

УДК 004.7

Статистический анализ трафика сети с беспроводным доступом

Бельков Д.В., Едемская Е.Н.

Донецкий национальный технический университет
belkov@telenet.dn.ua

Abstract

Belkov D.V., Edemskaya E.N. The statistical analysis of the traffic traces with the wireless access to the network. 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 with GPRS-access to the network at sender side and UMTS-access to the network at receiver side.

Введение

В связи с тенденциями объединения различных телекоммуникационных приложений на базе универсальной сетевой инфраструктуры актуальной проблемой является разработка методов анализа и синтеза информационно-управляющих систем. Данная работа посвящена одной из важных задач развития этого направления – исследованию процессов в компьютерных сетях с беспроводным доступом.

Использование в сетях интегрированных информационных приложений и пакетной коммутации сопровождается сложными явлениями, изучение которых может быть выполнено с помощью вероятностных подходов.

Характерным для процессов пакетной передачи является обнаруженное на практике свойство фрактальности. В процессе передачи возникают большие всплески при относительно низком среднем уровне трафика. Данное явление, которое значительно влияет на потери и задержки пакетов, не учитывается в теории телеграфика. Расчет параметров системы распределения информации по классическим формулам дает некорректные, неоправданно оптимистические результаты. Алгоритмы обработки трафика, созданные для работы с простейшим потоком неэффективны для фрактальных потоков [1,2].

Появление новых свойств сетевого трафика, необходимость обеспечения высокого качества обслуживания приложений, делают актуальным статистический анализ трафика сетей.

Целью настоящей работы является исследование структуры трафика сети с беспроводным доступом, 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.

1. Реализации сетевого трафика

Для изучения выбраны шесть реализаций сетевого трафика [3], полученные в 2007 году в университете города Наполи (Италия). Согласно лицензии данные свободно доступны для анализа. Условные обозначения, принятые в работе, показаны в таблице 1.

Таблица 1. Условные обозначения

Обозначение	Описание
TCP_d64	Ряд измерений задержки TCP-пакетов объемом 64 байт
TCP_d256	Ряд измерений задержки TCP-пакетов объемом 256 байт
TCP_d1024	Ряд измерений задержки TCP-пакетов объемом 1024 байт

Измерения проводились каждые 10 миллисекунд, получено свыше 3000 отсчетов. Отправитель имел GPRS-доступ, получатель - UMTS-доступ, скорость передачи 100 rps, операционная система Windows32, на каждой из сторон.

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 и складывается из времени обработки, времени распространения, времени

коммутации, времени ожидания в очереди. Изучаемые временные ряды показаны на рисунках 1-3. Нижний уровень задержки соответствует отсутствию очередей.

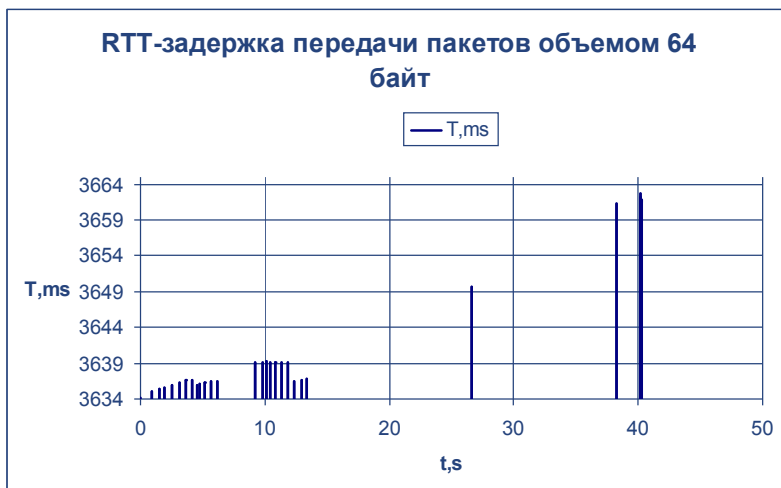


Рисунок 1 - Задержка при передаче пакетов объемом 64 байт



Рисунок 2 - Задержка при передаче пакетов объемом 256 байт

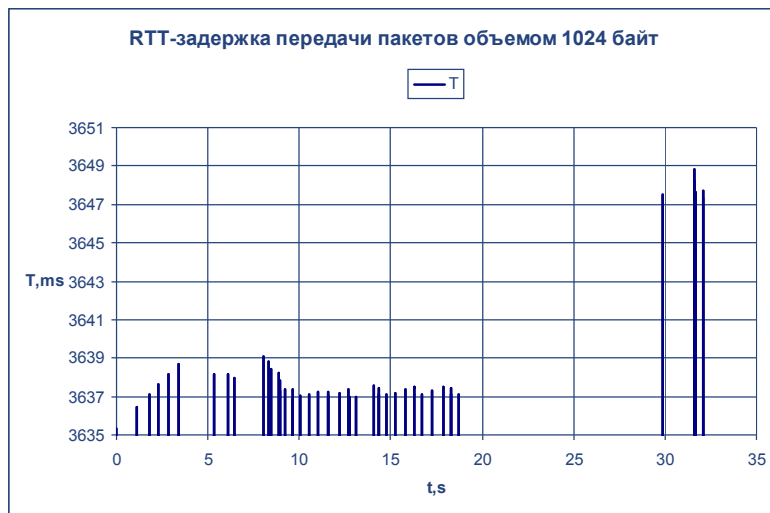


Рисунок 3 - Задержка при передаче пакетов объемом 1024 байт

2. Анализ плотности распределения

В этом разделе выполнен анализ плотности распределения. Оценка проводится на основании гистограмм относительных частот,

показанных на рисунках 4-6. Визуальный анализ позволяет сделать вывод, что исследуемые ряды не подчиняются нормальному распределению.

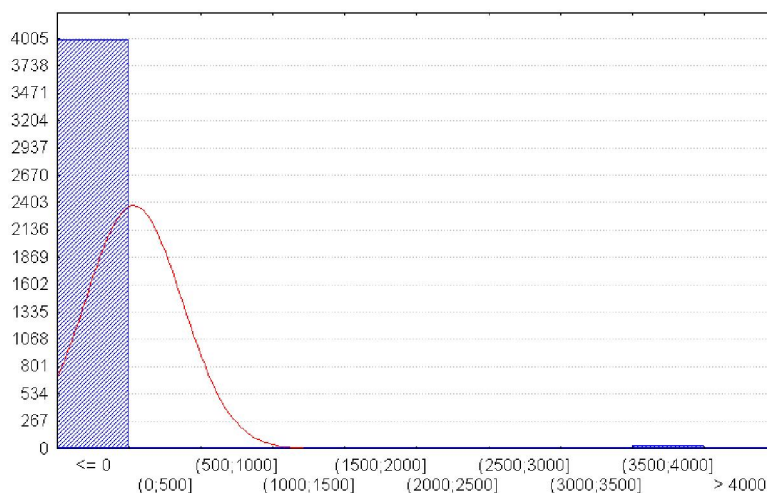


Рисунок 4 - Плотность распределения временного ряда TCP_d64

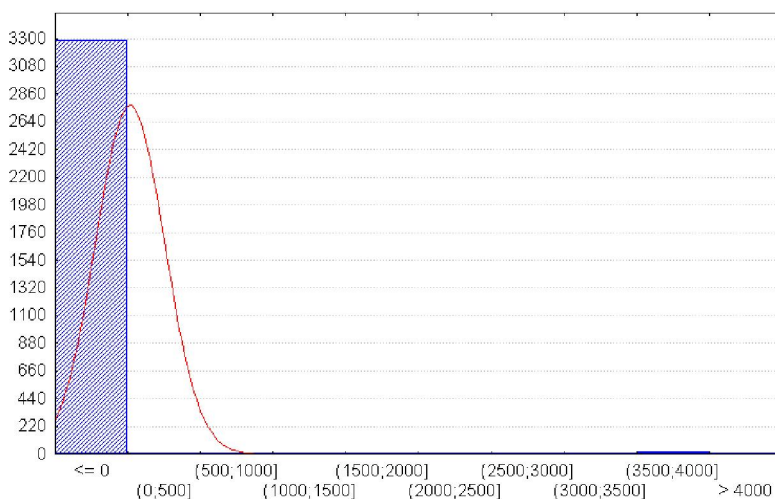


Рисунок 5 - Плотность распределения временного ряда TCP_d256

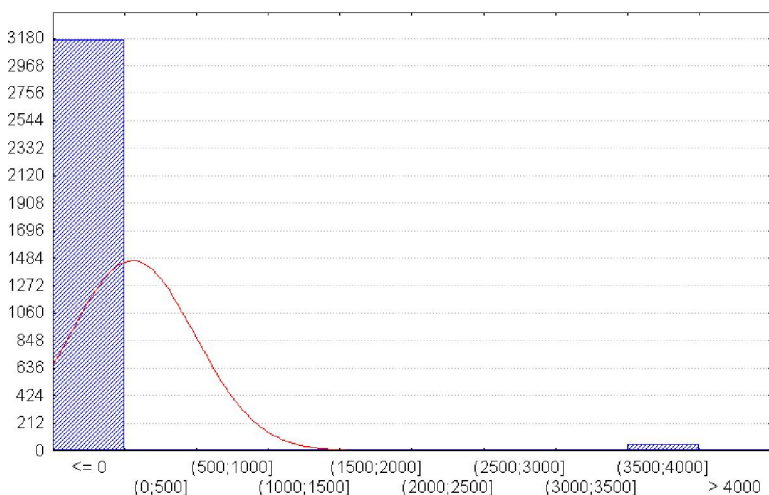


Рисунок 6 - Плотность распределения временного ряда TCP_d1024

3. Анализ автокорреляционных функций

В этом разделе необходимо определить обладают ли временные ряды медленно убывающей зависимостью или быстро убывающей зависимостью. По определению процесс X обладает медленно убывающей зависимостью (МУЗ) [long-range dependence], если для его АКФ выполняется условие

$r(k) \sim k^{-\beta} \cdot L_1(k), k \rightarrow \infty$, где $0 < \beta < 1$ и L_1 - медленно меняющаяся на бесконечности функция: $\lim_{t \rightarrow \infty} \frac{L_1(tx)}{L_1(t)} = 1$ для всех x . 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 $r(k) \sim e^{-k}, k \rightarrow \infty$.

Вычисления АКФ выполнены с помощью программы Fractan 4.4 по формуле

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

выборочное среднее ряда X , $\sigma^2(X)$ - выборочная дисперсия ряда X , $k=0,1,\dots$. Автокорреляционные функции исследуемых временных рядов приведены на рисунках 7-9. Их аппроксимация степенными функциями будет показана в разделе 5.

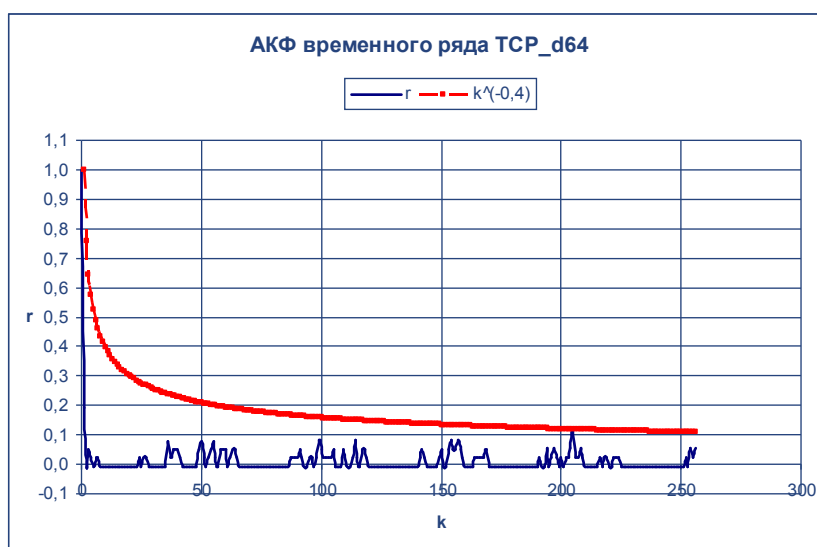


Рисунок 7 - АКФ временного ряда TCP_d64

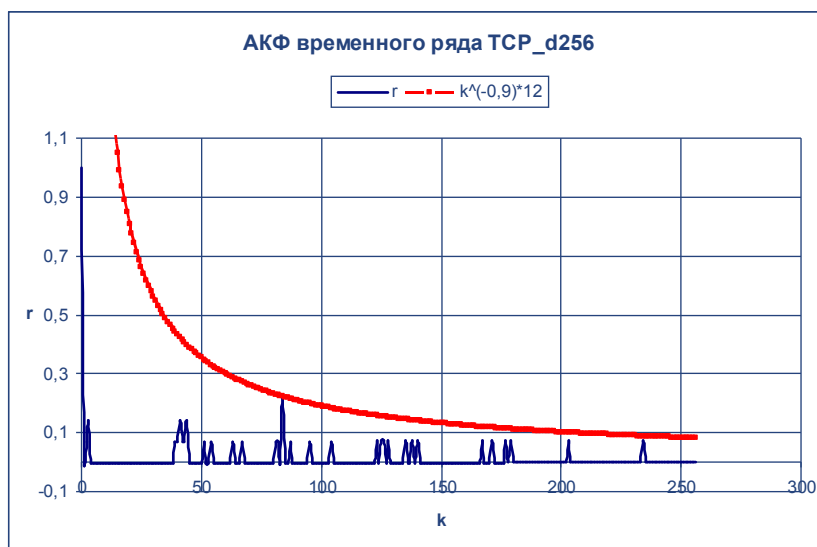


Рисунок 8 - АКФ временного ряда TCP_d256

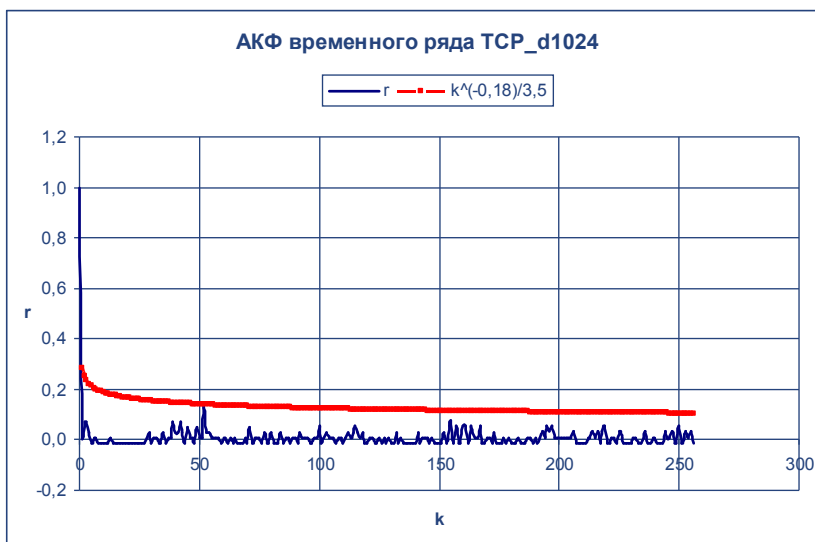


Рисунок 9 - АКФ временного ряда TCP_d1024

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 $S(f) \sim f^{-b} \cdot L_2(f)$, где $f \rightarrow 0$, $i = \sqrt{-1}$, $0 < b < 4$ и L_2 - медленно изменяющаяся в нуле функция.

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 10-12.

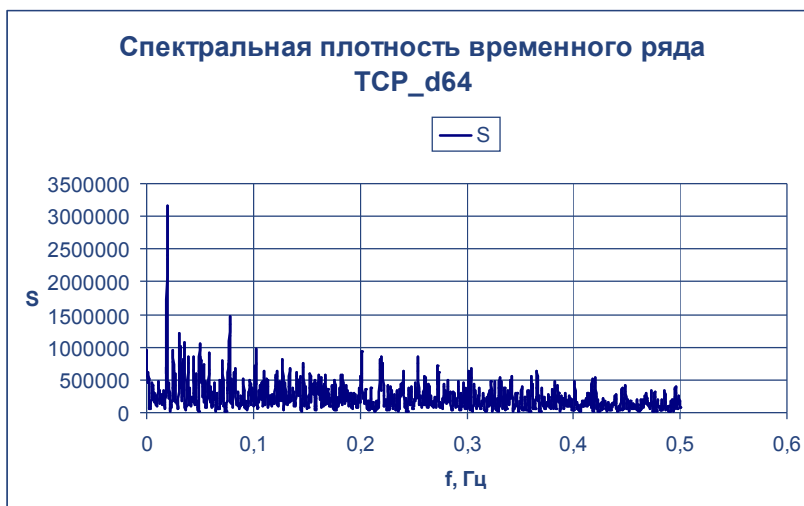


Рисунок 10 - Спектральная плотность временного ряда TCP_d64

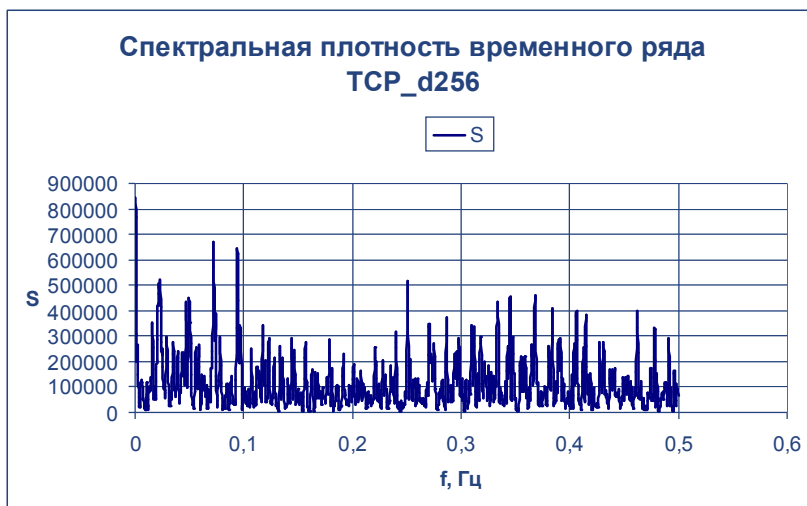


Рисунок 11 - Спектральная плотность временного ряда TCP_d256

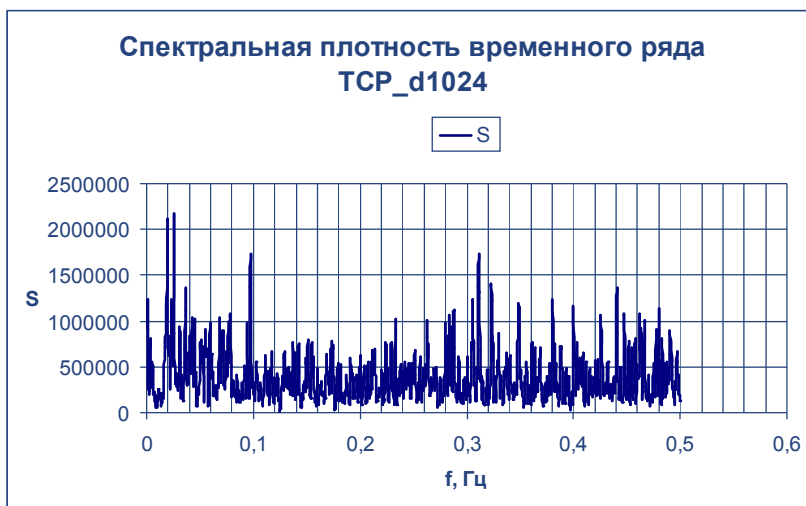


Рисунок 12 - Спектральная плотность временного ряда TCP_d1024

Визуальный анализ позволяет сделать вывод, что временные ряды RTT-задержки имеют на низких частотах самые мощные гармоники. Анализ спектральной плотности изучаемых рядов будет выполнен в разделе 5.

5. Исследование показателя Херста

В этом разделе оценивается показатель Херста H , который является мерой устойчивости статистического явления или мерой длительности долгосрочной зависимости процесса. 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 трендоустойчивым (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 2.

Таблица 2. Фрактальные параметры

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

Для определения значения H в работе используется алгоритм R/S-анализа временных рядов, реализованный в программе Fractan 4.4. 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 с помощью формулы $\frac{R}{S} = \left(\frac{\tau}{2}\right)^H$.

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 13 показан результат R/S-анализа для процесса задержки пакетов объемом 64 байт. Показатель Херста вычислен с погрешностью $\Delta = \pm 0,2846$ и лежит в диапазоне между 0,62 и 1,19. Для определения значения H по графику спектральной плотности предлагается следующий метод исследования спектральной плотности временного ряда. Как показано на рисунке 14, строим линию L , которая огибает максимальные значения плотности S в направлении слева направо, и выполняем ее аппроксимацию степенной функцией. Показатель степени, равный 0,6, дает значение спектрального показателя b . Искомое значение $H = 0,8$, вычисляется по формуле $H = (1 + b) / 2$.

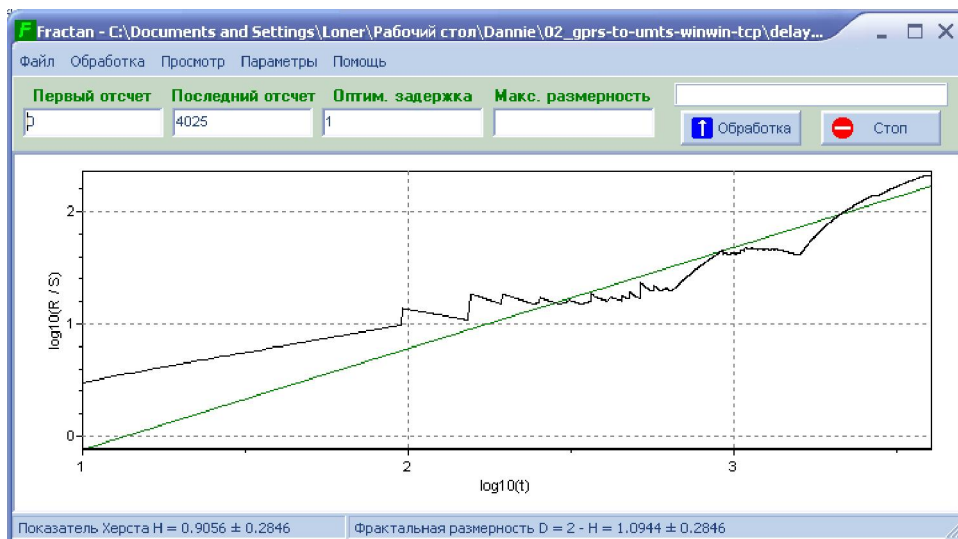


Рисунок 13 - Результат R/S-анализа временного ряда TCP_d64

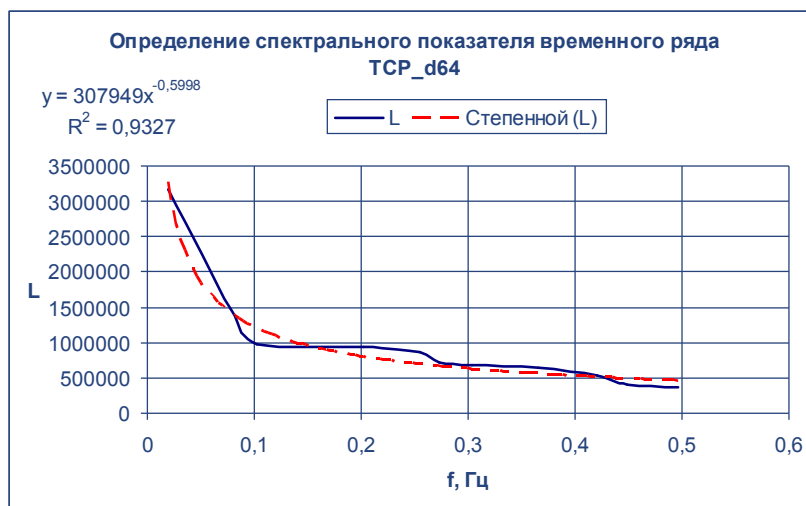


Рисунок 14 - Определение показателя b временного ряда TCP_d64

Поскольку значение H лежит в диапазоне $0,5 < H \leq 1$, то ряд TCP_d64 трендоустойчив. Для аппроксимации АКФ этого ряда вычислено значение $\beta = 0,4$ по формуле $\beta = 2(1 - H)$.

На рисунке 15 показан результат R/S-анализа для процесса задержки пакетов объемом 256 байт. Показатель Херста лежит в диапазоне

между 0,57 и 0,92. Определение спектрального показателя по графику спектральной плотности показано на рисунке 16. Значение $b=0,09$, $H=0,55$, $\beta = 0,9$. Ряд TCP_d256 близок к случайному.

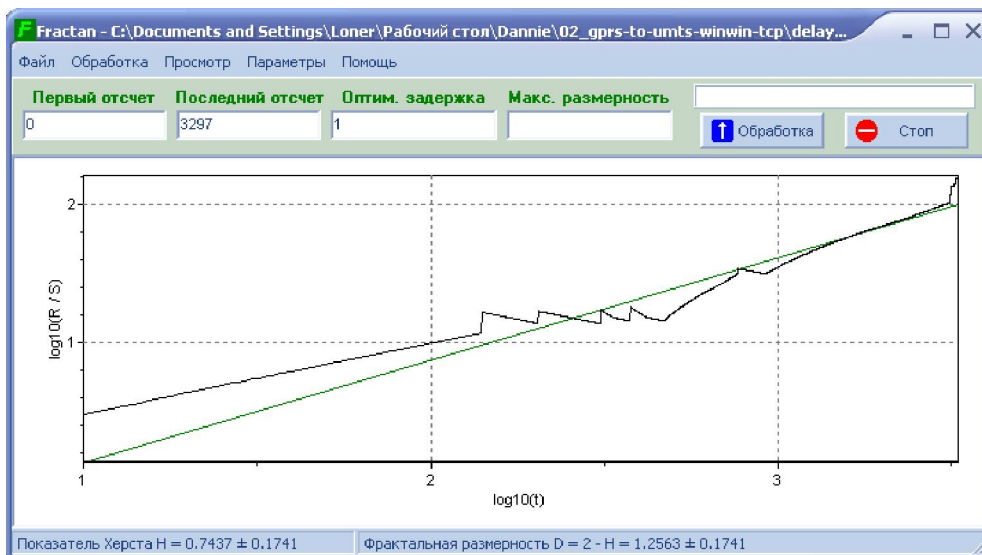


Рисунок 15 - Результат R/S-анализа временного ряда TCP_d256

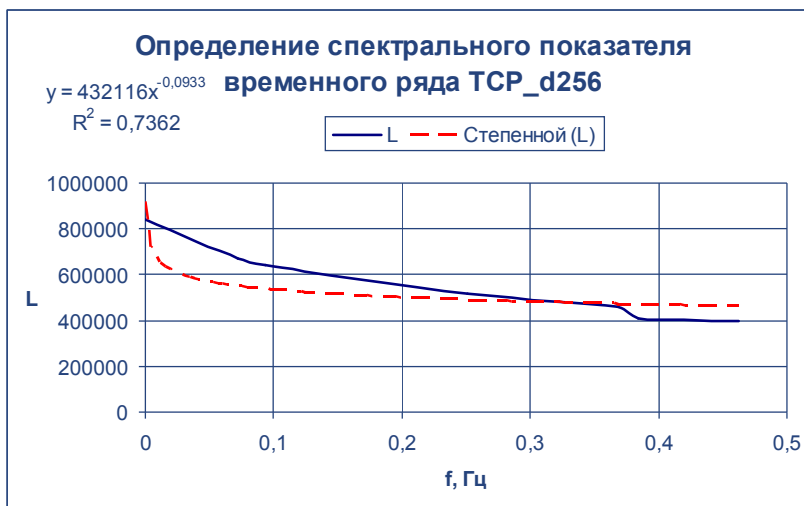


Рисунок 16 - Определение показателя b временного ряда TCP_d256

На рисунке 17 показан результат R/S-анализа для процесса задержки пакетов объемом 1024 байт. Показатель Херста лежит в диапазоне между 0,59 и 1,07. Определение спектрального показателя по графику спектральной плотности

показано на рисунке 18. Значение $b=0,18$, $H=0,59$. Аппроксимации АКФ этого ряда выполнена при $\beta = 0,82$.



Рисунок 17 - Результат R/S-анализа временного ряда TCP_d1024

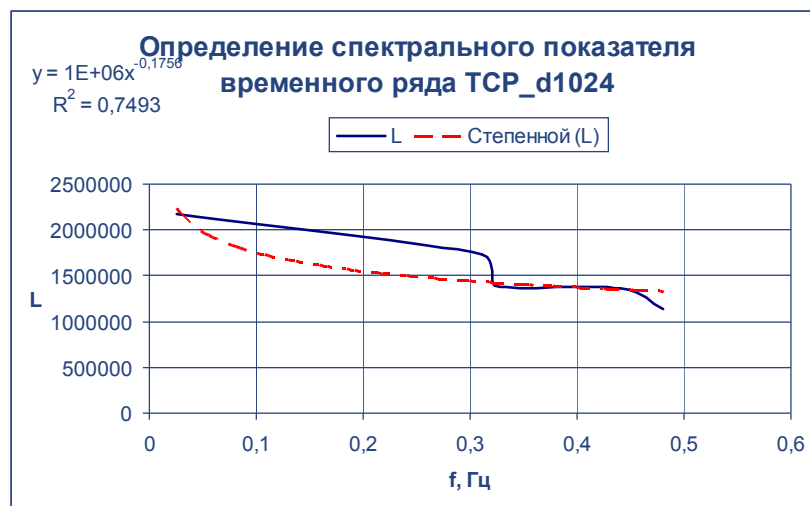


Рисунок 18 - Определение показателя ν временного ряда TCP_d1024

Выводы

В данной работе для сети с беспроводным доступом и реализаций [3] процесса задержки TCP-пакетов выполнен статистический анализ. Предложен метод определения спектрального показателя по графику спектральной функции.

Получены следующие результаты:

1. Изучаемые временные ряды не подчиняются нормальному распределению.
2. Изучаемые временные ряды имеют медленно убывающую зависимость АКФ, т.е. у процессов есть долговременная память.
3. Спектральная плотность временных рядов стремится к бесконечности, по мере того как частота стремится к нулю.

Перспективным направлением дальнейших исследований является изучение структуры трафика сетей с беспроводным доступом методами нелинейной динамики.

Литература

1. Городецкий А.Я., Заборовский В.С. Информатика. Фрактальные процессы в компьютерных сетях. – СПб.: СПбГУ, 2000. – 102 с.
2. Park K. Self-Similar Network Traffic: An Overview. [Электронный ресурс], 2003. – Режим доступа: <http://pi.314159.ru/park1.pdf>
3. Network tools and traffic traces. [Электронный ресурс], 2007. – Режим доступа: <http://www.grid.unina.it/Traffic/Traces/traces.php>