

УДК 004.7

Статистический анализ сетевого трафика

Бельков Д.В., Едемская Е.Н., Незамова Л.В.
Донецкий национальный технический университет
belkov@telenet.dn.ua

Abstract

Belkov D.V., Edemskaya E.N., Nezamova L.V. The statistical analysis of the network traffic traces. Recent network packet traffic investigations have demonstrated, that it is self-similar process and more conventional models, such a simple Markovian (memory-less) models have fault. Analysis the real traffic is an important task. This paper presents the results of the statistical analysis of RTT-delay traces.

Введение

Многочисленные исследования [1,2] трафика компьютерных сетей свидетельствуют о том, что он обладает свойством масштабной инвариантности - имеет особую фрактальную (самоподобную) структуру, сохраняющуюся на разных масштабах. В процессе передачи возникают большие всплески при относительно низком среднем уровне трафика. Данное явление, которое значительно влияет на потери, задержки и джиттер пакетов, не учитывается в классической теории телетрафика. Расчет параметров системы распределения информации, предназначенной для обработки сетевого трафика, по классическим формулам дает некорректные, неоправданно оптимистические результаты. Алгоритмы обработки трафика, созданные для работы с простейшим потоком неэффективны для фрактальных потоков.

Ситуация, сложившаяся в современных компьютерных сетях, наличие большого количества сетевых маршрутов на которых периодически возникают резкие колебания задержки в передаче данных и большие потери пакетов, появление новых свойств сетевого трафика, необходимость обеспечения высокого качества обслуживания приложений, делают актуальным статистический анализ сетевого трафика.

Целью настоящей работы является исследование структуры трафика, направленное на выявление его характерных особенностей. Изучается один из основных сетевых процессов – процесс RTT-задержки, который служит для получения информации о состоянии сети методом “черного ящика”, когда через сеть пропускается последовательность пакетов, и на основании времени их прохождения до удаленного узла и обратно делаются выводы о загрузке сети.

В работе решаются следующие задачи: оцениваются плотности распределения, автокорреляционные функции (АКФ), энергетические спектры изучаемых временных рядов, определяется степень фрактальности (показатель Херста). В исследовании использовались программы Fractan и Statistika. Алгоритм R/S-анализа выполнен в среде VBA.

1. Реализации сетевого трафика

Для изучения выбраны четыре реализации сетевого трафика [3], полученные в 2007 году в университете города Наполи (Италия). Согласно лицензии данные свободно доступны для анализа. Измерения проводились каждые 10 миллисекунд, получено свыше 12 000 отсчетов. В первом случае (ряд UDP_d64) измерялась задержка UDP-пакетов объемом 64 байт, во втором (ряд UDP_d512) - 512 байт. Третий ряд (UDP_j64) - измерения джиттера при передаче UDP-пакетов объемом 64 байт, четвертый ряд (UDP_j512) - измерения джиттера при передаче UDP-пакетов объемом 512 байт. Отправитель имел ADSL-доступ (640 Kbps), на стороне получателя - 100 Mbps Ethernet, операционная система Linux на каждой из сторон, скорость передачи 100 pps.

Необходимо отметить, что для процесса RTT-задержки на маршруте прохождения пакета можно выделить постоянную минимальную составляющую, обусловленную отсутствием очередей и переменную составляющую, возникающую из-за задержек в очередях. RTT-задержка измеряется на стороне источника. Она складывается из времени обработки, времени распространения, времени коммутации, времени ожидания в очереди. Изучаемые временные ряды показаны на рисунках 1-4. Нижний уровень задержки на рисунках 1 и 3, соответствует отсутствию очередей.

RTT-задержка пакетов объемом 64 байт

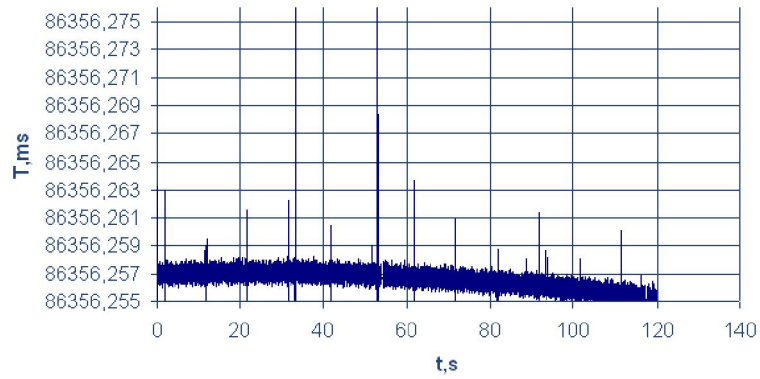


Рисунок 1.- Задержка при передаче пакетов объемом 64 байт

Джиттер при передаче пакетов объемом 64 байт

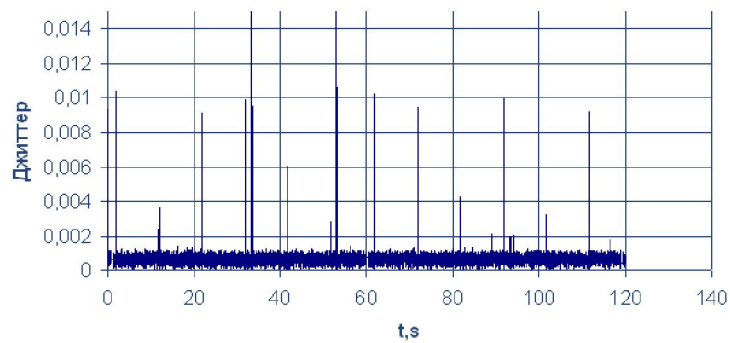


Рисунок 2.- Джиттер при передаче пакетов объемом 64 байт

RTT-задержка пакетов объемом 64 байт

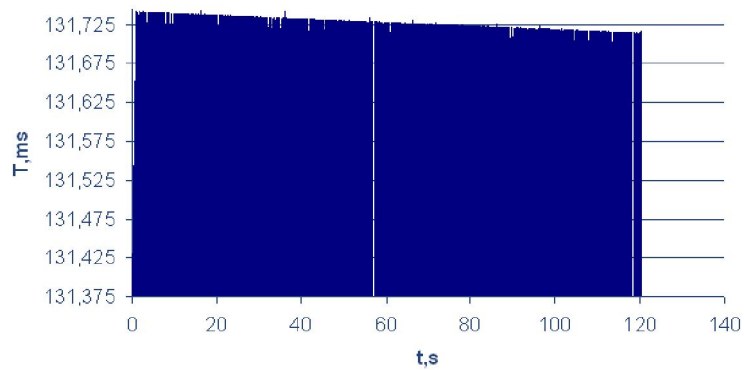


Рисунок 3.- Задержка при передаче пакетов объемом 512 байт

Джиттер при передаче пакетов объемом 512 байт

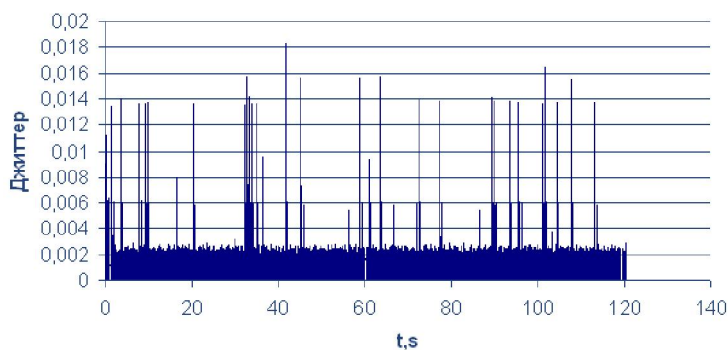


Рисунок 4.- Джиттер при передаче пакетов объемом 512 байт

2. Анализ плотности распределения

В этом разделе выполнен анализ плотностей распределения. Оценка проводится на основании гистограмм относительных

частот, показанных на рисунках 5-8. Визуальный анализ позволяет сделать следующие выводы: ряды UDP_d64, UDP_d512, UDP_j512 не подчиняются нормальному распределению. Ряд UDP_j64 близок к случайному.

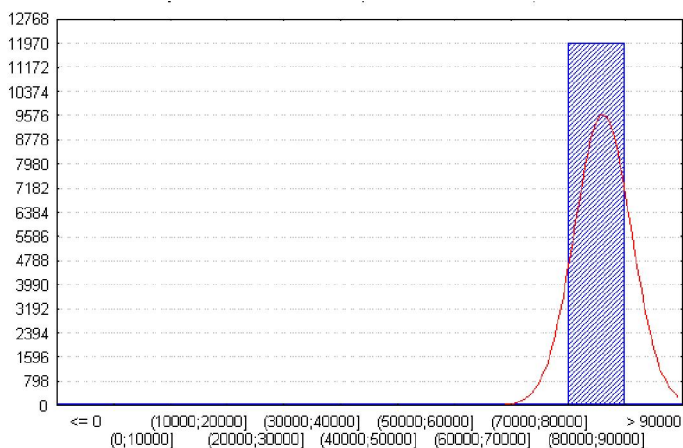


Рисунок 5.- Плотность распределения временного ряда UDP_d64

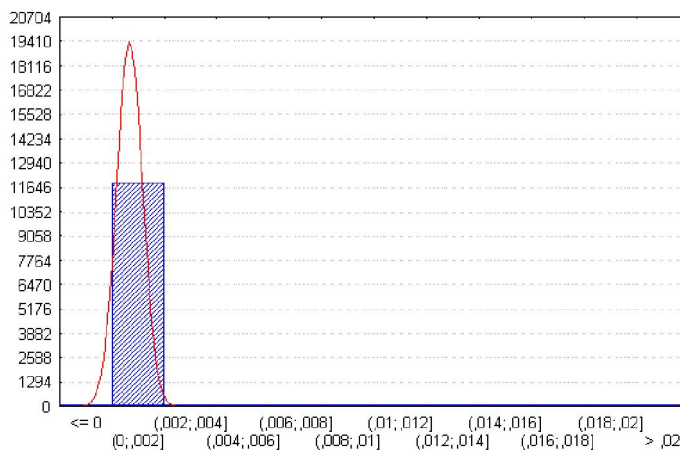


Рисунок 6.- Плотность распределения временного ряда UDP_j64

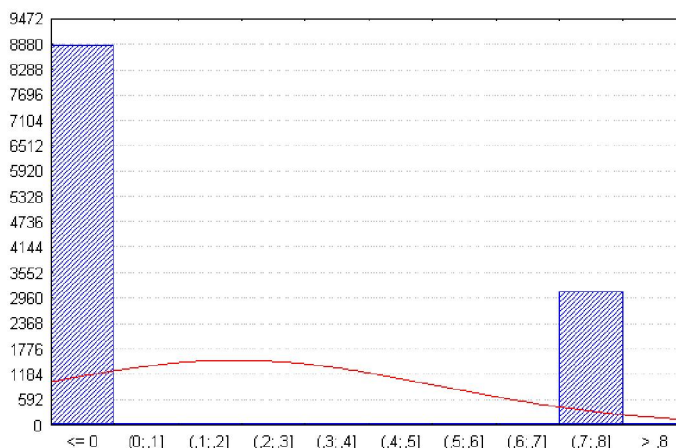


Рисунок 7.- Плотность распределения временного ряда UDP_d512

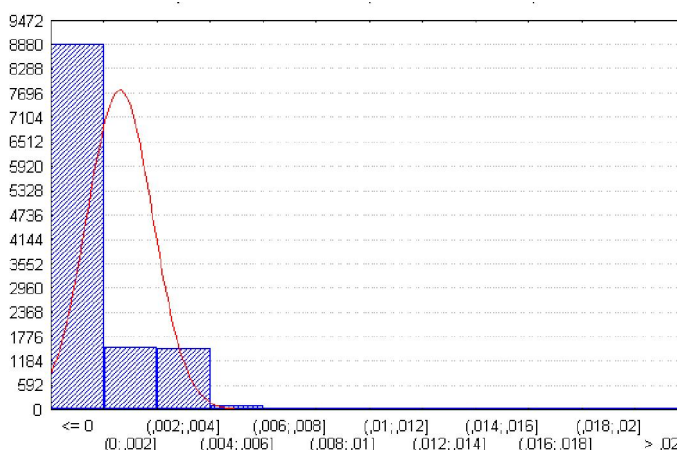


Рисунок 8.- Плотность распределения временного ряда UDP_j512

3. Анализ автокорреляционных функций

В этом разделе необходимо определить обладают ли временные ряды медленно убывающей зависимостью или быстро убывающей зависимостью.

Определение [4]. Процесс X обладает медленно убывающей зависимостью (МУЗ) [long-range dependence], если для его АКФ выполняется условие

$$r(k) \sim k^{-\beta} \cdot L_1(k) + C_1, k \rightarrow \infty, \text{ где } C_1 = const, 0 < \beta < 1 \text{ и } L_1 - \text{медленно меняющаяся на}$$

бесконечности функция: $\lim_{t \rightarrow \infty} \frac{L_1(tx)}{L_1(t)} = 1$ для всех

х. Процессы с МУЗ характеризуются автокорреляционной функцией, которая убывает по степенному закону при увеличении временной задержки (лага). В отличие от процессов с МУЗ, процессы с быстро убывающей зависимостью (БУЗ) [short-range

dependence] обладают экспоненциально спадающей АКФ вида $r(k) \sim \rho^k, k \rightarrow \infty$.

Вычисления АКФ выполнены с помощью программы Fractan 4.4 по формуле

$$r(k) = \frac{\sum_{i=1}^{N-\tau} (X_i - \bar{X})(X_{i+k} - \bar{X})}{(N - \tau)\sigma^2(X)}, \text{ где } \bar{X} -$$

выборочное среднее ряда X , $\sigma^2(X)$ - выборочная дисперсия ряда X , $k=0,1,\dots$. Автокорреляционные функции исследуемых временных рядов приведены на рисунках 9-12. Они не обращаются в ноль при больших значениях k , что говорит о медленном убывании АКФ и присутствии МУЗ во всех исследуемых реализациях трафика. Аппроксимация автокорреляционных функций будет рассмотрена в разделе 5.

АКФ задержки пакетов объемом 64 байт



Рисунок 9.- АКФ временного ряда UDP_d64

АКФ джиттера при передаче пакетов объемом 64 байт

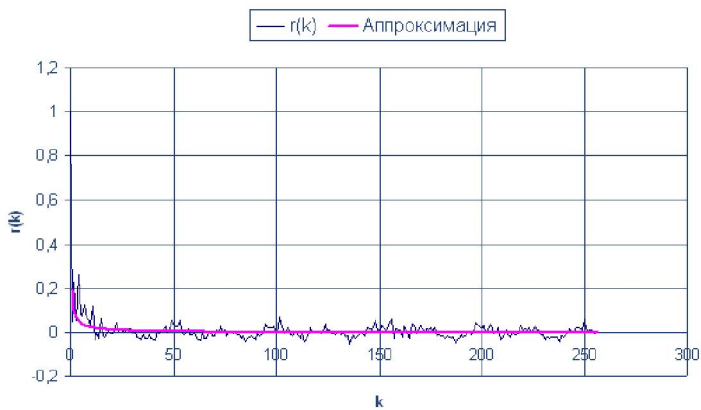


Рисунок 10.- АКФ временного ряда UDP_j64

АКФ задержки пакетов объемом 512 байт

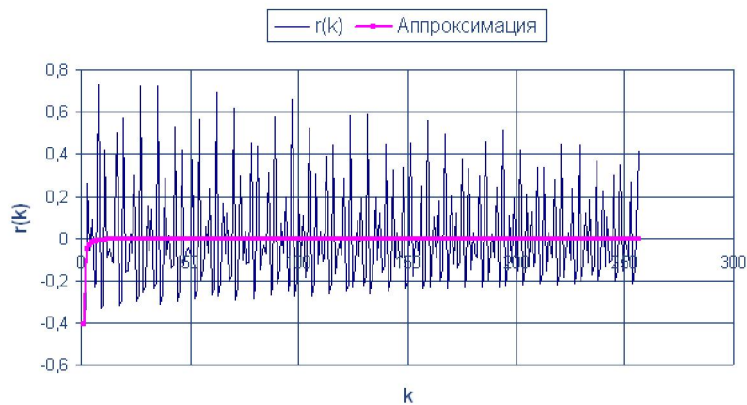


Рисунок 11.- АКФ временного ряда UDP_d512



Рисунок 12.- АКФ временного ряда UDP_j512

4. Спектральный анализ

В этом разделе выполнен расчет энергетических спектров исследуемых временных рядов. Известно [4], что в частотной области МУЗ приводит к степенному закону поведения спектральной плотности процесса. Процесс X обладает МУЗ, если для спектральной плотности $S(f) = \sum_k r(k) \cdot e^{ikf}$

выполняется условие $S(f) \sim f^{-b} \cdot L_2(f) + C_2$,

где $f \rightarrow 0$, $i = \sqrt{-1}$, $0 < b < 4$ и L_2 - медленно изменяющаяся в нуле функция, $C_2 = const$.

С точки зрения спектрального анализа процесс с МУЗ обладает спектральной плотностью с особенностью в нуле: спектральная плотность стремится к бесконечности, по мере того как частота f стремится к нулю. Энергетические спектры реализаций представлены на рисунках 13-16. Аппроксимация спектральных зависимостей для низких частот будет рассмотрена в разделе 5.

Спектральная плотность временного ряда UDP_d64



Рисунок 13.- Спектральная плотность временного ряда UDP_d64

Спектральная плотность временного ряда UDP_j64

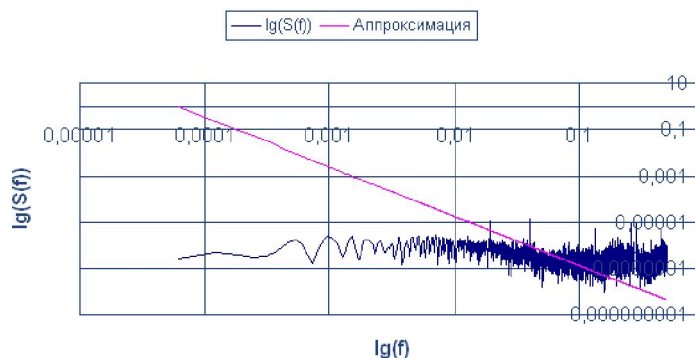


Рисунок 14.- Спектральная плотность временного ряда UDP_j64

Спектральная плотность задержки пакетов объемом 512 байт

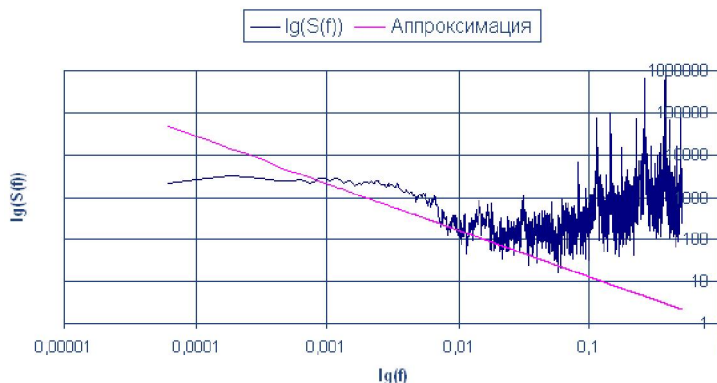


Рисунок 15.- Спектральная плотность временного ряда UDP_d512

Спектральная плотность временного ряда UDP_j512

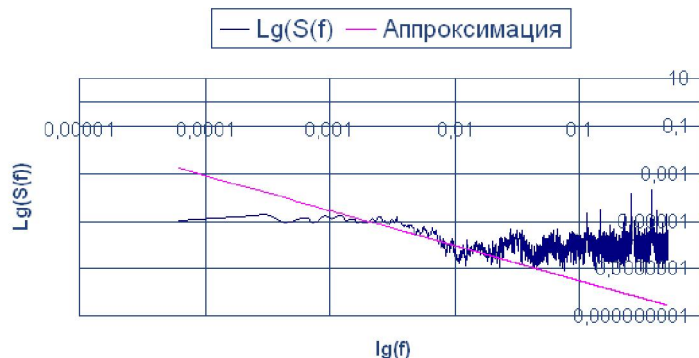


Рисунок 16.- Спектральная плотность временного ряда UDP_j512

5. Исследование показателя Херста

В этом разделе оценивается показатель Херста H , который является мерой устойчивости статистического явления или мерой длительности долгосрочной зависимости процесса [4]. Значение $H=0,5$ указывает на

отсутствие долгосрочной зависимости. Корреляция между событиями отсутствует. Ряд является случайным, а не фрактальным. Чем ближе значение H к 1, тем выше степень устойчивости долгосрочной зависимости. При $0 \leq H < 0,5$ временной ряд является трендонеустойчивым (антиперсистентным). Он более изменчив, чем случайный ряд, поскольку состоит из частых реверсов спад-подъем. При

$0,5 < H \leq 1$ ряд трендоустойчив. Тенденция его изменения может быть спрогнозирована. Показатель Херста связан с другими фрактальными параметрами формулами, показанными в таблице 1.

Таблица 1. Фрактальные параметры

Параметр	Расчетная формула
Фрактальная размерность D	$D = 2 - H$
Корреляционный параметр β	$\beta = 2(1 - H)$
Спектральный показатель b	$b = 2H + 1$
Фрактальный показатель α	$\alpha = 3 - 2H$

Для определения значения H в работе используется алгоритм R/S-анализа временных рядов, реализованный в среде VBA. Алгоритм состоит в следующем. В исходной реализации Z последовательно выделяем отрезки $Z_\tau = z_1, z_2, \dots, z_\tau$, $\tau = 1000, 2000, \dots, 12000$, для каждого из которых вычисляем текущее среднее

$$\bar{z}_\tau = \frac{\sum_{i=1}^{\tau} z_i}{\tau}$$

Далее для каждого фиксированного

Z_τ вычисляем накопленное отклонение для его отрезков длины t : $X_{\tau,t} = \sum_{i=1}^t (z_i - \bar{z}_\tau)$, $t = 1, 2, \dots, \tau$,

разность (размах) между максимальным и минимальным накопленными отклонениями $R = R(\tau) = \max_{1 \leq t \leq \tau} (X_{\tau,t}) - \min_{1 \leq t \leq \tau} (X_{\tau,t})$ и стандартное

отклонение $S = S(\tau)$. Показатель Херста

вычисляем по формуле $\frac{R}{S} = \left(\frac{\tau}{2}\right)^H$.

Логарифмируя обе части этого равенства, получаем последовательность координат точек (x_τ, y_τ) , где $x_\tau = 1000, 2000, \dots, 12000$,

$$y_\tau = H(\tau) = \frac{\lg(R(\tau)/S(\tau))}{\lg(\tau/2)}$$

Построив график зависимости $H(\tau)$ и линию тренда, определим аппроксимированное значение показателя Херста H .

На рисунке 17 показан результат R/S-анализа для процесса задержки пакетов объемом 64 байт. Показатель Херста равен 0,958.

R/S-анализ

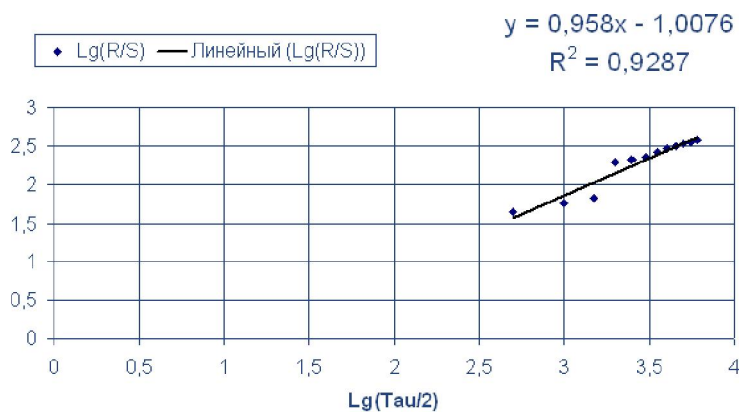


Рисунок 17.- Результат R/S-анализа временного ряда UDP_d64

Используя полученное значение $H = 0,958$, по формулам, указанным в таблице 1, определены фрактальные параметры процесса задержки пакетов объемом 64 байт: $D = 1,044$, $\beta = 0,088$, $b = 2,912$, $\alpha = 1,088$. Поскольку выполняется условие $0,5 < H < 1$, изучаемый ряд является трендоустойчивым (персистентным) и обладает долговременной памятью. Спектральный показатель b близок к 3, такой сигнал называют “черным шумом”. Линия аппроксимации на рисунке 9 построена по формуле

$$r(k) = k^{-\beta} \cdot L_1(k) + C, L_1 = 0,81, C = -0,52.$$

Линия

аппроксимации на рисунке 13 построена для $f \rightarrow 0$ по формуле $S(f) = f^{-b} \cdot 10^3 + 0,1$.

На рисунке 18 показан результат R/S-анализа для джиттера при передаче пакетов объемом 64 байт. Показатель Херста равен 0,5662, фрактальные параметры процесса: $D = 1,4338$, $\alpha = 1,8676$, $\beta = 0,8676$, $b = 2,1324$. Поскольку величина H близка к 0,5, то ряд UDP_j64 близок к случайному. Линия аппроксимации на рисунке 10 построена по формуле

$$r(k) = k^{-\beta} \cdot L_1(k) + C, L_1 = 0,195, C = -0,003.$$

Линия аппроксимации на рисунке 14 построена для $f \rightarrow 0$ по формуле $S(f) = f^{-b} \cdot 10^{-9}$.

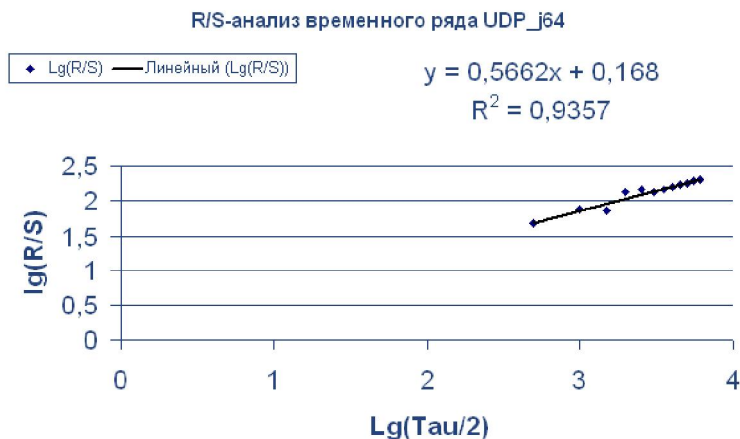


Рисунок 18.- Результат R/S-анализа временного ряда UDP_j64

На рисунке 19 показан результат R/S-анализа для процесса задержки пакетов объемом 512 байт, полученный в среде VBA. Показатель Херста равен 0,0506. Фрактальные параметры процесса задержки пакетов объемом 512 байт: $D = 1,9494$, $\beta = 1,8988$, $b = 1,1012$, $\alpha = 2,8988$. Спектральный показатель b близок к 1, т.е. исследуемый временной ряд является фликкершумом. Линия аппроксимации на рисунке 11 построена по формуле $r(k) = k^{-\beta} \cdot L_1(k) + C$, $L_1 = -0,4$, $C = 0,001$. Линия аппроксимации для $f \rightarrow 0$ на рисунке 15 построена по формуле $S(f) = f^{-b}$.

На рисунке 20 показан результат R/S-анализа для джиттера при передаче пакетов объемом 512 байт. Показатель Херста равен 0,2427. Ряд является антиперсистентным.

Фрактальные параметры процесса: $D = 1,7573$, $\beta = 1,5146$, $b = 1,4854$, $\alpha = 2,5146$. Спектральный показатель b близок к 1,5, временной ряд является “розовым шумом”. Линия аппроксимации на рисунке 12 построена по формуле $r(k) = k^{-\beta} \cdot L_1(k) + C$, $L_1 = -0,133$, $C = 0,012$. Линия аппроксимации для $f \rightarrow 0$ на рисунке 16 построена по формуле $S(f) = f^{-b}$.

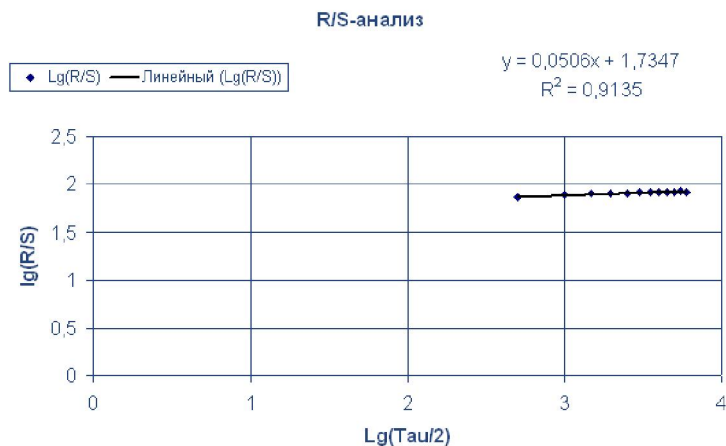


Рисунок 19.- Результат R/S-анализа временного ряда UDP_d512

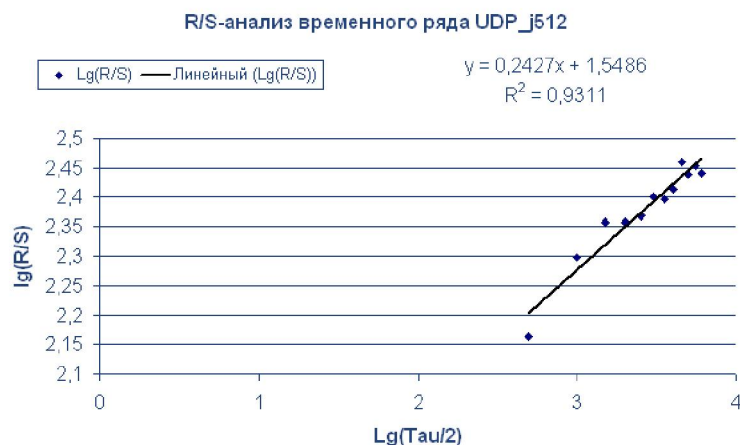


Рисунок 20.- Результат R/S-анализа временного ряда UDP_j512

Выводы

На современном этапе развития информационных технологий, который невозможен без эффективного использования компьютерных сетей, большое значение имеет моделирование трафика. Значительное количество работ, посвященное анализу сетей, основано на использовании теории очередей. Однако современный трафик имеет особенности, затрудняющие ее применение. Для процессов передачи данных пакетным трафиком, характерно обнаруженное на практике свойство масштабной инвариантности статистических характеристик, связанное с особым классом физических процессов – фрактальными процессами.

В связи с этой особенностью сетевых процессов особую актуальность приобретает разработка конструктивных методов исследования фрактальности и учет ее при передаче пакетного трафика.

В данной работе для реализаций [3] процесса задержки и джиттера UDP-пакетов объемом 64 и 512 байт соответственно, выполнен анализ плотности распределения, автокорреляционных функций и энергетических спектров. Найдены значения показателя Херста H . Получены следующие результаты:

1. Изучаемые временные ряды не подчиняются нормальному распределению.
2. Изучаемые временные ряды не обладают экспоненциально спадающей АКФ, свойственной случайным рядам.
3. Для процесса задержки пакетов объемом 64 байт выполняется условие $0,5 < H < 1$. Этот ряд является трендоустойчивым (персистентным) и обладает долговременной памятью. Спектральный показатель b близок к 3, ряд является “черным шумом”.
4. Для процесса задержки пакетов объемом 512 байт выполняется условие $0 < H < 0,5$.

Спектральный показатель b близок к 1, временной ряд является фликкер-шумом.

5. Для джиттера при передаче пакетов объемом 64 байт показатель Херста близок к значению 0,5 - ряд близок к случайному.

6. Для джиттера при передаче пакетов объемом 512 байт выполняется условие $0 < H < 0,5$. Спектральный показатель b близок к 1,5, временной ряд является “розовым шумом”.

Перспективным направлением дальнейших исследований может быть анализ сетевых процессов методами нелинейной динамики. Для повышения эффективности работы современных телекоммуникационных систем требуется создание математических моделей, наиболее полно отражающих фрактальные свойства сетевых процессов.

Литература

1. Park K. Self-Similar Network Traffic: An Overview. [Электронный ресурс], 2003. – Режим доступа: <http://pi.314159.ru/park1.pdf>
2. Willinger W., Taqqu M.S., Errimilli A. A bibliographical guide to self-similar traffic and performance modeling for modern high-speed networks. [Электронный ресурс], 2001. – Режим доступа: <http://linkage.rockefeller.edu/wli/reading/taqqu96.pdf>
3. Network tools and traffic traces. [Электронный ресурс], 2007. – Режим доступа: <http://www.grid.unina.it/Traffic/Traces/traaces.php>
4. Петров В.В. Структура телетрафика и алгоритм обеспечения качества обслуживания при влиянии эффекта самоподобия. Автореферат диссертации. Москва. – 2004. – 20 с.