

УДК 004.7

Исследование сетевого трафика с помощью функции Херста

Е.Н. Едемская, Д.В. Бельков
Донецкий национальный технический университет
belkov@telenet.dn.ua, botba@list.ru

Едемская Е.Н., Бельков Д.В. Исследование сетевого трафика с помощью функции Херста. Исследование сетевого трафика является важным, поскольку фрактальный трафик в современных компьютерных сетях ухудшает качество обслуживания. Целью настоящей работы является исследование реального трафика для выявления его характерных особенностей. Для исследования фрактальных процессов используется показатель Херста H , который является мерой длительности долгосрочной зависимости процесса. В статье при анализе трафика предлагается использовать функцию Херста. Ее график строится следующим образом. По оси абсцисс откладываем значения $N=2,3,\dots,N_m$, где N_m – максимальная длина временного ряда, N – текущее количество элементов временного ряда. Для каждого значения N определяем величину показателя Херста $H(N)$ и откладываем ее по оси ординат. Исследования выполнены в среде Matlab. Для изучения выбраны четыре реализации сетевого трафика, полученные в университете города Наполи (Италия). Согласно лицензии данные свободно доступны для анализа. Изучаемые временные ряды представляют собой измерения задержки (с.) UDP и TCP-пакетов. В первом случае (ряд UDP_d64) UDP-пакеты имеют объем 64 байт, во втором (ряд UDP_d512) – 512 байт. Третий ряд (TCP_d64) – передача TCP-пакетов объема 64 байт, четвертый ряд (TCP_d512) – передача TCP-пакетов объема 512 байт. Получены следующие результаты. Трафик UDP_d64 состоит из двух персистентных фаз и фазового перехода между ними в момент, когда функция Херста достигает значения $H=0,5$. Трафик UDP_d512 имеет фазовый переход из антиперсистентного состояния ($H=0,25$) к состоянию однородного трафика ($H=0,5$) и постепенный возврат в антиперсистентное состояние. TCP-трафик имеет постепенный переход из персистентного к однородному состоянию, однородный режим, фазовый переход в персистентное состояние и постепенный возврат из персистентного в однородное состояние.

Введение

Для повышения эффективности функционирования сетей важной задачей является анализ структуры трафика. В теории телетрафика используются математические методы и модели, ориентированные на независимые стационарные потоки. Однако, многочисленные исследования [1-6] трафика сетей свидетельствуют о его фрактальности. 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 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.

Фрактальные процессы обладают самоподобием (масштабной инвариантностью)

корреляционной функции, спектральной мощности, дисперсии. Масштабная инвариантность сетевого трафика позволяет разработать алгоритмы прогнозирования, которые смогут с помощью анализа трафика на относительно небольшом отрезке времени предсказать его поведение на длительных интервалах. Используя такие прогнозы, можно создавать эффективные методы управления пропускной способностью, что позволит сократить задержки передачи данных и потери пакетов.

Среди иностранных ученых, которые активно занимаются вопросом самоподобия трафика, можно выделить W. Leland, M. Taqqu, W. Wilinger, D. Wilson, V. Paxson, K. Park, Ertamilli и др. Среди российских и украинских исследователей нужно отметить работы О.И. Шелухина, Б.С. Цыбакова, В.В. Петрова, А.В. Осина, А.Г. Ложковского, В.С. Заборовского, О.Я. Городецкого. Однако исследование самоподобного трафика с целью прогнозирования его поведения и обеспечения QoS не потеряло своей актуальности.

Для анализа фрактальных процессов используется 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. Временной 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 антиперсистентным. Он более изменчив, чем случайный ряд, поскольку состоит из частых реверсов спад-подъем. При $0,5 < H \leq 1$ ряд называется персистентным. Тенденция его изменения может быть спрогнозирована [7].

Во многих работах [7-9] показатель Херста является индикатором: если исследуемый трафик имеет $H > 0,5$, то принимается решение о возможности прогнозирования трафика, иначе – такая возможность отвергается. При разработке метода прогнозирования найденное значение H не используется.

В данной статье для анализа трафика предлагается использовать функцию, которую назовем функцией Херста. Ее график строится следующим образом. По оси абсцисс откладываем значения $N=2,3,\dots,N_m$, где N_m – максимальная длина временного ряда, N – текущее количество элементов временного ряда. Для каждого значения N определяем величину показателя Херста $H(N)$ и откладываем ее по оси ординат.

Целью работы является исследование структуры сетевого трафика, направленное на выявление его характерных особенностей. Задача работы – анализ трафика задержки UDP и TCP-пакетов с помощью функции Херста.

Исследования выполнены в среде Matlab. Для изучения выбраны четыре реализации сетевого трафика [10], полученные 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
UDP_d64	Ряд измерений задержки (с.) UDP-пакетов объемом 64 байт
UDP_d512	Ряд измерений задержки (с.) UDP-пакетов объемом 512 байт
TCP_d64	Ряд измерений задержки (с.) TCP-пакетов объемом 64 байт
TCP_d512	Ряд измерений задержки (с.) TCP-пакетов объемом 512 байт

Исследование UDP-трафика

В первой серии экспериментов исследуются ряды UDP_d64 и UDP_d512. Измерения проводились каждые 10 миллисекунд, получено 12 000 отсчетов. Отправитель имел ADSL-доступ (640 Kbps), на стороне получателя – 100 Mbps Ethernet, операционная система Linux на каждой из сторон, скорость передачи 100 rps.

Временной ряд UDP_d64 и его функция Херста показаны на рисунках 1, 2.

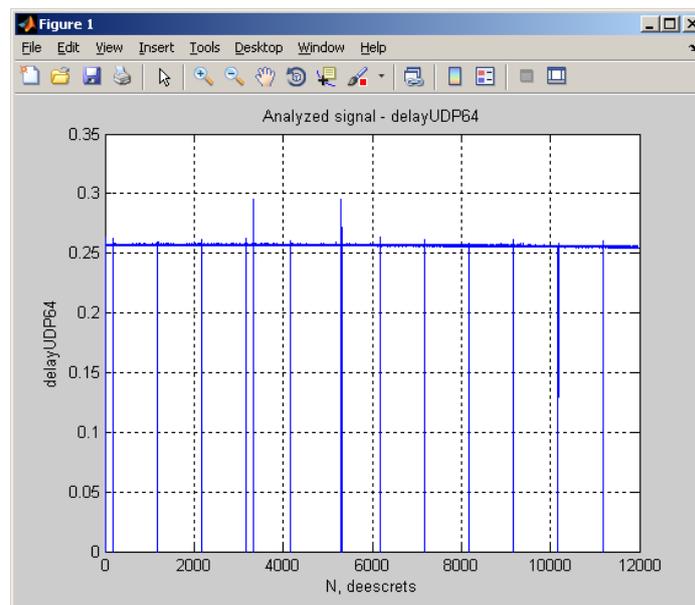


Рисунок 1 – Временной ряд UDP_d64

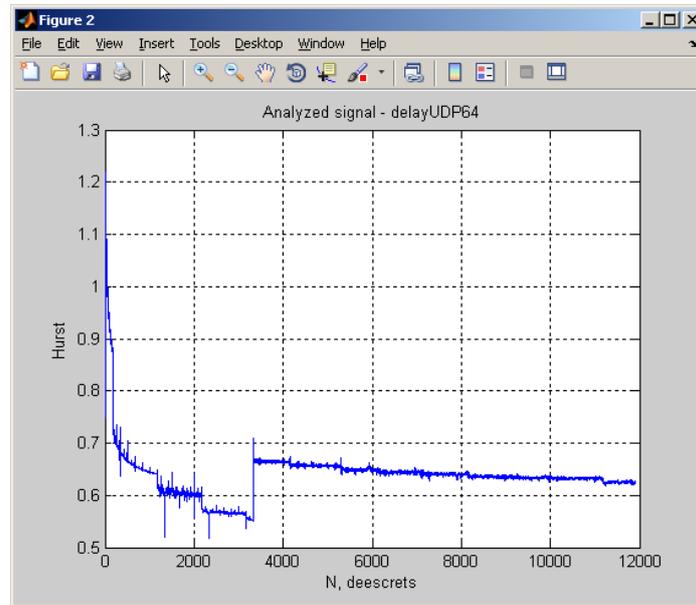


Рисунок 2 – Функция Херста ряда UDP_d64

На рисунке 1 заметен порядок, состоящий в том, что длительные (примерно по 1000 тактов) активные интервалы с положительной задержкой чередуются с кратковременными (1-2 такта) пассивными интервалами с нулевой задержкой. Этот порядок существует до значения $N=3300$, когда очередной активный интервал прерывается, происходит всплеск задержки, затем порядок восстанавливается. Еще один всплеск задержки при $N=5000$ не нарушает порядка. Функция Херста, изображенная на рисунке 2, имеет две ветви. Переход от одной ветви к другой происходит за один такт. Как показано на рисунке 3, на каждой ветви есть интервалы с примерно

постоянным значением функции, которые соответствуют активным интервалам задержки. Пассивным интервалам соответствует быстрое убывание функции Херста. Скачок функции совпадает по времени со всплеском задержки при $N=3300$ и $H = 0,55 \approx 0,5$. Для наглядности значения задержки на рисунке 3 увеличены в 4 раза.

Таким образом, трафик UDP_d64 состоит из двух персистентных фаз и фазового перехода между ними в момент, когда функция Херста достигает значения $H=0,55$.

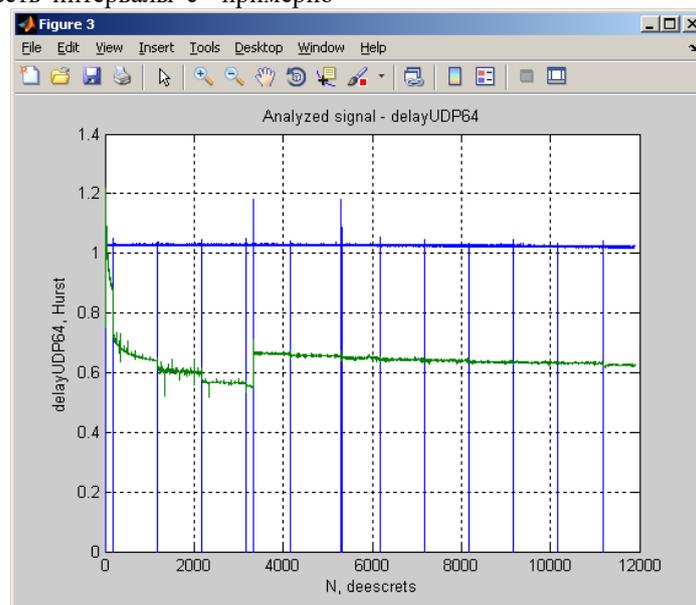


Рисунок 3 – Динамика функции Херста ряда UDP_d64

На рисунке 4 показан фрагмент временного ряда UDP_d512. Функция Херста показана на рисунке 5. Ее анализ позволяет сделать следующие выводы: при $N=150$ происходит фазовый переход из антиперсистентного

состояния ($H=0,25$) к состоянию однородного трафика ($H=0,5$). Затем происходит фазовый переход в антиперсистентное состояние при $H=0,4$ и постепенное уменьшение функции Херста.

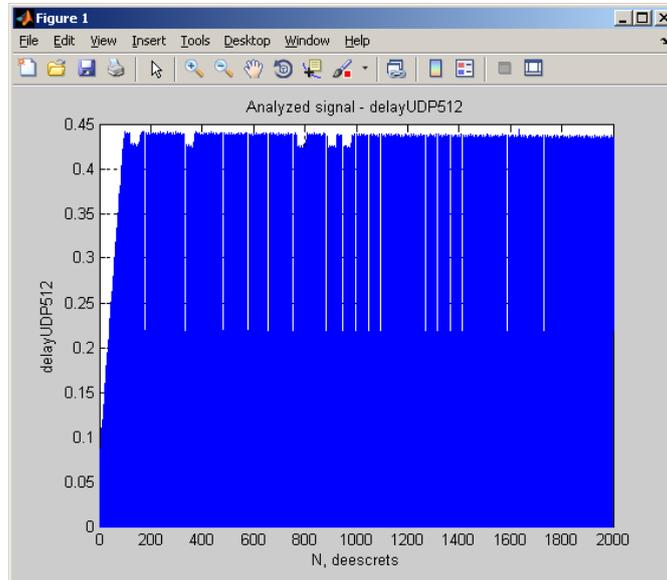


Рисунок 4 – Временной ряд UDP_d512

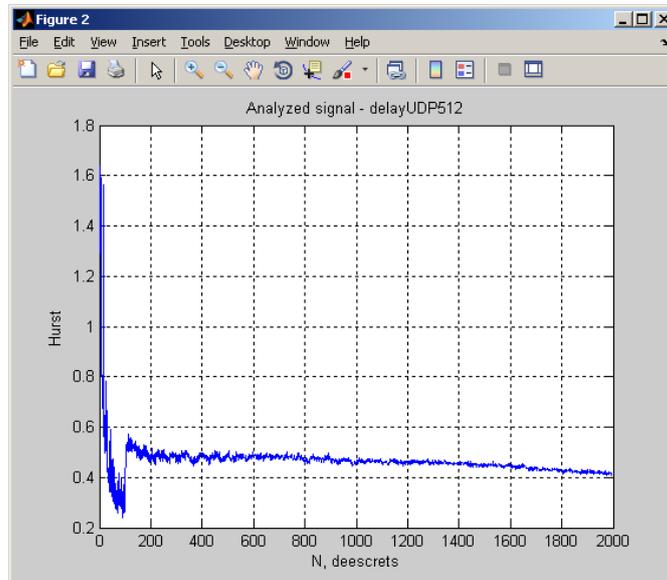


Рисунок 5 – Функция Херста ряда UDP_d512

Исследование TCP-трафика

Во второй серии экспериментов исследуются ряды TCP_d64 и TCP_d512. Измерения проводились каждые 10 миллисекунд, получено свыше 3000 отсчетов. Отправитель имел ADSL-доступ (640 Kbps), на стороне получателя - UMTS-доступ, скорость передачи 100 prps, операционная система Windows32, на каждой из сторон. Изучаемые временные ряды и функции Херста показаны на рисунках 6-9. Для каждого

ряда функция Херста принимает значение $H=0,5$ при $N=100$. Однородное состояние, соответствующее пассивному интервалу временного ряда, длится до значения $N=250$, при котором происходит фазовый переход в персистентное состояние. Затем происходит постепенный возврат в состояние однородного трафика, которое устанавливается при $N=1000$. Фазовый переход в персистентное состояние соответствует большому всплеску задержки.

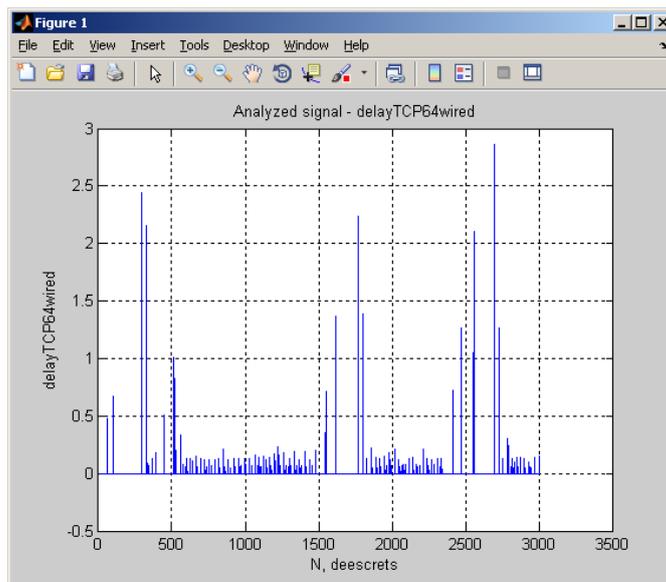


Рисунок 6 – Временной ряд TCP_d64

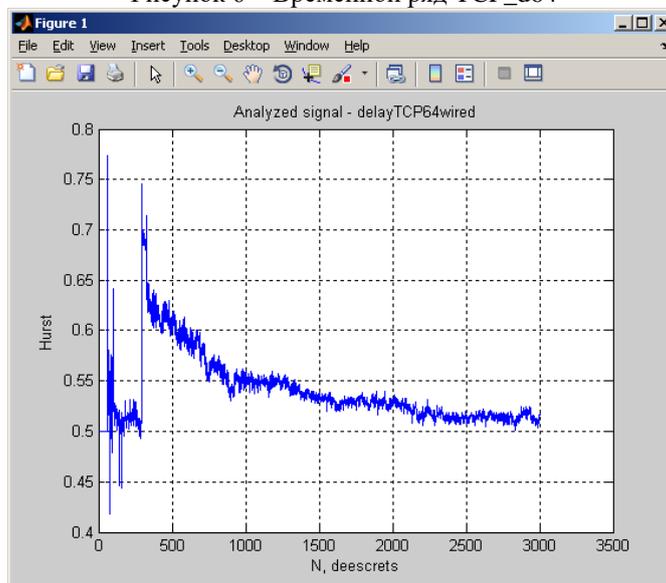


Рисунок 7 – Функция Херста ряда TCP_d64

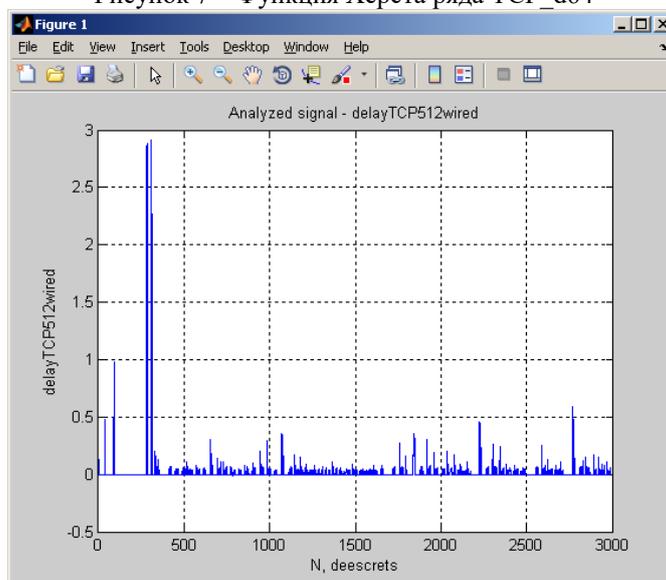


Рисунок 8.- Временной ряд TCP_d512

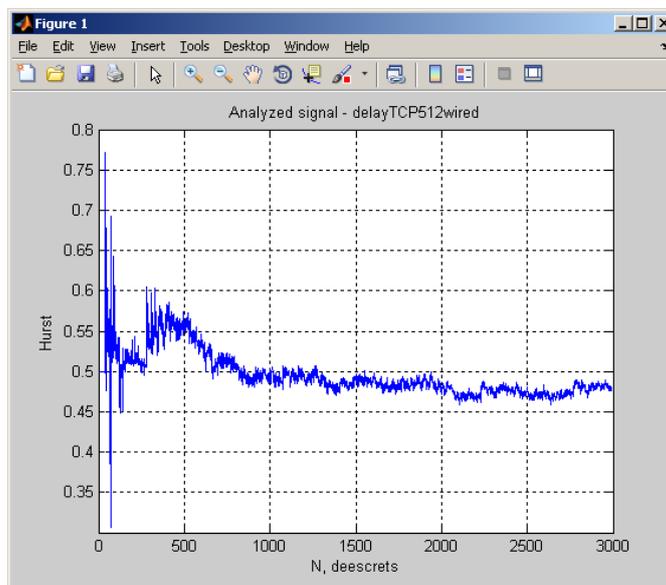


Рисунок 9 – Функция Херста ряда TCP_d512

Выводы

В данной работе для исследования сетевого трафика предложено использовать функцию Херста. Выполнен анализ реализаций [10] трафика задержки UDP и TCP-пакетов с помощью функции Херста. Получены следующие результаты:

Трафик UDP_d64 состоит из двух персистентных фаз и фазового перехода между ними в момент, когда функция Херста достигает значения $H=0,5$. Трафик UDP_d512 имеет фазовый переход из антиперсистентного состояния ($H=0,25$) к состоянию однородного трафика ($H=0,5$) и постепенный возврат в антиперсистентное состояние. TCP-трафик имеет постепенный переход из персистентного состояния в однородное состояние, фазовый переход в персистентное состояние и постепенный возврат в однородное состояние.

Перспективным направлением дальнейших исследований является использование функции Херста для анализа мультимедийного трафика.

Список литературы

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. Hae-Duck Joshua Jeong. Modeling of self-similar teletraffic for simulation. University of Canterbury, 2002. – 297 p.
4. Ложковський А.Г. Аналіз і синтез систем розподілу інформації в умовах мультисервісного трафіка. Автореферат дисертації. Одеса. - 2010. – 38 с.
5. Бельков Д.В., Едемская Е.Н., Незамова Л.В. Статистический анализ сетевого трафика. 36. Наук. праць ДонНТУ. Серія “Інформатика, кібернетика, обчислювальна техніка”. Вип. 13 (185): - Донецьк: ДонНТУ.- 2011.- С. 66 -75.
6. Бельков Д.В., Едемская Е.Н. Статистический анализ трафика сети с беспроводным доступом. 36. Наук. праць ДонНТУ. Серія “Інформатика, кібернетика, обчислювальна техніка”. Вип. 14 (188): - Донецьк: ДонНТУ.- 2011.- С. 113-122.
7. Петров В.В. Структура телетрафика и алгоритм обеспечения качества обслуживания при влиянии эффекта самоподобия. Автореферат диссертации. Москва. – 2004. – 20 с.
8. Урьев Г.А. Исследование фрактальных свойств потоков трафика реального времени и оценка их влияния на характеристики обслуживания телекоммуникационных сетей. Автореферат диссертации. Москва. – 2007. – 21 с.
9. Городецкий А.Я., Заборовский В.С. Информатика. Фрактальные процессы в компьютерных сетях. – СПб.: СПбГТУ, 2000. – 102 с.
10. Network tools and traffic traces. [Электронный ресурс], 2007. – Режим доступа: <http://www.grid.unina.it/Traffic/Traces/ttraces.php>

Є.М. ЄДЕМСЬКА, Д.В. БЕЛЬКОВ

Донецький національний технічний університет

ДОСЛІДЖЕННЯ МЕРЕЖЕВОГО ТРАФІКУ ЗА ДОПОМОГОЮ ФУНКЦІЇ ХЕРСТА

Дослідження мережевого трафіку є важливим, оскільки фрактальний трафік в сучасних комп'ютерних мережах погіршує якість обслуговування. Метою даної роботи є дослідження реального трафіку для виявлення його характерних особливостей. Для дослідження фрактальних процесів використовується показник Херста H , який є мірою тривалості довгострокової залежності процесу. У статті при аналізі трафіку пропонується використовувати функцію Херста. Її графік будується таким чином. По осі абсцис відкладаємо значення $N=2,3,\dots,N_m$, де N_m – максимальна довжина тимчасового ряду, N – поточна кількість елементів тимчасового ряду. Для кожного значення N визначаємо величину показника Херста $H(N)$ і відкладаємо її по осі ординат. Дослідження виконані в середовищі Matlab. Для вивчення вибрані чотири реалізації мережевого трафіку, одержані в університеті міста Наполи (Італія). Згідно ліцензії дані вільно доступні для аналізу. Тимчасові ряди, що вивчаються, є вимірюваннями затримки (с.) UDP і TCP-пакетів. У першому випадку (ряд UDP_d64) UDP-пакети мають об'єм 64 байт, у другому (ряд UDP_d512) – 512 байт. Третій ряд (TCP_d64) – передача TCP-пакетів об'єму 64 байт, четвертий ряд (TCP_d512) – передача TCP-пакетів об'єму 512 байт. Одержані наступні результати. Трафік UDP_d64 складається з двох персистентних станів і фазового переходу між ними в момент, коли функція Херста досягає значення $H=0,5$. Трафік UDP_d512 має фазовий перехід від антиперсистентного стану ($H=0,25$) до стану однорідного трафіку ($H=0,5$) і поступове повернення до антиперсистентного стану. TCP-трафік має поступовий перехід від персистентного до однорідного стану, однорідний режим, фазовий перехід до персистентного стану і поступове повернення від персистентного до однорідного стану.

Ключові слова: UDP- трафік, TCP- трафік, затримка, показник Херста, функція Херста, фазовий перехід.

E.N. EDEMSKAYA, D.V. BELKOV

Donetsk National Technical University

NETWORK TRAFFIC RESEARCH BY THE HURST FUNCTION

Recent studies of real traffic data in modern computer networks have shown that traffic exhibits self-similar (or fractal) properties over a wide range of time scales. The properties of self-similar traffic are very different from the traditional models of traffic based on Poisson, Markov-modulated Poisson, and related processes. The use of traditional models in networks characterized by self-similar processes can lead to incorrect

conclusions about the performance of analyzed networks. These include serious over-estimations of the performance of computer networks, insufficient allocation of communication and data processing resources, and difficulties ensuring the quality of service expected by network users. The self-similar network traffic can have a detrimental impact on network performance, including amplified queuing delay, retransmission rate and packet loss rate. Modern network traffic consists of more bursts than Poisson models predict over many time scales. This difference has implications for congestion control mechanisms and performance. Research of network traffic is important, as a fractal traffic in modern computer networks worsens quality of service. Research of the real traffic for the exposure of its characteristic features is the purpose of the work. For research of fractal processes the Hurst index H , which is the measure of duration of long-term dependence of process is used. In the article for the traffic analysis it is suggested to use the Hurst function. Its graph is built as follows. On abscising axis we put aside the values $N=2,3,\dots,N_m$, where N_m is maximal length of temporal row, N is current amount of elements of temporal row. For every value N we determine the size of the Hurst $H(N)$ index and put aside it on a y -axis. Researches are executed in the Matlab environment. For the study four realization of network traffic is chosen, got in the university of Napoly city (Italy). In obedience to a license information is freely accessible for the analysis. The studied temporal rows are measuring of the UDP delay and TCP-packets. In first case (time-series UDP_d64) UDP packets have a volume a 64 byte, in the second (time-series UDP_d512) is a 512 byte. Third time-series (TCP_d64) is the transmission of TCP-packets of volume is a 64 byte, fourth time-series (TCP_d512) is the transmission of TCP- packets of volume is a 512 byte. Next results are got. The traffic UDP_d64 consists of two persistent states and phase transition between states in moment when the Hurst function achieves the value $H=0,5$. The traffic UDP_d512 has phase transition from an antipersistent state ($H=0,25$) to the state of homogeneous traffic ($H=0,5$) and gradual return to antipersistent state. The TCP-traffic has gradual return from persistent to homogeneous state, homogeneous mode, phase transition to the persistent state and gradual return from persistent state to homogeneous state.

Keywords: *UDP- traffic, TCP- traffic, delay, Hurst index, Hurst function, phase transition.*

*Статья поступила в редакцию 20.09.2015
Рекомендована к публикации д-ром техн. наук В.Н. Павлышом*