

**ЗАДАЧИ АНАЛИЗА И ОПТИМИЗАЦИИ  
МУЛЬТИСЕРВИСНЫХ МОБИЛЬНЫХ СЕТЕЙ**

**А. Б. Мархасин**

*Сибирский государственный университет телекоммуникаций и информатики,  
г. Новосибирск  
E-mail: almar@rinet.su*

Развиваются методы анализа и оптимизации распределенных многомерных очередей с приоритетами и динамическим управлением качеством обслуживания (QoS – Quality of Services) для подуровней управления множественным доступом к среде MAC (Medium Access Control) в мультисервисных мобильных и беспроводных сетях 3G/4G. Предлагается теоретико-игровая постановка задачи распределенного динамического управления QoS. Полученные результаты основаны на развитии численно-аналитического метода баланса интенсивностей нагрузок.

**Введение.** Развитие мобильных сетей 3-го и 4-го поколений (3G/4G) сопровождается революционными изменениями в телекоммуникационных и сервисных технологиях [1]. Доминирующей идеей развития является переход от технологий коммутации каналов (КК) к всеобъемлющим IP (Internet Protocol) технологиям коммутации пакетов (КП). На смену моносервисным сетям 2-го поколения (2G), предназначенным для передачи речи на основе технологий КК, приходят мультисервисные сети новых поколений (3G/4G) с интегральной передачей голоса, данных, видео и многих других видов информации на принципах КП. Интеграция различных видов информации в общий мультисервисный IP-трафик поддерживается путем дифференциации и динамического управления гарантированным качеством услуг (QoS – Quality of Services). Стремительно увеличивается также объем трафика доступа к услугам сети Internet.

Новые технологии нуждаются в разработке адекватных методов анализа телетрафика, оптимизации и проектирования [2, 3]. Взамен одномерных моделей Эрланга для однородных трафиков в двухточечных каналах необходимо развивать теорию многомерных распределенных очередей с приоритетным обслуживанием интегральных потоков данных в многоточечных радиоканалах множественного доступа (МД), что поясняется сравнительными моделями обслуживания в мобильных сетях 2G (рис. 1, *a*) и 3G/4G (рис. 1, *b*).

Как следует из рис. 1, *a*, основная задача анализа вероятностно-временных характеристик (ВВХ) в сетях 2G с коммутацией каналов сводится к решению с помощью формулы Эрланга уравнения  $F(n, G, R) = 0$ , т. е. к нахож-

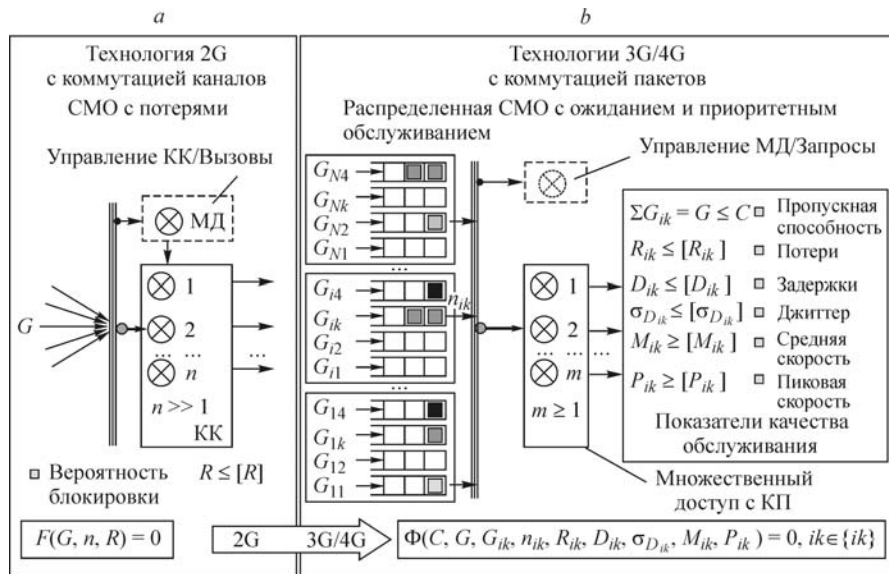


Рис. 1. Сравнительные модели систем обслуживания: в моносервисных (а) и мультисервисных (b) мобильных сетях

дению одной из трех переменных по заданным значениям двух других. Наиболее часто необходимо найти число каналов трафика  $n$  по заданной интенсивности нагрузки  $G$  и вероятности блокировки вызовов  $R$ .

В мобильных мультисервисных сетях 3G/4G предъявляются дифференцированные и изменяющиеся во времени требования к показателям качества обслуживания для каждого  $i$ -го пользователя и каждого  $k$ -го класса обслуживания [1]. Мобильные сети 3G/4G с управлением QoS моделируются распределенными в пространстве многомерными очередями с приоритетами и интегральным обслуживанием [3]. Основная задача анализа ВВХ в мультисервисных сетях новых поколений сводится к вычислению функционалов вида

$$\Phi(C, G, G_{ik}, n_{ik}, R_{ik}, D_{ik}, \sigma_{D_{ik}}, M_{ik}, P_{ik}) = 0, \quad (1)$$

$$ik \in \{ik \mid i=1,2,\dots, N, k=1,2,\dots, K\}.$$

Цель данной работы – развитие методов анализа и оптимизации распределенных многомерных очередей с приоритетами и динамическим управлением качеством обслуживания для подуровней управления множественным доступом к среде МАС (Medium Access Control) в мультисервисных мобильных и беспроводных сетях (ММБС) 3G/4G. Предлагаемые решения основаны на развитии численно-аналитического метода баланса интенсивностей нагрузок [4, 5].

**1. Задачи анализа протоколов МАС мультисервисных мобильных и беспроводных сетей.** 1.1. *Постановка задачи анализа ВВХ.* Пусть на некоторых отрезках времени  $[t_s, t_{s+1})$ ,  $s=1,2,\dots$ , в сети активны (инициализированы)  $N_t$   $i$ -х мобильных станций,  $i=1,2,\dots, N_t$ . Пусть далее на входы станций поступают для передачи через беспроводную среду с МД неоднородные пуассоновские потоки пакетов данных  $k$ -х классов приоритетного обслуживания

ния с входными нагрузками  $G_{ikt} = \lambda_{ikt} T_k$ ,  $k = 1, 2, \dots, K$ , где  $\lambda_{ikt}$  – интенсивность, а  $T_k$  – средняя длительность входных пакетов  $k$ -го приоритетного класса. Длительность удобно измерять в относительных единицах, равных среднему времени передачи одного пакета  $T = E[T_k]$ .

Такие сети моделируются многомерной (с размерностью  $KN_t$ ) системой массового обслуживания (СМО) вида  $M/G_{G_t, E, P}/1/\mathbf{n}_t$  с приоритетами и неоднородной распределенной векторной очередью

$$\mathbf{n}_t = (n_{11t}, n_{12t}, \dots, n_{1kt}, \dots, n_{i1t}, n_{i2t}, \dots, n_{ikt}, \dots, n_{N1t}, n_{N2t}, \dots, n_{NKt})$$

с емкостями  $n_{ikt}$  локальных составляющих, которые распределены в пространстве. Индексы при символе  $G_{G_t, E, P}$  закона распределения общего вида означают зависимость времени обслуживания заявок от величины входной

нагрузки  $G_t = \sum_{i=1}^{N_t} \sum_{k=1}^K \lambda_{ikt} T_k$ , параметров  $E$  внешней среды (т. е. вероятностей

$p_x$  обнаружения ошибок в  $x$  пакетах) и от системных параметров  $P$  протоколов MAC/LLC (LLC – Link Layer Control) канального уровня. Векторный символ  $\mathbf{n}_t$  означает, что состояния очереди, или чисел  $j_{ikt}$  занятых ячеек в  $i$ -х очередях  $k$ -х приоритетных классов обслуживания, задаются вектором

$$\mathbf{j}_t = (j_{11t}, j_{12t}, \dots, j_{1kt}, \dots, j_{i1t}, j_{i2t}, \dots, j_{ikt}, \dots, j_{N1t}, j_{N2t}, \dots, j_{NKt}). \quad (2)$$

Задачи анализа ВВХ ММБС, описываемых моделями СМО вида  $M/G_{G_t, E, P}/1/\mathbf{n}_t$ , не аналитичны в реалистической постановке, т. е. имеют аналитические решения лишь при сильных упрощающих допущениях и, как правило, не позволяют учитывать такие существенные характеристики реальных сетей, как потоки ошибок, неоднородности входных нагрузок, реальные входные потери, потоки переспросов передач, реальные системные параметры протоколов MAC и дисциплин приоритетного обслуживания (неопустошающее обслуживание, ограничение емкости очередей и т. п.). Для решения поставленной задачи в данной работе будет развит численно-аналитический метод баланса интенсивностей нагрузок [3].

Главная идея метода заключается в сведении модели общей многомерной очереди  $M/G_{G_t, E, P}/1/\mathbf{n}_t$  размерностью (2) к эквивалентной сумме  $NK$  линейно независимых, или парциальных, очередей  $M/M_{ik}^*/1/n_{ik}$ ,  $ik \in \{ik \mid i=1, 2, \dots, N, k=1, 2, \dots, K\}$ , с помощью эквивалентных законов  $M_{ik}^* \in \{M_{ik}^* \mid i=1, 2, \dots, N, k=1, 2, \dots, K\}$  распределения времени обслуживания в парциальных очередях.

1.2. Метод баланса интенсивностей нагрузок состоит из трех этапов решения задачи. На первом этапе определяются в параметрической форме интенсивности нагрузок в сечениях сети от входов до выходов распределенных очередей соответственно вероятностному графу процедур приоритетного обслуживания и передачи (рис. 2), в рассматриваемом случае (см. разд. 1.1) – для мультифункционального протокола адаптивного контролируемого доступа с распределенным управлением качеством сервисов (QoS) и приоритетами (например, [5]).

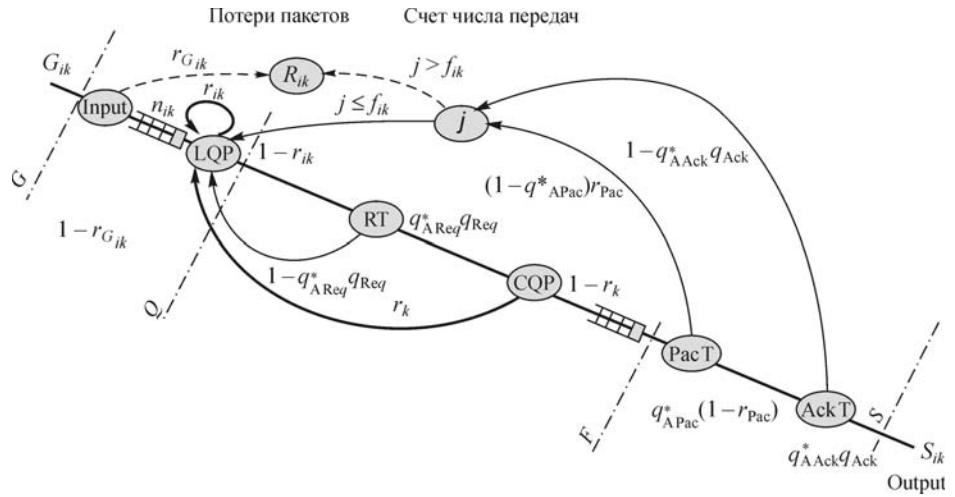


Рис. 2. Вероятностный граф протокола MAC адаптивного контролируемого доступа с приоритетами и управлением QoS: LQP (Local Queue Priority) – локальная приоритетная очередь; RT (Request Transmission) – передача запросов; CQP (Common Queue Priority) – общая приоритетная очередь; PacT (Packet Transmission) – передача пакетов; AckT (Acknowledgment Transmission) – передача подтверждений;  $G$  – входной трафик;  $Q$  – виртуальный трафик;  $F$  – канальный трафик;  $S$  – производительность канала

Интенсивности нагрузок в сечениях сети удобно выразить в функции от входной нагрузки  $G$  через промежуточный параметр  $F$  – интенсивность нагрузки в канале (суммарный трафик первичных и повторных передач пакетов данных измеряется в эрлангах). В соответствии с вероятностным графом можно найти:

– парциальную виртуальную интенсивность нагрузки (т. е. входной трафик запросов на обслуживание)  $ik$ -й парциальной очереди

$$Q_{ik}(F) = F / [(q_{AReq}^* q_{Req})(1 - r_k(F))], \quad (3)$$

– парциальную интенсивность нагрузки  $ik$ -й парциальной очереди

$$F_{ik}(F) = G_{ik} [1 - r_{G_{ik}}(F)] \alpha_{ik}, \quad (4)$$

– парциальную производительность (трафик успешно переданных пакетов)

$$S_{ik}(F) = F_{ik}(F) q_{APac}^* (1 - r_{Pac}) q_{AAck}^* q_{Ack}, \quad (5)$$

где коэффициент смешанных переспросов

$$\alpha_{ik} = \frac{1 - [1 - q_{APac}^* (1 - r_{Pac}) q_{AAck}^* q_{Ack}]^{f_{ik}}}{q_{APac}^* (1 - r_{Pac}) q_{AAck}^* q_{Ack}}; \quad (6)$$

$f_{ik}$  – максимально допустимое число попыток передачи пакета данных  $k$ -го класса обслуживания в  $i$ -й очереди;  $q_x^*$  – вероятности успешной передачи различных адресованных единиц данных, определяемые на множестве стан-

ций. Условием успеха является сложное событие – правильная адресация единицы данных  $i$ -й станцией и отсутствие ошибок адресации у остальных станций (т. е. исключение «шумовых» коллизий):

$$q_x^* = (1 - p_x) \exp(-p_x). \quad (7)$$

На графе определяются также вероятности:

$r_{G_{ik}}$  – переполнение  $i$ -й очереди  $k$ -го класса обслуживания,

$r_{ik}$  – прерывание обслуживания заявки  $k$ -го класса в  $i$ -й очереди при поступлении заявки более высокого приоритета,

$r_k$  – вытеснение из общей очереди обслуживания пакета  $k$ -го класса приоритета.

На втором этапе определяются параметры эквивалентных законов  $\{M_{ik}^* \mid i=1,2,\dots,N, k=1,2,\dots,K\}$  распределения времени обслуживания в парциальных очередях. В качестве параметров вводятся так называемые циклы обслуживания заявок  $V_{ik}(F)$  в локальных  $i$ -х очередях  $k$ -х классов [4]. Циклы обслуживания запросов в приоритетных очередях определяются как периоды времени, которые начинаются в моменты поступлений запросов  $k$ -х классов на обслуживание и заканчиваются, когда каналы освобождаются для обслуживания запросов того же  $k$ -го класса приоритета [6].

Понятие «цикл обслуживания» является обобщением термина «время обслуживания». Из графа (см. рис. 2) видно, что циклы обслуживания зависят от большого числа случайных редких событий (переспросов, конфликтов, искажений, поступления запросов и т. п.), т. е. выполняются условия сходимости к предельным распределениям. Можно показать [7], что распределение длительностей циклов обслуживания при слабых ограничениях сходится (точнее, может быть аппроксимировано) в пределе к экспоненциальному распределению. Поэтому в моделях  $M/M_{ik}^*/1/n_{ik}$ ,  $ik \in \{ik \mid i=1,2,\dots,N, k=1,2,\dots,K\}$ , эквивалентных парциальных очередей законы распределения обобщенного времени обслуживания (цикла обслуживания) приняты как экспоненциальные. Это допущение подтверждается также результатами имитационного моделирования.

На третьем этапе из полученных в параметрической форме характеристики очередей исключается промежуточный параметр. Для этого составляется уравнение баланса интенсивностей нагрузок и решается численным методом относительно промежуточного параметра  $F$  уравнение баланса интенсивностей парциальных нагрузок

$$\sum_{i=1}^{N_i} \sum_{k=1}^K G_{ik} [1 - r_{G_{ik}}(F)] \alpha_{ik} = F. \quad (8)$$

Далее нетрудно получить вероятностно-временные и качественные характеристики очередей в функции от входной нагрузки и системных параметров протоколов с помощью известных соотношений для моделей СМО вида  $M/M/1/n$ . Значения коэффициентов использования локальных очередей находятся на основании средних длительностей циклов обслуживания:

$$\rho_{ik}(F) = G_{ik} V_{ik}(F). \quad (9)$$

1.3. *Определение циклов обслуживания.* Средние длительности циклов обслуживания для мультисервисных мобильных сетей с управляемым QoS можно найти с помощью полученных в [4] соотношений. Пусть  $V_{ik}(F)$  – средняя длительность цикла обслуживания запроса  $k$ -го приоритета в  $i$ -й абонентской (парциальной) очереди:

$$V_{ik}(F) = E_{ik} / (1 - r_{ik}(F)), \quad (10)$$

где  $E_{ik}(F)$  – средняя длительность цикла обслуживания запроса  $k$ -го приоритета  $i$ -го абонента в общей очереди принятых к обслуживанию пакетов.

В случае, когда относительное время распространения  $v_p$  много меньше периода опроса  $i$ -го абонента  $P_i(F)$ , можно получить

$$E_{ik}(F) = P_i(F) \alpha_{ik} / [q_{AReq}^* q_{Req}(1 - r_k(F))], \quad (11)$$

а средний цикл обслуживания (10) выразить в виде

$$V_{ik}(F) = P_i(F) \alpha_{ik} / [q_{AReq}^* q_{Req}(1 - r_k(F))(1 - r_{ik}(F))], \quad (12)$$

где  $\alpha_{ik}$  – коэффициент смешанных переспросов (6);  $r_{ik}(F)$  – вероятность прерывания обслуживания запроса  $k$ -го приоритета в  $i$ -й абонентской очереди:

$$r_{ik}(F) = \sum_{j=1}^{k-1} [1 - r_{G_{ij}}(F)] \rho_{ij}(F); \quad (13)$$

$r_k(F)$  – вероятность вытеснения из общей очереди обслуживания пакета  $k$ -го класса приоритета:

$$r_k(F) = 1 - \left\{ 1 - v_c \sum_{j=1}^{k-1} F_j(F) / \left\{ q_{AReq}^* q_{Req} [1 - r_j(F)] [1 - (1 + v_a) z(F)] \right\} \right\}^{l_k}; \quad (14)$$

$l_k$  – параметр управления приоритетами [5]:

$$l_K > \dots > l_{k-1} > l_k > l_{k-1} > \dots > l_1 = 0; \quad (15)$$

$P_i(F)$  – период опроса  $i$ -й станции [4]:

$$P_{ik}(F) = (v_c M / m_i) / [1 - (1 + v_a) z(F)]; \quad (16)$$

$z(F)$  – интенсивность нагрузки канала во времени, или доля времени занятия канала информационными интервалами, которая может быть выражена в виде [4]

$$z(F) = \frac{1}{1 + v_a} - \frac{\frac{1}{1 + v_a} - bF \left\{ 1 - \sum_{k=1}^K \left[ \frac{F_k}{(1 - r_k)F} \right] \right\} \left( 1 - \sum_{j=1}^K \frac{r_j}{K} \right) \frac{P_{Req}}{q_{Req}}}{1 + bq_{AReq}^* P_{Req} \left( 1 - \sum_{j=1}^K \frac{r_k}{K} \right) \frac{1 + v_a}{v_c}}. \quad (17)$$



Здесь  $b$  – коэффициент,  $b = 1 + (1 - q_{ARec})p_{ARec} / [(1 - p_{ARec} / (N - 1))q_{ARec}]$ ;  $v_c$ ,  $v_a$ ,  $v_p$  – относительные (по отношению к средней длительности пакета  $E[T_{pac}]$ ) затраты ресурса времени на передачу команд управления множественным доступом  $E[T_{Req}]$  и квитирования  $E[T_{Ack}]$  и распространение сигналов  $E[T_p]$  соответственно;  $M$ ,  $m_i$  – управляющие параметры распределения ресурсов полосы: период рекуррентной последовательности (РП) и число РП-адресов, образующих MAC-адрес  $i$ -й станции [1, 5] соответственно.

1.4. Численный эксперимент. Для проверки точности метода баланса интенсивностей нагрузок были проведены численные эксперименты по сравнению ВВХ, полученных предлагаемым методом, с результатами имитационного моделирования [4]. Сравнения среднего времени доставки пакетов на примере характеристик мультисервисной сети, включающей 1024 станции, которые разделены на четыре группы по классам приоритетного обслуживания и нормированном весам входной нагрузки, представлены на рис. 3. Относительные затраты ресурса полосы приняты следующими: на управление МД  $v_c = 0,015$  отн. ед. времени, на распространение в среде –  $v_p = 0,379$  отн. ед. времени, на квитирование –  $v_a = 0,030$  отн. ед. времени; вероятность ошибки на символ –  $p = 0,0001$ . Для прослеживания влияния дисперсии адресной рекуррентной последовательности на время задержки пакетов рассмотрены два случая: с учетом дисперсии РП и без ее учета. Из представленных на рис. 3 графиков следует, что в рабочем диапазоне значений входной нагрузки (примерно до 0,9 эрланг) взвешенные среднеквадратические расхождения аналитических и модельных значений времени задержки/доставки пакетов составляют от  $\pm 4\%$  до  $\pm 8\%$  при доверительной вероятности 0,95. Дисперсия интервалов повторения одноименных адресных комбинаций на РП вызывает некоторое увеличение дисперсии периода опроса (16). Как показано в [8], флуктуации периода опроса приводят к возрастанию времени задержки, что подтверждается данными численного эксперимента.

2. Потенциальная пропускная способность MAC в беспроводных средах. В работе [9] рассматриваются реализованная и потенциальная пропускные способности протоколов управления МД в беспроводные среды. Под реализованной пропускной способностью понимается максимальное

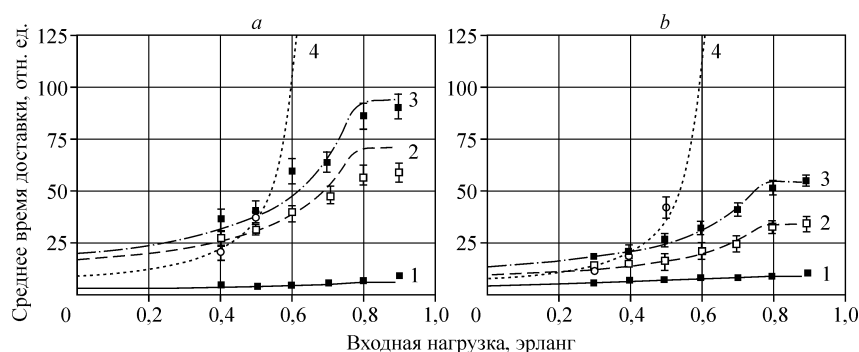


Рис. 3. Пример характеристик приоритетного обслуживания [5]: с учетом (а) и без учета (б) дисперсии адресной РП (кривая 1 – класс приоритета  $k = 1$ , пакетированная речь; 2 – команды в реальном масштабе времени,  $k = 2$ ; 3 – интерактивные данные и телеметрия,  $k = 3$ ; 4 – файлы,  $k = 4$ ; линии – метод баланса интенсивностей нагрузок, значки – имитационное моделирование)

значение выходной нагрузки  $S(G)$  при варьировании входной нагрузки и, возможно, при заданном ограничении времени задержки  $D$ :

$$C_{\text{MAC}} = \max_{\{G, D \leq [D]\}} S(G). \quad (18)$$

Реализуемые значения пропускных способностей различных протоколов MAC (см. таблицу) задаются их регламентными механизмами доступа, видами временных процессов доступа и системными параметрами [1]. Фактически такие значения определяются степенью информированности системы множественного доступа MAC о текущих состояниях (траекториях) процессов доступа в распределенных в географическом пространстве абонентских очередях [9]. Эта информация о состояниях очередей недоступна станциям беспроводной сети до свершившегося факта ее передачи в общую среду, поэтому возникают конфликтные ситуации и коллизии в общей среде из-за принципиальной неполноты информации о процессах МД.

По сути, снятие информационной неопределенности о процессах поступления заявок (о моментах времени их поступления, требуемых классах обслуживания, длительностях и т. п.) в распределенные в пространстве очереди и является основной функцией протоколов MAC. Чем больше информированность, тем выше значение реализуемой пропускной способности. По протоколу «Простая Aloha» информация о состояниях распределенных очередей полностью отсутствует. Процедура тактирования вносит сведения

Регламент доступа		Временной процесс доступа			
		Случайный процесс доступа			Детерминированный процесс доступа
		непрерывный	дискретный		
	фиксированный		адаптивный		
Контролируемый доступ	Централизованный	$\frac{1}{1 + v_l}$	$\frac{1}{1 + v_s N/J}$	$\frac{1}{1 + v_m}$	$\frac{1}{1 + v_s}$
	Распределенный	$\frac{1}{1 + v_m}$	$\frac{1}{1 + v_s N/J}$	$\frac{1}{1 + v_s}$	$\frac{1}{1 + v_s}$
Свободный (случайный) доступ	Без проверки носителя	$1/2e$	$1/e$	–	–
	С проверкой носителя	$\frac{1}{1 + a_{31} v_s}$ , $a_{31} > 1$	–	$\frac{1}{1 + a_{33} v_s}$ , $a_{31} > a_{33} > 1$	–
Смешанный доступ		$\frac{1}{1 + a_{41} v_l}$ , $a_{41} > 1$	$\frac{1}{1 + a_{42} v_m N/J}$ , $a_{41} > a_{42} > 1$	$\frac{1}{1 + a_{43} v_m N/J}$ , $a_{42} > a_{43} > 1$	–

*Примечание:*  $v_s, v_m, v_l$  – малые (до 0,02), средние (от 0,02 до 0,10), большие (более 0,10) затраты ресурса на управление доступом соответственно;  $a_{ij}$  – коэффициент, зависящий от вида протокола MAC и номеров строк и столбцов.



о границах нумерованных временных интервалов (протокол «Тактированная Aloha») или мини-интервалов (CSMA – Carrier Sense Multiple Access). Адресация (нумерация) и «персонализация» служебных и информационных интервалов в протоколах контролируемого доступа (КД) существенно увеличивают информированность очередей [1].

Потенциальные характеристики протоколов МАС могут быть найдены через вычисление предельно допустимых минимальных издержек (количеств) информационного ресурса коллективной среды на полное информирование распределенных станций о текущих (реализованных) траекториях процессов доступа. Такие потенциальные информационные характеристики можно найти с помощью вычисления энтропии траекторий доступа, моделируемых марковскими процессами в системах массового обслуживания с распределенными очередями, как это устанавливается в следующих теоремах [9]:

**Теорема 1.** Пусть сеть множественного доступа с временным разделением (МДВР) описывается моделью СМО с распределенными очередями вида  $M/M/1$  с точностью до дискретного приближения распределения времени обслуживания, тогда предельно достижимые относительные затраты на организацию управления множественным доступом в беспроводных каналах без ошибок будут

$$v_c^0 = [2 + H(k)] / (\Delta T_{\text{инф}} B), \quad (19)$$

а потенциальная пропускная способность идеального протокола управления множественным доступом (МАС)

$$C_{\text{МАС}}^0 = 1 / \{1 + [2 + H(k)] / (\Delta T_{\text{инф}} B)\}, \quad (20)$$

где  $\Delta T_{\text{инф}}$  – средняя длительность информационных интервалов для передачи ячеек, инкапсулированных пакетов и т. д.;  $B$  – битовая скорость;  $H(k)$  – энтропия распределения длительностей информационных интервалов. Для геометрического распределения

$$H(k) = -[x \log x + (1 - x) \log(1 - x)] / (1 - x), \quad (21)$$

где параметр  $x = \exp(-1 / \Delta T_{\text{инф}} B)$ .

**Теорема 2.** Пусть в условиях теоремы 1 модель СМО будет вида  $M/D/1$ , тогда предельно достижимые относительные затраты на организацию управления множественным доступом в каналах без ошибок будут

$$v_c^0 = 1,854 / (\Delta T_{\text{инф}} B), \quad (22)$$

а потенциальная пропускная способность идеального протокола управления множественным доступом (МАС)

$$C_{\text{МАС}}^0 = 1 / [1 + 1,854 / (\Delta T_{\text{инф}} B)]. \quad (23)$$

В [9] получены потенциальные характеристики также для каналов с ошибками. Доказанные теоремы устанавливают предельные потенциальные значения характеристик, но не указывают на конструктивные методы их достижения. Представляется наиболее целесообразным развитие механизмов

передачи служебной информации МАС, которые в максимальной степени могли снимать неопределенность информации о процессах множественного доступа с минимальными затратами информационных ресурсов коллективной среды. К таким МАС-технологиям принадлежит предложенный в [5] протокол маркерного широковещательного резервирования (МРШ-РП), по которому удается и существенно снизить информационную неопределенность, и одновременно минимизировать издержки управления благодаря всестороннему использованию возможностей рекуррентных М-последовательностей [1].

**3. Задачи оптимизации и динамического управления качеством.** По протоколу МРШ-РП пропорциональное динамическое управление распределением ресурсов полосы  $\{Y_{it}\}$  осуществляется путем дифференцированно-го регулирования интенсивностей передачи запросов (см. рис. 2, RT). Интенсивности запросов, в свою очередь, пропорциональны мощностям  $\{m_{it}\}$  подмножеств РП-адресов, образующих МАС-адреса  $i$ -х мобильных станций для каждого отрезка времени, вплоть до реального масштаба времени [5]. В математической модели это отображается множителем  $M/m_{it}$  в выражении (16) для периода опроса  $i$ -й станции на  $t$ -м отрезке, где  $M$  – период М-последовательности ( $M = 2^n - 1$ ),  $n$  – степень порождающего полинома М-последовательности.

Далее мощности  $i$ -х подмножеств  $\{m_{it}\}$  МАС-адресов находятся путем решения динамических задач оптимального распределения ресурсов, а реализуются посредством соответствующего кодирования МАС-адресов оптимальными неприводимыми кодами (ОНК), например, по методу Шеннона – Фано. Динамическое перекодирование дерева ОНК МАС-адресов может осуществляться практически в реальном времени. Размерность адресного пространства будет при этом достигать  $M = 2^n - 1$ ,  $n = 64-128$ , т. е. сопоставима с новой технологией IPv6 [1].

Стратегии управления и оптимизации могут строиться как на основе централизованного управления, так и на игровой основе. Рассмотрим пример постановки игровой задачи оптимизации распределения ресурсов полосы [5].

*Игровая задача резервирования ресурсов полосы.* При децентрализованной стратегии динамическое управление распределением ресурсов  $Y_{it}$ ,  $i = 1, 2, \dots, N_t$ ,  $t \in [t_s, t_{s+1})$ , может осуществляться на игровой основе [5]. Станции (игроки) путем решения своих локальных динамических задач оптимизации определяют и заказывают требуемые им ресурсы – делают «ставки»  $\mathfrak{R}_{it}$ , а сетевой брокер обрабатывает сделанные ставки посредством соответствующего перекодирования ОНК  $m_{it}$ -кратных МАС-адресов. При этом обеспечивается пропорциональность выделенных ресурсов сделанным ставкам:

$$Y_{it} \sim m_{it} \sim \mathfrak{R}_{it} / \sum_i \mathfrak{R}_{it}. \quad (24)$$

Положим, что задана некоторая мера  $H[G_{ikt}, \mathbf{W}_k(Y_{it})]$  ценности в единицу времени данных  $k$ -го приоритетного класса в зависимости от их трафика  $G_{ikt}$  и вектора показателей качества их доставки

$$\mathbf{W}_{ikt}(Y_{it}) = (D_k(Y_{it}), \sigma_{it}, R_k(Y_{it}), P_k(Y_{it}), M_k(Y_{it})),$$

а также тарифы  $T_Y$  и  $T_{G_k}$  на полосу и интенсивность трафика в единицу времени. Положим далее, что  $i$ -е станции могут отслеживать статистические оценки – скользящие средние – интенсивности нагрузки  $\langle G_{ikt} \rangle$  и качества обслуживания  $\langle \mathbf{W}_{ikt}(Y_{it}) \rangle$  при текущем значении ресурса полосы  $Y_{it}$ .

Некоторой  $i$ -й станции ( $i$ -му игроку),  $i \in \{1, 2, \dots, N_t\}$ , требуется найти (заказать) на заданном отрезке времени  $t \in [t_s, t_{s+1})$  такое допустимое значение ресурса полосы  $Y_{it}$ , которое максимизирует его «выигрыш»:

$$\sum_{k=1}^K \{H[\langle G_{ikt} \rangle, \langle \mathbf{W}_{ikt}(Y_{it}) \rangle] - T_{G_k} \langle G_{ikt} \rangle\} - T_Y Y_{it} \rightarrow \max. \quad (25)$$

Варьируя заказываемый ресурс  $Y_{it}$  и передаваемый трафик или делая ставки  $\mathfrak{R}_{it}$  и  $G_{ikt}$ , игроки (станции) стремятся максимизировать свои выигрыши (25). Такие процедуры моделируются бескоалиционными динамическими играми  $N$  лиц [10].

В свою очередь, администратор сети, изменяя тарифы за ресурсы  $T_Y$  и за трафик  $T_{G_k}$ , может вступать в игру с пользователями с целью максимизации своего выигрыша:

$$\sum_{i=1}^{N_t} \sum_{k=1}^K T_{G_k} \langle G_{ikt} \rangle + \sum_{i=1}^{N_t} T_Y Y_{it} \rightarrow \max. \quad (26)$$

Игры (25) и (26) обеспечивают саморегулирование входных нагрузок  $G_{ikt}$ , требуемых параметров трафика и качества обслуживания  $[\mathbf{W}_{ikt}] = ([D_{ikt}], [R_{ikt}], [P_{ikt}], [M_{ikt}])$  и динамического распределения ресурсов полосы для каждого  $k$ -го класса данных, каждой  $i$ -й станции и каждого малого интервала времени  $[t_s, t_{s+1})$ .

Задачи (25) и (26) составляют аналитическую базу для динамического управления распределением ресурсов полосы и дифференцированными показателями качества обслуживания QoS в сетях 3G/4G.

**Заключение.** Получены численно-аналитические решения задач анализа и оптимизации вероятностно-временных характеристик, а также динамического управления распределением ресурсов полосы и качеством обслуживания QoS на уровнях управления множественным доступом для реалистичных условий функционирования мультифункциональных и мультисервисных спутниковых, мобильных и беспроводных сетей 3G/4G.

Эти решения составляют теоретическую базу для создания MAC-технологий с расширенными функциональными возможностями и передовыми характеристиками, а именно мультифункциональных и мультисервисных MAC-технологий (MFMAC) [1] с возможностью динамического распределения ресурсов полосы и гарантированного поддержания дифференцированных ВВХ и динамических показателей качества сервисов для каждого  $k$ -го класса данных, каждой  $i$ -й станции и каждого малого интервала времени  $[t_s, t_{s+1})$ . Технологии MFMAC особенно актуальны для построения беспроводных коммуникаций новых поколений (3G/4G) с радикально распределенной (grid-образной) архитектурой в странах с большой территорией, труднодоступными регионами, малой и неравномерной плотностью населения (Россия, Китай, Бразилия, Австралия, Индонезия и т. п.).

## СПИСОК ЛИТЕРАТУРЫ

1. **Мархасин А. Б.** Перспективные широкополосные беспроводные технологии для будущих глобальных мобильных и спутниковых систем персональной связи 4G // Электросвязь. 2003. № 8. С. 20.
2. **Roberts J. W.** Traffic theory and the Internet // IEEE Commun. Magazine. 2001. 39, N 1. P. 94.
3. **Мархасин А. Б.** Анализ интегрального телетрафика и проектирование мобильных сетей 3G // Электросвязь. 2002. № 12. С. 3.
4. **Мархасин А. Б.** Об анализе вероятностно-временных характеристик интегральных радиосетей множественного доступа // Тез. докл. 2-й Всесоюз. конф. по информационным системам множественного доступа. Минск, 1991. Ч. 1. С. 146.
5. **Мархасин А. Б.** Множественный доступ с динамическим управлением трафиком и качеством обслуживания в широкополосных сетях ATM // Автометрия. 1996. № 3. С. 97.
6. **Джейсуол Н.** Очереди с приоритетами: Пер. с англ. М.: Мир, 1973.
7. **Мархасин А. Б.** О суммировании нестационарных потоков с последствием в информационных сетях // Системное моделирование: Сб. тр. Новосибирск: ВЦ СО АН СССР, 1984. С. 98.
8. **Watson K. S.** Performance evaluation of cyclic service strategies – A survey // Proc. of the X Intern. Symp. "Performance'84". N. Y., 1984. P. 521.
9. **Мархасин А. Б.** Потенциальная пропускная способность сетей множественного доступа с временным разделением // Тез. докл. 1-й Всесоюз. конф. по информационным системам множественного доступа. Минск, 1989. С. 102.
10. **Дюбин Г. Н., Суздаль В. Г.** Введение в прикладную теорию игр. М.: Наука, 1981.

*Поступила в редакцию 5 февраля 2007 г.*

---