

На правах рукописи

ФИНСТЕРБУШ

Штефан

**ПРИОРИТЕТНЫЕ ДИСЦИПЛИНЫ ОБСЛУЖИВАНИЯ
МУЛЬТИСЕРВИСНОГО ТРАФИКА СAMEL-УСЛУГ**

05.13.13 – Телекоммуникационные системы и компьютерные сети

Автореферат

диссертации на соискание ученой степени

кандидата технических наук

Санкт-Петербург

2009

Работа выполнена в Санкт-Петербургском государственном университете телекоммуникаций им. проф. М.А. Бонч-Бруевича на кафедре систем коммутации и распределения информации.

Научный руководитель: доктор технических наук,
профессор Б.С. Гольдштейн

Официальные оппоненты: доктор технических наук,
профессор А.Е. Кучерявый

кандидат технических наук
А.А. Зарубин

Ведущая организация: научно-технический центр Протей

Защита состоится «___» _____ 2009 г. в ___ часов на заседании диссертационного совета Д 219.004.02 при Санкт-Петербургском Государственном Университете Телекоммуникаций им. проф. М.А. Бонч-Бруевича по адресу: 191186, Санкт-Петербург, наб. р. Мойки, 61.

С диссертацией можно ознакомиться в библиотеке университета.

Автореферат разослан «___» _____ 2009 г.

Ученый секретарь диссертационного совета,
к.т.н., доцент

В.Х. Харитонов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. С начала коммерческой эксплуатации сетей подвижной связи наблюдалось постоянное увеличение количества и улучшение качества услуг мобильной связи. В это вносили свой вклад новые телекоммуникационные технологии и новые методы предоставления услуг.

Первоначальный спектр услуг GSM (Global System for Mobile communications) был аналогичен перечню услуг узкополосной фиксированной связи и отличался только спецификой мобильности пользователей. В процессе развития сетей подвижной связи, как, впрочем, и фиксированных сетей, начались разработки сервисных платформ VAS (Value Added Services), повышающих функциональную ценность традиционных услуг и обеспечивающих создание, дополнение и конфигурацию разно-образных новых инфокоммуникационных услуг. Эти платформы явились функциональной частью сетей подвижной связи 2.5G и, в частности, системы GSM, но они не стандартизированы и могут предоставлять услуги исключительно "своим" абонентам и тогда, когда они находятся в "родной" сети. Именно попытки снять это ограничение в предоставлении VAS породили технологию CAMEL (Customized Application for Mobile Network Enhanced Logic).

Технология CAMEL является частью концепции виртуальной домашней среды VHE (Virtual Home Environment) и гарантирует получение широкого спектра новых (не стандартизованных, уникальных для каждого Оператора связи) инфокоммуникационных услуг за границами той домашней сети, где они организованы. Благодаря CAMEL пользователю в любое время становится доступно множество персонализированных услуг, независимо от выбранного им в это время окончного устройства или Оператора сети. Далее предоставление этих услуг реализуется с учетом местонахождения участника.

Переход от стандарта GSM к UMTS (Universal Mobile Telecommunication System) включает в себя новую технологию IP Multimedia Subsystem (IMS), которая дает возможность предлагать клиентам сетей подвижной связи услуги, базирующиеся на приложениях, услугах и протоколах Интернет, а также на WEB-технологиях. В рамках IMS сохраняет свою актуальность и проблема роуминга услуг, решаемая CAMEL.

В процессе предоставлении услуг в сетях связи GSM/UMTS важную роль играет оборудование, называемое Home Location Register / Home Subscriber Server (HLR/HSS). В связи с тем, что в будущем продолжится расширение функций этих сетевых элементов, актуально исследование производительности и характеристик HLR/HSS. Более того, разнообразие услуг связи в мультисервисной сети приводит к тому, что алгоритмы обслуживания заявок в HLR/HSS при внедрении CAMEL, переходе к 3G и реализации концепции IMS должны ориентироваться на разнородный трафик с разными допустимыми значениями времени обслуживания разных заявок. Для обслуживания такого трафика требуется применение приоритетных дисциплин обслуживания, так как классические алгоритмы предоставления ресурсов не в полной мере учитывают своеобразие обслуживаемых запросов,

что оказывает отрицательное влияние на использование обслуживающего устройства HLR/HSS и, в конечном итоге, на производительность всей системы.

В связи с этим разработана модель для исследования вероятностно-временных характеристик (VBX) и методов организации обработки мультисервисного трафика CAMEL-услуг в HLR/HSS мобильной связи является актуальной и своевременной.

Цель работы и задачи исследования. Целью работы является исследование VBX и разработка методов организации обработки мультисервисного трафика CAMEL-услуг в HLR/HSS сети подвижной связи GSM/UMTS на базе приоритетных дисциплин.

Эта цель определила необходимость решения следующих задач:

1. Анализ трафика разных CAMEL-услуг в сетях подвижной связи.
2. Построение структурно-функциональной модели реализации CAMEL-услуг в регистре HLR на основе приоритетных дисциплин.
3. Разработка математической модели HLR/HSS при организации CAMEL-услуг со статической смешанной абсолютно-относительной приоритетной системой.
4. Исследования VBX систем HLR/HSS с использованием смешанных абсолютно-относительных приоритетных дисциплин обработки очередей и оптимизации их характеристик путем выбора числа уровней абсолютного приоритета и размещения среди них значений относительного приоритета при заданных ограничениях длительностей обслуживания.
5. Исследование VBX систем HLR/HSS с использованием динамических значений приоритета по методу Earliest Deadline First (EDF), который соблюдает, прежде всего, допустимые сроки обслуживания поступивших в систему заявок.
6. Сравнительный анализ VBX классических узлов услуг HLR и систем с использованием динамических приоритетных дисциплин класса EDF.

Состояние исследования. В ряде работ по CAMEL-услугам подробно анализируются спецификации протокола MAP (Mobile Application Part). Но все эти работы носят эмпирический (инженерный) характер и не предполагают оценок VBX.

Исследования приоритетных дисциплин с абсолютными и относительными уровнями приоритета выполнены целым рядом авторов в 1980-2000 годах и обобщены в диссертации, но эти исследования не отражают последних достижений компьютерной техники при построении HLR (а затем – HSS), позволяющих в быстродействующих серверах HLR реализовывать динамические назначения приоритета в реальном времени, чему посвящена глава 3. Там же анализируются опубликованные работы по EDF, но эти работы не учитывают специфику сценариев HLR/HSS в мультисервисной сети.

Методы исследования. В процессе исследования использованы методы теории телетрафика, включая приоритетные модели обслуживания.

Научная новизна диссертационной работы заключается в предложенной модели систем с приоритетными дисциплинами обработки мультисервисного трафика CAMEL-услуг в регистре HLR/HSS, в новых методах и результатах исследования процесса предоставления CAMEL-услуг и в анализе BBX задержек в регистре HLR (и, соответственно, в HSS) при предоставлении CAMEL-услуг в сети GSM/UMTS.

Личный вклад. Все результаты, составляющие содержание диссертационной работы, получены автором самостоятельно.

Практическая ценность и реализация результатов. Полученные формулы, методы, алгоритмы позволяют найти наилучшие параметры и вычисления характеристик, определяющих качество предоставления CAMEL-услуг в регистре HLR с использованием приоритетных дисциплин обслуживания. Результаты работы используются в Научно-техническом центре ПРОТЕЙ при разработке CAMEL-шлюза, в Санкт-Петербургском центре разработки ПО компании ООО «Нокиа Сименс Нетворкс» при разработке алгоритмов обслуживания заявок в узле управления услугами (Service Control Point) в сетях GSM/UMTS.

Апробация работы. Основные результаты работы докладывались и были одобрены на 59-й и 60-й научно-технических конференциях студентов, аспирантов и молодых специалистов СПбГУТ, на 59-й, 60-й и 61-й научно-технических конференциях профессорско-преподавательского состава научных сотрудников и аспирантов СПбГУТ, на конференции «Joint Advanced Student School: The Turbo Principle in Communications», на конференции IEEE EUROCON2009, а также на заседаниях кафедры Систем коммутации и распределения информации в 2005, 2006, 2007, 2008 и 2009 годах.

Публикации. Материалы, отражающие основное содержание и результаты диссертационной работы, опубликованы в изданиях научно-технических конференций и в журналах отрасли - всего в 7 работах.

Структура и объем диссертации. Диссертационная работа включает в себя содержание, список сокращений, список обозначений, введение, четыре главы, заключение и список литературы. Объем пояснительной записки - 110 страниц, 43 иллюстрации, 5 таблиц, 4 приложения. Список использованной литературы насчитывает 116 наименований.

Основные положения, выносимые на защиту:

1. Функциональная модель HLR/HSS, отражающая принципы реализации услуг связи при использовании дисциплины обработки очередей с динамическим назначением приоритета.

2. Математическая модель системы организация услуг, учитывающая свойства поступающей нагрузки.

3. Алгоритм вычисления характеристик статической приоритетной системы с фиксированными значениями приоритета (метод Rate Monotonic Scheduling, RMS) и результаты расчета с использованием метода EDF.

4. Сравнительный анализ характеристик двух методов приоритетного обслуживания заявок.

5. Инженерный алгоритм назначения приоритета в CAMEL-шлюзе.

СОДЕРЖАНИЕ ПОЯСНИТЕЛЬНОЙ ЗАПИСКИ

Во введении обосновывается актуальность темы, рассматривается состояние исследуемой проблемы, формулируется цель работы, перечисляются основные научные результаты диссертации и её краткое содержание. Приводятся основные положения, выносимые на защиту.

Первая глава работы посвящена обзору эволюции услуг связи в сетях GSM/UMTS и методам их исследования с использованием теории телетрафика.

Как показано в этой главе, в сетях подвижной связи постоянно растет объем административных данных, которые требуются для поддержки новых инфокоммуникационных услуг и которыми сетевые элементы HLR/HSS должны управлять, сохранять их, рассматривать во время реализации услуг и передавать в соответствующие объекты. В связи с постоянным ростом количества пользователей нагрузка обслуживаемого устройства увеличивается так, что обработка заявок в допустимом диапазоне времени может быть не гарантирована. Более того, взаимосвязь HLR/HSS с относительно большим количеством сетевых элементов, в том числе вследствие подключения ряда разных контент-провайдеров, ведет к непредсказуемому потоку разнородных заявок на HLR/HSS, что и обуславливает важность и актуальность выполненных исследований.

В настоящее время мощная вычислительная технология допускает применение сложных алгоритмов для составления такого плана обработки разнородных задач (разной значимости), чтобы оптимально воспользоваться ограниченными системными ресурсами.

Вводится математический аппарат для моделирования систем с приоритетным обслуживанием по широко используемому в операционных системах реального времени методу Rate-monotonic Scheduling (RMS). Проведенный обзор известных результатов для систем M/G/1 позволяет использовать их в главах 2 и 3 при разработке оригинальных моделей времени пребывания запросов в очереди к HLR/HSS, чему и посвящена данная диссертация.

В соответствии с проведенными исследованиями формулируются цель и задачи диссертационной работы.

Во второй главе вводится функциональная модель HLR/HSS. Согласно спецификации 3GPP TS 29.02 для всех объектов в сети подвижной связи стандарта GSM, пользующихся протоколом MAP, в частности, для регистра HLR, применяется следующий метод работы при перегрузке: если обнаружена перегрузка объекта MAP, то перегруженному отвечающему объекту, например, регистру HLR, разрешено игнорировать запросы определенных операций MAP. Решение, какие именно запросы можно игнорировать, принимает Оператор сети (MAP service provider).

Далее в спецификации 3GPP TS 29.02 предлагается назначение уровня приоритета по прикладному контексту (Application Context) MAP-сообщений для регистра HLR, который выступает в качестве отвечающего устройства. На

рис. 1 представлена идея связи источников запросов с регистром HLR и обслуживания разнородных запросов в порядке значимости запроса.



Рис. 1. Функциональная схема организации взаимодействия источников запросов с регистром HLR/HSS

При этом уровни приоритета, как правило, определяются по виду вызова, породившего запрос, а иногда и по тому, через какой интерфейс этот запрос поступает. В HLR/HSS запросы поступают через интерфейсы, как представлено на рис. 2. Разумеется, количество потоков запросов с разными уровнями приоритета может не совпадать с количеством интерфейсов, т.к. запросы одного и того же вида могут поступать в систему через множество входов или интерфейсов. Но определенная корреляция имеется; что же касается общего числа уровней приоритета для разных потоков запросов, то оно определяется из математической модели в этой главе.

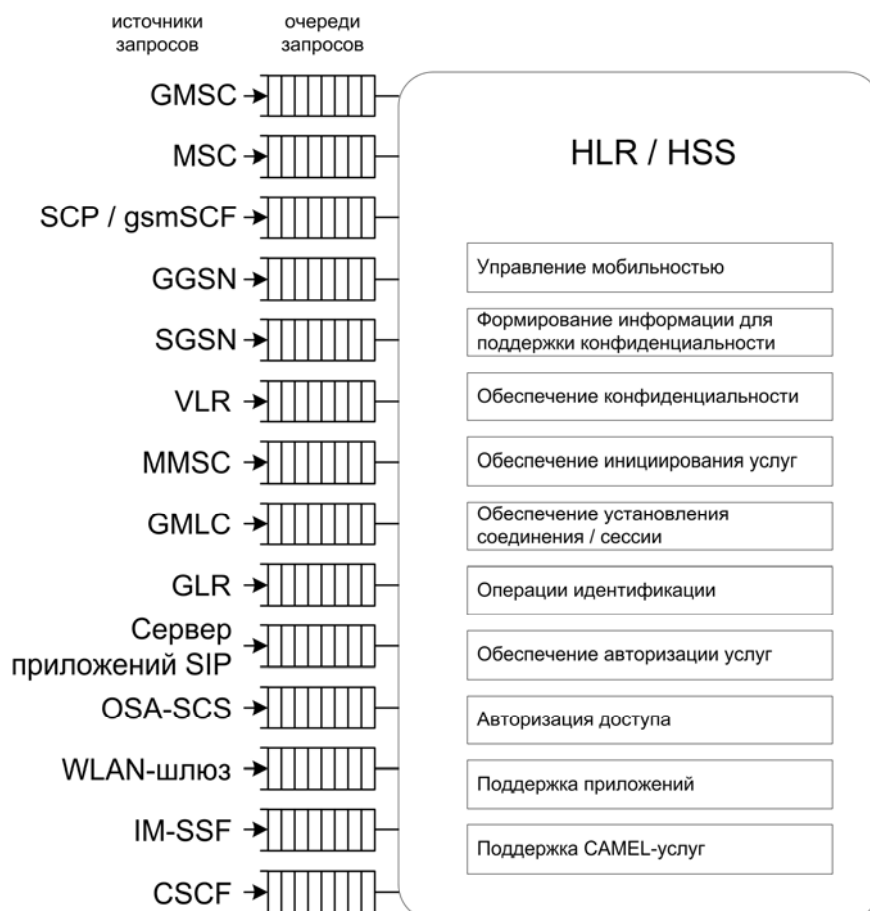


Рис. 2. Функциональная модель HLR/HSS и связанных с ним сетевых элементов

Рис. 2 еще раз иллюстрирует центральное положение HLR/HSS в сети сигнализации. Разумеется, Оператору сети целесообразно оптимально эксплуатировать ресурсы и именно сервера HLR/HSS. Привилегированное обслуживание запросов определенных сетевых элементов или всеобщего интерфейса могло бы входить в намерения Оператора сети, чтобы иметь возможность предоставить определенной услуге некие гарантии использования ресурса HLR/HSS.

На основе функциональной модели и анализа методов исследования предоставления CAMEL-услуг в регистре HLR/HSS, проведенного в первой главе, определяется математическая модель, позволяющая исследовать ВВХ HLR/HSS с использованием дисциплин обслуживания со статическими значениями приоритета.

Для описания дисциплины обслуживания со смешанными абсолютно-относительными уровнями/значениями приоритета представим значения приоритета в виде пары чисел (k, m) . Например, имеется N потоков запросов, которым поставлены в соответствие N значений приоритета.

Пусть эти N значений распределены каким-то образом по K уровням абсолютного приоритета. Обслуживание запроса любого приоритета уровня k ($k=2,3,\dots,K$) прерывается при появлении другого запроса, приоритет которого соответствует меньшему k ($k=1,2,\dots,K-1$). На каждом уровне предусмотрено M_k статических значений приоритета, и запросы с одним и тем же его значением не прерывают друг друга, а обслуживаются по принципу FIFO. Тогда каждое значение приоритета можно описывать в виде пары чисел (k, m) . На уровне k таких пар будет $(k, 1), (k, 2), \dots, (k, M_k)$.

$$\text{Очевидно, что } \sum_{k=1}^K M_k = N. \quad (1)$$

Разбиение вида $\Phi = \{M_1, M_2, \dots, M_K\}$ однозначно определяет приоритетную организацию обслуживания запросов. Приоритетная организация обслуживания в HLR/HSS может зависеть, например, от типа вызовов, порождающих запросы, и/или от интерфейсов с источниками запросов и поступающего потока.

На рис. 3 показана организация приоритетного обслуживания запросов в HLR/HSS. Здесь a – потоки запросов, прерывающих обслуживание запросов с приоритетом (k, m) , b – потоки запросов, не прерывающих обслуживание запросов с приоритетом (k, m) , v – потоки запросов, обслуживание которых не прерывается при появлении запроса с приоритетом (k, m) , z – потоки запросов, обслуживание которых прерывается при появлении запроса с приоритетом (k, m) .

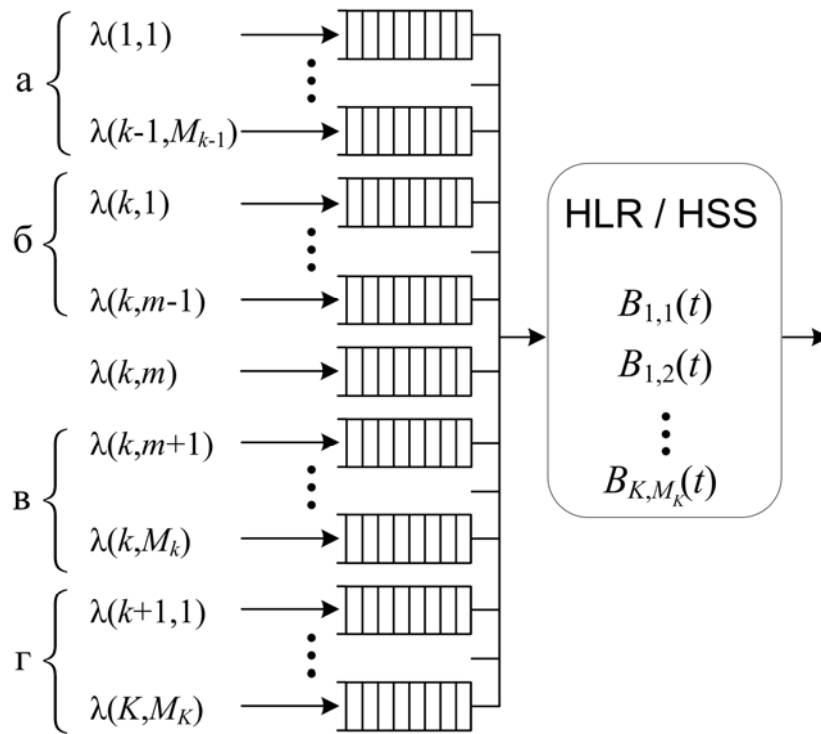


Рис. 3. Организация обслуживания запросов в HLR/HSS со смешанными абсолютно-относительными уровнями/значениями приоритета

Очевидно, что математическое ожидание времени $T(k, m)$ пребывания запросов с приоритетом (k, m) в HLR/HSS складывается из среднего времени $W(k, m)$ ожидания начала обслуживания запроса с приоритетом (k, m) и среднего времени обслуживания $V(k, m)$ запросов с приоритетом (k, m) , что выражается следующим образом:

$$T(k, m) = W(k, m) + V(k, m). \quad (2)$$

Потоки запросов являются независимыми пуассоновскими потоками с параметрами $\lambda(k, m)$. Время обслуживания запросов приоритета (k, m) – независимая величина с функцией распределения $B_{k, m}(t)$, математическим ожиданием $b(k, m)$ и вторым начальным моментом $b^{(2)}(k, m)$. Размеры очередей запросов выбираются таким образом, чтобы исключить возможность потери запросов из-за отсутствия мест для ожидания. Прерывание обслуживания запроса не приводит к потере уже полученных при его обслуживании результатов и данных.

Рассматривается только стационарный режим работы системы, т.е. суммарная интенсивность нагрузки обслуживающего устройства HLR/HSS

$$\rho = \sum_{i=1}^K \sum_{j=1}^{M_i} \lambda(i, j) b(i, j) = \sum_{i=1}^K \sum_{j=1}^{M_i} \rho(i, j) < 1. \quad (3)$$

Для этих условий получено выражение, позволяющее оценить среднее время ожидания начала обслуживания запроса с приоритетом (k, m) :

$$W(k, m) = \frac{\frac{1}{2} \sum_{i=1}^k \sum_{j=1}^{\varphi(k, m)} \lambda(i, j) b^{(2)}(i, j) + \left[1 - \sum_{i=1}^{k-1} \sum_{j=1}^{M_i} \rho(i, j) \right] \sum_{j=m+1}^{M_k} \rho(k, j) v(k, j)}{\left[1 - \sum_{i=1}^k \sum_{j=1}^{\varphi(k, m-1)} \rho(i, j) \right] \cdot \left[1 - \sum_{i=1}^k \sum_{j=1}^{\varphi(k, m)} \rho(i, j) \right]}, \quad (4)$$

где $v(k, j)$ – математическое ожидание времени обслуживания запроса с приоритетом (k, j) при наличии в HLR/HSS запроса с приоритетом (k, m) . $\varphi(k, m)$ – множество потоков от $\lambda(1, 1)$ до $\lambda(k, m)$ включительно.

Выражения (2) и (4) позволяют производить оценку характеристик функционирования HLR/HSS. Следует заметить, что полученные формулы справедливы и для систем с "чистыми" абсолютными или относительными значениями приоритета. Действительно, при $K = N$ и $M_i = 1, \forall i$, имеет место обслуживание с абсолютными, а при $K = 1$ и $M_i = N$ – с относительными значениями приоритета. Таким образом, изучаемые традиционно системы только с абсолютными и только с относительными значениями приоритета являются частными случаями рассмотренной системы.

На рис. 4 приведены графики зависимости математического ожидания времени T пребывания в системе запроса с приоритетом r для приоритетной стратегии $\Phi = \{3, 2, 5, 6, 4\}$ при разных значениях суммарной нагрузки ρ и при $1 \leq r \leq N$, причем $N = 3 + 2 + 5 + 6 + 4 = 20$.

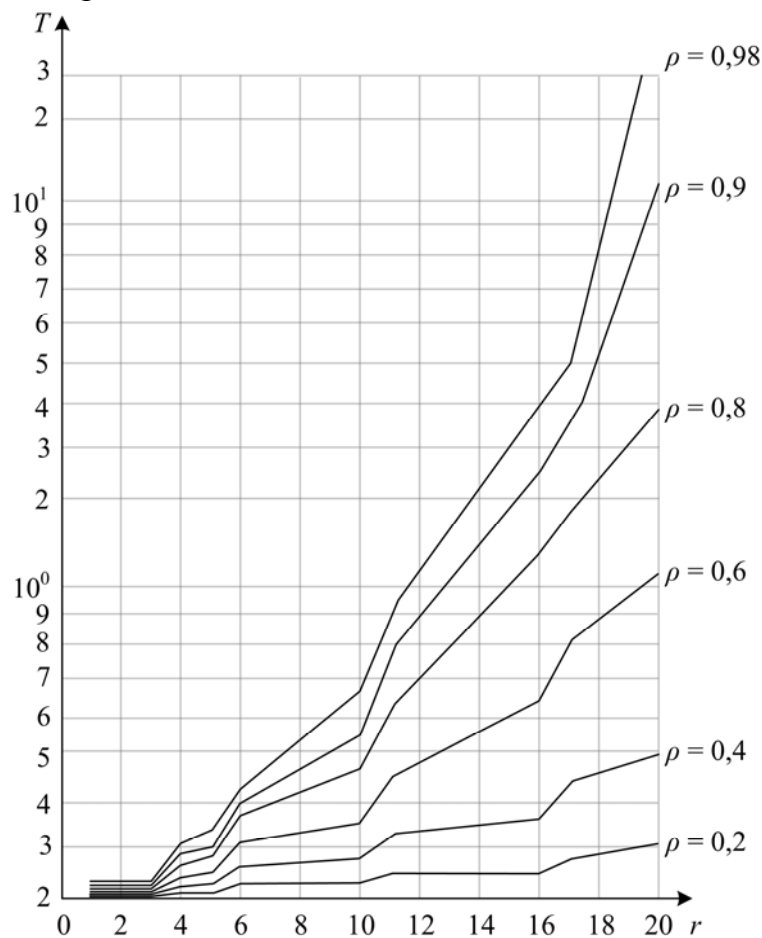


Рис. 4. Зависимость среднего времени T пребывания запроса в системе от значения r его приоритета

Основным преимуществом такой организации приоритетности является сохранение почти неизменным среднего времени пребывания в HLR/HSS запросов с высшими уровнями приоритета (в данном примере: $r \leq 3$) при увеличении суммарной нагрузки HLR/HSS.

Результатом анализа в главе 2 является получение формул для расчета среднего времени пребывания запроса в системе со смешанными абсолютно-относительными уровнями/значениями приоритета.

В третьей главе исследуется метод управления выбором запросов из очереди Earliest Deadline First (EDF), который оперирует динамическими значениями приоритета и соблюдает, прежде всего, допустимые сроки обслуживания поступивших в систему заявок.

EDF присваивает наибольший уровень приоритета запросу, у которого осталось наименьшее время до срока завершения обслуживания. Для этого необходимо определить директивный срок d , до истечения которого запрос обязательно должен быть обслужен. В дальнейшем, этот *срок допустимой задержки* будем называть "дедлайн". После любого события в системе, изменяющего набор не обслуженных запросов, планировщик задач назначает высший уровень приоритета задаче с *ближайшим окончанием срока ее исполнения* (англ.: *earliest deadline*).

Представлен математический подход к вычислению времени пребывания заявок в системе HLR/HSS с дисциплиной обслуживания EDF, базирующийся на анализе среднего значения. Этот математический аппарат позволяет быстро оценить поведение системы, в которую поступает произвольное количество потоков с ограничением по дедлайну. Для этого сначала исходим из предположения, что в систему поступает всего два потока запросов с разным дедлайном. На основе найденных правил формулируется общий вышеназванный математический аппарат.

Предполагаем, что классы запросов расположены по уровню приоритета $r \in \{1, 2, \dots, R\}$, причем $r = 1$ – это самый высокий уровень приоритета и, соответственно, $r = R$ – самый низкий. Если запросы относятся к одному и тому же классу, или, иными словами, к одному уровню приоритета, то внутри класса порядок обслуживания запросов – по принципу FIFO.

Среднее время W_r обслуживания запросов класса r состоит из трех компонентов:

$$W_r = W_0 + \sum_{i=1}^R \bar{N}_{i,r} \frac{1}{\mu_i} + \sum_{i=1}^R \bar{M}_{i,r} \frac{1}{\mu_i} \equiv W_0 + W_N + W_M \quad (5)$$

- а) среднее оставшееся время W_0 обслуживания запроса, находящегося в обслуживающем устройстве;
- б) среднее время W_N обслуживания запросов из очереди, которые обслуживаются до рассматриваемого запроса. Это запросы того же или более высокого приоритета относительно приоритета r рассматриваемого запроса;

с) среднее время W_M обслуживания запросов более высокого приоритета, которые поступают в систему во время ожидания обслуживания рассматриваемым запросом.

Эта формула справедлива также при методе EDF, так как она действительна в одноканальных системах массового обслуживания (СМО) для всех дисциплин обслуживания, не зависящих от времени обслуживания запросов. В связи со спецификой метода обслуживания EDF необходимо определить как среднее число $\bar{N}_{i,r}$ запросов приоритета i , которые уже находятся в очереди при поступлении рассматриваемого запроса приоритета r , так и среднее число $\bar{M}_{i,r}$ запросов приоритета i , которые поступают в систему во время ожидания обслуживания рассматриваемым запросом приоритета r . Символ μ_i обозначает интенсивность обслуживания запросов приоритета i .

В качестве примера использования алгоритма для определения переменных $\bar{N}_{i,r}$ и $\bar{M}_{i,r}$ представим систему с двумя потоками запросов ($R = 2$). Предполагаем, что запросы потока $r = 1$ имеют меньше, чем запросы второго потока, допустимого времени для обслуживания запроса или, вернее в этом контексте, до начала обслуживания запроса: $d_1 < d_2$, так как не предполагается прерывания обслуживания запросов. При методе EDF время обслуживания запросов не должно влиять на определение динамического значения приоритета каждого запроса в очереди (и в обслуживающем устройстве).

Поступающий запрос потока $r = 1$ должен ждать окончательного обслуживания всех запросов того самого потока $r = 1$, уже находящихся в системе:

$$\bar{N}_{1,1} = \lambda_1 W_1, \quad (6)$$

где, в общем виде, λ_r – интенсивность поступления запросов потока r и, сравнивая с формулой (1), W_r – среднее время обслуживания запросов потока r .

Соответственно, запросы, поступающие после рассматриваемого запроса потока $r = 1$, обслуживаются только тогда, когда этот запрос будет обслужен.

$$\bar{M}_{1,1} = 0, \quad \bar{M}_{2,1} = 0. \quad (7)$$

Из всех запросов потока $r = 2$, находящихся в системе (число их, в среднем, равно $\lambda_2 W_2$), обслуживаются только те, которые поступили раньше рассматриваемого поступающего запроса потока $r = 1$ и у которых абсолютный дедлайн меньше или равен дедлайну рассматриваемого запроса:

$$\bar{N}_{2,1} = \lambda_2 \cdot \max(0, W_2 - D_{2,1}). \quad (8)$$

В этом случае запросы потока должны были бы ожидать своего (окончательного) обслуживания не менее времени $D_{2,1} = d_2 - d_1$. Вероятность того, присутствуют ли вообще запросы потока $r = 2$ в системе после промежутка $D_{2,1}$, зависит в значительной степени от среднего времени пребывания запросов этого потока в очереди W_2 .

Рассмотрим теперь тот случай, когда поступает запрос потока $r = 2$. В соответствии с условием, что $d_1 < d_2$, поступающий запрос потока $r = 2$ должен ждать окончательного обслуживания всех запросов потока $r = 1$, уже находящихся в системе:

$$\bar{N}_{1,2} = \lambda_1 W_1. \quad (9)$$

Кроме того, запросы потока $r = 1$, которые поступают в систему не позже, чем в промежутке $D_{2,1}$ и до среднего времени W_2 , обслуживаются до этого запроса потока $r = 2$:

$$\bar{M}_{1,2} = \lambda_1 \cdot \min(W_2, D_{2,1}). \quad (10)$$

Аналогично с предыдущим случаем, получим:

$$\bar{N}_{2,2} = \lambda_2 W_2, \quad \bar{M}_{2,2} = 0. \quad (11)$$

Для $\bar{N}_{i,r}$ и $\bar{M}_{i,r}$ формируется общее выражение в зависимости от класса r рассматриваемого запроса:

$$\bar{N}_{i,r} = \begin{cases} \lambda_i W_i, & 1 \leq i \leq r \\ \lambda_i \cdot \max(0, W_i - D_{i,r}), & r < i \leq R \end{cases}, \quad (12)$$

$$\bar{M}_{i,r} = \begin{cases} \lambda_i \cdot \min(W_r, D_{r,i}), & 1 \leq i < r \\ 0, & r \leq i \leq R \end{cases}. \quad (13)$$

Интенсивность нагрузки ρ_r обслуживающего устройства при обработке запросов потока r составляет $\rho_r = \lambda_r / \mu_r$. Применяя вышеизложенные результаты к формуле (5), получим, что среднее время пребывания запросов потоков $r = 1$ и $r = 2$ в очереди при $R = 2$:

$$W_1 = W_0 + \rho_1 W_1 + \rho_2 \cdot \max(0, W_2 - D_{2,1}),$$

$$W_2 = W_0 + [\rho_1 W_1 + \rho_2 W_2] + \rho_1 \cdot \min(W_2, D_{2,1}).$$

Общее выражение для среднего времени пребывания запросов каждого потока r в очереди:

$$W_r = W_0 + \left[\sum_{i=1}^r \rho_i W_i + \sum_{i=r+1}^R \rho_i \max(0, W_i - D_{i,r}) \right] + \sum_{i=1}^{r-1} \rho_i \min(W_r, D_{r,i}) \quad (14)$$

С учетом закономерности $\min(a, b) = a - \max(0, a - b)$ среднее время пребывания запросов потока r в очереди выражается таким образом:

$$\begin{aligned} W_r &= W_0 + \left[\sum_{1 \leq i \leq R} \rho_i W_i - \sum_{r < i \leq R} \rho_i W_i \right] + \sum_{r < i \leq R} \rho_i \max(0, W_i - D_{i,r}) + \sum_{1 \leq i < r} \rho_i \min(W_r, D_{r,i}) \\ &= \left[W_0 + \frac{\rho W_0}{1 - \rho} \right] - \sum_{r < i \leq R} \rho_i [W_i - \max(0, W_i - D_{i,r})] + \sum_{1 \leq i < r} \rho_i \min(W_r, D_{r,i}) \\ &= \left[\frac{W_0}{1 - \rho} - \sum_{r < i \leq R} \rho_i \min(W_i, D_{i,r}) \right] + \sum_{1 \leq i < r} \rho_i \min(W_r, D_{r,i}) \end{aligned} \quad (15)$$

Таким образом, формула (14) стала приемлемой для итеративного вычисления. В предположении, что все W_k известны для $L < k \leq R$, выражение в квадратных скобках последней строки формулы (15) становится постоянной величиной C , а величины a_k и b_k определяются решением уравнения вида:

$$X = C + \sum_{1 \leq k < L} a_k \min(X, b_k) \quad \text{при } R \geq L > 1. \quad (16)$$

Уравнение (16) имеет единственное решение $Z = \lim_{n \rightarrow \infty} X_n$, где ряд $\{X_n\}_{n \in \mathbb{N}}$ определяется через $X_0 = 0$, и

$$X_{n+1} = C + \sum_{1 \leq k < L} a_k \min(X_n, b_k), \quad (17)$$

причем учитывается закон сохранения для СМО типа M/G/1:

$$\sum_{r=1}^R \rho_r W_r = \begin{cases} \frac{\rho W_0}{1 - \rho}, & \rho < 1 \\ \infty, & \rho \geq 1 \end{cases}, \quad (18)$$

где ρ – интенсивность суммарной нагрузки, составленной всеми ρ_r . Для любой системы M/G/1 и любой, сохраняющей результаты работы до ее прерывания, дисциплины обслуживания должно выполняться равенство (18). Иными словами, формула (17) служит для последовательного расчета W_r , начиная с W_R до W_2 , по итерационному методу. После этого, с помощью равенства (18), вычисляется последний элемент W_1 вектора решения $\{W_r\}_{r=1..R}$.

Вместо длинного математического доказательства того, что элементы ряда $\{X_n\}$ являются как положительными, так и стабильно возрастающими с увеличением индекса n и стремятся к верхнему порогу, здесь приводится график (рис. 5) итерационного процесса по выражению (17).

График на рис. 5 – результат расчета конкретного числового примера для метода EDF с 5-ю потоками разного приоритета. В каждом случае расчет содержит 9 итерационных циклов. Конечный результат итерационного цикла приблизительно представляет собой значение среднего времени пребывания запросов потока r в очереди при определенной нагрузке обслуживающего устройства. В виде примера, среднее время пребывания запросов высшего приоритета в очереди составляет приблизительно 0,9 единиц времени, а низшего – около 7,7 единиц времени.

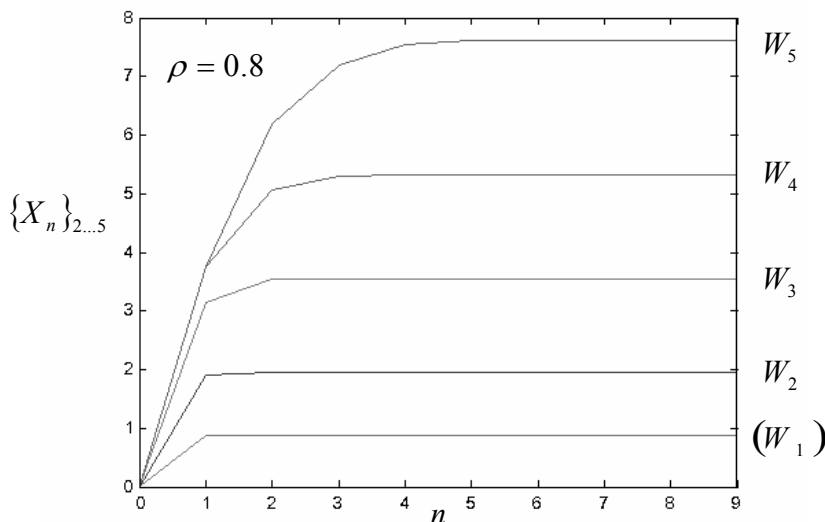


Рис. 5. Результат итерационного процесса для определения среднего времени пребывания в очереди запросов пяти потоков

Проведено сравнение двух фундаментальных методов обслуживания со статическими и с динамическими значениями приоритета: метод Rate

Monotonic Scheduling (RMS), проведен в параграфах 1.3 и 1.4, и метод Earliest Deadline First (EDF). Сравнение включает в себя признаки: способность к планированию задач, сложность технической реализации, стабильность при кратковременной и постоянной перегрузке.

В четвертой главе на основании формул, полученных в третьей главе, обсуждаются эффективность метода EDF, результаты научного эксперимента и вопросы практической реализации CAMEL-услуг.

С этой целью рассмотрено развитие функциональной структуры CAMEL по фазам и, в связи с этим, повышение нагрузки HLR. В процессах обслуживания запроса, при отправке SMS, при сессиях GPRS или при других специфических для CAMEL функциях используются необходимые функциональные единицы и описывается их взаимодействие (на высоком уровне). Названные функциональные единицы показаны на рис. 6.

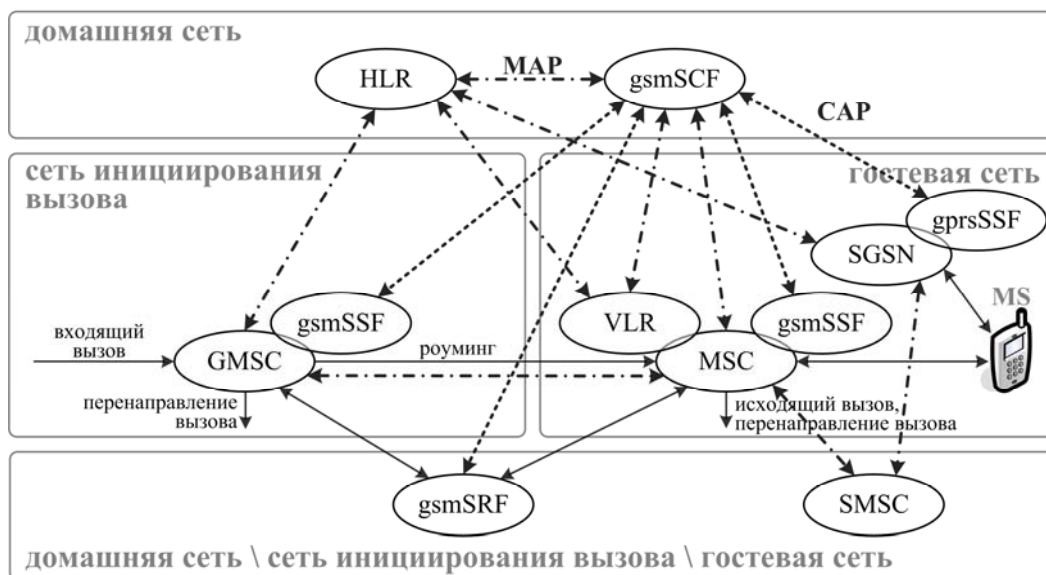


Рис. 6. Общая схема взаимодействия функциональных объектов CAMEL Phase 3

CAMEL Subscription Information (CSI) – это часть данных об абонентах, которая указывает, может ли абонент пользоваться услугами CAMEL, и если да, то какими. При перемещении абонента или при модификации данных CSI сетевой элемент HLR пересылает абонентские данные CSI к VLR гостевой сети, а GPRS-CSI и SMS-CSI – к SGSN. Непосредственно логика услуги сосредоточена в узле gsmSCF, имеющем интерфейсы с узлами gsmSSF, HLR, MSC (гостевой сети) и gsmSRF.

Один абонент может получать несколько CSI, например: O-CSI, T-CSI, GPRS-CSI, SS-CSI и до 20 U-CSI. Все CSI поддерживаются и администрируются Оператором домашней сети.

CSI содержат параметры:

- Trigger Detection Point
- gsmSCF address
- Service Key
- Default Call Handling/Default GPRS Handling/Default SMS Handling

gsmSCFaddress – это адрес E.164 для доступа к gsmSCF (Global Title of the SCP). Логика Service Key (в SCP) идентифицирует услугу gsmSCF. Default Call Handling устанавливает, должна ли продолжаться или прерываться связь в случае сбоя, распознанного в gsmSSF или gsmSCF. То же самое действует для сессий GPRS, контекста PDP, SMS-передачи, Default GPRS Handling и Default SMS Handling.

CSI сохранены в HLR, кроме N-CSI, находящегося в GMSC и, соответственно, в VMSC. На рисунке 7 представлено, какие CSI при определении местонахождения абонента или при общем обновлении данных посылаются в GMSC, VLR и SGSN.

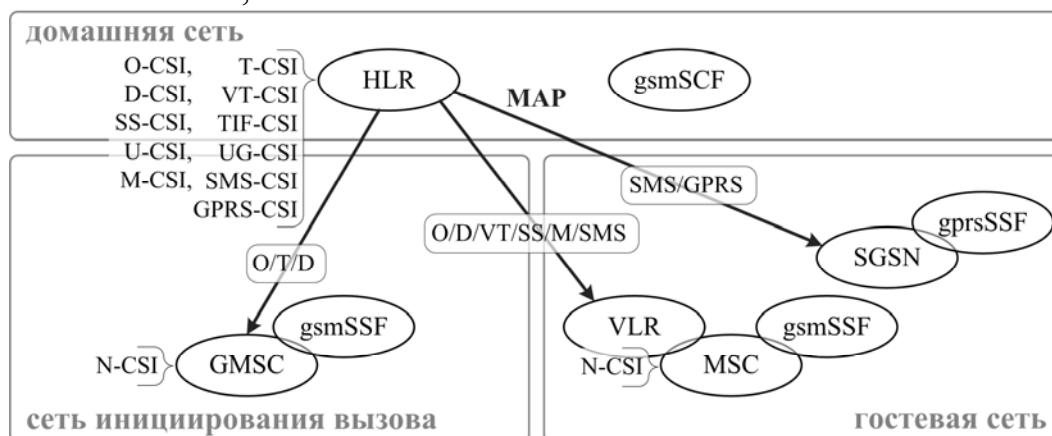


Рис. 7. Схематическое представление взаимодействия баз данных для хранения CSI и передача CSI между функциональными объектами CAMEL Phase 3

В спецификации 3GPP TS 29.02 представлено назначение уровней приоритета согласно прикладному контексту (Application Context) MAP-сообщений для регистра HLR, который выступает в качестве отвечающего устройства. Видно, что регистры учитывают операции MAP в зависимости от важности или приоритета прикладного MAP-контекста.

Как следует из назначения уровней приоритета MAP-сообщений, эти уровни определяются не по источнику запросов (в данном случае в качестве источников запросов выступают интерфейсы), а по виду запроса, независимо от того, через какой интерфейс этот запрос поступает. Иными словами, запрос одного и того же вида может поступать в систему через множество входов.

В названной спецификации выделяются пять функциональных категорий MAP-сообщений с разным приоритетным значением. На основе этого определения в дальнейшем исходим из пятиприоритетной классификации заявок, поступающих в систему или, соответственно, в HLR/HSS, для расчетного анализа характеристик моделей RMS и EDF приоритетного обслуживания.

Автором диссертационной работы с помощью компьютерной программы MATLAB создана программа для имитационного моделирования и сравнения различных методов с приоритетным обслуживанием запросов. Исходные данные для имитационного моделирования представлены в формате MATLAB и кратко описаны в табл. 1:

```

N           = 5;
lampda_rel_r = [.5 .3 .1 .075 .025];
X_mean_r    = [2 1 .5 1.5 3];
d_r         = [1 1.5 3 5 10];
no_tasks    = 20000;
roh         = [.2:.2:.6 .65:.05:.9 .92:.02:1];
Var_rel_mu  = .15;
    
```

Таблица 1: Описание исходных данных для имитационного моделирования

Величина	Описание
N	количество потоков или уровней приоритета
lampda_rel_r	отношение значений интенсивности поступления запросов
X_mean_r	среднее время обслуживания n разных запросов
d_r	дэдлайны
no_tasks	количество запросов для моделирования
roh	разрешение оси
Var_rel_mu	фактор дисперсии относительно X_mean_r

5 элементов векторов ($lampda_rel_r$, X_mean_r , d_r) отображают значения параметра соответствующих 5 потоков. При этом первым стоит элемент со значением, соответствующим потоку с наивысшим приоритетом ($r = 1$) и так далее.

На рис. 8 видно, что при методе EDF (сплошные кривые) запросы потоков $r = \{2,3,4,5\}$ быстрее обрабатываются и покидают системы раньше, чем при методе RMS (пунктирные кривые), но за счет характеристики потока с наивысшим приоритетом $r = 1$.

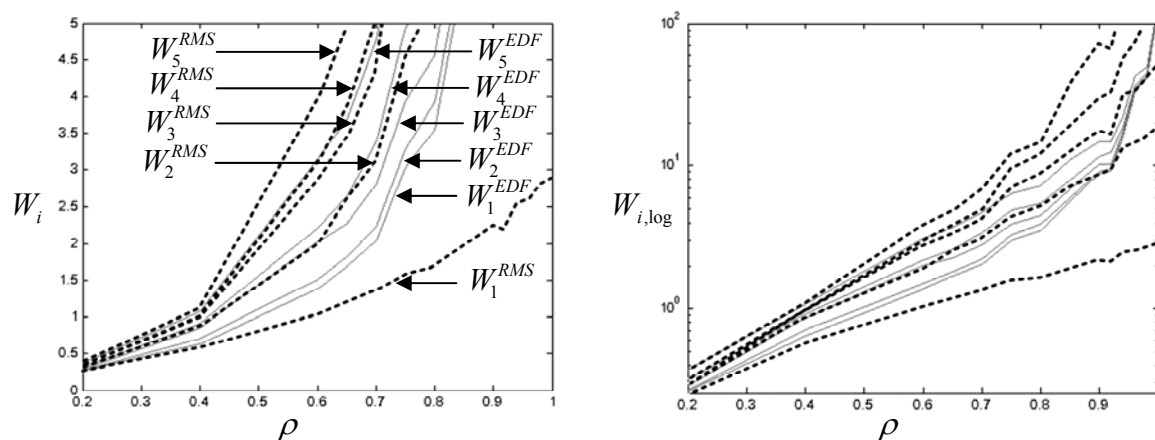


Рис. 8. Среднее время пребывания в очереди запросов разного приоритета или, для EDF, запросов разных потоков

При большой нагрузке системы ($\rho > 0,8$) метод EDF не исключает запросы низких приоритетов из обслуживания, как при методах со

статическими уровнями приоритета (RMS), потому что достигнутые дедлайны низкоприоритетных запросов требуют их обработку до обслуживания недавно поступивших запросов более высокой категории.

Из этого следует, что появление новых запросов в системе при большой нагрузке не приводит (или реже приводит) к приоритетному прерыванию текущего обслуживания. В противоположность этому у метода RMS количество приоритетных прерываний растет линейно с ростом нагрузки. Это видно на рис. 9, где иллюстрировано количество приоритетных прерываний запросов разной значимости $r = \{2,3,4,5\}$ в зависимости от нагрузки обслуживающего устройства по сравнению с суммарным количеством приоритетных прерываний всех запросов (пунктирные линии).

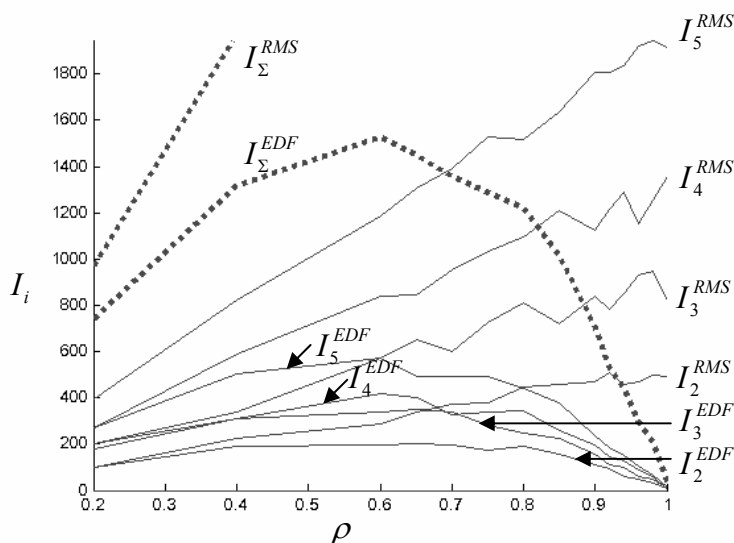


Рис. 9. Количество приоритетных прерываний обработки запросов

Метод EDF при увеличении нагрузки требует значительно меньше приоритетных прерываний обслуживания, чем метод RMS. Именно такое прерывание и сохранение состояния прерванного запроса для последующего продолжения обслуживания являются административными накладными расходами, которые тоже занимают ресурсы HLR/HSS и потому отрицательно влияют на производительность системы.

ЗАКЛЮЧЕНИЕ

В процессе проведенных в диссертационной работе исследований получены следующие основные результаты:

1. Произведен анализ трафика разных CAMEL-услуг в сетях подвижной связи.
2. Разработана структурно-функциональная модель реализации CAMEL-услуг в регистре HLR на основе приоритетных дисциплин.
3. Разработана математическая модель HLR/HSS при организации CAMEL-услуг со статической смешанной абсолютно-относительной приоритетной системой.
4. Проведено исследование ВВХ систем HLR/HSS, использующих смешанные абсолютно-относительные приоритетные дисциплины

обслуживания, и оптимизации их характеристик путем выбора числа уровней абсолютного приоритета и размещения на них значений относительного приоритета при заданных ограничениях длительностей обслуживания.

5. Проведено исследование ВВХ систем HLR/HSS с использованием динамических значений EDF-приоритета.

6. Проведен сравнительный анализ ВВХ классических узлов услуг HLR и систем с использованием динамических приоритетных дисциплин класса EDF.

7. Разработана программа имитационного моделирования алгоритмов RMS и EDF для сравнительного анализа их характеристик.

8. Проведенные исследования доказывают, что применение приоритетного алгоритма обслуживания запросов с дедлайнами в HLR/HSS является обоснованным подходом к оптимизации обслуживания запросов.

СПИСОК РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Финстербуш Ш. Имитационное моделирование систем очередей CAMEL-ориентированных услуг с помощью компьютерной программы SimEvents (MATLAB) // 59-я Научно-техническая конференция профессорско-преподавательского состава научных сотрудников и аспирантов: Тез. докл. – СПбГУТ. СПб, 2007. - с. 19.

2. Финстербуш Ш. Модель HLR в мультисервисной CAMEL-сети // 60-я Научно-техническая конференция профессорско-преподавательского состава научных сотрудников и аспирантов: Тез. докл. – СПбГУТ. СПб, 2008. - с. 22.

3. Финстербуш Ш. Модели СМО с приоритетной дисциплиной обслуживания // 60-я Научно-техническая конференция профессорско-преподавательского состава научных сотрудников и аспирантов: Тез. докл. – СПбГУТ. СПб, 2008. - с. 22-23.

4. Финстербуш Ш., Потапов Д.А., Ньянг Б.Д. 2-я попытка для организации услуг 3-й стороны // Вестник связи, М., 2008. - №4. (из перечня ВАК РФ)

5. Финстербуш Ш. Приоритетные варианты стратегии обслуживания заявок в HLR/HSS // Мобильные телекоммуникации, М., 2008. - №9.

6. Финстербуш Ш. Сравнение двух приоритетных дисциплин для CAMEL-услуг // 61-я Научно-техническая конференция профессорско-преподавательского состава научных сотрудников и аспирантов: Материалы – СПбГУТ. СПб, 2009. - с. 27-28.

7. Finsterbusch S., Mathematical analysis of priority model EDF for HLR / HSS // EUROCON 2009, International IEEE Conference, СПб, 2009.

Подписано к печати 30.10. 2009

Тираж 80 экз. Объем 1 печ. л. Заказ № 34

Тип. СПбГУТ, 191186 СПб, наб. р. Мойки, 61

