

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ОЦЕНКИ ПАРАМЕТРА ХЕРСТА САМОПОДОБНЫХ ПРОЦЕССОВ

Л.О. Кириченко, Т.А. Радивилова, М.И. Синельникова
(Харьковский национальный университет радиоэлектроники)

В работе представлены результаты численного расчета параметра Херста H для модельного самоподобного сетевого трафика методами R/S анализа и изменения дисперсии агрегированного ряда. Показано, что оценки H , полученные обоими методами при $H > 0,7$, дают заниженные значения, причем с увеличением параметра эта разница возрастает.

параметр Херста, модельный самоподобный сетевой трафик, метод R/S анализа, метод изменения дисперсии агрегированного ряда

Постановка задачи. Последние исследования различных типов сетевого трафика убедительно показывают, что трафик является самоподобным по своей природе. Традиционно самоподобие трафика как стохастического процесса распознается путем определения параметра Херста H . Обычно тот факт, что $0,5 \leq H \leq 1$, считается достаточным основанием для признания процесса самоподобным [1 – 3]. Свойство самоподобия и, вытекающая из него долгосрочная зависимость, позволяют повысить эффективность обработки трафика путем прогнозирования. Техника прогнозирования решает ряд важных задач: анализ очередей, планирование ресурсов, предсказание производительности, предсказание задержки доставки пакетов, размер буфера и т.д [4]. Исследование влияния самоподобия трафика на производительность сети показало, что чем выше нагрузка на сеть Ethernet, тем выше оцениваемый параметр Херста [6]. Точное определение показателя Херста является важной и актуальной задачей, так как вопрос производительности приобретает значимость как раз при высокой нагрузке в сети.

Целью данной работы является вычисление показателя Херста H по модельным данным и сравнительный анализ основных методов расчета H . Для оценки параметра Херста временного ряда существует множество методов: R/S анализ (метод нормированного размаха), изменение дисперсии агрегированного ряда, оценочная формула Уитла, изменение фактора Фано и др. Однако на практике наиболее используемыми являются R/S анализ и изменение дисперсии.

Анализ литературы. Метод нормированного размаха [5, 6]. При исследовании R/S –статистики временного ряда $x(t)$ длиной τ определяется отношение $\frac{R(\tau)}{S(\tau)}$, где $R(\tau)$ – размах кумулятивного ряда

$x^{\text{cum}}(t, \tau)$, $S(\tau)$ – среднее квадратичное отклонение исходного ряда:

$$R/S = \frac{R(\tau)}{S(\tau)} = \frac{\max(x^{\text{cum}}(t, \tau)) - \min(x^{\text{cum}}(t, \tau))}{\sqrt{\frac{1}{\tau-1} \sum_{t=1}^{\tau} (x(t) - \bar{x})^2}}, \quad t = \overline{1, \tau},$$

где $\bar{x}(\tau) = \sum_{t=1}^{\tau} x(t)$; $x^{\text{cum}}(t, \tau) = \sum_{i=1}^t x(i) - \bar{x}(\tau)$, $t = \overline{1, \tau}$.

В числителе этой дроби находится величина размаха процесса, а в знаменателе – выборочное среднеквадратическое отклонение. Для самоподобного процесса это отношение при больших значениях τ обладает следующей характеристикой:

$$M\left[\frac{R}{S}\right] = (c \cdot \tau)^H,$$

где c – некоторая постоянная величина.

Если прологарифмировать обе части этого выражения, мы получим:

$$\log\left(M\left[\frac{R}{S}\right]\right) = H \cdot \log(c) + H \cdot \log(\tau).$$

Значение H вычисляется как тангенс угла наклона прямой аппроксимирующей зависимость $\log \frac{R(\tau)}{S(\tau)}$ от $\log(\tau)$.

Метод дисперсии [6, 7]. Под агрегацией по шкале времени с параметром m понимается переход к такому процессу $x^{(m)}$, что $x_k^{(m)} = \frac{1}{m} \sum_{i=km-m+1}^{km} x_i$. Для агрегированных временных серий $x^{(m)}$ самоподобного процесса дисперсия при больших значениях m подчиняется следующей формуле:

$$\text{Var}(x^{(m)}) \sim \frac{\text{Var}(x)}{m^\beta}.$$

Здесь параметр самоподобия $H = 1 - \frac{\beta}{2}$. Если прологарифмировать эту формулу, то получим:

$$\log[\text{Var}(x^{(m)})] \sim \log[\text{Var}(x)] - \beta \cdot \log(m).$$

График зависимости $\log [\text{Var}(x^{(m)})]$ от $\log(m)$ в логарифмическом масштабе будет представлять собой прямую линию с наклоном, равным $-\beta$. График можно построить по набору данных $x(t)$, если сгенерировать агрегированный процесс на разных уровнях агрегации m , а затем вычислить дисперсию. После этого определяют значение параметра $H = 1 - \frac{\beta}{2}$.

Математическая модель. В качестве формальной математической модели сетевого трафика как самоподобного процесса обычно используется фрактальное броуновское движение [6 – 8]. При обычном случайном блуждании частицы для любой пары моментов времени t и t_0 приращение ΔX координаты частицы задается выражением

$$X(t) - X(t_0) \sim \xi |t - t_0|^H,$$

где значение $H = 1/2$; ξ – случайная величина со стандартным нормальным распределением. Обобщенное (фрактальное) броуновское движение $B_H(t)$ вводится путем замены показателя $H = 1/2$ на действительное число из интервала $(0,1)$. Фрактальное броуновское движение также может быть получено из обычного броуновского движения взятием интеграла:

$$B_H(t) = \frac{1}{\Gamma(H + 0,5)} \int_{-\infty}^t (t - u)^{H-0,5} dB(u),$$

где $\Gamma(x)$ – гамма-функция. Таким образом, значение $B_H(t)$ в момент времени t зависит от всех предыдущих ($u < t$) приращений $dB(u)$ простого броуновского движения.

Приращения этого процесса (фрактальный гауссовский шум) имеют нулевое математическое ожидание и дисперсию, зависящую от времени:

$$D(B_H(t) - B_H(t_0)) = D[\xi] \cdot |t - t_0|^{2H}.$$

Корреляционная функция фрактального гауссовского шума имеет вид

$$r(k) = \frac{1}{2} [(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}].$$

При $\frac{1}{2} < H < 1$ фрактальный шум является самоподобным процессом в широком смысле.

Результаты численного эксперимента. В работе представлены результаты численного эксперимента, в ходе которого моделировался

фрактальный гауссовский шум с заданным показателем H . Значения H для модельного ряда изменялись от 0,6 до 0,9, так как параметр Херста в реальных трафиках лежит в этом диапазоне.

Для полученного временного ряда рассчитывалась оценка H методами R/S анализа ($\hat{H}_{R/S}$) и дисперсии (\hat{H}_d). Оценки $\hat{H}_{R/S}$ и \hat{H}_d вычислялись как тангенс угла наклона прямой, аппроксимирующей зависимости $\log \frac{R(\tau)}{S(\tau)}$ от $\log(\tau)$ и $\log D(x^{(m)})$ от $\log(m)$ соответственно.

Для каждого значения H рассчитывались средние значения оценок $\bar{H}_{R/S}$ и \bar{H}_d , полученные по множеству различных реализаций фрактального шума.

На рис. 1 показана типичная зависимость средних оценок от значений показателя Херста. Сплошная прямая соответствует истинным значениям H . Средние квадратичные отклонения оценок практически не зависят от значения параметра H и равны $S_{R/S} \approx 0,08$ и $S_d \approx 0,05$.

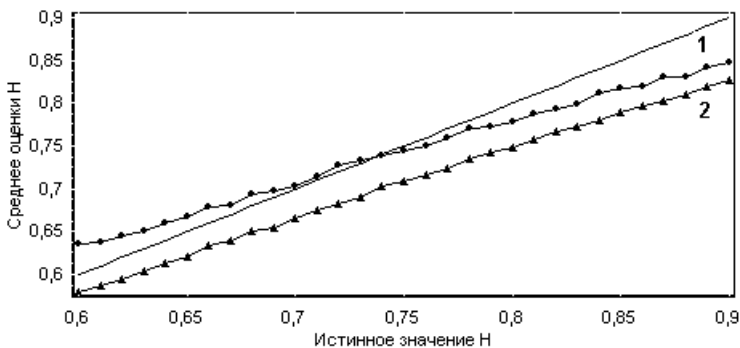


Рис. 1. Значения \bar{H}_d (линия 1) и $\bar{H}_{R/S}$ (линия 2) для заданных H

Результаты исследований показали, что, начиная со значений $H \approx 0,7 - 0,75$, оба метода дают заниженные значения, причем с увеличением H эта разница возрастает. На рис. 2 показаны отклонения $H - \hat{H}$ оценок для обоих методов.

В работе был рассмотрен вопрос о законе распределения оценок H . На рис. 3 приведена типичная гистограмма распределения оценки H . Была выдвинута гипотеза о нормальном распределении выборочных значений $\hat{H}_{R/S}$ и \hat{H}_d с соответствующими параметрами $N(\bar{H}_{R/S}, S_{R/S})$ и $N(\bar{H}_d, S_d)$. По критерию Пирсона гипотеза была принята с уровнем значимости $\alpha = 0,05$ практически для всех выборочных данных.

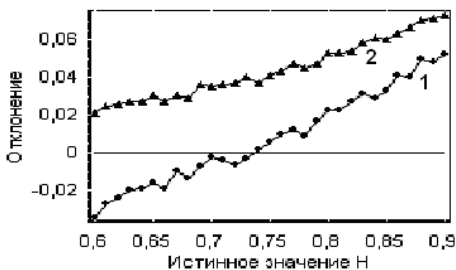


Рис. 2. Отклонения $H - \bar{H}_d$ (линия 1) и $H - \bar{H}_{R/S}$ (линия 2) для заданных H

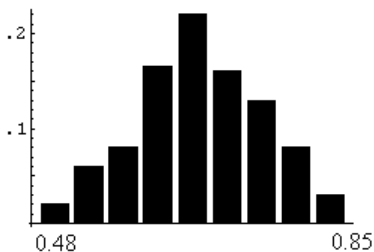


Рис. 3. Гистограмма распределения оценки $\hat{H}_{R/S}$, значение $H = 0,65$

Выводы. Таким образом, исходя из полученных результатов можно сделать вывод, что оценки показателя Херста, полученные методами R/S анализа и изменения дисперсии агрегированного ряда, имеют нормальное распределение, среднее значение которого значительно смещено относительно оцениваемого параметра. Особую важность это приобретает при расчете параметра Херста по реализации трафика как показателя нагрузки сети.

ЛИТЕРАТУРА

1. Петров В.В., Платов В.В. Исследование самоподобной структуры теле-трафика беспроводной сети // Радиотехнические тетради. – 2004. – № 30. – С. 58 – 62.
2. Цыбаков Б.С. Модель телетрафика на основе самоподобного случайного процесса // Радиотехника. – 1999. – № 5. – С. 24 – 31
3. Митилино С. Фрактальная катастрофа TCP/IP // Компьютерное обозрение. – 2001. – № 9. – С. 7.
4. Кучерявый Е. Управление трафиком и качество обслуживания в сети Интернет. – С.-Пб.: Наука и Техника, 2004. – 336 с.
5. Федер Е. Фракталы. – М.: Мир, 1991. – 254 с.
6. Столлингс В. Современные компьютерные сети. – С.-Пб.: Питер, 2003. – 784 с.
7. Городецкий А.Я., Заборовский В.С. Информатика. Фрактальные процессы в компьютерных сетях: Учебное пособие. – С.-Пб.: ГТУ, 2000. – 102 с.
8. Шелухин О.И., Теняшев А.М., Осин А.В. Фрактальные процессы в телекоммуникациях: Монография / Под ред. О.И. Шелухина. – М.: Радиотехника, 2003. – 480 с.

Поступила 2.09.2005

Рецензент: доктор технических наук, профессор В.И. Хаханов,
Харьковский национальный университет радиоэлектроники.