

УДК 621.39

Системы $M|G|1$ с групповым обслуживанием и их применение к анализу модели протокола управления потоковой передачей. Часть II

Н. В. Першаков, К. Е. Самуйлов

*Кафедра систем телекоммуникаций
Российский университет дружбы народов
ул. Миклухо-Маклая, 6, Москва, Россия, 117198*

В статье изложены результаты разработки модели функционирования протокола управления потоковой передачей (Stream Control Transmission Protocol, SCTP) в виде системы с групповым обслуживанием заявок. Исследована степень адекватности классических моделей типа $M|G|1$ с обслуживанием группами переменной и фиксированной длины [1] применительно к анализу показателей качества обслуживания протокола SCTP. Проведён численный анализ параметров качества функционирования протокола для случая передачи сигнального трафика сети сотовой подвижной связи.

Ключевые слова: протокол управления потоковой передачей, система массового обслуживания, групповое обслуживание, задержка в очереди на передачу.

1. Введение

Протокол управления потоковой передачей SCTP — это транспортный протокол стека TCP/IP, обеспечивающий надёжную передачу данных пользователя в независимых упорядоченных потоках. Данный протокол, специфицированный в международном стандарте RFC 2960 [2], снимает известные ограничения существующих транспортных технологий TCP и UDP при передаче данных, относящихся к приложениям реального времени таких, как речь и потоковое видео.

Протокол SCTP является важным элементом концепции построения сетей следующего поколения (Next Generation Network, NGN), использующих IP-транспорт и ориентированных на предоставление пользователям услуг с высокими требованиями к показателям качества обслуживания (Quality of Service, QoS). Качество функционирования сетей NGN во многом зависит от своевременной доставки и надёжности передачи сигнальной информации. Системы сигнализации применяются не только для управления процессом установления соединений пользователей традиционной телефонии, но также обеспечивают возможность предоставления услуг интеллектуальных сетей и сетей сотовой подвижной связи (СПС). Основной системой передачи сигнальной информации в этих сетях является общеканальная система сигнализации № 7 (ОКС7) [3]. На базе этой системы практически полностью построен стандарт GSM, где по ОКС7 помимо сигнальных сообщений (СС), необходимых для управления процессом установления соединения, передаются также данные роуминга и сообщения SMS.

Параметры QoS, например, время установления соединения, воспринимаются непосредственно пользователем, при этом качество большинства существующих услуг связи и пользовательских приложений в значительной мере определяется качеством функционирования сети сигнализации. Следует отметить, что сеть ОКС7, являясь фактически сетью пакетной коммутации, логически отделена от инфраструктуры базовой информационной сети связи и разделяет с ней лишь некоторые физические сетевые ресурсы, например, системы передачи. В технологии ATM (Asynchronous Transfer Mode) этот недостаток был практически преодолен за счёт использования уровня адаптации, который позволяет эмулировать каналы передачи данных — звенья сигнализации (ЗС) между узлами сети сигнализации (пунктами сигнализации) в виде виртуальных каналов сети ATM. В силу своей специфики сигнализация ATM применима только в широкополосных сетях

с интеграцией служб, что ограничивает возможности передачи сигнального трафика по IP-сетям.

Сети IP, обладая практически неограниченными транспортными возможностями, до последнего времени не обеспечивали требований, предъявляемых к передаче сигнального трафика. Рабочая группа IETF (Internet Engineering Task Force), получившая название Sigtran (Signaling Transport), разработала одноимённое решение, в котором в качестве транспорта был использован протокол SCTP, изначально призванный обеспечить надёжную доставку СС, управление потоками и защиту от ошибок по принципу «из конца — в конец». Следует отметить, что данное решение предполагает возможность совместной передачи сигнального трафика и других типов данных, в первую очередь, речи по IP-сети. Анализ функционирования протокола SCTP показывает, что, по существу, он является системой массового обслуживания (СМО) с групповым обслуживанием и таймером, ограничивающим время формирования группы заявок [4]. Результаты разработки модели протокола, изложенные в разделе 2 данной статьи, были опубликованы в [5, 6]. Построенная модель оказалась достаточно сложной для анализа существующими аналитическими методами, поэтому для её исследования, наряду с методом статистического моделирования, применены известные и некоторые новые результаты, изложенные в первой части данной статьи [1] и использованные в настоящей, второй части статьи при проведении численного анализа параметров QoS протокола SCTP. Изложению основных аспектов организации численного эксперимента и анализу адекватности аналитических моделей посвящены разделы 3 и 4 статьи, где в качестве объекта эксперимента выбран фрагмент сети СПС, актуальный для современных операторов мобильной связи [5, 7].

2. Построение модели протокола управления потоковой передачей

Протокол SCTP имеет ряд отличительных особенностей. Во-первых, он предусматривает возможность как надёжной (TCP), так и ненадёжной (UDP) доставки данных с возможностью поддержки порядка передачи внутри каждого из организуемых потоков (функция Multistreaming). За счёт этого достигается эффект независимой передачи несвязанных данных, например, относящихся к различным услугам. Основным преимуществом такого подхода является тот факт, что возникновение ошибки в одном из потоков и возможная потеря качества соответствующей ему услуги не оказывает влияние на своевременную доставку данных в остальных потоках и на показатели качества других услуг. Во-вторых, в случае необходимости надёжность передачи контролируется механизмами выборочных подтверждений, повторных передач и перенаправления потока с недоступного IP-адреса на резервный адрес (функция Multihoming), что особенно актуально в условиях дефицита времени на возобновление процесса передачи вследствие отказов оборудования и перегрузок.

Рассмотрим более подробно процесс передачи по протоколу SCTP [8]. Данные (сообщения в терминах протокола SCTP), поступающие от приложения пользователя, инкапсулируются в порции данных (Chunks), которые поступают в очередь на передачу (рис. 1). Приложение предварительно производит настройку параметров SCTP-соединения, включая необходимое количество потоков, организуемых между участвующими в сеансе связи сторонами. В том случае, если длина сообщения превышает максимальный размер блока передачи, оно разбивается на фрагменты, помещаемые в отдельные порции. В очереди порции данных собираются в SCTP-пакеты, причём время формирования каждого пакета ограничено таймером, который запускается в момент поступления в него первой порции данных. Таймер в соответствии с техническими характеристиками оборудования таких производителей, как Cisco и Ericsson, может принимать значения от 0 до 1000 мс. Передача сформированных пакетов, т.е. заполненных пакетов или пакетов, для которых истёк таймер сборки, осуществляется в независимых потоках по IP-сети.

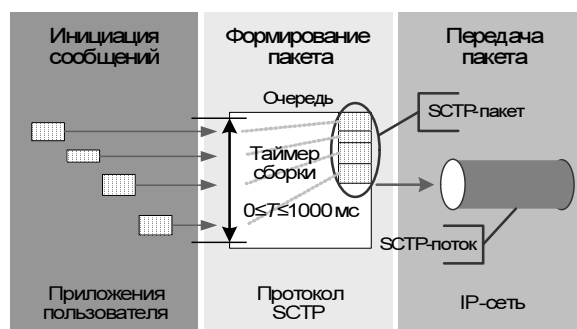


Рис. 1. Схема формирования SCTP-пакета

Проводя аналогии с системами массового обслуживания, легко видеть, что процесс передачи сообщения по протоколу SCTP может быть описан в терминах СМО с групповым обслуживанием [1]. Действительно, в такой системе под заявкой естественно понимать порцию данных, группы заявок — пакет, обслуживающего прибора — поток SCTP. В силу независимости SCTP-потоков количество приборов можно положить равным единице, а поскольку ресурсы современного оборудования в режиме нормального функционирования практически исключают переполнение буферной памяти, то ёмкость накопителя системы следует считать неограниченной.

Обобщая вышесказанное, рассмотрим СМО, кодируемую в виде $GI_{l_\lambda}|GI^{[X,T]}|1|(R, L)$, на которую поступает поток заявок с интенсивностью λ . Длина l_λ поступающей заявки является дискретной случайной величиной (СВ) с распределением $\gamma_m = P\{l_\lambda = m\}$, $m = \overline{1, M}$. Обслуживание заявок производится группами в порядке поступления, причём ёмкость накопителя рассчитана на R групп, а заявки, обслуживаемые прибором, не сохраняют за собой место в накопителе, одновременно покидая систему в момент окончания обслуживания группы. Время формирования каждой группы заявок ограничено таймером $0 \leq T \leq \infty$, а длина X группы является дискретной СВ, максимальное значение которой равно L . Обслуживание группы начинается либо в момент освобождения прибора, если группа сформирована, либо в момент завершения формирования группы, если прибор не занят обслуживанием. Времена обслуживания групп заявок распределены по произвольному закону с функцией распределения (ФР) $B(x)$.

Непосредственный анализ модели протокола SCTP в виде СМО $GI_{l_\lambda}|GI^{[X,T]}|1|(R, L)$ затруднён отсутствием известных авторам статьи аналитических методов, применимых для исследования систем данного типа. Тем не менее, как показано в первой части данной статьи [1], в некоторых важных частных случаях эта СМО была изучена в ряде работ зарубежных авторов. Заметим, что система $M|G^B|1$ с переменной длиной группы заявок во введённой нотации записывается как $M_1|G^{[K,0]}|1|(\infty, K) = M|G^{[K,0]}|1|\infty$, а система $M|G^k|1$ с фиксированной длиной группы — как $M_1|G^{[K,\infty]}|1|(\infty, K) = M|G^{[K]}|1|\infty$.

Обращаясь к современным задачам, стоящим перед операторами сетей СПС, необходимо отметить, что одной из ключевых проблем при переходе к сетям NGN является обеспечение качества передачи трафика реального времени по транспортной IP-сети. Сложность проблемы усугубляется значительным ростом объёмов трафика, обусловленного популярностью новых услуг, и трафика процедур, характерных для сетей СПС таких, как аутентификация и управление мобильностью. В этих условиях первоочередной становится задача передачи сигнального трафика с сохранением качества традиционной сети ОКС7 [3], развитие существующей инфраструктуры которой оказывается нецелесообразным по ряду известных технологических и экономических причин.

Как говорилось выше, в качестве решения данной задачи в IETF была разработана архитектура сигнального транспорта Sigtran, состоящая из протокола

управления потоковой передачей SCTP и уровней адаптации, например, адаптера M3UA, как показано на рис. 2. Процесс передачи СС состоит в следующем. Сообщения, поступающие от подсистемы ОКС7, например, с сетевого уровня (MTP3), преобразуются на уровне адаптации к формату протокола SCTP, который, в свою очередь, обеспечивает надёжную и своевременную передачу сигнального трафика. Таким образом, осуществляется замена традиционного транспорта сети ОКС7 — уровня звена данных сигнализации (MTP2) на IP-транспорт [7].

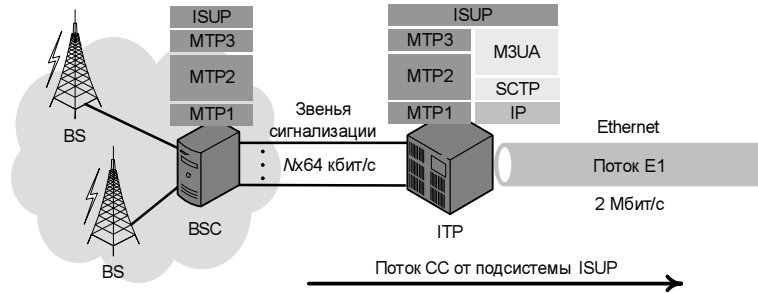


Рис. 2. Фрагмент сети СПС

В стандартах Международного союза электросвязи (International Telecommunications Union, ITU) для ОКС7 нормируются среднее значение и 95%-квантиль СВ времени передачи СС на уровне MTP2 [9]. Эта СВ соответствует времени передачи порций данных на уровне протокола SCTP, поскольку уровни адаптации в технологии Sigtran предоставляют интерфейс управления подсистемам, находящимся выше звена данных MTP2 в иерархии стека протоколов ОКС7. Таким образом, среднее значение и 95%-квантиль задержки передачи СС (порций данных) на уровне протокола SCTP и являются предметом дальнейшего анализа в данной статье.

Итак, мы показали, как при решении задачи анализа показателей QoS сети СПС строится модель протокола SCTP в виде СМО с групповым обслуживанием заявок. При анализе вероятностно-временных характеристик (ВВХ) этой системы может быть использован аналитический аппарат СМО типа $M|G|1$ [1]. В общем случае СМО $GI_{\lambda}|GI^{[X,T]}|1| \langle R, L \rangle$ применяется имитационная модель [4, 8], разработка которой не представляет существенной сложности для специалистов в области теории массового обслуживания.

Далее мы переходим к численному анализу, обращая внимание читателей на вычислительные аспекты ВВХ аналитических моделей, и показываем степень адекватности этих моделей исследуемому протоколу для приложений, возникающих в сетях СПС, с помощью результатов имитационного моделирования [4, 5, 7, 8].

3. Организация вычислений вероятностных характеристик

Перед тем, как перейти непосредственно к численному анализу, приведём алгоритм расчёта ВВХ системы $M|G^{[K,0]}|1|\infty$ с переменной длиной группы и $M|G^{[K]}|1|\infty$ с фиксированной длиной группы — среднего значения m_V и дисперсии σ_W^2 времени ожидания начала обслуживания. Алгоритм составлен с использованием результатов, полученных в первой части статьи [1], где также введены все необходимые обозначения.

Шаг 1. Исходные данные:

- максимальный размер группы K ;
- интенсивность входящего потока λ ;
- ФР времени обслуживания групп заявок $B(x)$.

Шаг 2. Расчёт значений промежуточных параметров:

- значений величин z_i , $i = \overline{1, K-1}$ в круге $|z| < 1$ как корней уравнения $z^K - \beta(\lambda - \lambda z) = 0$ по формуле

$$z_k = \sum_{l=1}^{\infty} c_l w_k^l, \quad k = \overline{0, K-1},$$

где

$$w_k = e^{2\pi k i / K} \text{ и } c_l = \frac{1}{2\pi i l} \int_{|z| \leq 1} \frac{\beta^{l/K} (\lambda - \lambda z)}{z^l} dz, \quad l \geq 1;$$

- коэффициентов r_i , $i = \overline{0, K}$ разложения произведения $\prod_{i=0}^{K-1} (z - z_i)$ в ряд по степеням z ;
- значений величин β_j , $j = \overline{0, K-1}$ по формуле $\beta_j = \int_0^{\infty} e^{-\lambda x} \frac{(\lambda x)^j}{j!} dB(x)$.

Шаг 3. Расчёт характеристик вложенной цепи Маркова (ЦМ) $\{x_n^-\}$:

- вероятностей q_j^- , $j = \overline{0, K-1}$ по формуле

$$q_j^- = \frac{1}{\beta_0} \left[\sum_{i=1}^j (\delta_{i,K} - \beta_i) q_{j-i}^- - \frac{K - \rho}{\prod_{i=1}^{K-1} (1 - z_i)} \sum_{i=0}^{\min(j,K)} \beta_{j-i} r_i \right],$$

где $\delta_{i,K}$ — символ Кронекера, $\rho = \lambda m_B$ — нагрузка на систему;

- среднего значения m_{X^-} и дисперсии $\sigma_{X^-}^2$ по формулам

$$m_{X^-} = \rho + \frac{\lambda^2 m_B^{(2)} - K(K-1)}{2(K-\rho)} + \sum_{i=1}^{K-1} \frac{1}{1 - z_i},$$

$$\begin{aligned} \sigma_{X^-}^2 = \sigma_A^2 + & \frac{\lambda^3 m_B^{(3)} - K(K-1)(K-2)}{3(K-\rho)} + \\ & + \frac{\lambda^2 m_B^{(2)} - K(K-1)}{2(K-\rho)} + \left(\frac{\lambda^2 m_B^{(2)} - K(K-1)}{2(K-\rho)} \right)^2 - \sum_{i=1}^{K-1} \frac{z_i}{(1 - z_i)^2}, \end{aligned}$$

где $m_B^{(n)}$ — n -й начальный момент распределения времени обслуживания $B(x)$.

Шаг 4. Расчёт вероятностных характеристик СМО $M|G^{[K,0]}|1|_{\infty}$ и $M|G^{[K]}|1|$:

- среднего значения m_W времени ожидания начала обслуживания по формулам

$$m_W = \frac{1}{q_0^- + \rho} \left[m_{X^-} \cdot m_B + \frac{f}{\lambda} \right], \quad \text{где } f = \frac{\rho^2}{2} (C_B^2 - 1),$$

для СМО $M|G^{[K,0]}|1|_{\infty}$ и

$$m_W = \frac{\lambda m_B^{(2)} - m_B(K-1)}{2(K-\rho)} + \frac{1}{\lambda} \sum_{i=1}^{K-1} \frac{1}{1 - z_i}$$

для СМО $M|G^{[K]}|1|$.

– дисперсии σ_W^2 времени ожидания начала обслуживания по формуле

$$\sigma_W^2 = \frac{1}{q_0^- + \rho} \left[\frac{1}{3} \lambda m_B^{(3)} - 2\rho \sigma_B^2 + \frac{m_B}{\lambda} (\sigma_{X^-}^2 - m_{X^-}) \right] + \frac{1}{\lambda^2 (q_0^- + \rho)^2} \times \\ \times [q_0^- m_{X^-} - (\rho m_{X^-} + 2f) - f^2]$$

для СМО $M|G^{[K,0]}|1|\infty$.

К алгоритму требуется сделать несколько замечаний. Во-первых, нетрудно показать, что матрицы переходных вероятностей вложенной ЦМ $\{x_n^-\}$ для обеих рассматриваемых систем совпадают. Это означает, что результаты, полученные на шаге 3 алгоритма, применимы на шаге 4 для вычисления вероятностных характеристик для СМО как с фиксированной, так и с переменной длиной группы. Во-вторых, на шаге 4 приведены выражения для характеристик времени ожидания заявки в очереди, хотя в ряде случаев необходимы характеристики её времени пребывания в системе. Это сделано для удобства представления результатов, так как вычисление данных характеристик при известных распределениях времени обслуживания групп заявок не представляет существенных сложностей.

Проиллюстрируем решение уравнения $z^K - \beta(\lambda - \lambda z) = 0$ на имеющем прикладное значение примере детерминированного обслуживания, которое будет использовано далее при расчёте параметров QoS протокола SCTP. На рис. 3 показаны корни этого уравнения при значениях параметров $K = 20$, $\rho = 2$ и $\rho = 8$, вычисленные с использованием формул шага 2 алгоритма, при разложении в круге $|z| \leq 1$ в ряд Фурье до $l = 25$ члена. Заметим, что в условиях детерминированного обслуживания коэффициенты c_l этого разложения принимают удобный для расчёта вид $c_l = e^{-l\theta} \frac{(l\theta)^{l-1}}{l!}$, $l \geq 1$, $\theta = \rho/K$. Легко видеть, что корни уравнения z_i , $i = \overline{1, 20}$ являются либо действительными, либо комплексно-сопряжёнными, что, как и следовало ожидать, приводит к действительным значениям вероятностных характеристик при подстановке в соответствующие выражения найденных комплекснозначных корней.

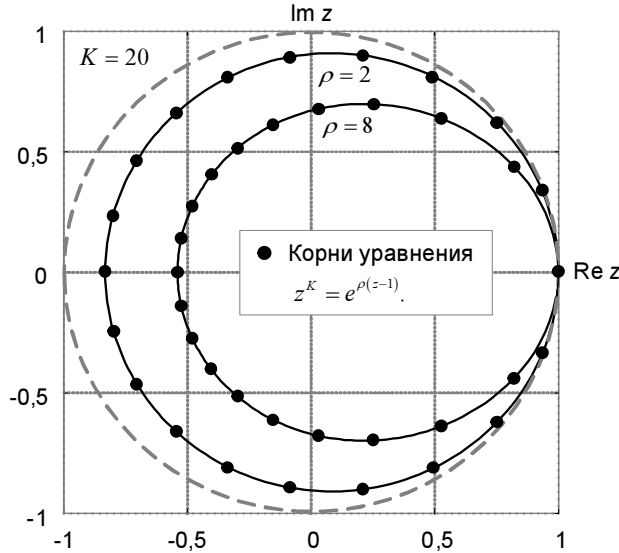


Рис. 3. Иллюстрация решения уравнения $z^K - \beta(\lambda - \lambda z) = 0$ в случае детерминированного обслуживания

4. Анализ адекватности аналитических моделей

Перейдём к описанию параметров модели протокола SCTP. Рассмотрим фрагмент сети на рис. 2, в котором установлен транзитный пункт сигнализации (IP Transfer Point, ИТР), поддерживающий технологию Sigtran. Данный узел предназначен для маршрутизации сигнального трафика от сегмента сети, состоящего из нескольких базовых станций (BS), управляемых контроллером (BSC), к узлам сети, относящимся к подсистеме коммутации СПС (MSC, VLR и HLR). Узлы BSC и ИТР соединены пучком ЗС, состоящим из N стандартных звеньев по 64 кбит/с каждое. Узел ИТР подключён к подсистеме коммутации по каналу E1 (2 Мбит/с), поверх которого реализована технология Ethernet, т.е. на этом участке транспорт СС реализован с использованием технологии Sigtran.

В настоящее время многими операторами, в том числе операторами СПС, эксплуатируются сети, первоначально созданные для передачи трафика коммутации каналов и построенные по технологии временного разделения канала и передачи сигнала, оцифрованного методом импульсно-кодовой модуляции. Как правило, такие операторы владеют обширной инфраструктурой сети, перевод которой на IP-технологию требует существенных инвестиций и временных затрат. Поэтому актуальной является задача обеспечения прозрачной доставки сигнального трафика, а также трафика услуг передачи голоса, данных и видео пользователю со сквозной гарантией качества независимо от транспортной сетевой инфраструктуры. В качестве одного из решений был предложен метод эмуляции сети Ethernet по каналам E1 первичного уровня цифровой иерархии, позволяющий сократить расходы и инвестиции в сетевую инфраструктуру, одновременно повышая конкурентоспособность и прибыльность компании.

Для вычисления времени обслуживания группы заявок, как следует из рекомендации ITU Q.706 [9], следует использовать детерминированное распределение. Известно, что длина заголовка кадра Ethernet H_{Ethernet} составляет 26 байт (для распространённых типов кадров Ethernet DIX и RAW 802.3), заголовка IP-датаграммы H_{IP} — 20 байт (для версии протокола IPv4 без учёта поля опций), заголовка SCTP-пакета H_{SCTP} — 12 байт и заголовка SCTP-порции H_{Chunk} — 16 байт. Тогда для вычисления параметра распределения времени обслуживания b (сек.) в аналитической модели можно использовать выражение

$$b = \frac{8 [(\bar{l}_{\text{Sigtran}} + H_{\text{Chunk}}) \bar{c} + H_{\text{SCTP}} + H_{\text{IP}} + H_{\text{Ethernet}}]}{2 \cdot 10^6}, \quad (1)$$

где \bar{c} — среднее число порций данных в SCTP-пакете и \bar{l}_{Sigtran} — средняя длина порции данных.

Поскольку максимальный размер блока передачи в сети Ethernet составляет 1500 байт, то для определения максимального числа K порций данных в SCTP-пакете (максимального значения длины группы заявок) можно воспользоваться соотношением

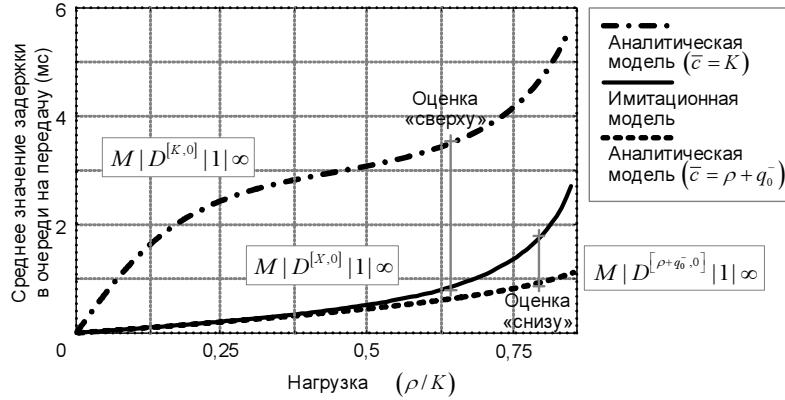
$$K = \left\lfloor \frac{1500 - (H_{\text{SCTP}} + H_{\text{IP}})}{\bar{l}_{\text{Sigtran}} + H_{\text{Chunk}}} \right\rfloor. \quad (2)$$

Без ограничения общности оценку области применения СМО $M|G^{[K,0]}|1|_{\infty}$ и $M|G^{[K]}|1|_{\infty}$ можно дать для модели трафика, в которой средняя длина порции данных равна 272 байтам — максимально возможной длине поля сигнальной информации подсистемы МТР в ОКС7. В соответствии с [2] при передаче по протоколу SCTP объём данных должен быть кратен 4 байтам. Следовательно, средняя длина порции данных \bar{l}_{Sigtran} без учёта её заголовка при передаче по протоколу SCTP также составляет 272 байта. Теперь из формулы (2) следует, что максимальное число порций в SCTP-пакете $K = 5$ и мы получили необходимые для дальнейших расчётов исходные данные.

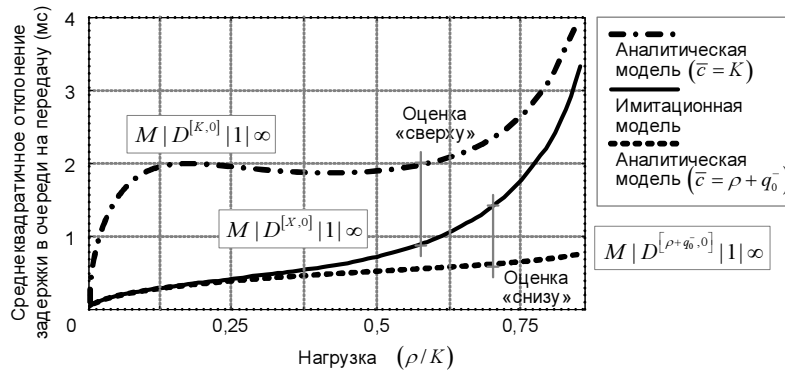
Графики на рис. 3–4 для анализа вероятностных характеристик, рассчитанные с помощью алгоритма предыдущего раздела данной статьи, приведены в сравнении с данными, полученными с помощью имитационной модели функционирования протокола SCTP, описание которой находится за рамками данной статьи.

Вкратце, основные результаты численного анализа моделей состоят в следующем:

- 1) для оценки вероятностных характеристик системы с таймером, равным нулю, при любом допустимом значении нагрузки ρ следует использовать аналитическую модель $M|G^{[K,0]}|1|\infty$ с переменной длиной группы заявок;
- 2) для оценки вероятностных характеристик системы с таймером, большим нуля, в диапазоне средних и высоких нагрузок следует использовать аналитическую модель $M|G^{[K]}|1|\infty$ с фиксированной длиной группы заявок.



(а) Среднее значение задержки в очереди на передачу



(б) Среднеквадратичное отклонение задержки в очереди на передачу

Рис. 4. Область применения модели $M|G^{[K,0]}|1|\infty$ с переменной длиной группы

Ниже содержится подробное обсуждение этих результатов. Итак, на рис. 3 для СМО $M|G^{[K,0]}|1|\infty$ показана зависимость среднего значения и среднеквадратичного отклонения времени ожидания в очереди от удельного значения нагрузки ρ/K . Для правильной трактовки представленных данных необходимо особо отметить различие в функционировании аналитической и имитационной моделей. На первый взгляд СМО $M|G^{[K,0]}|1|\infty$ должна соответствовать имитационной модели со значением таймера, равным нулю, т.е. СМО $M|G^{[X,0]}|1|\infty$. Тем не менее, время передачи пакета в канал в аналитической модели вычисляется по формуле (1) с заранее заданным числом \bar{c} порций в пакете, в то время как в имитационной адекватно протоколу SCTP число порций может изменяться от 1 до K в зависимости от длины сформированного к моменту начала передачи пакета.

Теперь очевидно, что СМО $M|G^{[K,0]}|1|\infty$ не может адекватно учитывать особенности формирования пакета по протоколу SCTP, что приводит к задаче определения нижней и верхней границы для оценки искомых характеристик путём выбора параметра \bar{c} и, соответственно, параметра b распределения длительности обслуживания в формуле (1).

В качестве значения параметра \bar{c} для нижней границы естественно принять среднее число СС в пакете, которое соответствует среднему числу заявок в группе и для СМО $M|G^{[K,0]}|1|\infty$ вычисляется по формуле $\bar{c} = \rho + q_0$. При низкой и средней нагрузке на сигнальное отношение аналитическая модель весьма точно описывает поведение протокола. При росте нагрузки все больше проявляется следующий эффект, приводящий к увеличению разницы между моделями. По протоколу SCTP пакеты, имеющие длину выше среднего, оказывают значительное влияние на время ожидания порций в очереди, поступивших за время передачи в канал, в то время как в аналитической модели этот эффект отсутствует, поскольку время передачи в канал одинаково для всех пакетов и не зависит от длины.

Для получения верхней границы достаточно положить число порций в пакете максимальным $\bar{c} = K$. При низкой нагрузке разница между моделями существенна, но при приближении нагрузки к максимально допустимой, и, следовательно, числа порций в пакете к значению параметра K вероятностно-временные характеристики имитационной модели стремятся к аналитическим, и разница между моделями быстро сокращается.

На рис. 4 показана зависимость среднего значения задержки в очереди на передачу от удельной нагрузки ρ/K для системы $M|G^{[K]}|1|\infty$ с фиксированной длиной группы. Анализ области применения этой системы следует проводить относительно имитационной модели с таймером, большим нуля. В условиях численного эксперимента для наглядности представления результатов значение таймера было выбрано $T=20$ мс. Это следует из очевидного факта: с некоторого значения нагрузки, а точнее с момента, когда в среднем завершение формирования пакета происходит раньше истечения таймера, аналитическая модель должна достаточно «хорошо» описывать поведение протокола SCTP. В этом случае основным фактором, сдерживающим начало передачи пакета в канал, является не значение таймера, а интенсивность поступления сообщений, что в полной мере отражает функционирование аналитической модели. Несмотря на все сказанное, аналитическая модель даёт лишь оценку снизу. Это связано с тем, что по протоколу SCTP передача пакета в канал при условии его доступности начинается не в момент поступления порции, которая оказывается последней в пакете, а следующей за ней с длиной, не позволяющей поместить её в формирующийся пакет. При этом очевидно, что разница между моделями не превышает среднего времени между последовательными поступлениями сообщений, т.е. λ^{-1} .

На рис. 5 показаны графики среднего значения и 95%-квантили задержки передачи СС (времени пребывания заявки в системе), полученные с помощью имитационного моделирования, где в качестве исходных данных использована модель сигнального трафика, разработанная в [10] на основе измерений в реальной сети связи. В этом случае средняя длина СС \bar{l}_{Sigtran} при передаче по протоколу SCTP составила 24 байта, а число порций в пакете $K = 36$. Значение таймера формирования пакета T , равное 11,8 мс, было получено экспериментальным путём с целью соответствия показателей качества передачи СС рекомендации ITU Q.706.

Из графика на рис. 6 видно, что среднее значение задержки имеет две экстремальные точки, определяющие три промежутка монотонности. При сравнительно низкой нагрузке (на рисунке эта область значений обозначена римской цифрой I) интенсивность поступления СС мала, время формирования пакета определяет таймер и, следовательно, задержка близка к значению таймера. С ростом нагрузки задержка быстро убывает до того момента, когда среднее число СС в пакете становится более одного. При последующем увеличении интенсивности поступления СС (область II) среднее время формирования пакета по-прежнему превышает среднее время передачи, которое ещё не достигло максимального значения. Этим обуславливается дальнейшее убывание задержки, что в сравнении с предыдущим

интервалом происходит медленнее, поскольку среднее число СС в пакете увеличивается, приводя к росту среднего значения времени передачи пакета в канал. После того как пакет достигает максимальной длины (область III), чем больше становится нагрузка, тем меньше на время формирования пакета влияет таймер. В данном промежутке аналогично предыдущему рассмотренному интервалу этот показатель превышает время передачи пакета в канал, что приводит к дальнейшему более быстрому уменьшению задержки.

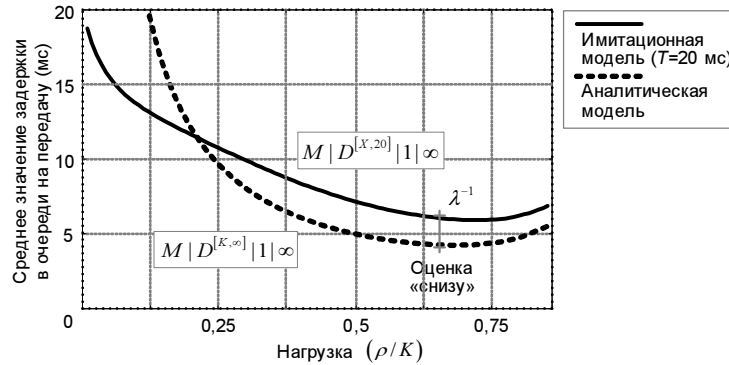


Рис. 5. Область применения модели $M | G^{[K]} | 1 | \infty$ с фиксированной длиной группы

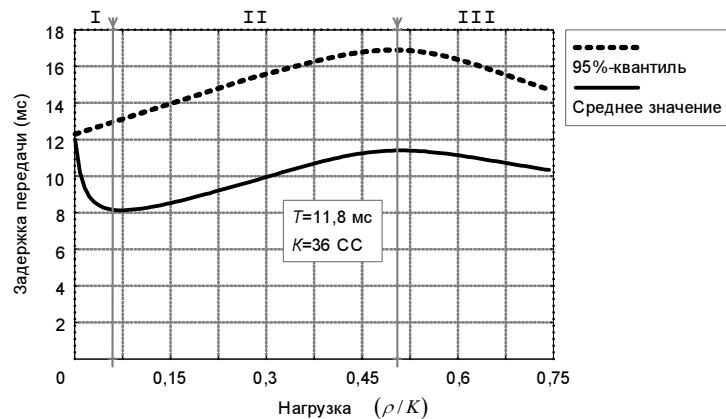


Рис. 6. Среднее значение и 95%-квантиль задержки передачи сигнальных сообщений

5. Заключение

В статье представлены результаты разработки математической модели протокола управления потоковой передачей в виде СМО с групповым обслуживанием и таймером ограничения времени формирования группы заявок. Разработан вычислительный алгоритм для расчёта вероятностных характеристик СМО с переменной и фиксированной длиной группы заявок. Исследована адекватность аналитических моделей к анализу показателей качества обслуживания протокола SCTP.

Область применения предложенных в статье аналитических моделей и методов для расчёта вероятностных характеристик протокола SCTP не ограничивается анализом возможностей реализации технологии Sigtran в сети сотовой подвижной связи. Полученные результаты могут быть без значительной модификации

перенесены на случай анализа любых других моделей трафика, передаваемого по протоколу SCTP, типов и пропускной способности канала, открывая широкие возможности для дальнейших исследований параметров качества функционирования современных и будущих мультимедийных приложений в сетях связи.

Литература

1. *Першаков Н. В., Самуйлов К. Е.* Системы $M|G|1$ с групповым обслуживанием. Часть I: обзор методов анализа вероятностных характеристик // Вестник РУДН. Серия «Математика. Информатика. Физика». — 2009. — Т. 1. — С. 32–44.
2. Stream Control Transmission Protocol / RFC 2960 / R. Stewart, Q. Xie, K. Morneault et al // IETF. — 2000.
3. *Самуйлов К. Е.* Методы анализа и расчета сетей ОКС 7. — М.: Изд-во РУДН, 2002. — 191 с.
4. *Самуйлов К. Е., Першаков Н. В., Гудкова И. А.* Построение и анализ моделей системы с групповым обслуживанием заявок // Вестник РУДН. Серия «Математика. Информатика. Физика». — 2007. — Т. 3–4. — С. 45–52.
5. *Гайдамака Ю. В., Першаков Н. В., Чукарин А. В.* Модель протокола SCTP и ее применение к анализу характеристик сигнального трафика в сетях сотовой подвижной связи // Электросвязь. — 2007. — № 8. — С. 4–7.
6. *Гудкова И. А., Першаков Н. В., Самуйлов К.* Имитационная модель для анализа показателей качества функционирования протокола SCTP // Труды LXII конференции РНТОРЭС. — 2007. — С. 231–233.
7. *Chukarin A. V., Pershakov N. V., Samouylov K. E.* Performance of Sigtran-Based Signaling Links Deployed in Mobile Networks // Proc. of the 9th International Conference on Telecommunications (ConTEL 2007, Zagreb, Croatia). — 2007. — Pp. 163–166.
8. *Самуйлов К. Е., Чукарин А. В., Першаков Н. В.* Разработка модели функционирования протокола управления потоковой передачей // Вестник РУДН. Серия «Прикладная и компьютерная математика». — 2005. — Т. 4, № 1. — С. 40–47.
9. ITU-T Recommendation Q.7xx Series: Specifications of Signaling System No. 7 // ITU-T White Book. — 1993.
10. *Андреанов 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.

UDC 621.39

$M|G|1$ Queues with Batch Service and its Application to the Stream Control Transmission Protocol Performance Analysis. Part II

N. V. Pershakov, K. E. Samuylov

*Telecommunication Systems Department
Peoples' Friendship University of Russia
Miklukho-Maklaya str., 6, Moscow, Russia, 117198*

In this paper we introduce an analytical model of the stream control transmission protocol — a queuing system with batch service and job bundling timeout. The applicability of the analytical models in which jobs are served in batches of fixed and variable size [1] is assessed based on numerical results. An algorithm for the calculation of key performance measures is obtained. We also consider the process of signaling messages transmission over IP-network and analyze protocol parameters to provide quality of service in mobile networks.