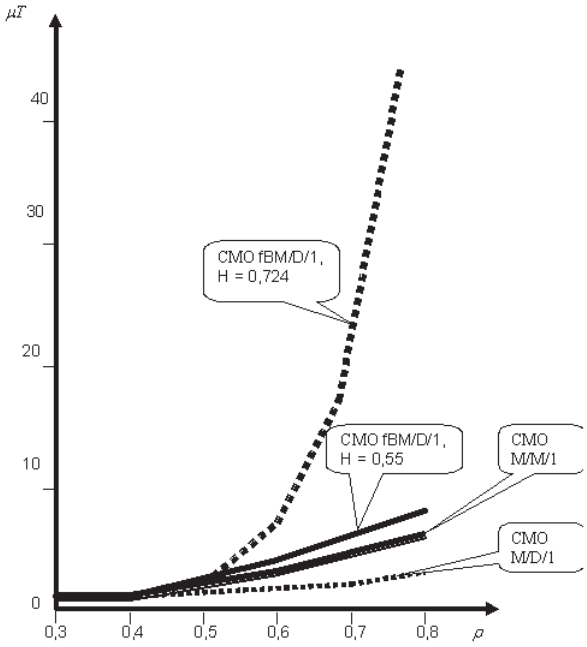


Рис. 1. Зависимость среднего времени пребывания требования в системе от интенсивности нагрузки

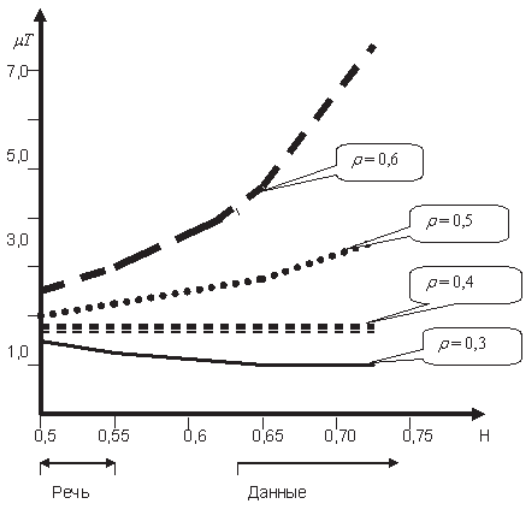


Рис. 2. Зависимость среднего времени пребывания требования в системе от параметра самоподобия Херста

На рис. 2 представлены зависимости среднего времени пребывания требования в системе μT от параметра самоподобия Херста H при фиксированных величинах ρ.

Анализируя полученные зависимости можно сделать такие выводы:

• если $\rho \leq 0{,}3$, для описания функционирования системы можно использовать ее представление в виде СМО M/M/1, учитывая, что такое представление дает завышенную оценку;

• если $0{,}3 < \rho \leq 0{,}4$, для описания функционирования системы также следует использовать ее представление в виде СМО M/M/1;

• начиная со значений $\rho = 0{,}5$ и выше, следует рассматривать функционирование системы как СМО BM/D/1, поскольку возникает и с ростом усугубляется расхождение между результатами оценки (например, при H = 0,724 имеет место резкий рост среднего времени пребывания требований в системе).

По результатам проведенных расчетов можно сделать общий вывод — интенсивность нагрузки на однолинейную систему передачи информации ρ не должна превышать 0,4. Это обеспечит защиту от негативного влияния фрактальных свойств потоков трафика.

Выводы

1. По мере развития корпоративных информационных систем сети пакетной коммутации приобретают роль транспортной инфраструктуры для различных приложений. Взаимодействие распределенных корпоративных структур может быть реализовано с использованием ресурсов глобальной сети или путем формирования виртуальных частных сетей на базе транспортных структур технологии IP/MPLS. Как при подключении пользователей корпоративных информационных систем, так и при формировании структур VPN следует ограничивать величину ρ значениями $\rho \leq 0{,}4$, что создает условия для использования методов и формул теории телетрафика при сравнении вариантов построения корпоративной сети связи.

2. Характерные особенности современных информационных систем — клиент-серверные архитектуры, тяготение к централизации на уровне центров обработки данных, географическая распределенность клиентских компонент — существенно влияют на характер сетевого трафика. Существенное влияние на качество обслуживания потоков информации оказывает самоподобие трафика, заключающееся в том, что пакеты данных, при высокой скорости их движения по сети, поступают на узел пачкой, что приводит к потерям из-за ограниченности буфера.

3.Мультиплексированный трафик от множества независимых источников сходится к α — устойчивым распределениям, которые удовлетворяют обобщенной центральной предельной теореме. Процедура ограничения скорости в шлюзах глобальной сети приводит к тому, что любой исходный α — устойчивый трафик в "ограниченном" канале сводится к Гауссовскому распределению ($\alpha = 2$), сохраняя при этом фрактальные свойства. Значение параметра Херста H не изменяется.

4. Расчет характеристик работы шлюзов в зависимости от пропускной способности глобальной сети показал, что использование нормального распределения дает излишне оптимистическую оценку. Перспективно использование для расчетов Гауссовского распределения. Задача проектирования VPN включает в себя совокупность задач рационального распределения ресурсов сети для совместной реализации различных VPN и маршрутизации в сети трафика каждой VPN.

5. Анализ тенденций развития корпоративных информационно-справочных структур показал перспективность перехода от локальных структур к территориально — распределенным контакт-центрам. Выбор системы массового обслуживания для описания функционирования корпоративной информационно-справочной структуры определяется дисциплиной обслуживания вызовов, принятым вариантом автоматического распределения вызовов между операторами, используемым набором регламентов и правил по взаимоотношению с клиентами.

Литература

1. **Попова А.Г., Амарян С.Р.** Методы децентрализации услуг в сетях VPN// Международный форум информатизации (МФИ-2007): Труды конференции "Телекоммуникационные и вычислительные системы". — М.: МТУСИ, 2007. — С.116.

2. **Leland W.E., Taqqu M.S., Willinger W.I. Wilson D.V.** On the Self-Similar Nature of Ethernet Traffic//Proceedings ACM SIGCOMM'93. — San Fransisco, CA, 1993. — P. 183-193.

3. **A.Feldmann.** Characteristics of TCP connection arrivals//Technical report, AT&T Labs Research, 1998.

4. **Нейман В.И., Ромашкова О.Н.** Два типа сетей связи // Автоматика, связь, информатика, 2002. — № 1. — С. 28-32.

5. **Треногин Н.Г., Соколов Д.Е.** Модели трафика корпоративных сетей на основе -устойчивых фрактальных процессов// Красноярск. Вестник НИИ СУВПТ, Сборник научных трудов, 2003. — вып.14.

6. **В.В. Крылов, С.С. Самохвалова.** Теория телетрафика и ее приложения. — Санкт-Петербург, 2005. — 288 с.

7. **Д. Комер.** Принципы функционирования Интернета. — СПб.Ж Питер, 2002. — 384 с.

8. **Ивченко Г.И., Каштанов В.А., Коваленко И.Н.** Теория массового обслуживания. — ГЛ.: Высшая школа, 1982. — 256 с.

9. **Саати Т.Л.** Элементы теории массового обслуживания и ее приложения. — М.: Советское радио, 1971. — 515 с.

10. **Джонатан Девидсон, Джеймс Питерс, Манож Бхатия, Сатиш Калидинди, Судипто Мукхержи.** Основы передачи голосовых данных по сетям IP. — Москва, Санкт-Петербург, Киев: Cisco Press. — 2007, 400 с.

Проектирование и внедрение информационных центров

Предоставление пользователям информационных услуг является одной из основных и широко используемых задач телекоммуникационной сферы. Для оказания подобных услуг используются операторские информационные центры. Процесс проектирования и внедрения современных центров обслуживания вызовов предусматривает решение ряда важных научно-практических задач. В современных контакт-центрах присутствует Web-сервер — базовый функциональный элемент, существенно отличающийся по всем ха-рактеристикам от систем взаимодействия пользователь-оператор.

Степанова И.В., Покровская М.А.,
МТУСИ

Информационная услуга — это любая услуга, предусматривающая прием информации от клиента, обработку и, если требуется, дальнейшее продвижение этой информации, передачу клиенту ответной информации, а также распределение заранее подготовленной в центре информации по списку, то есть передачу ее всем клиентам, занесенным в этот список. В операторских центрах используются информационные ресурсы, хранящиеся в соответствующих базах данных, обрабатывается и запоминается поступающая информация, автоматически протоколируется вся деятельность операторов, связанная с обслуживанием вызовов, а также выполняется ряд других функций. Главной задачей, которую решают операторские центры, является обеспечение ответов на возможно большее количество входящих вызовов и/или создание и завершение возможно большего количества исходящих вызовов, ибо каждый вызов, требующий обработки в информационном центре, либо является потенциальным источником дохода, либо несет в себе важную информацию. Потеря вызова нежелательна, а иногда и просто недопустима.

Процесс проектирования и внедрения современных центров обслуживания вызовов предусматривает выполнение следующих основных задач:

- проработка основных требований к центру обслуживания вызовов и разработка технического задания на проектирование;
- разработка технико-экономического обоснования проекта или технико-коммерческих предложений;
- разработка технического проекта центра обслуживания вызовов;
- установка технических средств центра;
- установка системного программного обеспечения центра;
- установка прикладного программного обеспечения центра;
- ввод и конвертация баз данных центра обслуживания вызовов;
- адаптация системы под конкретные особенности заказчика;
- обучение персонала центра обслуживания вызовов;
- опытная эксплуатация центра обслуживания вызовов;
- коммерческая эксплуатация центра.

Каждая из перечисленных задач содержит один или более этапов. Выполнение некоторых этапов взаимосвязано, в то время как другие могут осуществляться параллельно.

На этапе проектирования центра обслуживания вызовов необходимо определить его основные технические параметры и показатели качества функционирования, а также экономические показатели.

Существуют два основных параметра, которые необходимо определять при проектировании любых операторских центров обслуживания вызовов независимо от приложений: число операторов, которое требуется для качественного обслуживания вызовов;

- число телефонных линий и пропускная способность каналов операторского центра.

В качестве базовых показателей качества работы центров обслуживания вызовов в международной практике используются.

1. Среднее время ожидания (Average Speed of Answer, или ASA) — показывает, сколько в среднем абоненту приходится ждать в очереди до соединения с оператором. Показатель ASA пересчитывается, как правило, каждые полчаса.

2. Процент вызовов, прерванных абонентами во время ожидания (Abandon Rate) — процент абонентов, которые повесили трубку, не дождавшись соединения с оператором. Нормальным уровнем считается 2-3%. Например, операторский центр газеты USA Today в городе Арлингтон (США) обеспечивает среднее время ожидания 12 с, при числе прерванных вызовов, не превышающем 2%. Если показатель прерванных вызовов превышает 5%, надо принимать определенные меры, поскольку теряются вызовы, а значит, клиенты и продажи.

3. Уровень сервиса (Service Level) — процент вызовов, для которых время ожидания превысило заданного количества секунд. В международной практике существует стандарт, по которому не менее 80% поступающих вызовов должны быть соединены с опера-тором в течение 20 с.

Экономические показатели центров обслуживания вызовов включают капиталь-ные затраты, необходимые для строительства центра, эксплуатационные расходы, по-лучаемые доходы и параметры окупаемости вложенных инвестиций.

Главная особенность контакт-центра по сравнению с предшествующими системами — это способность обслуживать запросы разных типов, поступающие из разных телекоммуникационных сетей:

- запросы речевой связи из телефонной сети общего пользования (ТфОП);
- запросы речевой связи из Интернет, с использование технологии IP-телефонии;
- запросы связи по факсу, электронной почте;
- запросы связи в режиме текстового чата.

Необходимо отметить, что организация очередей и механизмы маршрутизации вызовов в контакт — центрах, интегрированных с Интернет, сложнее, чем в операторских центрах ТфОП. Это связано не только с совершенствованием алгоритмов распределения вызовов, но и с тем, что вызовы поступают от источников разного типа.

В случае контакт-центра запросы поступают с разной интенсивностью от источ-ников разного типа, допускают разную длительность ожидания, разную продолжительность и законы обслуживания, то есть различаются параметрами, которые определяют характеристики входящей нагрузки и на основании которых обычно производится распределение вызовов и организация очередей. Так, например, пороги длительности ожидания для заявок, допускающих отложенное обслуживание, могут измеряться десятками минут, а для традиционных телефонных вызовов — десятками секунд. Фактически, очередь ожидания превращается в буфер, выбор заявок из которого производится не в порядке их поступления, а на основе анализа нескольких параметров, характеризующих эти заявки.

Механизмы обслуживания разных заявок могут быть различными. Их могут обслуживать либо отдельные операторы или группы,

либо одни и же операторы. При этом запросы, допускающие отложенную обработку, операторы могут обрабатывать в периоды, когда интенсивность потока речевых вызовов снижается.

В контакт-центре присутствует базовый функциональный элемент, резко отличающийся по всем характеристикам от систем взаимодействия пользователь-оператор. Это — Web-сервер, исследование которого должно проводиться отдельно от остальных подсистем.

Характеристики трафика, поступающего от источников заявок, могут меняться во времени. В зависимости от типов сетей и вида предоставляемых контакт-центром услуг возникают периодические возрастания и спады интенсивностей поступающих запросов. Исследование контакт-центров требует учета поступающего трафика. В рассматриваемой работе раздельного внимания требуют операторская и Web подсистема.

В общем случае Операторская подсистема осуществляет обслуживание запросов на информационные услуги приходящие:

• в речевом виде от ТфОП, сетей подвижной связи (СПС), сетей iP-телефонии;

• в текстовом виде от систем интерактивного (диалогового) обмена сообщениями се-ти Интернет и СПС;

• в текстовом виде от систем обмена сообщениями позволяющих;

• отложенную обработку ТфОП, СПС, сетей IP-телефонии и Интернет (факсимильные сообщения и электронная почта). Следовательно, интерфейсы операторской подсистемы могут быть разделены на три группы: речевой, текстовый, диалоговый и текстовый с отложенной обработкой. Потоки запросов на информационные услуги, приходящие на каждую из групп интерфейсов могут отличаться своими характеристиками.

Выделим основные в контексте данной работы особенности процессов поступления и обработки потоков запросов по указанным группам интерфейсов. Для речевой группы интерфейсов особенности поступающих потоков известны ещеиз базовых работ по теории телетрафика, рассматривавших процессы, происходящие в ТфОП.

Установлено, что потоки речевых вызовов, приходящие на операторскую подсистему от большого числа источников, имеют показательное распределение интервалов времени между поступающими запросами и аналогичное распределение времени обслуживания. Данный факт имеет экспериментальное подтверждение во множестве классических работ по обслуживанию телефонной нагрузки [1,2].

Кроме речевых интерфейсов, операторская подсистема содержит, также два типа текстовых. Технической реализацией текстового диалогового интерфейса и текстового интерфейса с отложенной обработкой, например, могут быть системы интерактивного обмена текстовыми сообщениями Web-chat и электронная почта Интернет. Зависимости для процессов обслуживания запросов, поступающих через текстовые интерфейсы оператор-ской подсистемы контакт-центра, могут заметно отличаться от привычных зависимостей, используемых для расчетов телефонной нагрузки. Они могут являться медленно-затухающими распределениями [3, 4], иначе называемыми распределениями с "тяжелым хвостом" (heavy-tailed). В первую очередь, это проявляется у рассматриваемых процессов больших значений дисперсии.

Обратим внимание на Web подсистему контакт-центра, отметим особенности процессов, происходящих в ней.

Запросы на предоставление информационных услуг передаются пользователем на сервер web через браузер — клиентское программное обеспечение (ПО), являющееся терминалом WWW. Рассматриваемая web-подсистема включает интерфейс только для одно-го типа запросов.

Данным запросом является посылаемый браузером идентификатор (URI — Uniform Resource Identifier), определяющий нахождение определенного документа в сети в целом и его расположение непосредственно на сервере Web. Под документом будем понимать любой файл или динамически генерируемую информацию, которые могут быть запрошены браузером в процессе предоставления пользователю услуги. Пример URI — www.protei.ru/teor/indext.html, документом в данном случае является файл indext.html, интерпретируемый пользовательским терминалом, как HTML-страница,

Потоком запросов на web-подсистему контакт-центра является поток HTTP-запросов, содержащих идентификаторы запрашиваемых документов.

Предполагается, что время обслуживания запроса складывается из двух составляющих: во-первых, времени обработки принятого URI и поиска файла либо динамической генерации ответа, содержащего нужную информацию на сервере Web; во-вторых, времени передачи по исходящему каналу Web-IPCC.

В зависимости от области применения сервера web определяющей может оказать-ся первая, вторая, либо обе составляющие. Для web-сервера в составе контакт-центра, можно предположить, что соотношение времени подготовки ответа и его передачи позво-ляет пренебречь первым, по сравнению со временем передачи гипертекстовых страниц и различных файлов большого размера. Это очевидно для web-серверов, не осуществляющих динамическую подготовку ответов из баз данных большого объема. Следовательно, при определении характеристик web-подсистемы, предлагается ориентироваться только на время передачи документа в исходящий канал web. Это время зависит непосредственно от пропускной способности канала и объема передаваемых данных.

Исходя из известного среднего времени пребывания HTTP — запроса в web-подсистеме, может быть найдена средняя скорость передачи данных в ответ на один запрос. Данный параметр представляется более наглядно отражающим качество предоставления информационной услуги пользователю. Также, на основе социологических опросов, и исследований, для него могут быть определены минимальные значения, после которых качество предоставления информационной услуги может считаться неудовлетворительным.

Средняя скорость передачи данных в ответ на один запрос может быть получена исходя из следующего выражения:

$$S = F_s / T,$$

где S — средняя скорость передачи данных в ответ на один запрос, F_s — средний размер документов подсистемы web, T — среднее время пребывания запроса в системе для web-подсистемы контакт-центра.

В зависимости от полученной средней скорости передачи данных пользователям и возмож-ностей их терминалов, может быть принято решение об увеличении или уменьшении пропускной способности канала web-подсистемы. Далее выделим особенности процессов поступления и обработки потоков запросов для web-подсистемы исследуемого контакт-центра. Для оп-ределения процессов обслуживания запросов подсистемой web необходимо рассматри-вать медленно-затухающие распределения. Это является следствием соответствующих распределений размеров документов на Web-серверах [5].

Литература

1. **Гольдштейн Б.С., Фрейнкман В.А.** Call-центры и компьютерная телефония. СПб: БХВ, 2002.

2. **Лившиц Б. С., Пшеничников А. П., Харкевич А.Д.** Теория телетрафика. — М.: Связь, 1979.

3. **Цитович И.И.** Устойчивые модели трафика мультисервисных сетей // Труды Рос-сийского научно-технического общества радиотехники, электроники и связи имени А.С.Попова. Серия: научная сессия, посвященная Дню радио. Выпуск: LX-2. Москва, 2005. — Т2. — С. 271-273.

4. **Leland W.E., Taqqu M.S., Willinger W., and Wilson D.V.** On the self-similar nature of Ethernet traffic // IEEE/ACM Transactions of Networking, 2(1), 1994.

5. **Гольдштейн Б.С., Зарубин А.А.** Контакт-центры мультисервисных сетей связи. — Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2005.

Анализ подходов к расчету операторской системы контакт-центра

Одной из проблем расчета системы операторов, предоставляющих справочную информацию клиентам, является сложность прогнозирования интенсивности поступления вызовов, которая подвержена значительному изменению в течение суток и в течение часа наибольшей нагрузки (ЧНН). Наличие разнообразной статистической информации, доступной в режиме реального времени, позволяет оперативно управлять обслуживанием вызовов, гибко перераспределяя поток вызовов на различные группы. С помощью резервных групп операторов можно гарантировать приемлемое качество обслуживания вызовов при появлении повышенного трафика.

Попова А.Г.,
Чан Туан Минь,
МТУСИ

Современные контакт-центры обеспечивают возможность гибкой, интеллектуальной маршрутизации вызовов на основе следующих параметров:

- установленное максимальное время ожидания абонента;
- среднее время ожидания в очереди;
- расчетное время ожидания в очереди;
- средняя продолжительность ожидания ответа;
- среднее время разговора;
- количество вызовов в очереди;
- время суток, день недели;
- количество абонентов в очереди;
- наличие и количество свободных операторов;
- команды, переданные пользователем контакт-центру;
- номер вызывающего абонента, набранный абонентом номер.

Алгоритмы гибкой маршрутизации могут включать в себя все многообразие математических и логических операций, использование переменных, массивов, условных и безусловных переходов, циклов, обращение к базам данных.

Программа, реализующая алгоритмы маршрутизации, как правило, не имеет ограничений по длине и количеству переменных. Интерактивное информирование звонящих о времени ожидания в очереди позволяет удержать больше вызовов, в то время когда все операторы заняты. По статистике, клиенты ожидают в 2 раза дольше, если получают такую информацию. Например, абонент может сам оценить предполагаемое время ожидания, и если оно слишком велико для него, оставить свой номер телефона, для последующего контакта с оператором. Повысить пропускную способность системы операторов в часы наибольшей нагрузки позволяет увеличение количества работающих операторов. На рис.1 представлен вариант реализации интеллектуальной маршрутизации вызовов на примере частично распределенного контакт-центра. Предполагается, что контакт - центр должен обеспечивать обслуживание вызовов в стационарном режиме. При возникновении вероятности перегрузки часть вызовов будет маршрутизирована на рабочие места операторов, ранее выполнявших другие задачи. Это могут быть операторы различных групп поддержки, входящие в состав данного контакт-центра, или операторы на значительном удалении от контакт-центра.

Вызовы

Контакт-центр приобретает при этом вид многоуровневой структуры. Изменение алгоритма маршрутизации вызовов может производиться в реальном масштабе времени, без помех и ограничений на работу контакт - центра.

Наличие разнообразной статистической информации, доступной в режиме реального времени, позволяет проводить оперативное управление, обслуживанием вызовов, гибко перераспределяя поток вызовов на различные группы, а грамотная настройка контакт-центра позволит избежать вмешательства при возникновении разовых нагрузок. Например, с помощью резервных групп можно сбалансировать загрузку операторов при появлении аврального трафика, тем самым гарантировать качество обслуживания вызовов.

В таблицах 1 и 2 и на рис.2 представлены результаты оценки показателей качества обслуживания вызовов, полученные для дисциплины обслуживания с ожиданием. При проведении расчетов использовался принцип декомпозиции, введение которого в данном случае допустимо благодаря тому, что потери на входе центра исключаются значительным превышением числа каналов доступа над числом рабочих мест операторов. В качестве критерия перехода к распределенной структуре центра при обслуживании с ожиданием использовалось расчетное время ожидания в очереди. Предполагалось, что при достижении равенства $t_{расч} = t_{max}$, где $t_{расч}$ — расчетное время ожидания в очереди и t_{max} — максимальное установленное время ожидания абонента, к основной группе операторов контакт - центра подключаются операторские группы поддержки. Затем при достижении равенства $t_{расч} = t^*_{max}$, где t^*_{max} — максимальное установленное время ожидания абонента, при достижении которого подключаются удаленные подразделения. Расчет проводился по второй формуле Эрланга. Возможность использования этой формулы обусловлена тем, что современные системы коммутации и современные контакт - центры позволяют обеспечить очень длинные очереди с числом мест ожидания, которое значительно превышает потребности. Среднее время разговора принималось равным $t_{разговора} = 40$ секунд.

Предполагалось, что основная группа операторов контакт - центра имеет емкость $V_{операторы} = 12$ операторов. За счет маршрутизации избыточных вызовов на вспомогательную операторскую группу общее число работающих операторов может быть увеличено до $V_{конт.-центр\ операторы} = 16$ операторов. Маршрутизация на удаленные

Рис.1. Вариант организации многоуровневой системы операторов

Таблица 1

Результаты расчета вероятности задержки начала обслуживания вызова более P(>τ) в многоуровневой системе операторов контакт-центра с гибкой маршрутизацией

Интенсивность нагрузки Y, Эрл	$V_{операторы}$ = = 12 операторов	$V^{конт.-центр}_{операторы}$= =16 операторов	$V^{сумма}_{операторы}$ = =20 операторов
	P(>τ) при τ = 1, τ = $t_{ожидания}$ / $t_{разговора}$		
10	0,060818		
11(подключение вспомогательной группы)	0,255272	0,000772	
12		0,003747	
13		0,016642	
14		0,068816	
15(подключение удаленной группы)		0,26858	0,001081
16			0,004691
17			0,019171
18			0,074538
19			0,277897

Таблица 2

Результаты расчета вероятности задержки начала обслуживания вызова более P(>τ) в многоуровневой системе операторов контакт-центра с гибкой маршрутизацией

Интенсивность нагрузки Y, Эрл	$V_{операторы}$ = 12 операторов	$V^{конт.-центр}_{операторы}$= =16 операторов	$V^{сумма}_{операторы}$ = =20 операторов
	P(>τ) при τ = 2, τ = $t_{ожидания}$ / $t_{разговора}$		
10	0,008231		
11 (подключение вспомогательной группы)	0, 093909	0,000005	
12		0,000069	
13		0,000829	
14		0,009313	
15(подключение удаленной группы)		0,098805	0,000007
16			0,000086
17			0,000954
18			0,010088
19			0,102232

подразделения позволяет довести общую численность операторов до $V^{сумма}_{операторы}$ = 20 операторов. Предполагалось также, что число каналов на уровне доступа равно V = 90, что соответствует реализации доступа по 3 потокам E1. Если исходить из требования обеспечения потерь по вызовам на стыке телефонная сеть — контакт-центр не выше 0,002, то это требование будет выполняться при Y < 68,5 Эрл (расчет проводился по первой формуле Эрланга). То есть расчеты проводились для значений Y, которые существенно меньше этой величины. То есть на уровне доступа практически нет потерь, поэтому предполагалось наличие равенства Yпроп = Y.

Проанализируем зависимости, представленные на рис. 2. На них явно выражены моменты подключения дополнительных групп операторов, в которые происходит резкое улучшение характеристик работы центра.

Уменьшение величины P(>τ) на несколько порядков в данном примере обусловлено подключением групп, состоящих из четырех операторов. Если подключаться будут один-два оператора, то изменение характеристик будем менее выражено. Кроме того, изменение требований по продолжительности времени ожидания в очереди от τ = 1 (время пребывания в очереди равно времени обслуживания, в данном случае это — 40 секунд) до τ = 2 (время пребывания в очереди вдвое больше времени обслуживания, то есть составляет 80 секунд), вызывает значительное снижение величины P(>τ) при прочих равных условиях.

Вспомогательная и удаленная группы операторов могут отличаться друг от друга разной дисциплиной выбора операторов для обслуживания поступившего вызова. Вспомогательная группа операторов располагается в помещении контакт-центра, и при ее подключении может быть реализован псевдослучайный режим выбора свободного оператора. То есть речь идет о численном расширении операторской системы без изменения ее характеристик. В случае подключения удаленной группы малой емкости, скорее всего, будет задействовано серийное включение линий с упорядоченным занятием.

Проблема состоит в том, что организация и подключение удаленных групп операторов зависит от возможностей и потребностей той организации, для нужд которой используется данный контакт-центр. В частности, удаленные группы операторов могут формироваться с привлечением ресурсов различных филиалов.

Таким образом, при регулировании работы многоуровневой системы операторов предлагается использовать следующую методику.

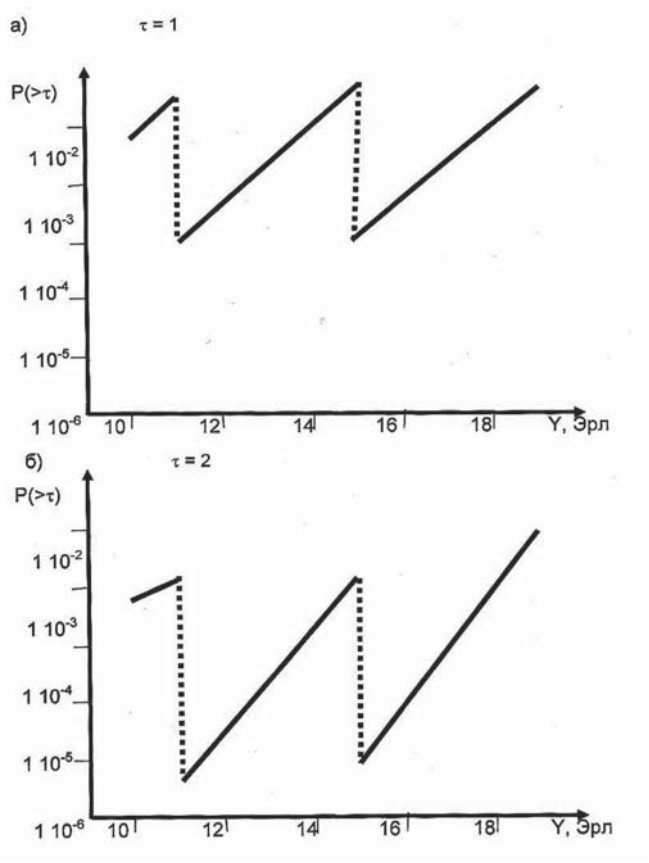


Рис. 2. Зависимость P(>τ) от интенсивности нагрузки на входе контакт-центра Y с учетом дополнительных групп операторов

1. Оценить по первой формуле Эрланга допустимую интенсивность нагрузки на входе контакт-центра $Y_{доп\ вход}$, исходя из допустимой вероятности потери по вызовам на стыке телефонная сеть — контакт-центр равной 0,002. При проведении оценки $Y_{доп\ вход}$ можно устанавливать и изменять число каналов $V_{доступа}$, с учетом реально имеющегося или предполагаемого числа потоков E1. Если доступ обеспечивается с использованием IP-технологии, то величину $Y_{доп\ вход}$ можно ориентировочно определить как

$$Y_{доп\ вход} = 0,9\ C_{стыка} / C_{соединения},$$

где 0,9 — использование линии доступа; $C_{стыка}$ — скорость линии доступа на стыке телефонная сеть — контакт-центр; $C_{соединения}$ — скорость одного соединения, которая зависит от выбранной технологии и средств связи, которыми пользуются абонент и оператор.

2. Для области значений интенсивности нагрузки, поступающей на вход контакт-центра $Y < Y_{доп\ вход}$, по второй формуле Эрланга произвести расчет зависимости P(>τ), определяя величину τ в зависимости от предполагаемого времени ожидания и времени обслуживания.

3. Оценить принципиальную возможность формирования дополнительных групп операторов, число рабочих мест операторов в этих группах и возможность (или не возможность) сохранения режима псевдослучайного выбора операторов.

4. Произвести анализ работы действующего контакт - центра с целью определения тенденций изменения $t_{расч}$ — расчетного времени ожидания в очереди, которое определяется системой управления контакт-центра на основании данных о длительности ожидания в очереди для трех последних вызовов, распределенных из очереди на операторов, и текущем количестве работающих операторов. Выбрать значения t_{max} и t^*_{max}, при достижении которых должно производит подключение дополнительных групп операторов. Выбор величин t_{max} и t^*_{max} определяется рядом субъективных факторов и зависит от тех задач, которые выполняет контакт-центр. Можно рекомендовать соблюдение равенства $t_{max} = t^*_{max}$.

5. По второй формуле Эрланга произвести расчет зависимости P(>τ) от интенсивности нагрузки Y с учетом подключения дополнительных групп операторов.

Литература

1. **Домбрин И.М., Колодников В.А.** Call-центры для российских железных дорог. Применение решений "Протон-ССС" для интеграции информационных систем ОАО "РЖД" с центрами обслуживания вызовов// Технологии и средства связи. — 2004. — №5. — С.38-39.

2 **Ефимов А.Р.** Три модели аутсорсинга для организации контакт-центра// Технологии и средства связи. — 2005. — №2. — С.84-86.

3. **Зарубин А.** Распределенный контакт-центр// CONNECT. — 2007. — №10. — С. 117-122.

4. **Грег Левин.** Измеряем качество работы контакт-центра// Сети и системы связи. — 2007. — №9. — С. 20-26.

5. **Самолюбов А.Б.** Аутсорсинговые центры обслуживания вызовов// Технологии и средства связи. — 2005. — №3 . — С. 36-38.

Мониторинг и контроль call-центров экстренных служб

Справочные центры экстренных служб (СЦ ЭС) ориентированы на прием и обработку вызовов от абонентов телефонной сети общего пользования. Систематизируем возможности эффективного мониторинга и контроля СЦ ЭС. К мониторингу относится часть функций управления по непрерывному наблюдению и анализу состояния оборудования, а также в сборе данных статистики с оценкой динамики изменений. Соответственно, контроль — это часть управления по однократному запросу состояния оборудования или значений различных величин.

Степанова И.В.,
Киселев И.В.,
МТУСИ

Одним из приоритетных направлений развития инфраструктуры связи в Российской Федерации является создание единой системы вызовов экстренных служб по телефонному номеру "112" на всей территории России. Экстренная служба (ЭС) — это служба, к которой абонент обращается за помощью в экстремальных ситуациях, требующих немедленного вмешательства (угроза жизни и здоровью человека, пожар, стихийные бедствия, противоправные действия, терроризм). Аппаратно-программные комплексы оборудования для приема и обслуживания с помощью операторов потока экстренных вызовов называются справочными центрами экстренных служб (СЦ ЭС).

Помимо стандартных задач, которые решают обычные Call — центры, таких как обеспечение максимального качества обслуживания потока входящих вызовов, минимизация потерь и обеспечение высокой эффективности работы операторов, комплексы обработки вызовов СЦ ЭС отличаются гибкой системой управления и мощной интеллектуальной системой анализа оперативной и статистической информации, накапливаемой в процессе функционирования системы.

В современных СЦ ЭС все функциональные возможности реализуются компьютерными серверами приложений, работающими с управляющей информацией и взаимодействующими в процессе обслуживания вызова с информационными и технологическими базами данных. Каждое из таких приложений отвечает за свой набор услуг (например, приложение автоматизированной среды распределения вызовов ACD и интерактивное голосовое меню IVR). Компьютерные серверы приложений мониторинга и контроля подключаются с помощью оборудования ввода — вывода к учрежденческим автоматическим телефонным станциям (УПАТС), интерфейс которых адаптируется к ее программному обеспечению (рис.1). Рассмотрим возможности средств мониторинга и контроля Call — центра экстренных служб на базе оборудования MD 110 ACD.

Средства мониторинга и контроля СЦ ЭС на оборудовании MD 110 ACD включают в себя (рис.2):

- средства мониторинга и контроля работы операторов службы;
- средства мониторинга и контроля поступающего трафика;
- средства мониторинга и контроля состояния оборудования MD 110 ACD;
- средства мониторинга и контроля состояния линейного тракта между узлом специальных служб электронного типа (УССЭ) и СЦ ЭС.

Средства мониторинга и контроля СЦ ЭС позволяют эффективно управлять и правильно организовывать работу операторов службы. Особая важность их функционирования — сбор информации во время чрезвычайных ситуаций (ЧС). Анализ статистических данных по трафику в период ЧС позволяет объективно оценивать работу операторов службы, определять недостатки в работе оборудования СЦ ЭС и находить оптимальные пути решения по уменьшению потерь вызов. Средства мониторинга и контроля работы операторов службы включают в себя: сервер регистрации переговоров операторов службы, позволяющий записывать разговоры операторов с абонентами на основании задаваемых параметров в режиме ре-

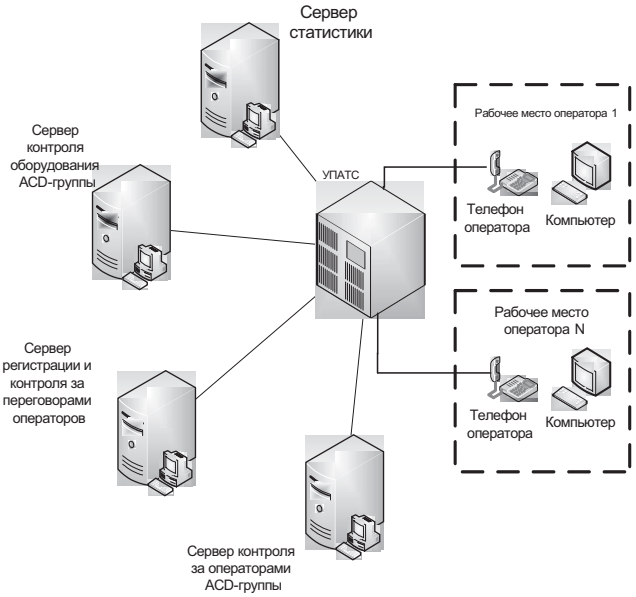


Рис. 1. Распределенная система средств мониторинга и контроля Call-центра экстренных служб

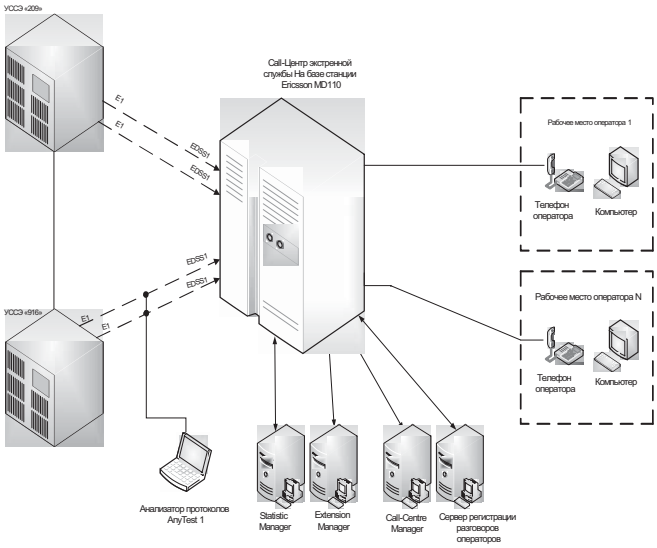


Рис. 2. Схема подключения средств мониторинга и контроля Call-центра экстренных служб на базе оборудования MD 110 ACD

ального времени; Call Center Manager (CCM) — программно-аппаратный комплекс, который используется для:

• задания пользовательской конфигурации, просмотра информации о текущей и прошедшей работе центра обслуживания вызовов, а также информации о рабочем персонале — операторах;

• просмотра информации в реальном масштабе времени, о работе группы АРВ, групп операторов и отдельных операторов;

• создания отчетов о выполненной работе;

• создания текстов и сообщений в реальном масштабе времени для вывода на внешние настенные дисплеи.

Основным средством мониторинга и контроля состояния оборудования MD 110 ACD является приложение Extension manager (EMG) — прикладная программа, упрощающая работу обслуживающего персонала по вводу и изменению данных об абонентах MD110. Приложение EMG поддерживает работу с аналоговыми и цифровыми номерами, номерами оборудования для передачи данных, беспроводных абонентов и абонентские номера ISDN, а также производит автоматическое наблюдение по свободным и используемым абонентским номерам, по информации о группах и услугах.

Программно-аппаратные средства EMG работают в режиме реального времени и обеспечивают следующие функции.

1. На абонентской линии (АЛ):

• ввод в эксплуатацию и выключение АЛ;

• распределение и изменение категорий обслуживания абонентов;

• изменение номеров телефонных аппаратов без переключения АЛ;

• запрос и выдача информации о состоянии, о категории АЛ или группы АЛ, о заблокированных линиях.

2. На соединительной линии (СЛ):

• изменения числа пучков линий в направлениях;

• создание новых пучков (направлений);

• изменение маршрутов пропускания нагрузки.

3.Управление нагрузкой с целью:

• указания о количестве включенных и работающих СЛ;

• вывода данных о нагрузке на отдельные направления связи, пучки, управляющие устройства.

4. Управление системой с целью:

• загрузки новых пакетов программ;

• запуска оборудования ввода-вывода.

С помощью директив технического обслуживания (ТО), используемых в EMG, можно в режиме диалога вывести на монитор следующие сообщения:

• сообщение о состоянии системы;

• сообщение об отказах и локализации неисправностей;

• функциональные проверки после ремонта;

• запуск периодических контрольных операций;

• автоматический перезапуск;

• сообщение о неисправностях различных категорий срочности.

В процессе ежедневной эксплуатации MD 110 ACD могут произойти какие-либо нарушения. Это могут быть ошибки, вносимые обслуживающим персоналом, либо отказы программного или аппаратного обеспечения. Кроме того, может возникнуть необходимость в изменении конфигурации системы в уже действующей станции. Станция располагает некоторым числом встроенных функций, облегчающих техническое обслуживание и выявление проблем, которые могут возникнуть в процессе эксплуатации.

К таким функциям относятся:

• функции наблюдения;

• функции и алгоритмы обнаружения неисправностей;

• уведомления с помощью аварийных сигналов, передаваемых на консоль оператора, или на панель аварийной сигнализации, о наличии проблем;

• анализ аварийных сигналов на EMG с помощью указания на дату и время, код отказа, блок с неисправностью, источник аварийного сигнала, состояние устройства, блока или соединения.

Рассмотрим в качестве примера технического обслуживания команды, связанные с аварийной сигнализацией — AL (ALLIP, ALLOP).

Информация обо всех ошибках хранится в регистраторе аварийных сигналов. Каждому типу ошибки присвоен свой собственный уникальный код ошибки. Всего существует пять классов, причем нулевой класс относится только к тем авариям, которые система устранила самостоятельно без вмешательства обслуживающего персонала. Можно определить число аварийных сигналов, которые могут регистрироваться и накапливаться для конкретного класса аварий до того, как система выдаст аварийный сигнал более высокого класса, уведомляющий о превышении установленного порога значения.

Для мониторинга и контроля состояния линейного тракта между УССЭ и СЦ ЭС используется система тестирования Linkbit AnyTest — многоцелевой инструмент тестирования систем связи, основанных на протоколах линий E1 (T1/J1) и IP. Система тестирования Linkbit AnyTest состоит из следующих компонентов: карта PCMCIA AnyTest (рис.3); проводник AnyTest (рис. 4).

Карта AnyTest (PCMCIA) имеет: два независимых E1 (T1/J1) приемника/передатчика, маркированных Port 1 и Port 2; процессор цифровых сигналов (DSP) и встроенный микропроцессор. Имеется возможность проводить мониторинг и анализировать два двунаправленных T1/E1 потока одновременно. Предусмотрены два режима тестирования — режим монитора и режим терминала.

Рис. 3. Карта PCMCIA AnyTest

Рис. 4. Проводник AnyTest

В первом случае осуществляется мониторинг линии, соединяющей два устройства. Входное сопротивление устанавливается высоким, а передатчик отключается для того, чтобы устранить возмож-

Рис. 5. Подсоединение к тестируемой линии в режиме монитора

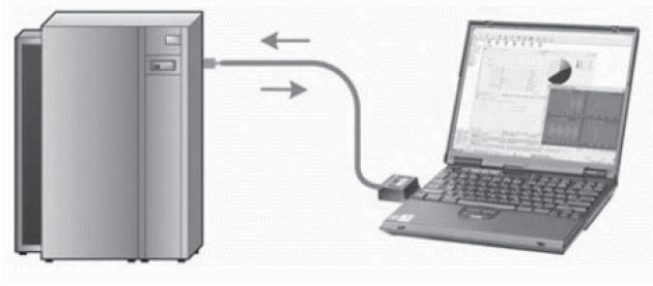


Рис 6. Использование устройства AnyTest в режиме терминала

ность наведения помех в линии.

В памяти EEPROM карты хранится лицензионная информация о активированных опциях программы AnyTest. На карте выполняются задачи обработки в реальном времени, а задачи связанные с IP, пользовательский интерфейс и другие компоненты, не требующие режима реального времени, передаются компьютеру. Проводник AnyTest обеспечивает уникальный графический пользовательский интерфейс (GUI). Он превращает сложный, чреватый ошибками процесс настройки тестов в простой и быстрый выбор нескольких параметров с помощью мыши. Результаты тестов собираются, фильтруются, упорядочиваются, анализируются компьютером и отображаются на его экране. Данные теста можно сохранить на жестком диске компьютера или передать в общую базу данных через сетевое соединение. Программа генерации отчетов позволяет быстро сформировать отчет о тесте в том виде, который нужен пользователю.

Подсоединение к тестируемой линии в режиме монитора показано на рис. 5. Такое подсоединение наиболее часто используется для мониторинга и контроля состояния линейного тракта между УССЭ и СЦ ЭС. В режиме терминала AnyTest эмулирует работу устройства (станции), подсоединяется к другому устройству напрямую и генерирует тестовый поток телефонных вызовов (рис.6).

Интерфейсы карты AnyTest настраиваются в диалоговом окне Настройки Линии. Интерфейсы настраиваются независимо друг от друга, можно делать их одинаковыми или разными. Разделение на сетевую и пользовательскую стороны в настройках интерфейсов позволяет передавать сообщения непосредственно с одного интерфейса на другой. Для этого нужно соединить Port 1 и Port 2 реверсивным кабелем.

Контроль линейного тракта между УССЭ и СЦ ЭС обычно проводится в режиме монитора (Any Test подключается параллельно к потоку E1) при подключении новых операторов связи к Call-центру экстренной службы для контроля параметров канальных настроек абонентской сети оператора.

Между УССЭ и СЦ ЭС используется протокол EDSS1 — Европейская Цифровая Система Сигнализации №1. Эта сигнализация, состоящая из трех уровней, часто применяется на участке абонентского доступа. В информации по результатам изменений, представленной в виде так называемых "трэйсов" контролируются следующие параметры 3-го уровня:

Bearer capability — идентифицирует возможность передачи данных. В обязательном порядке присутствует в разделе SETUP на стыках абонента А и абонента В, содержит информацию для процедуры проверки совместимости;

Channel identification — указывает пользовательский канал, по которому в рамках установленного соединения может передаваться речевой сигнал;

Progress indicator — может присутствовать в различных сообщениях, информируя о событии, которое может косвенно влиять на процесс установления соединения (например, доступность тонового сигнала в канале В), либо определить степень ожиданий абонента (например, не полностью цифровое соединение);

Calling party number — служит для идентификации источника вызова. (задается вызывающим абонентом или инициирующей станцией);

Called party number — адрес вызываемого абонента, задаваемый абонентом А и используемый абонентом В при проверке совместимости;

Calling party subaddres — передача данных о местоположении (географические координаты для абонентов сотовых сетей связи) вызывающего абонента.

Перечислим основные параметры, которые контролируются в процессе снятия "трейсов": категория и номер вызывающего абонента; номер вызываемого абонента; вид передаваемой информации; номер канала в потоке, по которому пришел вызов; тип передаваемого номера; разрешение передачи номера вызывающего абонента; определение местоположения абонента.

В заключение следует отметить, что перспективы развития средств мониторинга и контроля СЦ ЭС будут определяться развитием IP-телефонии и сетей подвижной связи, которые позволяют использовать не только голосовые, но и мультимедийные способы взаимодействия — передачу изображений, видео, SMS-сообщений, MMS-сообщений. Их использование позволит увеличить эффективность взаимодействия граждан с операторами СЦ ЭС при возникновении экстренных ситуаций, но определяет потребность в разработке средств мониторинга и контроля для работы с мультимедийной информацией.

Литература

1. Распоряжение Правительства РФ от 25.08.2008 г. №1240-р. Об утверждении концепции создания системы обеспечения вызова экстренных оперативных служб через единый номер "112".

2. **Лагутин В.С., Попова А.Г., Степанова И.В.** Эволюция интеллектуальных служб в конвергентных сетях. — М., 2009. — 97 с.

3. **Брэд Кливленд.** 12 особенностей работы call — центров, о которых должно знать руководство// Сети и системы связи. — 2007. — №4. — С.52-54.

Результаты разработки специализированного программного обеспечения обработки трафика экстренных служб

Эффективность решения проблем, прямо и косвенно касающихся безопасности граждан, во многом зависит от функциональных возможностей справочных центров экстренных служб (СЦ ЭС). Программное обеспечение для обработки поступающего трафика обычно не поставляется. Это объясняется различными требованиями к обработке статистических данных в Call-центрах. В СЦ ЭС должно использоваться специализированное программное обеспечение.

Степанова И.В.,
Киселев И.В.,
МТУСИ

Обеспечение доступа населения к экстренным службам (ЭС) следует рассматривать как одну из важнейших задач государства. К СЦ ЭС предъявляются особые требования по четкой и бесперебойной работе. Обслуживающий персонал современных СЦ ЭС имеет возможность наблюдения за параметрами поступающего трафика и возможность вывода результатов по отдельным абонентским линиям (АЛ), группам АЛ, группам АЛ заданных категорий источников нагрузки.

Рассмотрим результаты использования специализированного программного обеспечения для обработки поступающего трафика экстренных служб (СППО ЭС), которое было разработано техническими специалистами одного из СЦ ЭС, использующего оборудование MD 110 ACD. Контрольные измерения трафика абонентов сетей операторов фиксированной и мобильной связи выполнялись с помощью анализатора протоколов AnyTest. СППО ЭС позволяет формировать более 300 разновидностей отчетов о работе комплекса. СППО ЭС позволяет создать отчеты в целях определения:

- числа, дня недели, часа и значения наибольшей нагрузки (ЧНН) за любой период времени в течение месяца;
- количества вызовов, обработанных различными группами операторов;
- количества вызовов, не дождавшихся ответа операторов;
- количества транзитных вызовов;
- общего количества входящих вызовов;
- общего количества внутренних вызовов по различным типам соединения;
- количества вызовов по потокам E1 от разных узлов специальных служб;
- количества вызовов по любому из каналов в потоке E1;
- количества вызовов по различным операторам связи;
- количества обслуженных вызовов;
- процента обслуживания;
- среднего времени обслуживания.

На рис.1 показана экранная форма СППО ЭС, которая состоит из командного меню и информационного поля. Рассмотрим пример обработки статистических данных СЦ ЭС с помощью СППО ЭС.

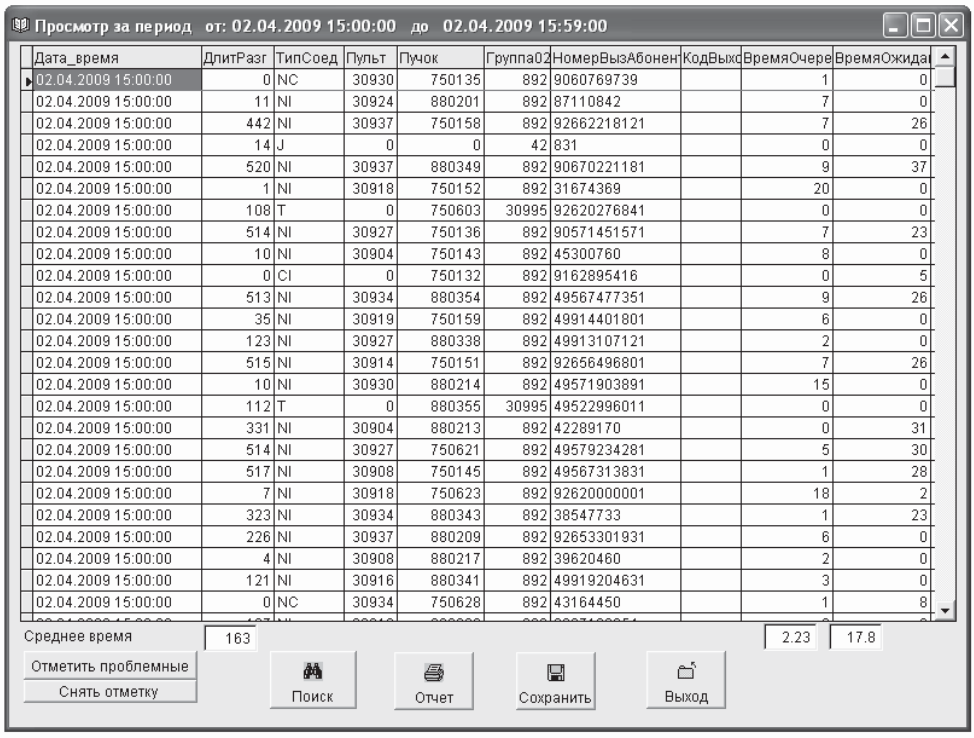
1. Определяем ЧНН — 2 апреля, четверг, 15-00 до 16-00.
2. Общее количество вызовов в ЧНН составило 827.
3. Общее количество обслуженных вызовов в ЧНН — 823.
4. Среднее время обслуживания вызова в ЧНН в секундах T равно 243,2 с.
5. Интенсивность нагрузки Y составила 55,6 эрл.

При помощи средств выборки определялось количество обслуженных вызовов, среднее время обслуживания и значение интенсивности нагрузки в ЧНН от различных сетей операторов связи.

Полученные данные представлены в таблице 1. Существенную долю всех запросов составляет трафик от подвижных пользователей. Суммарная нагрузка от операторов фиксированной связи составила 30,0 Эрл. Суммарная нагрузка от операторов сотовой связи составила 25,2 Эрл.

Для наглядности, по результатам выборок был сделан ряд графических отчетов, некоторые из которых приведены на рис. 2-3.

Анализ полученных данных позволяет производить обоснованный выбор структуры СЦ ЭС и вносить изменения в его структуру. Например, такой показатель, как общее количество не обслуженных вызовов позволяет судить о качестве работы центра в целом. А изучение тенденций изменения ЧНН позволяет организовать труд



Дата_время	ДлитРазг	ТипСоед	Пульт	Пучок	Группа02	НомерВызАбонент	КодВыхо	ВремяОчере	ВремяОжида
02.04.2009 15:00:00	0	NC	30930	750135	892	9060769739		1	0
02.04.2009 15:00:00	11	NI	30924	880201	892	87110842		7	0
02.04.2009 15:00:00	442	NI	30937	750158	892	92662218121		7	26
02.04.2009 15:00:00	14	J	0	0	42	831		0	0
02.04.2009 15:00:00	520	NI	30937	880349	892	90670221181		9	37
02.04.2009 15:00:00	1	NI	30918	750152	892	31674369		20	0
02.04.2009 15:00:00	108	T	0	750603	30995	92620276841		0	0
02.04.2009 15:00:00	514	NI	30927	750136	892	90571451571		7	23
02.04.2009 15:00:00	10	NI	30904	750143	892	45300760		8	0
02.04.2009 15:00:00	0	CI	0	750132	892	9162895416		0	5
02.04.2009 15:00:00	513	NI	30934	880354	892	49567477351		9	26
02.04.2009 15:00:00	35	NI	30919	750159	892	49914401801		6	0
02.04.2009 15:00:00	123	NI	30927	880338	892	49913107121		2	0
02.04.2009 15:00:00	515	NI	30914	750151	892	92656496801		7	26
02.04.2009 15:00:00	10	NI	30930	880214	892	49571903891		15	0
02.04.2009 15:00:00	112	T	0	880355	30995	49522996011		0	0
02.04.2009 15:00:00	331	NI	30904	880213	892	42289170		0	31
02.04.2009 15:00:00	514	NI	30927	750621	892	49579234281		5	30
02.04.2009 15:00:00	517	NI	30908	750145	892	49567313831		1	28
02.04.2009 15:00:00	7	NI	30918	750623	892	92620000001		18	2
02.04.2009 15:00:00	323	NI	30934	880343	892	38547733		1	23
02.04.2009 15:00:00	226	NI	30937	880209	892	92653301931		6	0
02.04.2009 15:00:00	4	NI	30908	880217	892	39620460		2	0
02.04.2009 15:00:00	121	NI	30916	880341	892	49919204631		3	0
02.04.2009 15:00:00	0	NC	30934	750628	892	43164450		1	8

Рис.1. Экранная форма СППО ЭС

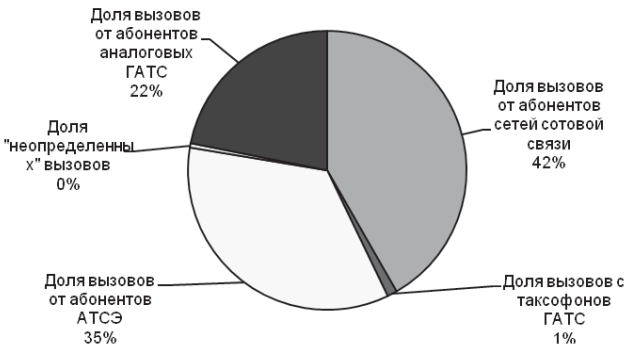


Рис.2. Распределение долей вызовов в ЧНН (02.04.09г.с 15-00 до 16-00)

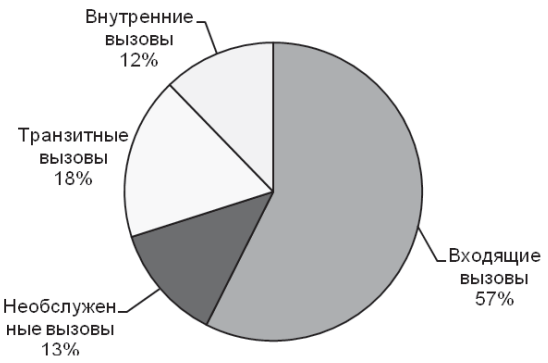


Рис. 3. Виды поступающих на СЦ ЭС вызовов

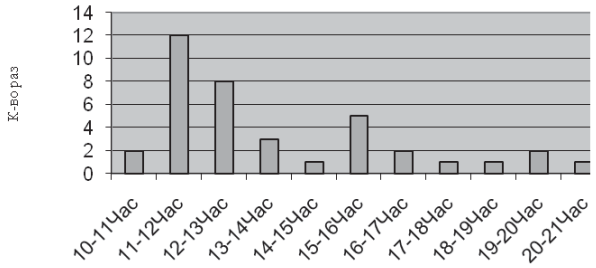


Рис. 4. Распределение ЧНН для экстренных служб за период с 2006 по 2009 гг.

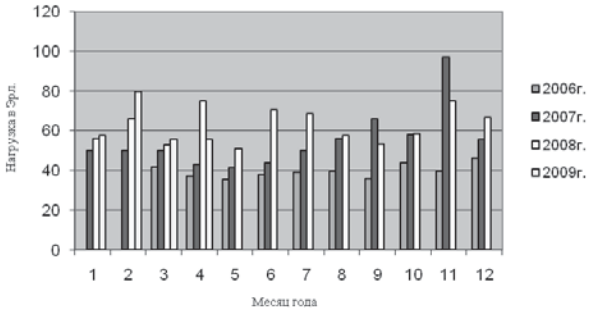


Рис. 5. Распределение интенсивности нагрузки в ЧНН в эрлангах по месяцам года

операторов центра таким образом, чтобы способствовать уменьшению процента отказов.

На рис. 4-5 показаны результаты анализа данных статистики по СЦ ЭС за несколько лет.

Возможность проведения такого анализа обоснована действием нормативными документами Министерства связи и массовых коммуникаций России относительно хранения статистических данных о работе СЦ не менее трех лет. Анализ данных статистики за несколько лет показал, что наибольший объем трафика был отмечен в ноябре 2007 г. Следует также отметить, что в 2008 и 2009 гг. уровень тревожности по сравнению с 2006 годом существенно вырос (ориентировочно — на 20 ... 50%).

Таким образом, система СППО ЭС является мощным инструментом для текущего контроля качества обслуживания, а также средством прогнозирования изменения объемов трафика.

Результаты анализа статистических данных СЦ ЭС

Источники нагрузки	Количество обслуженных вызовов	T, c	Y, Эрл
Аналоговые АТС	172	220,7	10,5
АТСЭ	272	254,7	19,2
Таксофоны	9	79,7	0,2
МТС	34	137,4	1,3
БиЛайн	129	307,6	11,0
Мегафон	166	263,7	12,2
«Прямые» сотовые	1	7	0,002
Сотовый роуминг	19	132,8	0,7
Без SIM карты	17	20,7	0,1
Без номера абонента «А»	4	271,8	0,3

Литература

1. **Дюбанов А.В., Зарубин А.А., Поташов А.И.** Мультисервисные контакт-центры в сетях связи МВД // "Вестник связи". — 2005. — №9.

2. **Лагутин В.С., Попова А.Г., Степанова И.В.** Эволюция интеллектуальных служб в конвергентных сетях. М.: Полиграфический центр МЭИ. — 2008. —120 с.

3. **Лагутин В.С., Попова А.Г., Степанова И.В.** Развитие телекоммуникаций мегаполиса. Монография. — М.: Полиграфический центр МЭИ. — 2008. — 190 с.

4. **Масленников И.О., Лыкошин А.С.** Тенденции развития рынка решений для организации центров обслуживания вызовов и контакт-центров// Каталог технология и средства связи. — 2005. — С.76-83.

К вопросу оценки качества обслуживания в сети NGN

Рассматривается проблема оценки качества обслуживания сети. Показано, что эта задача применительно к сетям NGN в силу большой структурной сложности объекта становится трудно разрешимой, даже с применением известных методов теории телетрафика. Предлагается использовать новый подход к построению моделей функционирования сетей NGN и получения соответствующих качественных и количественных оценок.

Костин А.Н.,
аспирант кафедры АЭС МТУСИ,
kstin@list.ru

Ершова Э.Б.,
профессор кафедры АЭС, к.т.н. МТУСИ

Сложившаяся в настоящее время тенденция перехода к сетям нового поколения определила более широкий круг проблем, связанных с оценкой качества обслуживания пользователей телекоммуникационных сетей и определением объема сетевых ресурсов, необходимых для его поддержания. Это обусловлено не только мультисервисным характером сетей, связанным с многопрофильным трафиком и маркированными потоками требований на обслуживание, но и с появлением новых дисциплин обслуживания, связанных со специфическим сочетанием свойств различных режимов коммутации в рамках технологий предоставления постоянно растущего спектра инфокоммуникационных услуг.

Как известно, задача оценки качества обслуживания в телекоммуникационной сети требует построения математической модели, позволяющей описать соответствующий процесс. Учитывая большую структурную сложность объекта, эта задача, особенно применительно к сетям NGN, становится трудно разрешимой, даже с применением известных методов теории телетрафика.

Исходя из этого, становится очевидным необходимость использования новых подходов к построению моделей функционирования сетей NGN и получения соответствующих качественных и количественных оценок. В настоящее время получает распространение теория Network calculus (NC), рассматривающая вопросы граничных оценок параметров качества обслуживания, [1]. В NC используется аппарат min-plus и max-plus алгебры, так называемой идемпотентной алгебры, которая позволяет свести сложные нелинейные системы к линейным системам. Теория NC имеет два направления — детерминированное и вероятностное. В детерминированной NC входящий поток ограничивается некой детерминированной кривой, называемой кривой входящего потока, а дисциплина обслуживания ограничивается кривой, называемой кривой обслуживания. Таким образом, в основе лежит представление входного потока и дисциплины обслуживания системы в виде функций, ограничивающих нагрузку.

Рассмотрим особенности применения детерминированной теории NC для решения задачи оценки QoS в сетях NGN. В общем случае в детерминированной NC трафик рассматривается как неизвестный, но ограниченный некой кривой, и казалось бы не учитывается его вероятностный характер, что возможно приведет к грубой оценке параметров системы. Однако такое условие является вполне допустимым, если учесть, что в реальных сетях трафик всегда ограничен полосой пропускания канала, трафик также ограничивается, если на сети реализован контроль доступа и происходит сглаживания поступающей нагрузки.

В соответствии с [2] рассмотрим 5 основных свойств математических моделей, которыми они должны обладать, чтобы на их основе получать аналитическое описание системы с приемлемой сложностью:

1. Возможность гарантировать с заданной точностью получение значений параметров системы при выбранных моделях потока и дисциплины обслуживания в единичном элементе сети.
2. Возможность представления выходного потока на основе модели входного потока.
3. Возможность представления потока на выходе группы элементов на основе модели единичного элемента системы.
4. Возможность представления и описания системы при наличие многих потоков на основе модели системы с одним потоком.
5. Возможность представления и описания суперпозиции потоков на основе модели системы с одним потоком.

Основными понятиями в NC являются: кривая входящего потока, кривая обслуживания, бэклог и виртуальная задержка.

Следуя теории NC, для абстрагирования от вероятностного характера трафика воспользуемся кривой входящего потока. Рассмотрим некоторую функцию A(t), которая будет определять объем трафика, поступившего на вход рассматриваемой системы за время t :

$$A(t) = \int_0^t r(t)dt, \qquad (1)$$

где r(t) мгновенное значение трафика. Таким образом, A(t) - это кумулятивная функция входящего потока. По определению, A(t) неубывающая функция. Примем, что значение функции в начальный момент времени равно нулю: A(0)=0.

Кривая входящего потока α(t) ограничивает функцию A(t) и определяет максимально возможный объем трафика, который может поступить в систему за время t:

$$\mathrm{A(t) - A(s)} \le \alpha(\mathrm{t - s})$$

или

$$\mathrm{A(t)} \le \inf\{0 \le \mathrm{s} \le \mathrm{t}:\ \alpha(\mathrm{t - s}) + \mathrm{A(s)}\}. \qquad (2)$$

В качестве функции α(t) может служить, например, функция, соответствующая алгоритму Token Bucket. Для алгоритма Token Bucket кривая входящего потока имеет вид:

$$\alpha(t) = rt + b. \qquad (3)$$

Следуя NC, введем кривую обслуживания β(t). Кривая обслуживания отражает внутренние механизмы работы системы и гарантирует для обслуживания трафика A(t) минимальные ресурсы Примем, что система обладает кривой обслуживания, если трафик на выходе узла D(t) удовлетворяет неравенству:

$$D(\mathrm{t})\text{-}\mathrm{A(s)} \ge \beta(\mathrm{t\text{-}s})$$
$$\Leftrightarrow \qquad (4)$$
$$D(\mathrm{t}) \ge \inf\{\le s \le t : A(s) + \beta(\mathrm{t\text{-}s})\}.$$

Рассмотрим в качестве примера кривую обслуживания:

$$\beta(t) = \begin{cases} 0, t \le T, \\ C(t-T), t > T, \end{cases} \qquad (5)$$

здесь T — задержка, вносимая системой, C- постоянная скорость обслуживания исходящего потока.

На основе функций $\alpha(t)$ и $\beta(t)$ можно определить кривую исходящего потока. Если объем трафика A(t) ограничен кривой $\alpha(t)$, а кривая обслуживания $\beta(t)$, то кривая исходящего потока $\delta(t)$ по аналогии с $\alpha(t)$ ограничена [3]:

$$\delta(t) = \sup\{s \geq 0 : \alpha(t+s) - \beta(s)\}. \quad (6)$$

Как видно, неравенство (6) подтверждает выполнение второго условия — свойство 2.

Проиллюстрируем сказанное примером. Пусть входящий поток ограничен согласно условию (3), система выполняет обслуживание согласно условию (5), тогда согласно (6) на выходе имеем:

$$\delta(t) = \sup\{s \geq 0 : \alpha(t+s) - \beta(s)\} = b + rt + rT.$$

Рассмотрим параметры системы, характеризующие задержку и емкость буфера.

1. Бэклог B(t). Поскольку бэклог физически равен объему трафика, обслуживаемому системой, то при наличии в системе одного обслуживающего устройства емкость буферов будет равна длине очереди. Следовательно, бэклог в момент времени t равен :

$$B(t) = D(t) - A(t).$$

2. Виртуальная задержка V(t). В момент времени t виртуальная задержка определяет ту задержку, которую испытает поступающий трафик в момент t при условии, что в системе к моменту времени t отсутствует необслуженный трафик. Следовательно, виртуальная задержка в момент времени t:

$$V(t) = \inf\{s \geq 0 : A(t) \leq D(t+s)\}.$$

Графически B(t) ограничено максимальным вертикальным отклонением функций $\beta(t)$ и $\alpha(t)$ и в нашем примере имеем: Bmax=b+rT. Vmax — это горизонтальное отклонение функций $\beta(t)$ и $\alpha(t)$ и для нашего случая имеем: T+b/C. На рис.1 отображены соотношения функций $\beta(t)$, $\alpha(t)$,$\delta(t)$ и Vmax,Bmax. Возможность определения граничных параметров системы поясняет свойство I.

Поскольку теория NC использует инструменты идемпотентной алгебры, то в отличии от привычной алгебраической структуры {R,+,x} здесь используются структуры Rmin и Rmax:

$$R\min = \{R \cup +\infty, \oplus, *\}, где \begin{cases} \oplus = \min(x, y), \\ * = x + y; \end{cases}$$

$$R\max = \{R \cup +\infty, \overline{\oplus}, \overline{*}\}, где \begin{cases} \overline{\oplus} = \max(x, y), \\ \overline{*} = x + y. \end{cases}$$

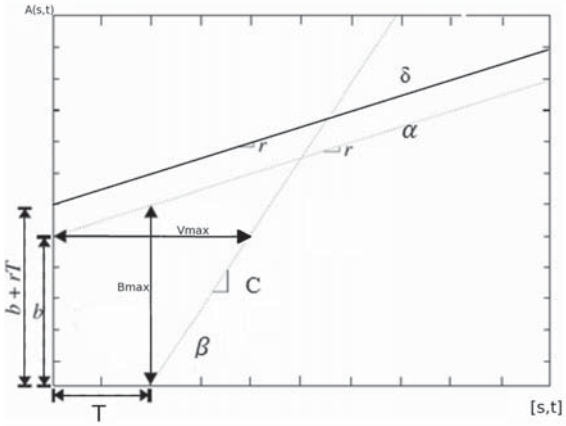


Рис.1. Соотношения между функциями

Таким образом в идемпотентной алгебре max-plus используются две основные операции: новое сложение, определяемое как максимум значений, и новое умножение, которое совпадает с обычным сложением. Аналогично определяются и заменяются операции в алгебре min-plus.

Анализ этих алгебраических структур показывает, что операции в Rmin и Rmax обладают следующими свойствами:

- Сложение $\oplus$ и умножение * ассоциативны: $x \oplus (y \oplus z) = (x \oplus y) \oplus z$
- Сложение $\oplus$ умножение * коммутативно: $x \oplus y = y \oplus x$
- Сложение $\oplus$ идемпотентно: $x \oplus x = x$
- Умножение * относительно сложения дистрибутивно: $x * (y \oplus z) = (x * y) \oplus (x * z)$

1. Операции идемпотентной алгебры иногда [3] интерпретируются в терминах теории фильтров, когда сигнал на выходе системы равен свертке функции входного сигнала и импульсной характеристики. В теории NC входному сигналу сопоставляется кривая входного потока, а импульсной функции — кривая обслуживания.

Свертка в {R,+,x} имеет следующий вид:

$$(f \otimes g)(t) = \int_0^t f(t-s)g(s)ds.$$

Хотя очевидно, что идемпотентным аналогом операции интегрирования является сложение, но доказано это в [5] . Таким образом, свертка функций в Rmax имеет следующий вид:

$$(f \overline{\otimes} g)(t) = \sup(0 \leq s \leq t : f(t-s) + g(s)).$$

Аналогично свертка функций в Rmin имеет следующий вид:

$$(f \otimes g)(t) = \inf(0 \leq s \leq t : f(t-s) + g(s)).$$

Операции, обратные свертке, имеют следующий вид:

$$(f \widehat{\otimes} g)(t) = \sup(0 \leq s \leq t : f(t+s) - g(s));$$

$$(f \widehat{\overline{\otimes}} g)(t) = \inf(0 \leq s \leq t : f(t+s) - g(s)).$$

С учетом приведенных соотношений, перепишем выражения основных параметров.

$A(t) \leq \inf(0 \leq s \leq t : \alpha(t-s) + A(s)) = A \otimes \alpha(t)$ — входящий поток, ограниченный кривой $\alpha(t)$;

$D(t) \geq \inf(0 \leq s \leq t : A(s) + \beta(t-s)) = A \otimes \beta(t)$ — исходящий поток, гарантированно обслуживаемый по закону с функцией $\beta(t)$.

Теперь перейдем от одиночного элемента системы к цепочке элементов сети S, приведенной на рис. 2. Покажем, как выполняется расчет задержки на сети путем последовательного расчета от узла к узлу.

Ради простоты рассмотрения примем, что кривые входных потоков и кривые обслуживания для всех элементов сети одинаковы и равны :

$$\alpha(t) = rt + b \quad и \quad \beta(t) = \begin{cases} 0, t \leq T, \\ C(t-T), t > T. \end{cases}$$

Как выше показано, кривая исходящего потока для этого случая имеет следующий вид:

$$\delta_1(t) \leq b + rt + rT.$$

Задержка V(t)ограничена:

$$V_1(t) \leq V\max = b / C + T.$$

Рис. 2. Фрагмент сети

Кривая $\delta_1(t)$ является кривой входящего потока для второго узла, т.е. $\delta_1 = \alpha_2$, следовательно :

$$\alpha_2(t) \le b + rt + 2rT \quad V_2(t) \le V\max = b/C + 2T.$$

Для k-го узла получаем:

$$\alpha_k(t) \le b + rt + 2krT;$$

$$V_S(t) = \sum_1^k D_k \le (\frac{b}{C} + 1T + \frac{b}{C} + 2T .. + \frac{b}{C} + (k-1)T + \frac{b}{C} + kT) =$$

$$= k \cdot \frac{b}{C} + kT + \frac{k(k-1)}{2}. \qquad (7)$$

Теперь для сравнения приведем расчет задержки другим способом. Согласно выражению (4), поток на выходе первого узла имеет следующий вид:

$$D_1(\mathrm{t}) \ge A \otimes \beta(\mathrm{t}).$$

На выходе второго узла, для которого $D_1(t)$ является входным потоком $A_2(t)$, с учетом свойства ассоциативности имеем:

$$D_2(\mathrm{t}) \ge A_2 \otimes \beta_2(\mathrm{t}) = (\mathrm{A}_1 \otimes \beta_1) \otimes \beta_2 = \mathrm{A}_1 \otimes (\beta_1 \otimes \beta_2).$$

На выходе k-го элемента, получаем :

$$D_k(\mathrm{t}) \ge \mathrm{A}_1 \otimes (\beta_1 \otimes \beta_2 \otimes \beta i ..... \otimes \beta_k) = \mathrm{A}_1 \otimes \beta_S,$$

здесь β_S — кривая обслуживания системы.

Полученный результат доказывает выполнение условия 3.

Отсюда вытекает наличие важного свойства системы:

— кривая обслуживания системы равна свертке функций обслуживания отдельных компонентов системы и поэтому сигнал на выходе группы элементов может быть описан той же моделью, что и на выходе одного элемента.

Тогда кривая обслуживания:

$$\beta_S = r(t - kT),$$

задержка ограничена:

$$V_S(t) \le \frac{b}{c} + kT. \qquad (8)$$

Сравнивая результаты расчета, полученные на основе соотношения (7) и соотношения (8), можно отметить, что соотношение (7) дает более грубые результаты. Проведем анализ расчетов подробнее. Отношение *b/C* характеризует задержку, возникающую за счет всплеска трафика. Мы видим, что в (7) член *b/C* посчитал k раз в отличие от (8), а это позволяет говорить, что за всплески в сети приходится "расплачиваться один раз". Также в (8) отсутствует член $\frac{k(k-1)}{2}$, это указывает на то, что увеличения берстности исходящего потока на rT не влияет на суммарную задержку.

Можно показать, что рассматриваемый подход позволяет анализировать системы, модели которых должны обладать свойствами 4 и 5.

Так, если есть система с кривой обслуживания $\beta(t)$, на которую поступает два потока, причем второй поток ограничен кривой $\alpha_2(t)$, то для первого потока справедливо: $D_1(\mathrm{t}) \ge A_1 \otimes (\alpha_2 - \beta)(\mathrm{t})$.

Таким образом, подтверждается свойство 4.

Теперь рассмотрим агрегированный трафик, состоящий из двух потоков:

$$\mathrm{A(s,t) = A_1(s,t) + A_2(s,t) \le \alpha_1(t-s) + \beta_2(t-s) \le \alpha s.}$$

Таким образом кривая агрегированного трафика образуется суммированием кривых потоков, составляющих трафик. Таким образом, подтверждается свойство 5.

В заключение отметим, что хотя в статье исследована цепочечная структура сети, рассмотренные методы расчета можно применять для сетей с другой топологией. Так, в [2] приведен пример применения теории NC при расчете максимальной задержки для произвольной сети и расчете необходимой емкости буферов при условии отсутствия потерь. Через каждый элемент сети проходит n_k потоков, дисциплина обслуживания в узлах FIFO, функции $\alpha(t)$ и $\beta(t)$ аналогичны рассмотренным выше. Окончательные результаты имеют вид:

$$V_k(t) \le \frac{k}{1-\mu_k(k-1)} \cdot (n_k \frac{b}{C(k-1)} + T);$$

$$B_k(t) \le \frac{C}{1-\mu_k(k-1)} \cdot (n_k \frac{b}{C(k-1)} + T),$$

где n_k — число потоков, проходящих через k-й узел,

$\mu_k = \frac{n_k r}{C}$ — использование канала.

Литература

1. **R.L. Cruz.** A Calculus for Network Delay. Part I: Network Elements in Isolation and Part II: Network Analysis. IEEE Transactions on Information Theory, 37(1):114-141p, Jan. 1991.

2. **Jiang, Yuming, Liu.** Stochastic Network Calculus. Springer-Verlag 2008. — 229 p.

3. **Литвинов Г.Л., Маслов В.П., Соболевский А.Н.** Идемпотентная математика и интервальный анализ // Вычислительные технологии, 2001. — Т.6. — № 6. — С. 47-70.

4. **Клейнрок Л.** Теория массового обслуживания. М.: Машиностроение, 1979. — 432 с.

Прогноз развития сетей связи в Республике Таджикистан

Рассматривается прогноз развития сетей связи в Республике Таджикистан до 2015 г. с использованием метода наименьших квадратов. Основной целью настоящей статьи является разработка прогноза развития сетей связи в Республике Таджикистан (РТ) до 2015 г. с использованием метода наименьших квадратов [1,2]. При построении прогнозов развития использованы данные Исполнительного комитета РСС за 2003 — 2007 гг. [3 — 7].

Данилов А.Н.,
Кайюмов С.Т.,
МТУСИ

При разработке прогнозов развития сетей связи показана динамика изменения с 2003 по 2015 гг.:

1) протяженности телефонных каналов на междугородных линиях связи;

2) числа пользователей телематических служб и передачи данных с использованием Интернета;

3) числа терминалов службы IP-телефонии.

При разработке аналитического прогноза развития сетей связи по вышеприведенным показателям методом наименьших квадратов с 95%-ным доверительном интервалом, воспользуемся существующими прогнозами МСЭ-Т, полученными с использованием данных Всемирного банка по следующим параметрам [8]:

1) телефонной плотности стационарных телефонов (ТП);

2) плотности пользователей Интернета (ИП);

3) плотности пользователей персональных компьютеров (ПКП).

Параметры плотности рассчитывались на 100 жителей страны.

В табл. 1 приведены аппроксимирующие уравнения для параметров плотностей, где y — плотность, а x — душевой валовый национальный продукт (ДВНП) [8].

Таблица 1

Вид плотности	Аппроксимирующее уравнение
ТП	$y=5{,}4x^{0.95}$
ИП	$y=2{,}2x^{1{,}003}$
ПКП	$y=1{,}6x^{1{,}014}$

При построении прогнозов ДВНП и численности населения РТ после 2007 г. использовалась линейная аппроксимация

$$u(t)=a_0+a_1t, \qquad (1)$$

где t — параметр, соответствующий годам наблюдения (t = 0 соответствует 2003 г., t = 1 — 2004 г. и т.д.), a_0 — значение параметра (ДВНП или количество населения) при t = 0.

Для определения протяженности междугородных телефонных каналов учитывалась как длина линий связи, так и число каналов ИКМ по направлениям. Число междугородных каналов ИКМ рассчитывалось из нормы 0,003 Эрл на один стационарный телефонный номер для входящей и исходящей нагрузки с использованием первой формулы Эрланга.

Длина междугородных линий связи РТ аппроксимировалась по следующей формуле

$$y=6697{,}1x^{1{,}7}, \qquad (2)$$

где y — длина междугородных линий связи, x — ДВНП.

Данные аппроксимирующих уравнений для прогноза приведены в табл. 2.

Таблица 2

ДВНП, тыс. дол.	Протяженность междугородных линий связи, км	Протяженность междугородных телефонных каналов, тыс. кан-км	Количество стационарных телефонных аппаратов, шт.	Годы
0,21	472	264,6	81407	2003
0,272	732	257,2	106311	2004
0,334	1038	914,6	131826	2005
0,396	1386	1445,9	157450	2006
0,458	1776	2161,0	185184	2007
0,52	2203	3067,1	212860	2008
0,582	2668	4197,4	241530	2009
0,644	3170	5572,0	270891	2010
0,706	3706	7218,4	301088	2011
0,768	4276	8730,8	331800	2012
0,83	4879	10943,6	363730	2013
0,892	5515	13494,0	396240	2014
0,954	6182	16431,8	429645	2015

Число терминалов IP-телефонии рассчитывалось исходя из числа ПКП с учетом того, что не все пользователи ПК пользуются IP-телефонией. Число пользователей IP-телефонии с учетом данных 2003-2007 гг. РСС по РТ и линейной аппроксимации прогноза (форм. 1) изменяется от 0,65% до 2,33% от общего числа ПКП.

Аппроксимационное уравнение для плотности ИП МСЭ-Т не подходит ввиду большого значения среднеквадратического отклонения фактических наблюдений от расчетных значений S_y ($S_y > 1$), что приводит к недопустимо большому значению границ 95%-го доверительного интервала. Ввиду этого, предлагаем другое уравнение для аппроксимации, используя данные МСЭ-Т (табл. 1), при которых $S_y < 1$

$$y = 0{,}9\,x^{1{,}92}, \qquad (3)$$

В табл. 3 приведены данные аппроксимирующих уравнений по прогнозам числа пользователей телематических служб и передачи данных с использованием Интернета и числа терминалов службы IP-телефонии.

Для аналитических прогнозов выполнено логарифмирование по натуральному логарифму аппроксимирующих уравнений.

При расчете доверительного интервала аналитических прогнозов воспользуемся следующим выражением [2]:

$$e^{(\ln \hat{y}_{i+L} \pm S_y K^*)}, \qquad (4)$$

где $\hat{y}_{i+L}$ — расчетное значение параметра сети связи в период прогноза (с 2008 по 2015 гг.).

S_yK^* и S_y определяется по следующим формулам:

$$S_yK^* = t_\alpha S_y \sqrt{1+\frac{1}{n}+\frac{1}{\sum t^2}t_L^2}, \; t=\overline{1,n},$$

Таблица 3

ДВНП, тыс. дол.	Число пользователей телематических служб и передачи данных с использованием Интернета, чел.	Число терминалов службы IP-телефонии, шт.	Годы
0,21	2986	143	2003
0,272	5011	229	2004
0,334	7585	339	2005
0,396	10731	470	2006
0,458	14469	632	2007
0,52	18822	817	2008
0,582	23813	1030	2009
0,644	29462	1272	2010
0,706	35794	1544	2011
0,768	36440	1847	2012
0,83	50598	2184	2013
0,892	59115	2553	2014
0,954	68407	2957	2015

где t_α — значение t-статистики Стьюдента; t_L — время, для которого производится экстраполяция; n — число наблюдений; $K^* = t_\alpha K$, K обозначает квадратный корень в выражении (5).

$$S_y = \sqrt{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2 / t_\alpha}, \qquad (6)$$

где $y_t, \hat{y}_t$ — фактическое и расчетное значения параметров сетей связи соответственно.

Доверительный интервал прогноза общей протяженности междугородных телефонных каналов приведем в табл. 4.

Таблица 4

Доверительный 95%-ный интервал, тыс. кан-км	Годы
12,7÷44355,9	2008
304,0÷57699,3	2009
392,3÷79221,3	2010
470,6÷110635,9	2011
514,4÷146825,5	2012
572,5÷208981,3	2013
626,4÷290686,3	2014
665,1÷408399,0	2015

На рис. 1 показана динамика изменения протяженности телефонных каналов на междугородных линиях связи $L_{МГК}$ с 2003 по 2015 гг. Масштаб по оси ординат — логарифмический (ln LМГК).

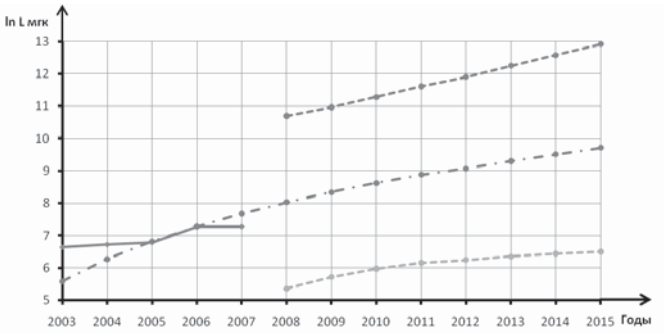


Рис. 1. Протяженность телефонных каналов на междугородных линиях связи

Как видно из рис.1, табл. 2 и 4, общая протяженность телефонных каналов на междугородных линия связи должна вырасти в соответствии с прогнозом по верхней границе в 282,45 раза в 2015 г. по отношению к 2003 г.

На рис. 2 проследим динамику роста числа пользователей $N_{инт}$ телематических служб и передачи данных с использованием Интернета. По оси ординат отложим значения ln $N_{инт}$.

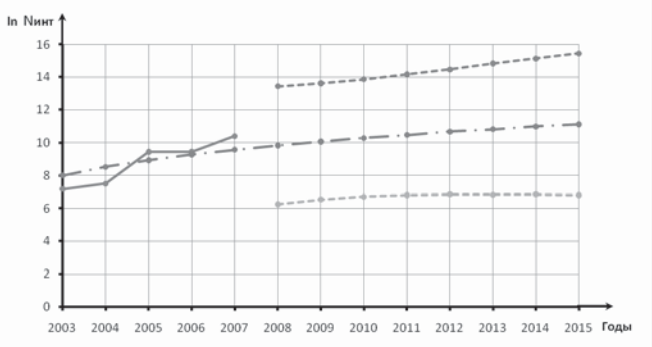


Рис. 2. Число пользователей телематических служб и передачи данных с использованием Интернета

В качестве формулы для расчета доверительного интервала прогноза использованы выражения (4-6).

Доверительный интервал прогноза числа пользователей телематических служб и передачи данных с использованием Интернета приведен в табл. 5.

Таблица 5

Доверительный 95%-ный интервал	Годы
517÷685566	2008
695÷816678	2009
824÷1056002	2010
906÷1414095	2011
948÷1931872	2012
927÷2760715	2013
954÷3760265	2014
904÷5178366	2015

Как видно из рис. 2, табл. 3 и 5, число пользователей телематических служб и передачи данных с использованием Интернета к 2015 г. растет практически по экспоненциальной зависимости и должно составлять по верхней границе 51788366, что в 154,4 раза больше по сравнению с 2003 г.

Проследим динамику роста числа терминалов N_{IP} службы IP-телефонии в РТ (см. рис. 3, табл. 3 и 6). Доверительный интервал (формулы 5-7) прогноза числа терминалов служб IP-телефонии в РТ покажем в табл. 6.

Таблица 6

Доверительный 95%-ный интервал	Годы
747÷894	2008
943÷1124	2009
1162÷1391	2010
1408÷1693	2011
1679÷2030	2012
1976÷2414	2013
2301÷2833	2014
2655÷3295	2015

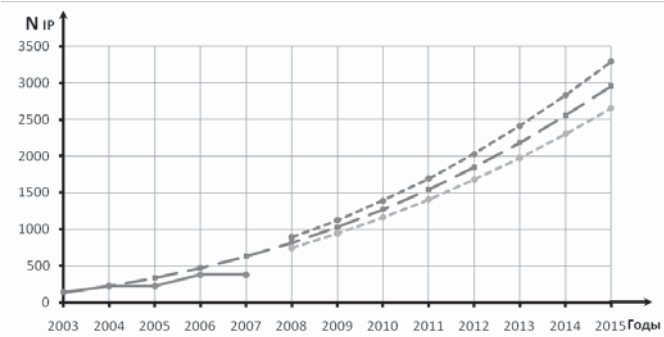


Рис. 3. Число терминалов служб IP-телефонии в РТ

Прогнозируемый рост числа терминалов службы IP-телефонии в РТ в 2015 г. превосходит показатель 2003 г. в 8,67 раза (по экспоненциальной зависимости).

На основе приведенных выше прогнозов по развитию сетей связи в РТ, можно сделать выводы:

1) о резком росте числа пользователей современных видов связи по экспоненциальной зависимости и необходимости дальнейшего серьезного развития существующей мультисервисной сети связи РТ в рамках концепции NGN;

2) допустимости использования аппроксимаций степенными функциями с использованием данных МСЭ-Т (табл. 1).

Литература

1. **Кендалл М., Стюарт А.** Многомерный статический анализ и временные ряды. — М.: Наука, 1976. — 736 с.

2. **Четыркин Е.М.** Статистические методы прогнозирования. Изд. 2-е, перераб. и доп. — М.: Статистика, 1977. — 200 с.

3. Статистический сборник о деятельности администраций связи в области связи и информатики за 2004 год. — М.: Исполнительный комитет РСС, 2005. — 164 с.

4. Статистический сборник о деятельности администраций связи в области связи и информатики за 2005 год. — М.: Исполнительный комитет РСС, 2006. — 174 с.

5. Статистический сборник о деятельности администраций связи в области связи и информатики за 2006 год. — М.: Исполнительный комитет РСС, 2007. — 194 с.

6. Статистический сборник о деятельности администраций связи в области связи и информатики за 2007 год. — М.: Исполнительный комитет РСС, 2008. — 184 с.

7. **Кайюмов С.Т.** Анализ состояния современных телекоммуникационных сетей в Республике Таджикистан // Труды Московского технического университета связи и информатики. — М.: "ИД Медиа Паблишер", 2007. — С. 388-395.

8. **Варакин Л.Е.** Информационно-экологический закон. Взаимосвязь инфокоммуникационной инфраструктуры и экономики. — М.:МАС, 2006. — 160 с.