

ОБ ИССЛЕДОВАНИИ СВОЙСТВА САМОПОДОБИЯ ТРАФИКА МУЛЬТИСЕРВИСНОЙ СЕТИ

М. Л. Федорова*, Т. М. Леденева**

* Воронежский государственный технический университет

** Воронежский государственный университет

Поступила в редакцию 05.04.2010 г.

Аннотация. В статье рассматривается взаимосвязь различных характеристик случайного процесса, связанных с эффектом самоподобия.

Ключевые слова: случайный процесс, самоподобие, мультисервисная сеть.

Abstract. The interrelation of various characteristics of the casual process connected with effect of self-similarity is considered in the article.

Key words: stochastic process, self-similarity, a multiservice network.

ВВЕДЕНИЕ

Трафик является одним из важнейших показателей мониторинга мультисервисной сети. Продолжительное время считалось, что трафик локальной сети описывается классическим Пуассоновским распределением. Затем в печати стали появляться публикации группы ученых [1], которые исследовали Ethernet-трафик в сети корпорации Bellcore и обнаружили, что методы расчета компьютерной сети, основанные на Марковских моделях и формулах Эрланга, с успехом использующиеся при проектировании телефонных сетей, дают неоправданно оптимистические решения и приводят к недооценке нагрузки. Исторически сложилось, что телефонные сети изначально строились по принципу коммутации каналов, а в основе компьютерных сетей, как правило, лежит принцип коммутации пакетов, но методики расчетов остались практически теми же. Пакеты при высокой скорости их движения по сети поступают на узел не по отдельности, а целой пачкой. Трафик в таких сетях имеет явно выраженный всплесковый характер, что повышает вероятность перегрузок в узлах сети, которые ведут к переполнению буферов и вызывают потери и /или задержки. Пульсации приводят к перепадам скорости информационных потоков, при которых отношение максимального значения к минимальной скорости составляет десятки раз. Однако, как

оказалось, в мультисервисных сетях (к которым относятся, в частности, компьютерные сети) число событий на заданном временном интервале зависит от прежних, весьма отдаленных событий. Это означает, что при больших масштабах мультисервисной сети трафик обладает свойством самоподобия, т.е. выглядит качественно одинаково при любых достаточно больших масштабах временной оси.

Цель данной статьи заключается в исследовании взаимосвязей между свойствами самоподобных потоков и разработке методов, позволяющих выявлять их в мультисервисных сетях.

1. ХАРАКТЕРИСТИКИ И СВОЙСТВА САМОПОДОБНЫХ ПРОЦЕССОВ

Концепция самоподобия тесно связана с получившим большую известность понятием фракталов и теорией хаоса. Определение фрактала, данное Б. Мандельбротом, звучит так: «Фракталом называется структура, состоящая из частей, которые в каком-то смысле подобны целому». С математической точки зрения фрактальный объект, прежде всего, обладает дробной размерностью, которая определяется в виде $d = \frac{\log N}{\log 1/r}$, где N — количество равных частей, на которое нужно разделить объект, а каждая часть будет уменьшенной копией целого в $1/r$ раз. Фрактальная размерность может рассматриваться как мера неровности изрезанной поверхности объекта $d \in [n, n + 1)$ в n -мерном

пространстве, причем более неровные, “шероховатые” поверхности соответствуют более высоким значениям d .

Другим параметром характеризующим самоподобие, является коэффициент Хэрста $H \in [0.5; 1]$. Чем ближе этот параметр к 1, тем более ярко проявляются фрактальные свойства. Напротив, равенство $H = 0.5$ свидетельствует об отсутствии самоподобия. Связь между параметрами стохастических фрактальных процессов изображена на рис. 1.

Качественным свойством, характеризующим самоподобие случайного процесса, является масштабная инвариантность по времени, так что при изменении временной шкалы корреляционная структура самоподобного процесса остается неизменной.

Пусть случайный процесс $X = (x_1, x_2, \dots)$ определяется множеством значений $x_i = x_i(t)$ дискретного аргумента (времени) $t \in N$. Пусть $\mu = M(X)$ — математическое ожидание, $D = \sigma^2 = D(X)$ — дисперсия процесса X . Определим автокорреляционную функцию, автоковариацию и спектр мощности процесса X соответственно формулами

$$r(k) = r(t, t+k) = \frac{M[(x_{t+k} - \mu)(x_t - \mu)]}{\sigma^2}, \quad (1)$$

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

$$S(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-ik\omega}, \quad -\pi \leq \omega \leq \pi, \quad (3)$$

где $\omega = 2\pi\lambda$ — угловая частота в радианах, $-\frac{1}{2} \leq \lambda \leq \frac{1}{2}$.

Случайный процесс стационарен в узком смысле, если статистические свойства этого процесса не меняются с течением времени. Если же предположить, что математическое ожидание и дисперсия процесса не зависят от времени, а корреляция зависела лишь от разности $\tau = |t_2 - t_1|$, то такой процесс будет стационарным в широком смысле. Таким образом, для стационарного в широком смысле процесса X $\mu = const$, $\sigma^2 = const$, $r(t, t+k) = r(k)$ и, следовательно, $b(t, t+k) = b(k)$, причем $r(k) = r(-k)$, $b(k) = b(-k)$.

Определим агрегированный процесс с коэффициентом сегментации $m \in N$ (рис. 2) следующим образом:

$$\forall m \left(X^{(m)} = \{X_k^{(m)} : k \geq 1\} \right),$$

где
$$X_k^{(m)} = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} x_i, \quad (4)$$

Обозначим через $r_m(k)$, $b_m(k) = \text{cov}(X_t^{(m)}, X_{t+k}^{(m)})$ и $D_m = b_m(0)$ коэффициент корреляции, автоковариацию и дисперсию процесса $X^{(m)}$ соответственно. Отметим $b(k) = b_1(k)$.

Понятие медленно убывающей зависимости имеет ключевое значение в теории самоподобных процессов. На интуитивном уровне ее можно интерпретировать следующим образом:

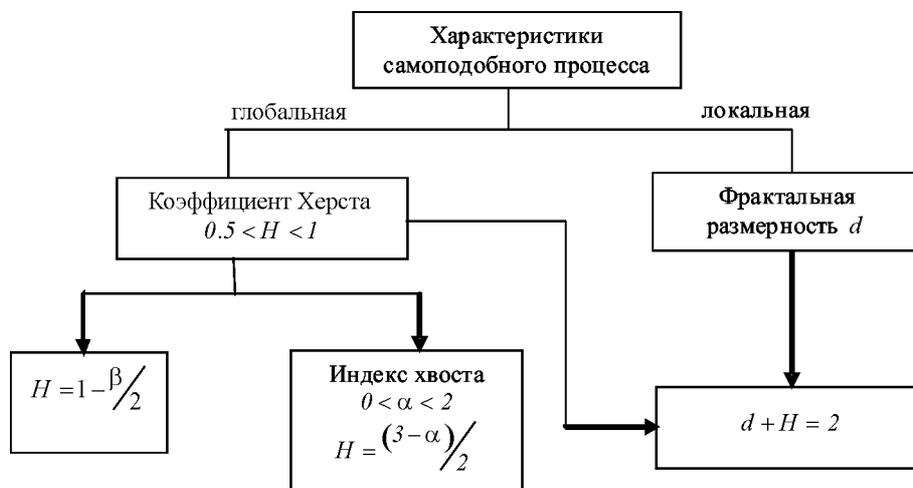


Рис. 1.

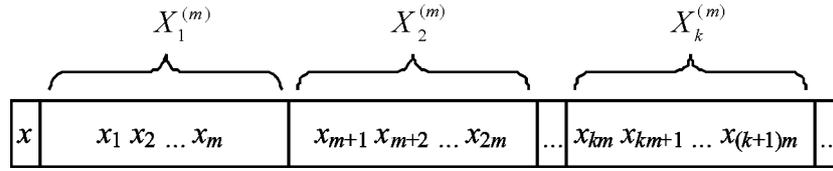


Рис. 2.

будущее процесса определяется его прошлым, причем с убывающей степенью влияния по мере того, как прошлое удаляется от настоящего.

Говорят [3, 4], что процесс X обладает медленно убывающей (или долговременной) зависимостью (МУЗ), если при $k \rightarrow \infty$ выполняется условие:

$$r(k) = k^{-\beta} \cdot L_1(k),$$

где $0 < \beta < 1$, $L_1(x) \in Z_+$ — медленно меняющаяся функция [5], такая что

$$\forall x > 0 \quad \forall \lambda > 0 \left(\lim_{x \rightarrow \infty} (L_1(\lambda x)/L_1(x)) = 1 \right),$$

Заметим, что из данного определения вытекает, что ряд, образованный последовательными значениями корреляционной функции, расходится, т.е.

$$\sum_{k=0}^{\infty} r(k) = \infty.$$

Процесс X обладает быстро убывающей (или кратковременной) зависимостью, если при $k \rightarrow \infty$ выполняется условие [5]

$$r(k) = \alpha^k, \quad (0 < \alpha < 1).$$

Заметим, что характерная особенность корреляции для стохастических фрактальных процессов — полиномиальное, а не экспоненциальное убывание.

Будем говорить, что случайная величина X имеет распределение с тяжелым «хвостом» (РТХ) [4], если

$$P(X \geq x) = 1 - F(x) = x^{-\alpha} L_1(x), \quad (5)$$

где $L_1(x)$ — медленно меняющаяся функция на бесконечности, параметр $0 < \alpha < 2$ называется индексом «хвоста».

В отличие от распределений с тяжелым «хвостом» распределения с легкими «хвостами» (например, экспоненциальное или нормальное) имеют экспоненциально спадающий «хвост».

Часто при анализе статических закономерностей пренебрегают возможностью крупных событий, лежащих на «хвосте» распределения. Но есть такие распределения, «хвост» которых нельзя отрезать, т. е. нельзя пренебрегать

крупными, пусть и редкими событиями. Под крупными событиями понимаются явления, ущерб от которых может превосходить ущерб от всех остальных событий этого класса, вместе взятых.

Заметим, что для распределений с тяжелым хвостом математического ожидания не существует, т.е. $M(X) = \infty$. Кроме того, параметр α из формулы (5) связан с параметром Херста H следующим соотношением [3]:

$$H = (3 - \alpha) / 2. \quad (6)$$

В [4] приведена функциональная зависимость между параметрами распределения с тяжелым хвостом и медленно убывающей зависимости

$$L(\tau) = (1 + n / \tau)^{-\alpha}, \quad (7)$$

где $0 < \alpha < 2$ — индекс «хвоста», $\tau \geq 1$ — время активного периода, $n \in N$ — номер шага для прогнозирования переменной $L(\tau)$. Таким образом, при достаточно больших τ ошибка предсказания может быть сколь угодно малой.

2. ОСНОВНЫЕ ТИПЫ САМОБОДОБНЫХ ПРОЦЕССОВ И ИХ ВЗАИМОСВЯЗИ

Пусть $X = (X_t : t > 0)$ — непрерывный случайный процесс со стационарными приращениями и плотностью распределения вероятностей

$$f(X_1, X_2, \dots, X_n) (n \in N),$$

$(a^{-H} X_{at} : t > 0, a \in Z_+, H \in [0.5; 1])$ — случайный процесс с масштабом времени at и плотностью распределения вероятностей $f(a^{-H} X_a, a^{-H} X_{2a}, \dots, a^{-H} X_{na})$.

Процесс X называется статически самоподобным [3, 4] с показателем $H \in [0.5; 1]$, если

$$\begin{aligned} & \forall a \forall n (f(X_1, X_2, \dots, X_n) \approx \\ & \approx f(a^{-H} X_a, a^{-H} X_{2a}, \dots, a^{-H} X_{na})). \end{aligned}$$

Иначе, в соответствии с [5] данное свойство можно определить выражением

$$a^{1-H} X_a = X, \quad (8)$$

которое понимается в смысле равенства конечномерных распределений.

Конечномерным распределением случайного процесса $X(t)$ в моменты t_1, \dots, t_n называется распределение многомерной случайной величины, составленной из сечений случайного процесса в моменты t_1, \dots, t_n :

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = P\{X(t_1) < x_1, \dots, X(t_n) < x_n\}.$$

Самоподобие в широком смысле слова означает выполнение следующих свойств [3, 4]:

$$M[X_t] = a^{-H} M[X_{at}],$$

$$D[X_t] = a^{-2H} D[X_{at}],$$

$$r(t, \tau) = \frac{r(at; a\tau)}{a^{2H}}.$$

В соответствии с [4], процесс X называется строго самоподобным в широком смысле (ССШС-процессом) с коэффициентом $H = 1 - (\beta/2)$, где $0 < \beta < 1$, если

$$r(k) = \frac{1}{2} \left[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta} \right] = g(k). \quad (9)$$

Согласно (9), функция $g(k)$ при больших k ведет себя как центральный разностный оператор 2-го порядка $\delta^2(\varphi(x))$, действующий на функцию $\varphi(x) = x^{2-\beta}$, при этом $\delta(\varphi(x)) = \varphi(x+1/2) - \varphi(x-1/2)$. Тогда функция $g(k)$ при $k \rightarrow \infty$ убывает асимптотически по закону

$$g(k) \sim \frac{1}{2} (2-\beta)(1-\beta) k^{-\beta} = H(2H-1)k^{-\beta}, \quad (10)$$

где $g(x) \sim f(x)$ означает, что $\lim_{x \rightarrow \infty} \frac{g(x)}{f(x)} = 1$.

Из равенства (10) видно, ССШС-процесс обладает МУЗ.

Приведем основные свойства ССШС.

Теорема 1 [8]. Процесс X с коэффициентом $0 < \beta < 1$ является ССШС-процессом, если для любого $k > 0$ при $m \rightarrow \infty$

$$r_m(k) = r(k), \quad (11)$$

т.е. коэффициент корреляции практически инвариантен к параметру агрегирования и сохраняет в асимптотическом смысле свою структуру. Последнее означает, что ССШС-процесс не меняет свой коэффициент корреляции после его усреднения по блокам любой длины m .

Для статистически самоподобного процесса $r(k) = g(k)$. Заметим, что равенства, которые

верны для статистически самоподобного процесса, также справедливы и для СШСС-процесса. Однако, обратное утверждение, в общем случае не верно. Связь между процессами этих типов аналогична связи между процессами, стационарными в широком и узком смыслах.

Следствие. Для СШСС процесса X с коэффициентом $0 < \beta < 1$ справедливо следующее утверждение:

$$D(X^{(m)}) = m^{-\beta} \text{ при } m \rightarrow \infty \quad (12)$$

т.е. дисперсия выборочного среднего этого процесса затухает медленнее, чем величина, обратная размеру выборки. Это факт [3] называют свойством медленно затухающей дисперсией (МЗД). Напротив, для традиционных процессов стационарных случайных процессов, обладающих кратковременной зависимостью, параметр $\beta = 1$ и

$$D(X^{(m)}) = m^{-1} \text{ при } m \rightarrow \infty \quad (13)$$

Свойство медленно затухающей дисперсии говорит о возможности существенных, не сглаживаемых усреднением, «всплесков» в случайном процессе, и связывает самоподобие с таким понятием, как РТХ.

Будем говорить, что процесс X является асимптотически самоподобным процессом в широком смысле (АСШС-процессом) [3] с параметром $H = 1 - (\beta/2)$, $0 < \beta < 1$, если для $k \in N$

$$\lim_{m \rightarrow \infty} r_m(k) = g(k), \quad (14)$$

Смысл этого определения состоит в том, что X является АСШС-процессом, если после усреднения по блокам длины m при $m \rightarrow \infty$, он сходится к ССШС-процессу. Обратим внимание на то, что здесь имеет место сходимости уже не к исходному процессу X (до усреднения), а к ССШС-процессу. Так как X является АСШС-процессом, то $r(k) \rightarrow L(k)k^{-\beta}$, $r_m(k) \rightarrow r(k)$, ($m \rightarrow \infty, k \rightarrow \infty$), где $L(k)$ — медленно меняющаяся функция. Согласно [6], если $X^{(m)}$ представляет собой СШСС-процесс при $m \rightarrow \infty$, то X является АСШС-процессом.

Теорема 2 [6]. Справедливы следующие утверждения:

а) Если процесс X с параметром $0 < \beta < 1$ обладает МУЗ, т.е.

$$b(k) \sim cH(2H-1)k^{-\beta}, \quad (15)$$

$$k \rightarrow \infty, 0 < c < \infty,$$

то агрегированный процесс $X^{(m)}$ с параметром $0 < \beta < 1$ обладает МЗД, т.е.

$$b_m(0) \sim cb(0)m^{-\beta} \text{ при } m \rightarrow \infty \quad (16)$$

б) Если агрегированный процесс $X^{(m)}$ с параметром $0 < \beta < 1$ обладает МЗД, то

$$\forall m \exists \varepsilon_m (b_m(0) = cm^{-\beta} + \varepsilon_m), \quad 0 < c < \infty, \quad (17)$$

при этом если $m^{2+\beta}\varepsilon_m \rightarrow 0$ при $m \rightarrow \infty$, то процесс X с параметром $0 < \beta < 1$ обладает МУЗ т.е. выполняется равенство (15).

в) Если для процесса X выполняются (15–17), то X является АСШС-процессом.

Эта теорема является необходимым и достаточным условием асимптотического самоподобия случайных процессов второго порядка. Согласно теореме, если случайный процесс обладает свойствами ДВЗ или МЗД, то он является АСШС-процессом. В свою очередь, согласно [4], РТХ влечет за собой МУЗ. А так как модель обобщенного трафика можно получить в результате агрегирования процессов с РТХ, то сетевой трафик обладает фрактальными свойствами и является процессом РТХ. Связи между рассмотренными понятиями представлены на рис. 3. Обозначения на рисунке 4: $L(\tau)$ — вероятность того, что соединение поддерживает активность в течении $1 \leq t \leq \tau$, τ — время активного периода источника генерируемого пакета, $0 < \alpha < 2$ — индекс «хвоста», $\delta \in N$ — номер шага для прогнозирования переменной $L(\tau)$.

Теорема 3 [7]. Пусть даны два процесса X' и X'' , такие что $r(k) = c_1 k^{-\beta_1}$ при $k \rightarrow \infty$ для процесса X' ; $r(k) = c_2 k^{-\beta_2}$ при $k \rightarrow \infty$ для процесса X'' , причем c_i, β_i ($i = 1, 2$) — константы, такие что $0 < c_i < \infty$, $0 < \beta_i < 1$. Тогда процесс $(X' + X'')$ является АСШС-процессом с параметром $H = 1 - \beta/2$, где $\beta = \min(\beta_1, \beta_2)$.

Таким образом, в результате объединения АСШС-процессов получается АСШС-процесс.

Эта теорема имеет большое практическое значение. Так как трафик сети является интегрированным, мультисервисным, т. е. комбинацией нескольких потоков с различными статистическими характеристиками (в частности, коэффициентами самоподобия), то при расчете вероятностных характеристик мониторинга трафика возможно рассматривать только один поток, что значительно упрощает расчеты.

3. ОЦЕНКА ПАРАМЕТРА САМОПОДОБИЯ Н

Для итога воспользуемся алгоритмом CoLoRade [8], который дает наиболее точную оценку параметра самоподобия.

Для ССШС-процесса, согласно (9),

$$r(k) = \frac{1}{2} \left((k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta} \right).$$

Учитывая, что $H = 1 - \beta/2$, получим

$$r(k) = 0.5 \left((k+1)^{2H} - 2k^{2H} + (k-1)^{2H} \right). \quad (18)$$

Для $r(k)$ можно определить статистическую оценку $\hat{r}(k)$ по формуле

$$\hat{r}(k) = \frac{\frac{1}{n} \sum_{i=1}^{n-k} [x_i - M][x_{i+k} - M]}{S^2}.$$

Из того, что $\hat{r}(1) = 2^{2H} - 1$, следует, что для $k = 1$ оценка H имеет вид

$$\hat{H} = \frac{1}{2} + \frac{1}{2 \ln 2} \ln (1 + \hat{r}(1)). \quad (19)$$

Пусть

$$f(H) = \hat{r}(k) - r(k) = \hat{r}(k) - 0.5 \left((k+1)^{2H} - 2k^{2H} + (k-1)^{2H} \right).$$

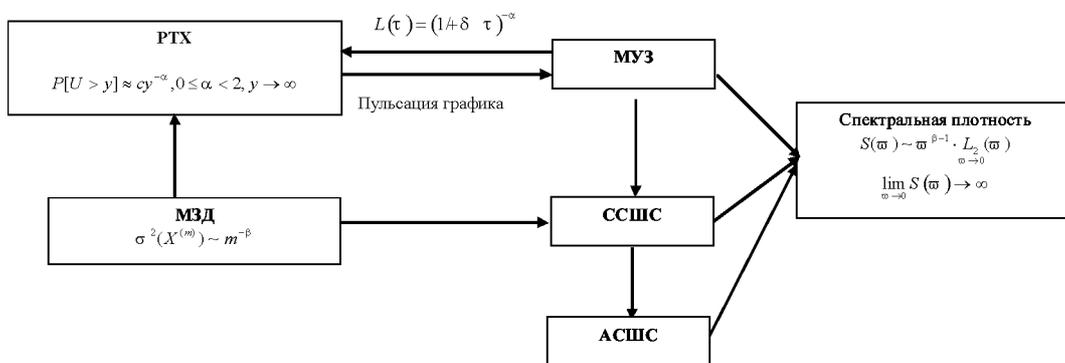


Рис. 3.

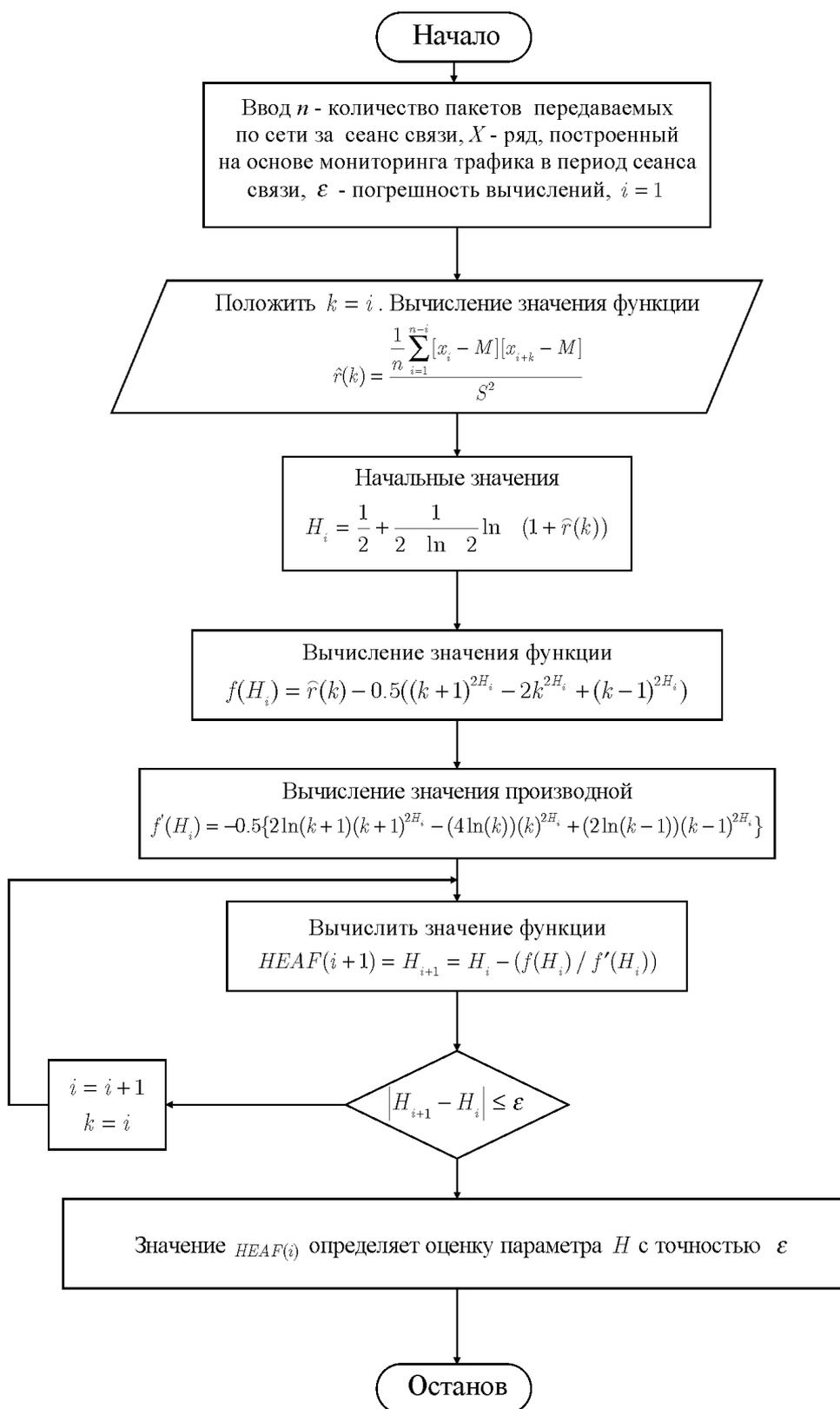


Рис. 4.

Тогда для нахождения оценки параметра Херста необходимо решить уравнение $f(H) = 0$. Для этого нами использовался итерационный метод Ньютона-Рафсона, позволяющий определить функцию HEAF(k), являющуюся оценкой параметра H , полученной в результате решения следующей задачи $\forall \varepsilon \forall n \exists N (n \geq N \rightarrow |H_n - H_{(n-1)}| < \varepsilon)$, k — номер итерации, на которой достигается заданная точность ε . Этот метод используется, если первая и вторая производные непрерывны и не меняют знаков на отрезке отделения корня. Найдем производные функции $f(H)$:

$$f'(H) = -0.5 \left(2 \ln(k+1)(k+1)^{2H} - 4 \ln(k)(k)^{2H} + 2 \ln(k-1)(k-1)^{2H} \right),$$

$$f''(H) = -0.5 \left(4 \ln^2(k+1)(k+1)^{2H} - 16 \ln^2(k)k^{2H} + 4 \ln^2(k-1)(k-1)^{2H} \right).$$

Заметим, что при $0 < H < 1$ и $k > 1$ $f'(H) > 0$, при $H = 0$ и $k > 1$ $f(H) < 0$ и $f''(H) < 0$.

Алгоритм CoLoRade для вычисления функции HEAF(k) представлен на рис. 4. Значения функции HEAF(k) для различных видов трафика сети Wi-Fi 802.11g представлены в табл. 1. На рис. 5 представлена графическая зависимость оценки параметра самоподобия H от значения лага k для асинхронного трафика; на рис. 6 а-б) — для синхронного трафика. В расчетах использовалось значение $\beta = 0.36$, которое соответствует параметру аудиотрафика.

СПИСОК ЛИТЕРАТУРЫ

1. Leland W. E., Taqqu M. S., Willinger W., Wilson D. V. On the self-similar nature of ethernet traffic (extended version). IEEE / ACM Transactions of Networking, 2(1):1–15, 1994.
2. Шелуфин О. И., Тенякшев А. М., Осин А. В. Фрактальные процессы в телекоммуникациях. — М.: Радиотехника, 2003.
3. Klemm A., Lindemann C., Lohmann M. Traffic Modeling of IP Networks Using the Batch Markovian Arrival Process // Performance Evaluation 54 (2) (2003) 149–173.
4. Столингс В. Современные компьютерные сети. — СПб. Питер, 2003. — 783 с. ил.

Таблица 1.

Вид трафика	Оценка HEAF(k) параметра самоподобия H	Погрешность ε	Количество шагов k для достижения точности
Трафик данных	0,7	0,01	23
Трафик потокового видео	0,8	0.01	17
Аудиотрафик	0,68	0.001	24
Трафик транзакций	0.75	0.01	21

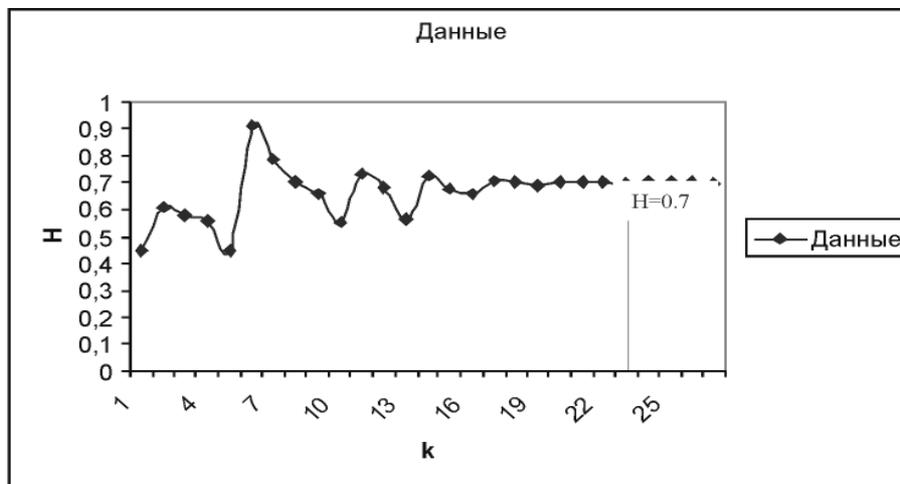
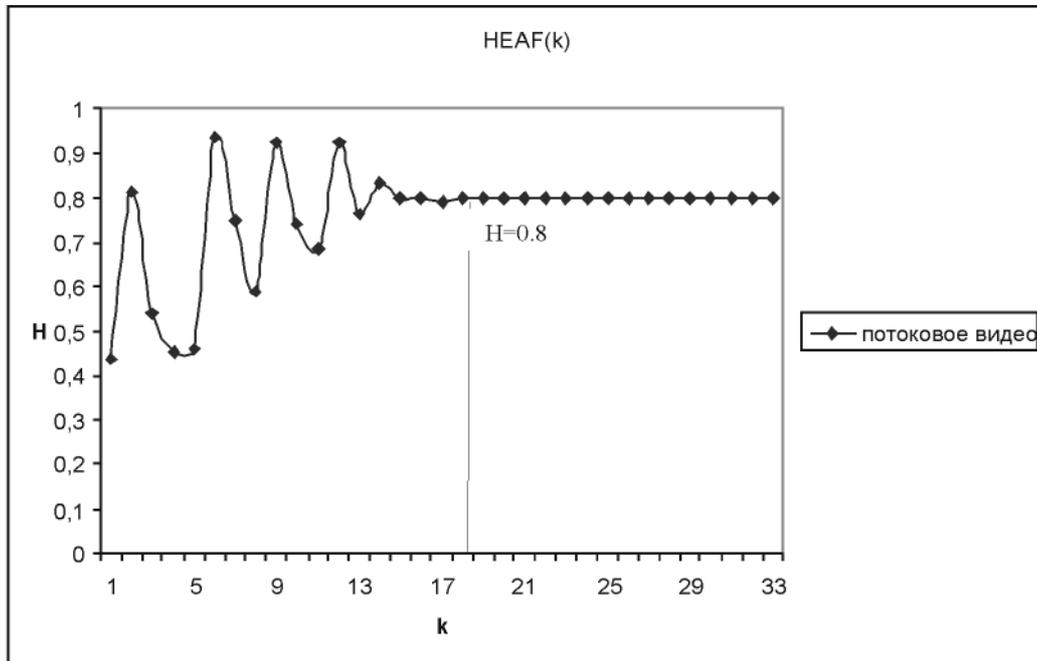
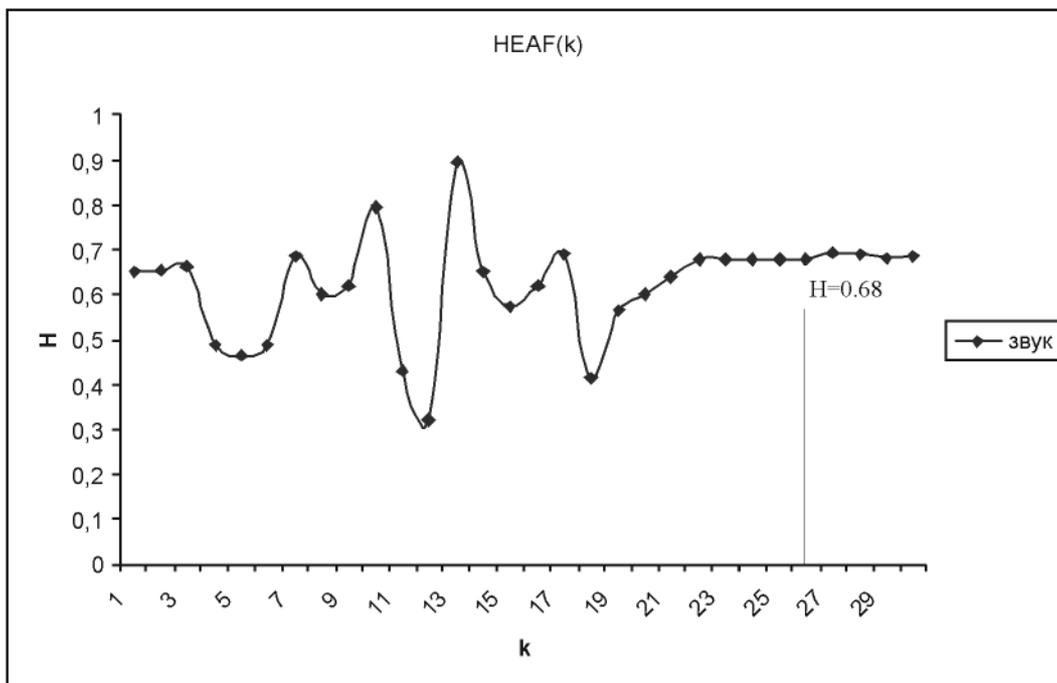


Рис. 5.



(а)



(б)

Рис. 6.

5. Сенета Е. Правильно меняющиеся функции. Пер. с англ. / Под ред. В. М. Золоторева. — М.: Наука. Главная редакция физико-математической литературы. — 1985. 144 с.

6. Tsybakov B., Georganas N. Self-similar processes in communications // IEEE Trans. Inform. Theory. vol. 44, pp. 1713—1725, Sep. 1998.

7. Park K, Tuan T. Performance Evaluation of Multiple Time Scale TCP Under Self-Similar Traffic

Conditions // IEEE / ACM Transactions on Modeling and Computer Simulation, Vol. 10, No. 2, April 2000, pp. 152—177.

8. Rezual K. M., Grout V. CoLoRaDe: A Novel Algorithm for Controlling Longrange Dependent Network Traffic // Sixth International Conference on Networking (ICN 2007), 22—28 April 2007, Sainte-Luce, Martinique, France. IEEE Computer Society 2007, pp. 57—63

Федорова Марина Львовна — старший преподаватель, Российский Химико-технологический университет им. Д. И. Менделеева, кафедра Вычислительной техники и информационных технологий, г. Новомосковск, E-mail: fed1232@bk.ru

Fedorova M. L. — senior teacher of the dept Information Systems and Technologies, D. Mendeleev University of Chemical Technology of Russia. E-mail: fed1232@bk.ru

Леденева Татьяна Михайловна — д. т. н., проф., каф. Математических методов исследования операций, факультет прикладной математики, информатики и механики, Воронежский государственный университет. Тел. (4732) 208-282. E-mail: dean@amm.vsu.ru

Ledenyova Tatyna Michaylovna — Doctor of Technic Sciences, Professor, The dept. of the Mathematical Methods of Oration Research, Voronezh State University. Tel. (4732) 208-282. E-mail: dean@amm.vsu.ru