УДК 519.216

ГИПОТЕЗА О СМЕШАННОМ FBM-ТРАФИКЕ[1]

Сидорова О.И.*, **Бурдина Е.М.****, **Сергеева Е.В.****, **Тихомирова А.Ю.***
*Тверской государственный университет, г. Тверь
**ПАО МТС, г. Тверь

Поступила в редакцию 31.08.2020, после переработки 05.10.2020.

В данной статье предлагается алгоритм проверки гипотезы о наличии в трафике двух независимых компонент, описываемых фрактальными броуновскими движениями с разными параметрами Хёрста H. Тестовая статистика основана на сумме по частоте и масштабу логарифмов модулей вейвлет–коэффициентов и имеет в пределе нормальное распределение при нулевой (монофрактальный трафик) и альтернативной (мультифрактальный процесс) гипотезах.

Ключевые слова: долговременная зависимость, распределения с тяжёлыми хвостами, дробный гауссовский шум, дискретное вейвлет-разложение.

Вестник ТвГУ. Серия: Прикладная математика. 2020. № 3. С. 14–26.
https://doi.org/10.26456/vtpmk597

Введение

Современные телекоммуникационные сети не поддаются моделированию с помощью процессов Пуассона или марковских моделей, которые долгое время успешно применяли для расчета характеристик качества обслуживания в голосовых системах связи. Причиной этого являются такие характерные особенности трафика компьютерных сетей, как

– самоподобие при широком диапазоне агрегирования трафика;

– медленно убывающая корреляционная зависимость наблюдений или долгая память;

– тяжелые хвосты распределений нагрузки, приходящей от источников.

Эти особенности в первую очередь связаны с изменившимися алгоритмами передачи данных. Вместо коммутации каналов, использовавшейся в обычной телефонии, в компьютерных сетях стали применять коммутацию пакетов. Это позволило сетевым операторам оказывать пользователям интегральные услуги по передаче голоса, текста, видео не только в рамках одной сети, но и в рамках одного соединения.

[1]Работа выполнена при финансовой поддержке РФФИ (проект №18-07-00678).

Самоподобие приводит к тому, что на любых временных масштабах трафик не усредняется, а остается случайным и трудно предсказуемым, а долгая память означает, что передача одного «тяжелого» сообщения способна вызвать задержки или отказ в обслуживании других клиентов с более «легкими» заявками. Серьезная недооценка реальной нагрузки, в свою очередь, может привести к значительному ухудшению качества обслуживания, отказам пользователей от услуг данного оператора и потере дохода для пользователей и операторов.

С момента обнаружения [2,3] этих особенностей исследователи пытаются предложить и обосновать новые более адекватные модели процессов трафика [4,6]. Одной из них стало фрактальное броуновское движение (FBM) — самоподобный процесс с долгой памятью, важнейшей характеристикой которого служит показатель Хёрста H. В литературе соответствующий трафик часто называют β–**трафиком**. К настоящему моменту это хорошо изученный процесс, для которого предложены методы оценки для ряда важных характеристик качества обслуживания.

Оценка параметра самоподобия H является одной из основных задач при моделировании современных телекоммуникационных процессов. К настоящему времени разработано достаточно много методов для оценки параметра H для «чистых» процессов, включая R/S–анализ, метод агрегированной дисперсии, метод максимального правдоподобия. Особый интерес у исследователей вызывают подходы, основанные на методах вейвлет–декомпозиции, в силу свойства декорреляции изучаемых процессов в области вейвлет–коэффициентов. «Классический» алгоритм оценки H для дробного броуновского движения, основанный на анализе выборочной дисперсии вейвлет–коэффициентов, описан в работе [7].

Однако реальные процессы трафика не всегда поддаются качественной подгонке в рамках «чистых» моделей. Более адекватной представляется модель, основанная на сумме процессов с разными H. В связи с этим возникает интересная и актуальная задача проверки реального трафика на присутствие в нем разных компонент.

Следует отметить, что в работе [1] предложен состоятельный и несмещённый алгоритм оценки экспонент Хёрста для трафика с разными фрактальными броуновскими компонентами, обобщающий алгоритм Абри–Вейча. В определённом смысле этот подход уже проверяет гипотезу о мультифрактальном характере трафика. Однако этот метод не лишён некоторых серьёзных недостатков, в том числе:

1. оценки параметров H_1 и H_2 строятся не по всем наблюдениям, а лишь по части октав, причём число октав для оценки доминируемой компоненты должно быть меньше чем для доминирующей. Хорошо известно, что чем меньше наблюдений при оценке регрессии, тем менее достоверными являются полученные результаты;

2. для более устойчивых результатов из рассмотрения приходится исключать уровни с малым числом коэффициентов, что опять же негативным образом влияет на достоверность результатов;

3. оценки H могут получиться выше единицы, что неприемлемо;

4. кроме того, как показывают результаты моделирования, при сохранении медианного значения близким к заданным экспонентам Хёрста H_1 и H_2, конкретные оценки тем не менее могут существенно отклоняться от них.

Поэтому в данной работе мы рассматриваем более формальный подход к тестированию гипотезы о мультифрактальном трафике, также основанный на вейвлет–разложении входящего потока.

1. Фрактальное броуновское движение

Определение 1. *Случайный процесс* $Y = \big(Y(t),\ t \geqslant 0\big)$ *называется* ***процессом Леви****, если выполнены условия:*

1. $Y(0) = 0$ *почти наверное;*
2. Y *имеет независимые и однородные (по времени) приращения;*
3. Y *является стохастически непрерывным;*
4. *траектории* Y *непрерывны справа и имеют конечные пределы слева при* $t > 0$.

В силу независимости и однородности приращений, распределение процесса Y полностью и единственным образом определяется распределением с.в. $Y(1)$, которое обладает свойством безграничной делимости.

Наиболее известным примером процесса Леви является броуновское движение (Винеровский процесс).

Определение 2. *Процесс Леви* $B = (B(t),\ t \geqslant 0)$ *называется* ***броуновским движением*** *(Brownian Motion = BM), если для любых* $t \geqslant 0,\ h > 0$ *случайная величина* $B(t+h) - B(t)$ *имеет гауссовское распределение со средним 0 и дисперсией* $\sigma^2 \cdot h$.

Если $\sigma^2 = 1$, то говорят, что соответствующее броуновское движение является стандартным. Нетрудно показать, что

$$K(t,s) = Cov(B(t), B(s)) = \sigma^2 \min(t,s).$$

Приращения броуновского движения имеют нормальное распределение. В силу центральной предельной теоремы такие распределения получаются асимптотически для нормированных сумм независимых и одинаково распределенных случайных величин с конечной дисперсией.

Определение 3. *Случайный процесс* $X = (X(t),\ t \geqslant 0)$ *называется* ***самоподобным*** *с параметром Херста* $0 < H < 1$, *если он удовлетворяет условию*

$$X(t) \stackrel{d}{=} c^{-H} X(ct),\ \forall t \geqslant 0,\ \forall c > 0, \tag{1}$$

где символ $\stackrel{d}{=}$ *означает равенство конечномерных распределений.*

Можно показать, что ковариационная функция самоподобного процесса с конечными моментами второго порядка имеет вид:

$$\gamma(t,s) = \frac{\sigma^2}{2}\Big(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\Big), \quad 0 < H < 1. \tag{2}$$

Одним из наиболее известных и наиболее популярных примеров таких процессов является дробное броуновское движение.

Определение 4. *Дробным броуновским движением (fractal brownian motion, FBM) с параметром H называется гауссовский процесс* $(B_H(t), t \geqslant 0)$ *с нулевым средним и ковариационной функцией, заданной в (2).*

Однородность ковариационной функции обуславливает самоподобие дробного броуновского движения

$$B_H(at) \sim |a|^H \cdot B_H(t).$$

Из (2) также следует, что при $H = 0.5$ приращения процесса независимы (обычное броуновское движение); при $0.5 < H < 1$ — приращения процесса положительно коррелированы; при $0 < H < 0.5$ — приращения процесса отрицательно коррелированы.

Асимптотически

$$\gamma_H(k) \sim k^{-\beta} \cdot L(k), \quad 0 < \beta = 2 - 2H < 1, \quad k \to \infty,$$

где $L(k)$ есть медленно меняющаяся на бесконечности функция, что даёт несуммируемую корреляционную функцию

$$\sum_{n=0}^{\infty} \rho(n) = \infty,$$

характеризующую свойство **долгой памяти**.

2. Гипотеза о наличии в трафике разных β–компонент

2.1 Постановка задачи

Предполагается, что входящий сигнал есть композиция двух независимых фрактальных броуновских движений (FBM) с разными параметрами Херста H, т.е.

$$X(t) = B_{H_1}(t) + B_{H_2}(t), \quad t = 1, 2, \ldots, N, \tag{3}$$

где $0.5 < H_2 < H_1 < 1$, $N = 2^J$ — длина сигнала.

В силу самоподобия имеем

$$X(ct) = c^{H_1} B_{H_1}(t) + c^{H_2} B_{H_2}(t), \quad c > 0. \tag{4}$$

Мы проверяем гипотезу «существенности» компонент разного типа, т.е. рассматриваем предположение вида

$$\begin{array}{lll} Hyp_0: & X(t) = B_{H_1}(t) & - \text{ монофрактальный трафик,} \\ Hyp_1: & X(t) = B_{H_1}(t) + B_{H_2}(t) & - \text{ мультфрактальный трафик.} \end{array} \tag{5}$$

Построим тестовую статистику, опираясь на дискретное вейвлет–преобразование (ДВП). ДВП основано на ортонормированном базисе функций вида

$$\psi_{j,k}(t) = 2^{-j/2} \psi_0(2^{-j} t - k), \quad j, k \in \mathrm{Z},$$

известным в литературе как **диадное вейвлет–преобразование**. Здесь $\psi_0(t)$ и $\psi_{j,k}(t)$ есть материнский и дочерний вейвлет, $1 \leqslant j \leqslant J$ отвечает за глубину разложения или уровень декомпозиции сигнала, причём за нулевой уровень $j = 0$ обычно принимается уровень максимального временного разрешения сигнала, т.е. сам сигнал, $J = \lfloor \log_2 N \rfloor$ — число октав, $k = \overline{1, n_j}$ — номер коэффициента, а $n_j = \lfloor 2^{-j} \cdot N \rfloor$ — число доступных вейвлет–коэффициентов на масштабе j.

Частотно–временная декомпозиция *приращений* процесса с помощью вейвлетов приводит нас к коэффициентам вида

$$d_{j,k} = \int\limits_{-\infty}^{\infty} \psi_{j,k}(t)\, dX(t) = \int\limits_{-\infty}^{\infty} 2^{-0.5 \cdot j} \cdot \psi_{0,k}(2^{-j} \cdot t - k)\, dX(t) =$$

$$= \int\limits_{-\infty}^{\infty} 2^{-0.5 \cdot j} \cdot \psi_{0,k}(u - k)\, dX(2^j \cdot u) = \int\limits_{-\infty}^{\infty} 2^{j \cdot (H - 0.5)} \cdot \psi_{0,k}(u - k)\, dX(u) =$$

$$= \begin{cases} 2^{j \cdot \nu_1} \cdot \theta_1 \cdot d_{0,k}, & \text{при верной } H_0, \\ 2^{j \cdot \nu_1} \cdot \theta_1 \cdot d_{0,k} + 2^{j \cdot \nu_2} \cdot \theta_2 \cdot d_{0,k}, & \text{при верной } H_1, \end{cases}$$

где $\nu_1 = H_1 - 0.5$ и $\nu_2 = H_2 - 0.5$.

С учётом самоподобия фрактального броуновского движения получаем следующую структуру вейвлет–коэффициентов:

$$\begin{aligned} Hyp_0: &\ d_{j,k} = 2^{j \cdot \nu_1} \cdot \theta_1 \cdot B(1), \\ Hyp_1: &\ d_{j,k} = \sqrt{2^{2j \cdot \nu_1} \cdot \theta_1^2 + 2^{2j \cdot \nu_2} \cdot \theta_2^2} \cdot B(1), \end{aligned}$$

где $B(1) \sim N(0,1)$.

В силу декоррелирующих свойств ортогонального ДВП можно далее полагать, что:

1. для фиксированного j с.в. $\{d(j,k)\}_{k \in Z}$ независимы;
2. последовательности $\{d(j,k)\}_{k \in Z}$ при разных j независимы.

Это приводит к тому, что

$$|d_{j,k}| \stackrel{d}{=} \begin{cases} 2^{j \cdot \nu_1} \cdot c_1 \cdot |B_k|, & \text{при верной } H_0, \\ 2^{j \cdot \nu_1} \cdot c_2 \cdot |B_k|, & \text{при верной } H_1, \end{cases} \tag{6}$$

где $c_1 = \theta_1$, $c_2 = \theta_1 \cdot \sqrt{1 + \lambda^2 \cdot 2^{2j \cdot (\nu_2 - \nu_1)}}$, $\lambda = \theta_2 / \theta_1 > 0$ — так называемое отношение «сигнал–шум» (signal–to–noise ratio).

2.2 Анализ тестовой статистики

Рассмотрим величину

$$T_j = \frac{1}{n_j} \sum_{k=1}^{n_j} \log_2 \bigl| d_{j,k} \bigr| = \begin{cases} j \cdot \nu_1 + \log_2 c_1 + \frac{1}{n_j} \sum\limits_{k=1}^{n_j} \log_2 |B_k|, & \text{при верной } H_0, \\ j \cdot \nu_1 + \log_2 c_2 + \frac{1}{n_j} \sum\limits_{k=1}^{n_j} \log_2 |B_k|, & \text{при верной } H_1. \end{cases} \tag{7}$$

Вычислим характеристики случайной величины $X = \ln|B|,\ B \sim (0,1)$, опираясь на свойства производящих функций моментов

$$M\left(X^n\right) = \frac{d^n}{dt^n} M\left(e^{tX}\right)\Big|_{t=0}, \quad n = 1, 2, \ldots.$$

Имеем

$$M\bigl(e^{tX}\bigr) = M\bigl(|B|^t\bigr) = \int\limits_{-\infty}^{\infty} \frac{|x|^t}{\sqrt{2\pi}} \cdot e^{-x^2/2}\, dx = \frac{2^{t/2}}{\sqrt{\pi}} \int\limits_{0}^{\infty} u^{\frac{t+1}{2}-1} \cdot e^{-u}\, du = \frac{2^{\frac{t}{2}} \cdot \Gamma\left(\frac{t+1}{2}\right)}{\sqrt{\pi}}, \tag{8}$$

где $\Gamma(x)$ — гамма–функция Эйлера.

С учётом $\Gamma(0.5) = \sqrt{\pi}$, имеем

$$M(X) = \frac{d}{dt} M\left(e^{tX}\right)\Big|_{t=0} = \left[\frac{2^{t/2} \cdot \ln 2 \cdot \Gamma\left(\frac{t+1}{2}\right)}{2 \cdot \sqrt{\pi}} + \frac{2^{t/2} \cdot \Gamma'\left(\frac{t+1}{2}\right)}{2 \cdot \sqrt{\pi}}\right]\Bigg|_{t=0} =$$

$$= \frac{\ln 2 \cdot \Gamma\left(0.5\right)}{2 \cdot \Gamma(0.5)} + \frac{\Gamma'\left(0.5\right)}{2 \cdot \Gamma(0.5)} = \frac{\ln 2 + \psi(0.5)}{2} = -\frac{\ln 2 + \gamma}{2}, \tag{9}$$

где $\psi(x) = \psi^{(0)}(x) = \dfrac{\Gamma'(x)}{\Gamma(x)}$ — дигамма–функция, $\psi(0.5) = -2\ln 2 - \gamma$, $\gamma \approx 0.5772$ — постоянная Эйлера.

$$M(X^2) = \frac{d^2}{dt^2} M\left(e^{tX}\right)\Big|_{t=0} = \left[\frac{2^{t/2} \cdot \ln^2 2 \cdot \Gamma\left(\frac{t+1}{2}\right)}{4 \cdot \sqrt{\pi}} + \frac{2^{t/2} \cdot \ln 2 \cdot \Gamma'\left(\frac{t+1}{2}\right)}{2 \cdot \sqrt{\pi}} + \right.$$

$$\left. + \frac{2^{t/2} \cdot \Gamma''\left(\frac{t+1}{2}\right)}{4 \cdot \sqrt{\pi}}\right]\Bigg|_{t=0} = \frac{\ln^2 2}{4} + \frac{\ln 2 \cdot \psi(0.5)}{2} + \frac{\psi_1(0.5) + \psi^2(0.5)}{4},$$

где $\psi_1(x) = \dfrac{\mathrm{d}^2}{\mathrm{dx}^2} \ln\Gamma(x) = \dfrac{\Gamma''(x)}{\Gamma(x)} - \psi^2(x)$ — тригамма–функция, $\psi_1(0.5) = 0.5 \cdot \pi^2$.

Вычисляя дисперсию, получаем

$$D(X) = M(X^2) - (MX)^2 = \frac{\psi_1(0.5)}{4} = \frac{\pi^2}{8}. \tag{10}$$

В силу независимости $d_{j,k}$ для разных k, при верной гипотезе H_0 мы получаем

$$\mu_j = M\bigl(T_j|H_0\bigr) = j \cdot \nu_1 + \log_2 c_1 - \frac{\gamma + \ln 2}{2 \cdot \ln 2}, \quad \sigma_j^2 = D\bigl(T_j|H_0\bigr) = \frac{\pi^2}{8 \cdot n_j \cdot \ln^2 2}.$$

Поскольку $\mu_j < \infty$ и $\sigma_j < \infty$ в силу центральной предельной теоремы

$$T_j|Hyp_0 \sim N(\mu_j,\, \sigma_j^2), \quad n_j \to \infty.$$

Таким образом, по крайней мере на уровнях декомпозиции $1 \leqslant j \leqslant J_0 < J$ с числом коэффициентов $n_j \geqslant 30$ можно ожидать нормального распределения для T_j.

Аналогичные рассуждения дают нам

$$T_j|Hyp_1 \sim N(\widetilde{\mu}_j, \widetilde{\sigma}_j^2), \quad n_j \to \infty,$$

где

$$\widetilde{\mu}_j = M\big(T_j|Hyp_1\big) = j \cdot \nu_1 + \log_2 c_2 - \frac{\gamma + \ln 2}{2 \cdot \ln 2},$$

$$\widetilde{\sigma}_j^2 = D\big(T_j|Hyp_1\big) = \sigma_j^2 = \frac{\pi^2}{8 \cdot n_j \cdot \ln^2 2}.$$

Определим тестовую статистику

$$T_F = \frac{1}{J_0} \sum_{j=1}^{J_0} T_j. \tag{11}$$

В силу независимости $d_{j,k}$ при разных j для статистики T_F при верной H_0 справедливо

$$\mu = M\big(T_F|Hyp_0\big) = \frac{J_0+1}{2} \cdot \nu_1 + \log_2 c_1 - \frac{\gamma + \ln 2}{2 \ln 2}.$$

В силу равенства $n_j = 2^{J-j}$, $1 \leqslant j \leqslant J$, имеем бесконечно убывающую геометрическую прогрессию с $b_1 = 1$ и $q = 0.5$. Следовательно,

$$\sum_{j=1}^{J_0} \frac{1}{2^{J-j}} = \frac{1}{2^{J-J_0}} \sum_{j=1}^{J_0} \frac{1}{2^{J_0-j}} \to \frac{1}{2^{J-J_0-1}}, \quad J_0 \to \infty$$

и, следовательно,

$$\sigma^2 = D\big(T_F|H_0\big) = \frac{\pi^2}{8 \cdot J_0^2 \cdot \ln^2 2} \cdot \sum_{k=1}^{J_0} \frac{1}{2^{J_0-j}} \approx \frac{\pi^2}{J_0^2 \cdot 2^{J-J_0+2} \cdot \ln^2 2}.$$

Для гипотезы H_1 соответствующие характеристики будут равны

$$\widetilde{\mu} = M\big(T_F|H_1\big) = \frac{J_0+1}{2} \cdot \nu_1 + \log_2 c_2 - \frac{\gamma + \ln 2}{2 \cdot \ln 2}$$

и

$$\widetilde{\sigma}^2 = D\big(T_F|H_1\big) = \sigma^2 \approx \frac{\pi^2}{J_0^2 \cdot 2^{J-J_0+2} \cdot \ln^2 2}.$$

При известных значениях параметров H_1, H_2, c, c_1 средние μ и $\widetilde{\mu}$ могут быть вычислены в явном виде. Поэтому гипотезу о «существенности» разных компонент в трафике можно свести к тестированию гипотезы о среднем для нормального распределения при известной дисперсии σ^2, т.е.

$$\begin{array}{l} Hyp_0 : \ mean = \mu, \\ Hyp_1 : \ mean = \widetilde{\mu}, \end{array} \quad \widetilde{\mu} > \mu. \tag{12}$$

Основанное на статистике $T_F|Hyp_0 \sim N(\mu, \sigma^2)$ **правило принятия решения** имеет вид

$$\varphi(x) = \begin{cases} 1, & \text{если} \quad T_F > q_\alpha = \mu + \sigma \cdot u_\alpha, \\ 0, & \text{иначе}, \end{cases} \tag{13}$$

где $u_\alpha = \Phi^{-1}(1-\alpha)$ — верхняя $\alpha\%$ точка стандартного нормального закона.

Вероятность «правильного решения» (мощность критерия)

$$\beta = P\big(Hyp_1|Hyp_1\big) = P\big(T_F > \mu + \sigma \cdot u_\alpha | H_1\big) = 1 - \Phi\left(\frac{\mu + \sigma \cdot u_\alpha - \widetilde{\mu}}{\sigma}\right) =$$

$$= 1 - \Phi\left(\frac{\log_2 c_1 - \log_2 c_2}{\sigma} + u_\alpha\right).$$

Замечание 1. Для построения критерия можно было использовать статистику, основанную на сумме квадратов вейвлет–коэффициентов. Из [7] следует, что при верной гипотезе Hyp_0 асимптотически при $n_j \to \infty$ случайная величина

$$\log_2(\tau_j) \overset{d}{\sim} N\left(\beta \cdot j + log_2(c_1^2) + g_j, \; \frac{2}{n_j \cdot \ln^2 2}\right), \quad \tau_j = \frac{1}{n_j}\sum_{k=1}^{n_j} d_{jk}, \tag{14}$$

где $\beta_1 = 2H_1 - 1$, $g_j \sim -\dfrac{1}{n_j \cdot \ln 2}$.

Определим тестовую статистику

$$\widetilde{T}_F = \frac{1}{J_0}\sum_{j=1}^{J_0} \log_2(\tau_j). \tag{15}$$

Тогда при верной гипотезе Hyp_0 случайная величина $\widetilde{T}_F$ будет иметь нормальное распределение с параметрами

$$\mu_1 = \frac{J_0 + 1}{2} \cdot \beta_1 + \log_2 c_1^2 - \frac{1}{J_0 \cdot 2^{J - J_0 - 1}}, \qquad \sigma_1^2 = \frac{1}{J_0^2 \cdot 2^{J - J_0 - 2} \cdot \ln^2 2}.$$

Основанное на статистике $\widetilde{T}_F$ **правило принятия решения** имеет вид

$$\varphi(x) = \begin{cases} 1, & \text{если} \quad \widetilde{T}_F > q_\alpha = \mu_1 + \sigma_1 \cdot u_\alpha, \\ 0, & \text{иначе}, \end{cases} \tag{16}$$

где $u_\alpha = \Phi^{-1}(1-\alpha)$ — верхняя $\alpha\%$ точка стандартного нормального закона.

В [5] было показано, что метод Абри–Вейча неустойчив к присутствию в сигналах нестационарных компонент. Поэтому статистика T_F представляется более предпочтительной в силу робастности.

2.3 Бутстрап–анализ

Для реальных процессов трафика значения параметров нам естественно неизвестны. Использование оценок $\widehat{H}_1$, $\widehat{H}_2$, $\widehat{lc1} = \widehat{\log_2(c_1)}$, $\widehat{lc2} = \widehat{\log_2(c_2)}$ для вычисления средних μ и μ_1 может повлиять на распределение статистики T_F и величину квантили q_α соответственно. Во избежание этого можно воспользоваться параметрическим бутстраппом:

1. по исходной выборке оценить параметры фрактальных броуновских движений с помощью обобщенного алгоритма Абри–Вейча;

2. смоделировать K процессов фрактального броуновского движения при заданном значении $\widehat{H}_1$, что соответствует трафику при нулевой гипотезе;

3. построить для каждого из K процессов вейвлет–разложение;

4. вычислить K значений статистики T_{Fboot} и найти эмпирическую квантиль $\widehat{q}_\alpha$ порядка α;

5. опираясь значение T_F, полученное по исходному процессу, построить критерий

$$\varphi_{boot}(x) = \begin{cases} 1, & \text{если} \quad T_F > \widehat{q}_\alpha, \\ 0, & \text{иначе.} \end{cases} \tag{17}$$

3. Численное моделирование

Программная реализация предложенной методики осуществлялась в системе Matlab. Результаты численной апробации метода представлены в таблице ниже.

При моделировании полагалось $J = 16$, $J_0 = 11$, $K = 1000$. Для оценки параметров использовался модифицированный алгоритм Абри–Вейча, дающий несмещённые и состоятельные оценки для H_1 и $\log_2(c_1^2)$. Вейвлет–разложение строилось с помощью семейства *Daubechies 2*. Для расчёта теоретического значения q_α использованы усреднённые значения оценок H_1 и $\log_2(c_1)$ по $K = 1000$ смоделированным траекториям монофрактального процесса. Заметим также, что в системе Matlab число коэффициентов на каждом уровне несколько выше теоретического значения $n_j = 2^{J-j}$. Поэтому при расчёте дисперсии использовалась величина $\sum\limits_{j=1}^{J_0} \frac{1}{n_j}$, вычисленная по реальному числу коэффициентов.

Исходные параметры	Оценки	Стат–ка теста	Теор. модель		Бутстрап	
			квантиль	решение	квантиль	решение
$H_1 = 0.8$, $H_2 = 0.6$	$\widehat{H}_1 = 0.8099$, $\widehat{lc1} = -2.8745$	$T_F = -1.5148$	$q_\alpha = -1.8904$	1	$\widehat{q}_\alpha = -1.9012$	1
$H_1 = 0.9$, $H_2 = 0.7$	$\widehat{H}_1 = 0.9047$, $\widehat{lc1} = -3.3521$	$T_F = -1.4063$	$q_\alpha = -1.7804$	1	$\widehat{q}_\alpha = -1.7921$	1
$H_1 = 0.7$	$\widehat{H}_1 = 0.7102$, $\widehat{lc1} = -2.3792$	$T_F = -2.1173$	$q_\alpha = -1.9933$	0	$\widehat{q}_\alpha = -2.0088$	0
$H_1 = 0.55$	$\widehat{H}_1 = 0.5565$, $\widehat{lc1} = -1.5970$	$T_F = -2.1589$	$q_\alpha = -2.1333$	0	$\widehat{q}_\alpha = -2.1456$	0

Результаты моделирования показывают, что теоретическая квантиль q_α, рассчитанная по усреднённым оценкам, получается близкой эмпирической величине $\widehat{q}_\alpha$. Однако конкретные оценки могут отличаться от усреднённых значений, поэтому эмпирическая квантиль выглядит более предпочтительной.

Заключение

В настоящей статье рассматривается метод тестирования самоподобного трафика на присутствие в нём компонент с разными свойствами. Критерий проверки основан на сумме по частоте и масштабу логарифмов модулей вейвлет–коэффициентов. Полученная статистика имеет асимптотически нормальное распределение, что позволяет легко проверять соответствующую гипотезу.

Необходимость оценки параметров трафика для расчёта теоретической квантили с одной стороны и ее асимптотический характер с другой стороны требуют дальнейшего исследования свойств статистики критерия. В качестве альтернативы можно использовать критические точки, полученные с помощью методов статистического моделирования.

Список литературы

[1] Галактионова О.В., Хохлов Ю.С. Оценка параметров неоднородного трафика // Вестник ТвГУ. Серия: Прикладная математика. 2010. № 18. С. 87–102.

[2] Crovella M., Bestavros A. Self-similarity in world wide web traffic: evidence and possible cases // Proceedings of the 1996 ACM SIGMETRICS International Conference on Measurement and Modelling of Computer Systems. Vol. 4. 1996. Pp. 160–169.

[3] Leland W.E., Taqqu M.S., Willinger W., Willson D.V. On the self-similar nature of Ethernet traffic (Extended version) // IEEE/ACM Transactions on Networking. 1994. Vol. 2. Pp. 1–15.

[4] Mikosch Th., Resnick S., Rootzen H., Stegeman A. Is network traffic approximated by stable Levy motion or fractional Brownian motion? // Annals of Applied Probability. 2002. Vol. 12, № 1. Pp. 23–68.

[5] Park J., Park C. Robust estimation of the Hurst parameter and selection of an onset scaling // Statistica Sinica. 2009. Vol. 19. Pp. 1531–1555.

[6] Taqqu M., Willinger W., Sherman R. Proof of a fundamental result in self-similar traffic modeling // Computer Communications Review. 1997. Vol. 27, № 2. Pp. 5–23.

[7] Veitch D., Abry P. A wavelet-based joint estimator of the parameters of long-range dependence // IEEE Transactions on Information Theory. 1999. Vol. 45, № 3. Pp. 878–897.

Образец цитирования

Сидорова О.И., Бурдина Е.М., Сергеева Е.В., Тихомирова А.Ю. Гипотеза о смешанном FBM-трафике // Вестник ТвГУ. Серия: Прикладная математика. 2020. № 3. С. 14–26. https://doi.org/10.26456/vtpmk597

Сведения об авторах

1. **Сидорова Оксана Игоревна**
   доцент кафедры математической статистики и системного анализа Тверского государственного университета.

   Россия, 170100, г. Тверь, ул. Желябова, д. 33, ТвГУ.
   E-mail: *oksana.i.sidorova@yandex.ru*

2. **Бурдина Елизавета Михайловна**
   специалист по тестированию программ и услуг мобильного бизнеса ПАО МТС.

   Россия, 170024, г. Тверь, пр-т 50 лет Октября, 3Б, ПАО МТС.
   E-mail: *lizburdina@yandex.ru*

3. **Сергеева Евгения Вадимовна**
   специалист по тестированию программ и услуг фиксированного бизнеса ПАО МТС.

   Россия, 170024, г. Тверь, пр-т 50 лет Октября, 3Б, ПАО МТС.
   E-mail: *evgeniyasergeeva69@mail.ru*

4. **Тихомирова Александра Юрьевна**
   студентка факультета прикладной математики и кибернетики Тверского государственного университета.

   Россия, 170100, г. Тверь, ул. Желябова, д. 33, ТвГУ. E-mail: *al-town@yandex.ru*

MIXED FBM-TRAFFIC HYPOTHESIS

Sidorova Oksana Igorevna
Associate Professor in the Department of Mathematical Statistics
and System Analysis, Tver State University
Russia, 170100, Tver, 33 Zhelyabova str., TverSU.
E-mail: oksana.i.sidorova@yandex.ru

Burdina Elizaveta Mikhailovna
QA for programs and services for mobile business, MTS
Russia, 170024, Tver, avenue 50 let Oktyabrya, 3B, MTS.
E-mail: lizburdina@yandex.ru

Sergeyeva Evgeniya Vadimovna
QA for programs and services for fixed business, MTS
Russia, 170024, Tver, avenue 50 let Oktyabrya, 3B, MTS.
E-mail: evgeniyasergeeva69@mail.ru

Tikhomirova Aleksandra Yurievna
Student at Applied Mathematics and Cybernetics faculty,
Tver State University
Russia, 170100, Tver, 33 Zhelyabova str., TverSU.
E-mail: al-town@yandex.ru

Received 31.08.2020, revised 05.10.2020.

In this article we consider some test for presence in traffic two different independent *FBM*–components with different Hurst parameters H. Test statistics are based on the sum of frequency and scale logarithms of wavelet coefficients absolute values and asymptotically converge to a normal distribution under null (monofractal traffic) and alternative (multifractal process) hypotheses.

Keywords: long-range dependence, fractal brownian noise, Hurst parameter.

Citation

Sidorova O.I., Burdina E.M., Sergeyeva E.V., Tikhomirova A.Yu., "Mixed FBM-traffic hypothesis", *Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics]*, 2020, № 3, 14–26(in Russian). https://doi.org/10.26456/vtpmk597

References

[1] Galaktionova O.V., Khokhlov Yu.S., "Estimation of heterogeneous traffic parameters", *Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics]*, 2010, № 18, 87–102 (in Russian).

[2] Crovella M., Bestavros A., “Self-similarity in world wide web traffic: evidence and possible cases”, *Proceedings of the 1996 ACM SIGMETRICS International Conference on Measurement and Modelling of Computer Systems*. V. 4, 1996, 160–169.

[3] Leland W.E., Taqqu M.S., Willinger W., Willson D.V., “On the self-similar nature of Ethernet traffic (Extended version)”, *IEEE/ACM Transactions on Networking*, **2** (1994), 1–15.

[4] Mikosch Th., Resnick S., Rootzen H., Stegeman A., “Is network traffic approximated by stable Levy motion or fractional Brownian motion?”, *Annals of Applied Probability*, **12**:1 (2002), 23–68.

[5] Park J., Park C., “Robust estimation of the Hurst parameter and selection of an onset scaling”, *Statistica Sinica*, **19** (2009), 1531–1555.

[6] Taqqu M., Willinger W., Sherman R., “Proof of a fundamental result in self-similar traffic modeling”, *Computer Communications Review*, **27**:2 (1997), 5–23.

[7] Veitch D., Abry P., “A wavelet-based joint estimator of the parameters of long-range dependence”, *IEEE Transactions on Information Theory*, **45**:3 (1999), 878–897.