

УДК 519.216

Модель телетрафика на основе самоподобного случайного процесса

Б.С.Цыбаков

Дан обзор некоторых современных достижений в области самоподобного телетрафика; определен самоподобный случайный процесс; описана модель самоподобного трафика пакетов и приведены выражения для асимптотической вероятности потерь пакетов в системе связи, передающей такой трафик.

A review of some modern achievements in the field of self-similar teletraffic is given. The self-similar random process is defined, a model of self-similar traffic of packets is described, and expressions for the asymptotic probability of the loss of such packets in a communications system transferring such a traffic are found.

1. Введение

Измерения на реальных локальных сетях передачи данных, измерения потока информации при передаче изображений с помощью кодов с переменной скоростью, измерения в WWW-сети интернет и в других сетях [1—6] привели к открытию того, что трафик в них является самоподобным случайным процессом. На качественном уровне самоподобие проявляется в том, что имеется медленно убывающая зависимость между величинами трафика в разные моменты времени, а также в том, что трафик сбивается в пачки данных и эти пачки выглядят статистически подобными в широком диапазоне изменения масштаба по шкале времени.

Запись измерений, выполненных компанией Bellcore в локальной сети Ethernet, воспроизведена в [1, раздел 12.2.3]. Кроме того, она доступна анонимному FTP по адресу flash.bellcore.com в директории pub/lan.traffic, файл rOct.TL. В качестве иллюстрации на рис. 1 показан отрезок трафика X локальной сети Ethernet. Горизонтальная ось — ось времени t в миллисекундах; вертикальная — логарифм от числа байт X_t плюс 1, т.е. $\log(X_t + 1)$, а на рис. 2 — зависимость дисперсии V_m усредненного трафика $X^{(m)}$ в сети Ethernet от длины блока усреднения m . По осям использован логарифмический масштаб. Сплошная прямая дает наклон функции $\log V_m$, соответствующий параметру Хёрста $H=0,85$. Для сравнения штриховая прямая соответствует наклону с $H=0,5$, который имел бы $\log V_m$ при трафике

с быстро убывающей зависимостью. Результаты статистической обработки записанных данных дают основание утверждать, что измеренный трафик является самоподобным.

Здесь трудно объяснить правомерность такого утверждения, так как для этого потребуются использовать некоторые понятия и определения самоподобного процесса. Однако уже в следующем разделе будут даны необходимые объяснения.

Запись измерений при передаче изображений можно найти в [1, раздел 12.2.2]. На рис. 3 представлен отрезок трафика X при передаче изображения кодами переменной длины (по горизонтальной оси отложены номера кадров, по вертикальной — логарифм числа пакетов АТМ (асинхронного режима передачи)), а на рис. 4 — зависимость от m логарифма дисперсии $\log V_m$ усредненного трафика $X^{(m)}$ при передаче изображения. Сплошная прямая дает наклон функции $\log V_m$, соответствующий параметру Хёрста $H=0,75$. Для сравнения штриховая прямая показывает наклон с $H=0,5$, который имел бы $\log V_m$ при трафике с быстро убывающей зависимостью. Результаты статистической обработки послужили основанием для вывода относительно самоподобия измеренного трафика.

Самоподобный характер трафика, наблюдаемого в современных высокоскоростных сетях связи, обусловлен тем, что сети являются интегральными и используются для передачи речи, данных, изображений (в том числе и по факсу), файлов и других

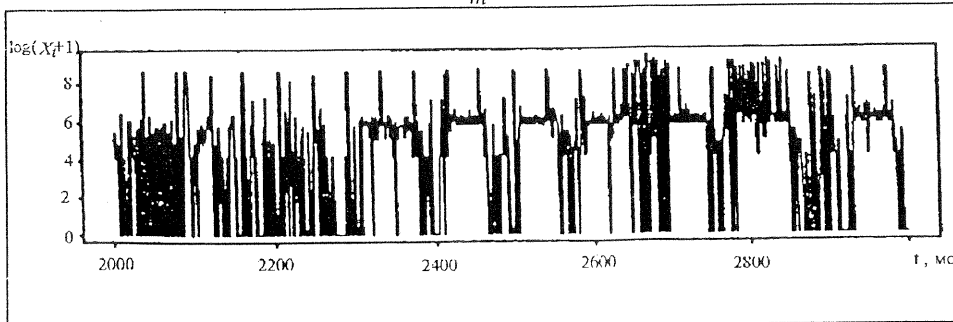


Рис. 1

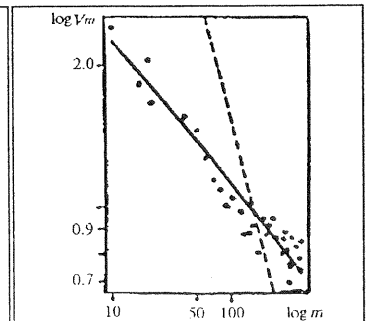


Рис. 2

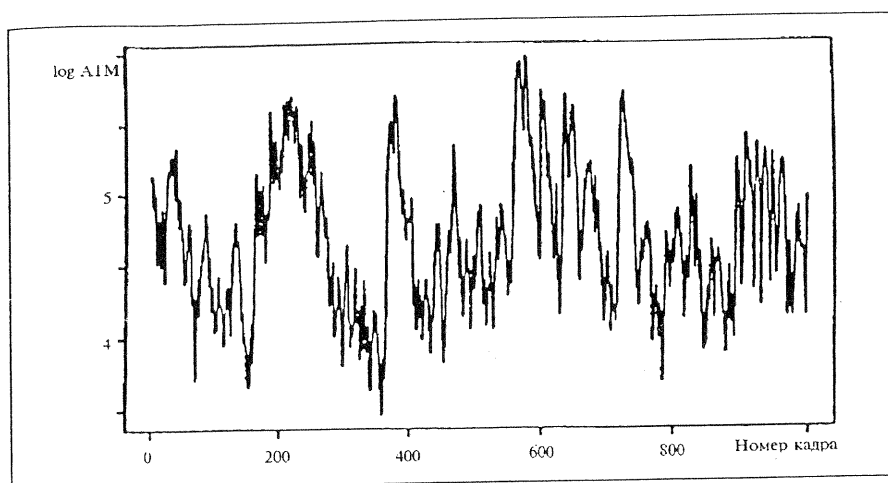


Рис. 3

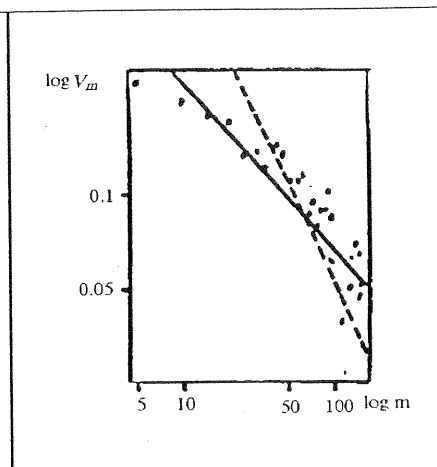


Рис. 4

при видов информации, представляемых в форме стандартизированных пакетов. Такой трафик оказывает значительное влияние на характеристики систем связи. В частности, как установлено с помощью измерений [4], при увеличении размера буфера на входе канала вероятность потерь падает значительно медленнее, чем по экспоненциальному закону, свойственному широко используемым классическим моделям телеграфика.

В статье представлены результаты недавних исследований в области теории самоподобного телеграфика; даются математическое определение самоподобного процесса и его свойства; рассмотрена модель самоподобного трафика как суперпозиция независимых источников; описана система связи, передающая пакеты, для которой приводится асимптотическая вероятность потерь пакетов, медленно снижающаяся по степенному закону с ростом размера буфера, расположенного на входе канала, и быстро падающая по показательному закону с ростом пропускной способности канала.

(m) Принимая во внимание цифровой и пакетный характер современных высокоскоростных сетей связи, здесь рассмотрены лишь процессы с дискретным временем и дискретным множеством состояний.

2. Самоподобие в широком смысле

Дадим определения строго и асимптотически самоподобных в широком смысле случайных процессов дискретного аргумента и укажем их связь с процессами, самоподобными в узком смысле, и с процессами с медленно убывающей зависимостью. Приводятся условия асимптотического самоподобия в широком смысле.

Обозначения. Пусть $X=(X_1, X_2, \dots)$ — полубесконечный отрезок стационарного, в широком смысле действительного случайного процесса дискретного аргумента (времени) $t \in N \triangleq \{1, 2, \dots\}$. Обоз-

начим через $\mu = EX_t < \infty$ и $\sigma^2 \triangleq DX_t < \infty$ среднее и дисперсию процесса X соответственно, а через

$$r(k) \triangleq \frac{E(X_{t+k} - \mu)(X_t - \mu)}{\sigma^2},$$

$$b(k) \triangleq \sigma^2 r(k), k \in Z_+ \triangleq \{0, 1, 2, \dots\} \quad (1)$$

— коэффициент корреляции и автоковариацию процесса X . Среднее μ , дисперсия $\sigma^2 = b(0)$, коэффициент корреляции $r(k)$ и автоковариация $b(k)$ не зависят от времени t и $r(k) = r(-k)$, $b(k) = b(-k)$.

Вместе с процессом X для введения самоподобного процесса и обсуждения его свойств и связи с другими процессами нам понадобятся еще два процесса, связанные с X , а именно, $X^{(m)} = (X_1^{(m)}, X_2^{(m)}, \dots)$ — усредненный по блокам длины m процесс X и $\check{X}^{(m)} = (\check{X}_1^{(m)}, \check{X}_2^{(m)}, \dots)$ — нормированный процесс X .

Компоненты процессов $X^{(m)}$ и $\check{X}^{(m)}$ определяются равенствами

$$X_t^{(m)} \triangleq \frac{1}{m} (X_{tm-m+1} + \dots + X_{tm}), m, t \in N; \quad (2)$$

$$\check{X}^{(m)} \triangleq \frac{1}{m^H} (X_{tm-m+1} + \dots + X_{tm}), H = 1 - \frac{\beta}{2}, 0 < \beta < 1. \quad (3)$$

Обозначим через $r_m(k)$, $b_m(k)$ и $V_m = b_m(0)$ коэффициент корреляции, автоковариацию и дисперсию $X^{(m)}$ соответственно.

Строгое самоподобие в широком смысле. *Определение* [7, 8]. Процесс X называется строго самоподобным в широком смысле (ССШС) с параметром $H = 1 - (\beta/2)$, $0 < \beta < 1$, если его коэффициент корреляции

$$r(k) = \frac{1}{2} [(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}] \triangleq g(k), k \in N. \quad (4)$$

Причина, по которой процесс X назван самоподобным, кроется в том, что процесс X с $r(k)=g(k)$ удовлетворяет условию $r_m(k)=r(k)$, $k \in Z_+$, $m \in \{2,3,\dots\}$, которое означает, что ССШС процесс не меняет свой коэффициент корреляции после усреднения по блокам любой длины m .

Смысл функции $g(k)$ состоит в том, что она является единственным решением уравнения $b_m(0)=b(k)m^{-\beta}$ относительно $b(k)$ [8]. Это уравнение (или уравнение $b_m(0)=b(0)m^{-\beta}$) может быть вместо (4) положено в основу определения ССШС процесса, давая определение, эквивалентное принятому здесь.

Функция $g(k)$ может быть записана как $g(k)=\frac{1}{2}\delta^2(k^{2-\beta})$ в терминах центрального разностного оператора 2-го порядка $\delta^2(f(x))$, действующего на функцию $f(x)=x^{2-\beta}$. Напомним, что

$$\delta(f(x))=f(x+1/2)-f(x-1/2). \quad (5)$$

На рис. 5 приведена зависимость коэффициента корреляции $r(k)=g(k)$ ССШС процесса (смысл имеют лишь точки при $k \in \{1,2,\dots\}$, а не соединяющие их линии), а на рис. 6 — спектральной плотности

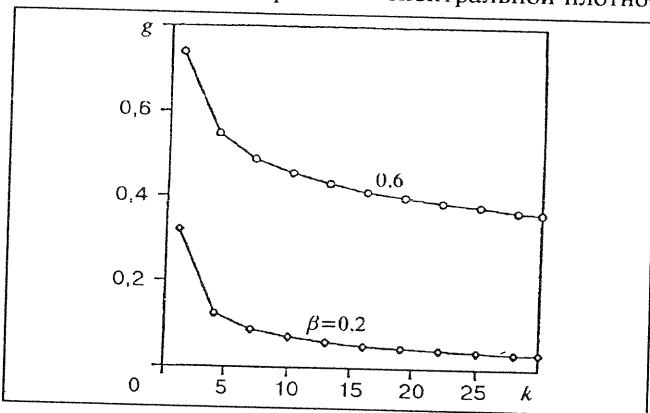


Рис. 5

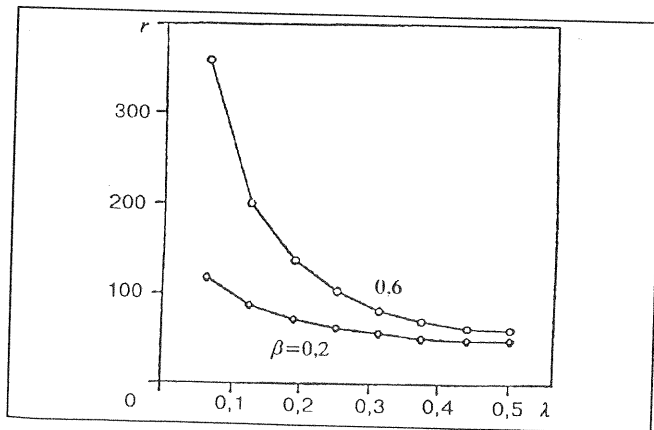


Рис. 6

сти $f(\lambda)$, соответствующей коэффициенту корреляции $r(k)=g(k)$.

Асимптотически с ростом k функция $g(k)$ убывает по степенному закону

$$g(k) \sim \frac{1}{2}(2-\beta)(1-\beta)k^{-\beta}, \quad k \rightarrow \infty, \quad (6)$$

где $g(x) \sim f(x)$, $x \rightarrow \infty$, означает, что $g(x)/f(x) \rightarrow 1$, $x \rightarrow \infty$. Из равенства (6) видно, что зависимость в ССШС процессе (соответствующая его корреляции) убывает очень медленно по сравнению, например, с экспоненциальным убыванием, свойственным марковским и авторегрессионным процессам, широко используемым в теории телетрафика.

Существует понятие процесса с медленно убывающей зависимостью (МУЗ). Согласно Берану [1], X называется процессом с МУЗ, если у него $r(k) \sim ck^{-\beta}$, $k \rightarrow \infty$, где $c > 0$ постоянная и $0 < \beta < 1$. Таким образом, ССШС процесс является процессом с МУЗ.

Какую роль играет параметр β ? Условие $1 < \beta < 2$ показывает, что X — процесс с быстро убывающей зависимостью (БУЗ). Этот процесс имеет $V_m \sim cm^{-1}$ при $m \rightarrow \infty$. Условие же $0 < \beta < 1$ показывает, что X — это процесс с МУЗ. Кроме того, когда $0 < \beta < 1$, значение β указывает, как медленно убывает зависимость в X : чем меньше β , тем сильнее зависимость в X . Аналогичную интерпретацию можно дать параметру Хёрста H .

Имеется связь между ССШС процессом и процессом, самоподобным в узком смысле, введенным Колмогоровым [8], Пинскером [10], Мандельбротом [11], Синаем [12]. Процесс X называется самоподобным в узком смысле (СУС) с параметром $H=1-(\beta/2)$, $0 < \beta < 1$, если

$$m^{1-H}X^{(m)} \doteq X, \quad m \in N \quad (7)$$

(или эквивалентно, $\overset{V}{X}^{(m)} \doteq X$), где \doteq обозначает равенство конечномерных распределений. Связь между ССШС процессами и СУС процессами аналогична связи между процессами, стационарными в широком и узком смысле.

Необходимое и достаточное условие того, что процесс является ССШС дано в [7, 8].

Асимптотическое самоподобие в широком смысле. Определение [7, 8]. Процесс X называется асимптотически самоподобным в широком смысле (АСШС) с параметром $H=1-(\beta/2)$, $0 < \beta < 1$, если

$$\lim_{m \rightarrow \infty} r_m(k) = g(k), \quad k \in N. \quad (8)$$

Смысл этого определения состоит в том, что X является АСШС процессом, если после усреднения

по блокам длины m и $m \rightarrow \infty$ он сходится к ССШС процессу. Обратим внимание на то, что здесь сходимость уже не к исходному доусредненному процессу X , а к ССШС процессу.

Ниже приведена теорема в терминах дисперсии V_m усредненного процесса, которая дает необходимое и достаточное условие (b) того, что X является АСШС процессом, а также достаточное условие (c) в терминах коэффициента корреляции $r(k)$ процесса X .

Теорема 1 [8]. Для процесса X и $H=1-(\beta/2)$, $0 < \beta < 1$, следующие утверждения являются эквивалентными:

(a) X есть АСШС процесс, т.е. справедливо (8);

(b) $(V_{km}/V_m) \sim k^{-\beta}$, целое $m \rightarrow \infty$, $k \in N$.

Асимптотическое равенство

(c) $r(k) \sim H(2H-1)L(k)k^{-\beta}$, целое $k \rightarrow \infty$

6) влечет асимптотическое равенство

(d) $V_m \sim \sigma^2 L(m)m^{-\beta}$, целое $m \rightarrow \infty$,

где $L(k)$ — функция, медленно меняющаяся на бесконечности, и каждое из (c) и (d) влечет (a) и (b).

Напомним, что измеримая функция $f(x) > 0$ называется медленно меняющейся на бесконечности, если при каждом $u > 0$ $f(ux) \rightarrow 1$ при $x \rightarrow \infty$.

Теперь можно пояснить смысл рис.2 и 4. На них в логарифмическом масштабе по осям отложены построенные по выборочным функциям реального трафика оценки дисперсии V_m в зависимости от m ; на графиках эти оценки указаны точками. Сплошная прямая, аппроксимирующая функцию V_m по оценочным точкам, имеет наклон, соответствующий параметру Хёрста, лежащему в диапазоне $\frac{1}{2} < H < 1$.

Это дает основание утверждать, что V_m ведет себя как $m^{-\beta}$ с $0 < \beta < 1$ и согласно теореме 1 (где сказано, что (d) влечет (a)) трафик можно считать АСШС процессом.

3. Модель трафика пакетов

В этом разделе дадим описание математической модели трафика пакетов и условия того, что эта модель — самоподобная. Материал является обзором работ [13—15, 8].

Общая модель трафика. Рассматриваемый трафик Y — это поток пакетов, имеющих длину, принятую за единицу времени. Пакеты генерируются источниками так, что трафик Y представляет собой суперпозицию пакетов, генерируемых источниками, которые имеют номера $s \in Z = \{\dots, -1, 0, 1, \dots\}$. Источник s начинает генери-

ровать свои пакеты в момент ω_s ($\omega_s \leq \omega_{s+1}$) и генерирует $\theta_s(i) \in Z_+$ пакетов в момент $\omega_s + i - 1$ в интервале $\omega_s, \dots, \omega_s + \tau_s - 1$, $i \in \{1, \dots, \tau_s\}$. Последовательность $(\theta_s(1), \dots, \theta_s(\tau_s))$ называется активным периодом источника s , а τ_s — его длиной. До момента ω_s и после момента $\omega_s + \tau_s - 1$ источник s пакеты не генерирует, $\theta_s(i) = 0$ при $i \leq 0$ и $i \geq \tau_s + 1$. Таким образом, $\theta_s(t - \omega_s + 1)$, $t \in Z$, — последовательность чисел пакетов, генерируемых источником s в последовательные моменты времени.

Частными случаями активных периодов, например, могут быть: 1) константа $\theta_s(i) = R \in N$, $1 \leq i \leq \tau_s$, где R — скорость источника (этот трафик рассматривается в пп. 4 и 5); 2) случайная константа $\theta_s(i) = R$, где $R = R(\tau_s)$; 3) независимые и одинаково распределенные (НОР) $\theta_s(i)$, принимающие значения 0 и 1 с вероятностями p_0 и p_1 соответственно; 4) НОР $\theta_s(i)$, принимающее значения из множества $\{0, 1, \dots, k\}$ с биномиальным распределением или из Z_+ с геометрическим, пуассоновским или каким-либо другим заданным распределением; 5) марковские, полумарковские, стационарные или другие хорошо известные последовательности $\theta_s(i)$ и др.

Момент t может быть моментом ω_s сразу для нескольких источников. Пусть ξ_t обозначает число источников, для которых $\omega_s = t$.

Рассматриваемый трафик $Y = (\dots, Y_{-1}, Y_0, Y_1, \dots)$ есть суперпозиция пакетов, генерируемых различными источниками,

$$Y_t = \sum_{s \in Z} \theta_s(t - \omega_s + 1), \quad t \in Z. \quad (9)$$

Это значит, что Y_t — общее число пакетов, генерируемых источниками, активными в момент t .

Сужение модели. Предполагается, что: активные периоды $(\theta_s(1), \dots, \theta_s(\tau_s))$ — НОР при различных s . При условии $\tau_s = l$ такой период является отрезком стационарного в широком смысле неотрицательного случайного процесса со значениями из множества Z_+ ;

числа ξ_t , $t \in Z$, независимы и одинаково распределены по закону Пуассона, $\text{Pr}\{\xi_t = k\} = e^{-\lambda} \lambda^k / k!$, $0 < \lambda < \infty$;

активные периоды $(\theta_s(1), \dots, \theta_s(\tau_s))$ независимы от чисел ξ_t и моментов ω_s .

Условие самоподобия модели трафика. Необходимое и достаточное условие того, что Y является ССШС процессом представлено в следующей теореме, использующей обозначения $\mu^{(l)} = E\theta(t)$ и $B^{(l)}(k) = E\theta(t)\theta(t+k)$, $k \in Z$, где $(\dots, \theta(-1), \theta(0), \theta(1), \dots)$ — зависящий от l случайный стационарный процесс, распределение которого на длине l совпадает с условным распределением активного периода источника при $\tau=l$ (здесь и далее τ — случайная величина, имеющая такое же распределение, как τ_s).

Теорема 2 [8]. Процесс Y будет ССШС с $0 < \beta < 1$ тогда и только тогда, когда $\Pr\{\tau=l\}$, $\mu^{(l)}$ и $B^{(l)}(k)$ удовлетворяют условиям

$$\frac{\sum_{l=k+1}^{\infty} (l-k) \Pr\{\tau=l\} B^{(l)}(k)}{\sum_{l=1}^{\infty} l \Pr\{\tau=l\} B^{(l)}(0)} = \frac{1}{2} \delta^2 (k^{2-\beta}), \quad k \in N, \quad (10)$$

$$\sum_{l=1}^{\infty} l \Pr\{\tau=l\} \mu^{(l)} < \infty. \quad (11)$$

В частности, если $\mu^{(l)}$ не зависит от l , а $B^{(l)}(k)$ постоянная величина, не зависящая от l и k , то Y будет процессом ССШС тогда и только тогда, когда

$$\Pr\{\tau=1\} = \frac{3^{2-\beta} - 2^{4-\beta} + 7}{4 - 2^{2-\beta}}; \quad (12)$$

$$\Pr\{\tau=k\} = \frac{\delta^4 (k^{2-\beta})}{4 - 2^{2-\beta}}, \quad k \in \{2, 3, \dots\}, \quad (13)$$

где

$$\delta^4 (k^{2-\beta}) = (k+2)^{2-\beta} - 4(k+1)^{2-\beta} + 6k^{2-\beta} - 4(k-1)^{2-\beta} + (k-2)^{2-\beta}, \quad k \in \{2, 3, \dots\}, \quad (14)$$

и распределение $\Pr\{\tau=k\}$, которое определяется (12) и (13), имеет конечное среднее значение $E\tau = (2 - 2^{1-\beta})^{-1}$, бесконечную дисперсию и следующую неэкспоненциальную асимптотику

$$\Pr\{\tau=k\} \approx \frac{(\beta+1)\beta(\beta-1)(\beta-2)}{4 - 2^{2-\beta}} k^{-(\beta+2)}, \quad k \rightarrow \infty. \quad (15)$$

Достаточное условие того, что Y — процесс АСШС, выражено в следующей теореме.

Теорема 3 [8]. Процесс Y является АСШС с $0 < \beta < 1$ тогда, когда $\Pr\{\tau=l\}$, $\mu^{(l)}$ и $B^{(l)}(k)$ такие, что

$$\Pr\{\tau=l\} B^{(l)}(l) \sim L(l) l^{-(\beta+2)}, \quad l \rightarrow \infty, \quad (16)$$

$$\sum_{l=1}^{\infty} \Pr\{\tau=l\} B^{(l)}(0) < \infty,$$

$$\sum_{l=1}^{\infty} l \Pr\{\tau=l\} \mu^{(l)} < \infty, \quad (17)$$

где $L(x)$ — любая функция, медленно меняющаяся на бесконечности.

В частности, согласно теореме 3, процесс Y станет АСШС с $H = (3 - \alpha)/2$, если $\mu^{(l)}$ не зависит от l , $B^{(l)}(k)$ не зависит от l и k , а

$$\Pr\{\tau=l\} = c_0 l^{-\alpha-1}, \quad 1 < \alpha < 2, \quad l \in N, \\ c_0 = \left[\sum_{l=1}^{\infty} l^{-\alpha-1} \right]^{-1}. \quad (18)$$

Здесь (18) называется распределением типа Парето.

4. Модель системы связи и определение вероятности потери пакетов

Рассмотрим систему связи, входной поток пакетов в которой является самоподобным. Экспериментально установлено [4], что вероятность потери пакетов в такой системе падает значительно медленнее, чем по экспоненте при росте размера буферной памяти на входе канала. Это прямое следствие того, что самоподобный трафик обладает медленно убывающими зависимостями между числами пакетов, поступающих в разные моменты времени.

Задача состоит в том, чтобы дать теоретическое объяснение этому экспериментально наблюдаемому явлению.

Здесь система связи описывается как система массового обслуживания, даются определения вероятности переполнения буфера и вероятности потери пакета, а также приводятся известные соотношения между этими вероятностями.

Модель системы связи с дискретным временем

$t \in \{\dots, -1, 0, 1, \dots\} \triangleq Z$, которая состоит из конечной буферной памяти (буфера) и канала. На вход буфера поступает трафик $Y = (\dots, Y_{-1}, Y_0, Y_1, \dots)$, где, как и ранее, Y_t обозначает число пакетов, приходящих в момент t . Считаем, что Y — это трафик, описанный в разделе 3 и имеющий источники, генерирующие пакеты с постоянной скоростью R ($R \in N$). В этом случае Y_t/R имеет пуассоновское распределение такое, что $EY_t = \lambda REt$.

Конечный размер буфера означает, что он может хранить не более чем h пакетов в любой момент времени t . В каждом окне t (окно — это интервал времени $[t, t+1)$, содержащий только один момент времени t дискретной оси Z) канал может передать (обслужить) не более чем C пакетов, которые берутся либо из буфера, либо из Y_t новых пакетов. Пакет

из буфера, который передается в окне t , уходит из канала и в целом из рассматриваемой системы в момент $t+1$. Величина $C \in N$ называется пропускной способностью канала.

Система рассматривается как система массового обслуживания (СМО). В каждый момент t предполагается, что события происходят в следующем порядке: {конец обслуживания в окне $t-1$ } \rightarrow {конец окна $t-1$ } \rightarrow {поступление новых пакетов (быть может, поступает 0 пакетов)} \rightarrow {выбор следующих пакетов для обслуживания} \rightarrow {потеря пакетов, если этого требует дисциплина обслуживания} \rightarrow {помещение непотеряных пакетов в буфер} \rightarrow {начало окна t } \rightarrow {начало обслуживания в окне t }. Рассматриваемую СМО обозначим как $Y/D/C/h/d$, где Y означает, что входной трафик есть Y ; D — детерминированное время обслуживания, равное единице; C — число обслуживающих приборов; h говорит, что размер буфера есть h , а d отмечает, что принимается во внимание дисциплина d , действующая в системе.

Д и с ц и п л и н а. Предполагается, что в СМО в каждый момент t дисциплина решает, какая из следующих альтернатив должна быть применена к каждому пакету, находящемуся в системе: 1) начать передачу (обслуживание) пакета в момент t ; 2) хранить пакет в буфере до момента $t+1$; 3) сбросить (потерять) пакет в момент t .

Наиболее интересным классом дисциплин d является класс $D_C(h)$, удовлетворяющий следующим условиям [16]: 1) если $Y_t + Z_t > 0$ (где Y_t — число новых пакетов, поступивших в момент t , Z_t — число пакетов, которые находились в системе в момент t до поступления новых пакетов), то $\min\{Y_t + Z_t, C\}$ пакетов идут на передачу (обслуживание) в момент t ; 2) если $Y_t + Z_t \leq h + C$, то ни один из пакетов не теряется в момент t , если $Y_t + Z_t > h + C$, то $Y_t + Z_t - h - C$ пакетов теряются в момент t . Какие пакеты теряются и какие передаются зависит от конкретной дисциплины $d \in D_C(h)$.

Вероятность переполнения и вероятность потерь. Определим вероятность переполнения буфера P_{over} и вероятности потери пакета P_{loss} .

Событие $\{Y_t + Z_t - h - C > 0\}$ называется переполнением буфера в момент t . Момент времени t называется моментом переполнения, если в этот момент теряется по крайней мере один пакет. Пусть A_t — индикатор момента переполнения; $A_t = 1$, если t — момент переполнения, и $A_t = 0$, если это не так.

Вероятность переполнения в СМО $Y/D/C/h/d$ определяется равенством

$$P_{over} \stackrel{\Delta}{=} \limsup_{t \rightarrow \infty} \Pr\{A_t = 1\}. \quad (19)$$

Это означает, что P_{over} — стационарная вероятность переполнения буфера, если существует предел $\Pr\{A_t = 1\}$ при $t \rightarrow \infty$.

Для определения вероятности потерь P_{loss} нужно занумеровать пакеты. Пусть пакеты нумеруются числами $n \in Z$ в порядке их появления. Пакеты, появляющиеся в один и тот же момент времени, при нумерации упорядочиваются произвольно. Номер 1 приписывается первому пакету, появляющемуся в первый момент времени $t \geq 0$, когда $\xi_t > 0$.

В СМО $Y/D/C/h/d$ положим $B_n = 0$, если пакет n в конце концов получает передачу (обслуживание), и $B_n = 1$, если пакет n в конце концов теряется.

Вероятность в СМО $Y/D/C/h/d$ определяется равенством

$$P_{loss} \stackrel{\Delta}{=} \limsup_{n \rightarrow \infty} \Pr\{B_n = 1\}. \quad (20)$$

Это значит, что P_{loss} — стационарная вероятность потери пакета, если существует предел $\Pr\{B_n = 1\}$ при $n \rightarrow \infty$.

Связь между вероятностями переполнения и потерь. Известно несколько связывающих P_{over} и P_{loss} соотношений, некоторые из которых справедливы для СМО с более общим по сравнению с Y входным трафиком G .

Так, пусть G_t — число пакетов в общем входном трафике в момент t ($0 < EG_t < \infty$), а L_t — число пропавших в момент t пакетов в СМО $G/D/C/h$ с $d \in D_C(h)$, и пусть $G = (\dots, G_{-1}, G_0, G_1, \dots)$ и $L = (\dots, L_{-1}, L_0, L_1, \dots)$ — стационарные и эргодические последовательности. Тогда [15, 17]

$$P_{loss} \geq \frac{P_{over}}{EG_t}. \quad (21)$$

Если $G = Y$, где Y — тот входной трафик, который описан в п. 4 с $R=1$, то в $Y/D/C/h/d$ с $d \in D_C(h)$, как показано в [15, 17],

$$P_{loss} \leq \frac{a(\lambda, C) P_{over}}{\lambda E\tau}, \quad (22)$$

где

$$a(\lambda, C) = e^{2\lambda} + \sum_{s=1+C}^{\infty} \frac{[(\lambda E\tau) - \lambda]^s}{s!} \leq$$

$$\leq e^{2\lambda} + e^{(\lambda E\tau) - \lambda} \frac{[(\lambda E\tau) - \lambda]^{1+C}}{(1+C)!}. \quad (23)$$

В частности, если $C=1$, то

$$a(\lambda, 1) = e^{2\lambda} + e^{(\lambda E\tau) - \lambda} - 1 - (\lambda E\tau) + \lambda. \quad (24)$$

Для $Y/D/C/h/d$ с $d \in D_C(h)$ из (21) и (22) следует, что

$$\lim_{\lambda \rightarrow 0} \frac{(\lambda E\tau) P_{\text{loss}}}{P_{\text{over}}} = 1. \quad (25)$$

Это означает, что $P_{\text{over}} \approx (\lambda E\tau) P_{\text{loss}}$ при малой интенсивности $\lambda E\tau$ трафика Y .

5. Границы для вероятностей переполнения и потерь

Здесь даем обзор результатов, демонстрирующих поведение вероятностей переполнения и потерь как функции размера буфера и других параметров системы связи и трафика. При этом считается, что входной трафик Y является таким, каким он был принят в п. 4, т.е. считается, что источники имеют постоянную скорость R генерирования пакетов в течение своих активных периодов. Часть результатов относится к наиболее интересному случаю, когда длины активных периодов источников имеют распределение $\text{Pr}\{\tau=l\}$ типа Парето (18). В этом случае входной трафик Y является асимптотически самоподобным в соответствии с теоремой 4.

Верхняя граница для вероятности переполнения. В [18] доказано, что для системы $Y/D/C/h$, где Y — трафик с τ , распределенным по закону типа Парето (18), а $d \in D_C(h)$, справедлива следующая асимптотическая верхняя граница для вероятности переполнения буфера:

$$P_{\text{over}} \leq c_1 h^{(-\alpha+1)k}, \quad k = 1 + [(C/R) - \lambda E\tau],$$

$$h \rightarrow \infty, \quad C > \lambda R E\tau, \quad (26)$$

где $[x]$ — целая часть числа x , неравенство $f(x) \leq g(x)$ при $x \rightarrow \infty$ понимается как $\limsup_{x \rightarrow \infty} f(x)/g(x) \leq 1$; а c_1 — некоторая функция скорости источников пакетов R , пропускной способности канала C , интенсивности появления источников λ , параметра α распределения

типа Парето (18) и средней длины активного периода источника $E\tau = c_0 \sum_{l=1}^{\infty} l^{-\alpha}$. Отметим, что c_1 не зависит от h .

Нижняя граница для вероятности переполнения. Для той же самой системы $Y/D/C/h$, для которой получена граница (26), в [19] установлена нижняя асимптотическая граница для вероятности переполнения буфера в виде

$$P_{\text{over}} \geq c_2 h^{(-\alpha+1)k}, \quad h \rightarrow \infty, \quad C \geq \lambda R E\tau, \quad (27)$$

где k — то же самое, что и в (26), неравенство $f(x) \geq g(x)$ понимается как $\liminf_{x \rightarrow \infty} f(x)/g(x) \geq 1$, а c_2 — некоторая функция R, C, λ, α , которая не зависит от h .

Таким образом, границы (26) и (27) с точностью до множителя, не зависящего от h , дают скорость истинного убывания вероятности переполнения буфера с ростом размера буфера h .

Вероятность P_{over} убывает не по привычному в теории телетрафика показательному (экспоненциальному) закону, а по степенному закону от h . При этом показатель степени пропорционален величине $C - \lambda R E\tau$, превышению пропускной способности C над скоростью входного потока $EY_t = \lambda R E\tau$. Последнее означает, что вероятность переполнения буфера все-таки убывает (при больших h) по показательному закону, но в зависимости не от размера буфера, а от превышения C над EY_t . Это показывает, что легче бороться с переполнениями буфера путем увеличения пропускной способности, чем путем увеличения размера буфера.

Для иллюстрации этого вывода, рассмотрим систему, в которой $C/R > 1$, C/R — целое, и $\lambda E\tau < 1$. В этой системе $k = C/R$. Увеличение в $b > 1$ раз пропускной способности C (для простоты предполагается, что C/R — целое число) дает такое же уменьшение функции $h^{(-\alpha+1)k}$, которое образуется при увеличении размера буфера в h^{b-1} раз. Если, например, $h = 10^4$, то увеличение пропускной способности в 2 раза эквивалентно увеличению размера буфера в 10^4 раз (до размера 10^8). Заметим, что увеличение пропускной способности сопровождается уменьшением задержки переданных по каналу пакетов, а увеличение размера буфера — увеличением такой задержки. Следует отметить, что при этом не принимаются во внимание такие важные факторы как стоимость и другие, которые существенны на практике.

В [18, 19] приводятся выражения для c_1 и c_2 . Из них следует, например, что в случае $R=C=1$,

$\lambda=0,2$ имеем $c_1/c_2=11,08$, если $\alpha=1,3$, $c_1/c_2=13,18$, если $\alpha=1,5$, и $c_1/c_2=22,32$, если $\alpha=1,8$.

Получение более близких друг к другу выражений для c_1 и c_2 является открытой задачей.

Следует отметить, что близкой задачей к той, которая рассматривалась выше, является отыскание функции распределения числа пакетов, находящихся в буфере бесконечного размера. Интересные результаты решения этой задачи получены в [20—26].

В заключение отметим, что, как следует из (21) и (22), границы (26) и (27) остаются в силе и для P_{loss} вместо P_{over} . Однако функции c_1 и c_2 в границах для P_{loss} отличны от тех, которые участвуют в (26) и (27).

Данный обзор некоторых последних достижений относится к исследованию свойств и передачи самоподобного трафика. Приведенные асимптотические формулы для вероятности потерь пакетов показывают, что она убывает по степенному закону в зависимости от размера буфера и по показательному закону в зависимости от пропускной способности канала.

Теория самоподобного телетрафика проходит сейчас раннюю стадию своего развития. Многие стоящие перед ней задачи еще не имеют не только своего решения, но и приемлемой для рассмотрения постановки.

Работа получила частичную поддержку Франко-русского центра по прикладной математике и информатике им. А.М.Ляпунова, проект 06-97.

Литература

1. Beran J. Statistics for Long-Memory Processes. — New York: Chapman and Hall, 1994.
2. Beran J., Sherman R., Taqqu M.S. and Willinger W. Long-Range Dependence in Variable-Bit-Rate Video Traffic. — IEEE Trans. Commun., 1995, v. 43, №2-4, p. 1566-1579.
3. Crovella M.E. and Bestavros A. Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes. — Proceedings of the 1996 ACM SIGMETRICS. International Conference on Measurement and Modeling of Computer Systems, May, 1996 and IEEE/ACM Trans. on Networking, 1997, v. 5, №6, p. 835-846.
4. Leland W.E., Taqqu M.S., Willinger W. and Wilson D.V. On the Self-Similar Nature of Ethernet Traffic. — Proc. ACM SIGCOMM'93, San Francisco, CA, 1993, p. 183-193.
5. Leland W.E., Taqqu M.S., Willinger W., and Wilson D.V. On the Self-Similar Nature of Ethernet Traffic (Extended version). — IEEE/ACM Trans. on Networking, 1994, v. 2, №1, p.1-15.
6. Willinger W., Taqqu M.S. and Erramilli A. A Bibliographical Guide to Self-Similar Traffic and Performance Modeling for Modern High-Speed Networks in "Stochastic networks: Theory and applications". — Ed. F.P.Kelly, S.Zachary, and I.Ziedins, Clarendon Press (Oxford University Press), Oxford, 1996, p.339-366.

7. Cox D.R. Long-Range Dependence: A Review, in Statistics: An Appraisal, H.A.David and H.T.David, Eds. Ames, IA: Iowa State Univ. Press, 1984, p.55-74.
8. Tsybakov B. and Georganas N.D. Self-Similar Processes in Communications Networks. — IEEE Trans. on Information Theory, 1998, v.44, №.5, p.1713-1725.
9. Колмогоров А.Н. Спираль Винера и некоторые другие инвариантные кривые в гильбертовом пространстве. — ДАН СССР, 1940, т.26, с. 115-118.
10. Линскер М.С. — Теория кривых в гильбертовом пространстве со стационарными n -ми приращениями. — Изв. АН СССР, Сер. Математическая, 1955, т.19, №5, с.319-345.
11. Mandelbrot B.B. Long-Run Linearity, Locally Gaussian Processes, H-spectra, and Infinite Variances. — Int. Econ. Rev., 1969, v. 10, p.82-113.
12. Синай Я.Г. Автомодельные распределения вероятностей. — Теория вероятностей и ее применения, 1976, т.21, вып.1, с. 63-80.
13. Likhanov N., Tsybakov B. and Georganas N.D. Analysis of an ATM Buffer with Self-Similar ("Fractal") Input Traffic. — Proc. IEEE INFOCOM'95, Boston, MA, 1995, p.985-992.
14. Smith W.L. Regenerative Stochastic Processes. — Proc. Roy. Soc. Ser. A, 1955, №232, p. 6-31.
15. Tsybakov B. and Georganas N.D. Self-Similar Traffic and Upper Bounds to Buffer-Overflow Probability in ATM Queue. — Performance Evaluation., 1998, v.32, p.57-80.
16. Цыбаков Б.С., Папантони-Казакос П. Наилучшая и наихудшая дисциплины передачи пакетов. — Проблемы передачи информации, 1996, т.32, вып.4, с.72-92.
17. Tsybakov B. and Georganas N.D. Overflow and Loss Probabilities in a Finite ATM Buffer Fed by Self-Similar Traffic. — Queueing Systems, 1999.
18. Tsybakov B. and Georganas N.D. — Conference "Performance and Control of Networks Systems". — Boston, USA, 1999.
19. Tsybakov B. — IEEE Journal of Selected Areas in Communications (Submitted, 1999).
20. Duffield N.G. On the Relevance of Long-Tailed Durations for the Statistical Multiplexing of Large Aggregations. — Proc. of the 34-th Annual Allerton Conf. on Communication, Control, and Computing, 1996, Oct. 2-4, p. 741-750.
21. Duffield N.G. Queueing at Large Resources Driven by Long-Tailed M/G/ ∞ -modulated Processes, a manuscript, December 30, 1996.
22. Duffield N.G. and O'Connell N. Large Deviations and Overflow Probabilities for the General Single-server Queue with Applications. — Math. Proc. Cam. Phil. Soc., 1995, v.118, p.363-374.
23. Liu Z., Nain P., Towsley D. and Zhang Z.-L. Asymptotic Behavior of a Multiplexer Fed by a Long-Range Dependent Process. — CMPSCI Technical Report 97-16, University of Massachusetts at Amherst, 1997.
24. Parulekar M. and Makowski A.M. Tail Probabilities for a Multiplexer with Self-Similar Traffic. — Proc. of the IEEE INFOCOM'96 Conf., Mar. 26-28, 1996, p. 1452-1459.
25. Parulekar M. and Makowski A.M. Tail Probabilities for M/G/ ∞ Input Processes (I): Preliminary Asymptotics. — Preprint, University of Maryland, 1996.
26. Parulekar M. and Makowski A.M. M/G/ ∞ Input Processes: A Versatile Class of Models for Network Traffic. — Preprint, University of Maryland, 1996.

Поступила 12 декабря 1998 г.