

ИССЛЕДОВАНИЕ САМОПОДОБНОЙ СТРУКТУРЫ ТЕЛЕТРАФИКА БЕСПРОВОДНОЙ СЕТИ

Введение

Благодаря многочисленным экспериментам и исследованиям, проведенным за последнее десятилетие учеными различных стран, сегодня можно с уверенностью утверждать, что трафик современных телекоммуникационных сетей с коммутацией пакетов обладает особой структурой, не позволяющей использовать при проектировании привычные методы, основанные на марковских моделях и формулах Эрланга и которые хорошо себя зарекомендовали в телефонных сетях с коммутацией каналов. Особенности, о которых идет речь, принято называть проявлением эффекта самоподобия телетрафика [1], [2] и в первую очередь они проявляются в специфическом для этого случая профиле (форме) трафика: в реализации всегда присутствует некоторое количество достаточно сильных выбросов на фоне относительно низкого среднего уровня. Описанное явление значительно ухудшает характеристики (увеличивает потери, задержки, джиттер) при прохождении самоподобного трафика через сеть даже в случаях, когда средняя интенсивность трафика намного ниже потенциально достижимой скорости передачи в данном канале.

К настоящему времени показано, что самоподобной структурой обладает телетрафик в проводных сетях при работе широко распространенных протоколов *Ethernet* [1], *ОКС7* [3], *VOIP* [4], *TCP* [5] и пр. Также обнаружены аналогичные эффекты в сотовых телефонных сетях с коммутацией пакетов.

Основные широкоизвестные реализации сетевого трафика, полученные *Vern Paxson*, *Will Leland* и *Dan Wilson* доступны в *Internet Traffic Archive* [6]. Однако данные реализации, в основном датируемые началом 90 годов, благодаря появлению новых технологий, протоколов, приложений к настоящему времени могли несколько устареть. Кроме того, все большее распространение получают беспроводные телекоммуникационные технологии, данные о степени самоподобия трафика в которых пока отсутствуют.

В настоящей работе описывается эксперимент по снятию реализаций трафика в беспроводной сети одного из крупных Интернет-провайдеров масштаба города (ЗАО "Вилайн", Воронеж), а также приводятся результаты анализа структурных особенностей данного трафика.

Постановка эксперимента

Все реализации трафика были получены в процессе мониторинга магистрального канала беспроводного Интернет-провайдера. Логическая схема организации сети отображена на рис. 1. На коммуникационном узле провайдера установлена шести-секторная антенна, к каждому сектору которой подключена точка доступа (*Access Point*) стандарта *IEEE 802.11b*.

Всего ко всем шести точкам доступа провайдера подключено порядка 40 беспроводных клиентов. Беспроводной клиент представляет собой точку доступа аналогичную провайдерской, к которой по локальной сети *Fast Ethernet* подключен один или несколько хостов. Все точки доступа работают как прозрачные мосты для протоколов верхних уровней. Маршрутизацию всех соединений, а также контроль и управление пропускной способностью (*Traffic Shaping*) осуществляет маршрутизатор. С помощью него каждому беспроводному клиенту предоставлена пропускная способность от 64 до

512 Кбит/с. Механизмы раннего обнаружения перегрузок, такие как *RED* или *WRED* не используются.

Для сбора трафика использовался персональный компьютер под управлением операционной системы *Windows 2000 Professional* с установленным на нем анализатором протоколов (сниффер) *Ethereal*, который обеспечивает точность определения времени получения пакета 10^{-9} сек.

Сниффер был подключен к сети таким образом, чтобы фиксировать трафик, проходящий через точку "1" и одновременно через точку "2".

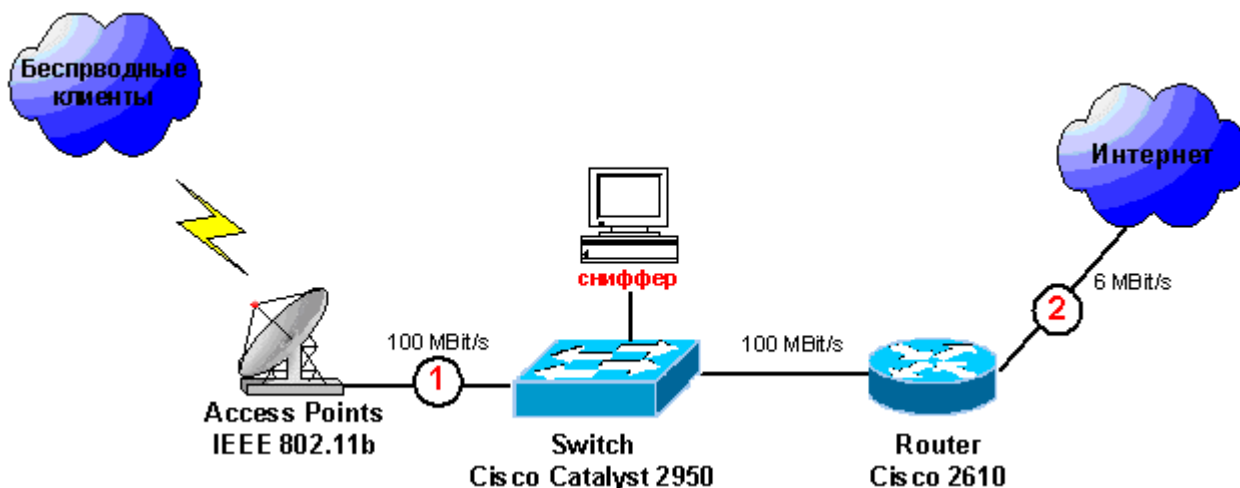


Рис. 1. Логическая схема организации эксперимента по снятию трафика.

Обратим внимание, что через точку "1" проходит трафик обмена информацией между беспроводными клиентами и вместе с ним трафик обмена информацией между беспроводными клиентами и сетью Интернет. Через точку "2" проходит лишь последний из них.

Все зафиксированные пакеты записывались в файл формата *tcpdump*. Разделение общего трафика, собранного сниффером, на трафики, проходящие через точку "1", точку "2", входящие и исходящие производилось в ходе его дальнейшей обработки на основе адресной информации.

Всего в результате эксперимента было собрано более 12,7 миллионов пакетов. Из них 67% пришлось на протокол TCP.

Характеристика реализаций

Процедура записи трафика началась 18 марта 2004 года (четверг) в 10.00 и закончилась 18 марта в 17.00 по московскому времени. Общая длительность составила 7 часов. В результате эксперимента были получены следующие реализации трафика:

Weth04.dat – общий (входящий + исходящий) трафик уровня протокола *Fast Ethernet*, зафиксированный в точке "1". Включает в себя трафик между клиентам, а также трафик обмена с сетью *Internet*. Содержит информацию о 12 712 076 Ethernet-кадрах.

Weth04.src – трафик уровня протокола *Fast Ethernet*, зафиксированный в точке "2", с направлением от беспроводных клиентов в сеть *Internet*. Содержит информацию о 4 554 668 Ethernet-кадрах.

Weth04.dst – трафик уровня протокола *Fast Ethernet*, зафиксированный в точке "2", с направлением из сети *Internet* к беспроводным клиентам. Содержит информацию о 5 586 556 Ethernet-кадрах.

Wtcp04.dat – общий (входящий + исходящий) трафик уровня протокола *TCP*, зафиксированный в точке "1". Включает в себя трафик между клиентами, а также трафик обмена с сетью *Internet*. Содержит информацию о 8 468 548 TCP-пакетах.

Wtcp04.src – трафик уровня протокола *TCP*, зафиксированный в точке “2”, с направлением от беспроводных клиентов в сеть *Internet*.

Wtcp04.dst – трафик уровня протокола *TCP*, зафиксированный в точке “2”, с направлением из сети *Internet* к беспроводным клиентам.

Данные представлены двумя колонками чисел в ASCII-формате: первая колонка содержит временные метки (в сек.), а вторая – общий размер *Ethernet*-кадра в байтах, либо размер поля данных *ip*-пакета в случае *TCP*-пакета. Заметим, что отсчеты не эквидистантны, а соответствуют моментам прихода пакетов. Для приведения реализаций трафика к эквидистантному виду по оси времени, была использована процедура агрегирования с уровнем агрегирования $dT = 0,05$ сек. Суть данной процедуры заключается в следующем: ось времени разбивается на равные интервалы dT , каждому из которых ставится в соответствие отношение количества прошедшей за данный интервал времени информации (в байтах) к длительности интервала (в секундах). При этом получают соответствующие реализации *Weth04_0.05.dat*, *Weth04_0.05.src*, *Weth04_0.05.dst*, *Wtcp04_0.05.dat*, *Wtcp04_0.05.src*, *Wtcp04_0.05.dst*, однако смысл данных, представленных в последних столбцах реализаций, изменяется и теперь представляет собой интенсивность передачи информации (байт/сек).

Особенности технологии IEEE 802.11b

Стандарт *RadioEthernet IEEE 802.11* определяет порядок организации беспроводных сетей на уровне управления доступом к среде (*MAC*-уровне) и физическом (*PHY*-уровне). Данная технология использует методы расширения спектра сигнала с помощью прямой последовательности (*DSSS*). Стандарт предусматривает работу на частоте 2.4 ГГц и обеспечивает скорость передачи данных до 11 Мбит/с.

Согласно спецификации *IEEE 802.11b* [7] в качестве алгоритма доступа к среде используется метод *CSMA/CA* (*Carrier Sense Multiple Access with Collision Avoidance*). Протокол *CSMA* широко распространен и применяется, например, в локальных проводных сетях *Ethernet* (вариант *CSMA/CD*). Принцип его работы очень прост: станция, которая намерена начать передачу, должна убедиться в том, что среда не занята. В противном случае передача данных будет отложена на более позднее время. Этот протокол хорошо работает в малозагруженных сетях. В сетях с довольно интенсивным обменом данными существенно возрастает вероятность возникновения коллизий. Коллизия – это одновременная передача данных несколькими станциями. В случае *Ethernet*-сетей возникновение коллизии определяют передающие станции (*CD*).

Однако, протокол *CD* (*Collision Detection*), хорошо работающий в проводных сетях, плохо подходит для сетей беспроводных по следующим причинам:

- использование механизма *CD* потребует применения полнодуплексных радиоустройств, способных одновременно передавать и принимать данные, что существенно увеличит цену оборудования;
- в случае беспроводной сети нельзя гарантировать того, что каждая станция будет иметь возможность взаимодействовать со всеми остальными станциями сети (что является главным условием механизма *CD*).

Для решения этих проблемы в стандарте *IEEE 802.11b* применен механизм предотвращения коллизий (*Collision Avoidance*) в сочетании с обязательными подтверждениями приема всех пакетов. Данный механизм работает следующим образом: станция перед передачей данных проверяет среду передачи, если она занята, то передача откладывается. Если среда свободна на требуемое время (называемое *DIFS*, *Distributed Inter Frame Space*), тогда станция передает данные. Принимающая станция проверяет *CRC* у принятого пакета, и если все нормально, то шлет пакет с подтверждением (*ACK*). Получение *ACK* говорит о том, что коллизии не произошло. Если передающая сторона не получит подтверждения, будет произведена повторная передача.

Для того, чтобы уменьшить вероятность коллизии двух станций, которые находятся вне зоны покрытия друг друга, стандартом определен механизм определения состояния виртуальной несущей (*Virtual Carrier Sense*): станция, желающая начать передачу данных, сначала передает короткий управляющий пакет, называемый *RTS* (*Request To Send*), в котором указывается источник, получатель и продолжительность основной передачи. Получатель (если среда передачи свободна) посылает ответный управляющий пакет, называемый *CTS* (*Clear To Send*), в котором содержится информация о длительности будущей передачи. Все станции, получившие либо *RTS*, либо *CTS* (либо и то, и другое), устанавливают свои *Virtual Carrier Sense* индикаторы (названные *NAV*, *Network Allocation Vector*) на указанное время. Станции будут использовать эту информацию вместе с физической проверкой среды передачи перед передачей данных.

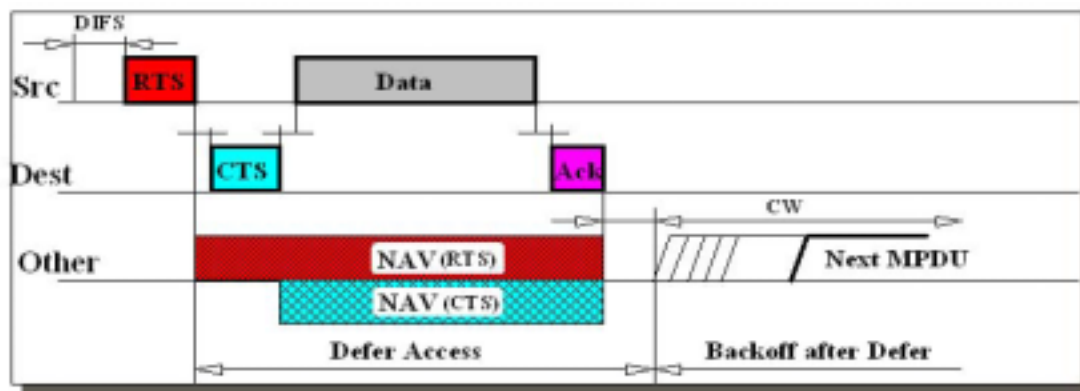


Рис. 2. Временная диаграмма функционирования алгоритма *CSMA/CA*

Результаты анализа трафика

На рис.3 представлен график реализации трафика *Weth04_0.05.dat*, которая представляет собой агрегированный по уровню 0,05 сек временной ряд, соответствующий общему трафику протокола *Fast Ethernet*, зафиксированному в сети (в точке “1”). Напомним, что данная реализация включает в себя трафик обмена информацией беспроводных клиентов друг с другом, а также трафик обмена с сетью *Internet*.

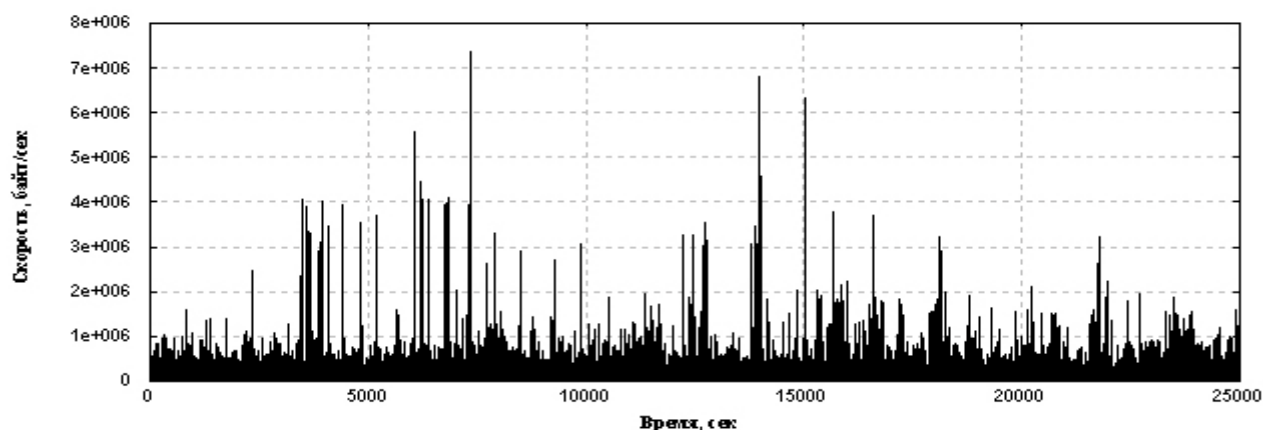


Рис. 3. Временной ряд *Weth04_0.05.dat*

В таблице 1 приведены основные характеристики данной реализации.

Анализируя визуально график, приведенный на рис. 3, можно заметить, что сетевой трафик является значительно неравномерным процессом: имеются выбросы достаточно сильной амплитуды, в то время как среднее значение трафика – относительно мало. Описывая такое поведение, часто говорят, что данный процесс подчиняется одному из

законов распределения с “тяжелыми хвостами” (типа Парето) или имеет “берстную” структуру (высокую пачечность).

Табл. 1. Основные характеристики реализации *Weth04_0.05.dat*

Параметр/Временной ряд	<i>Weth04_0.05.dat</i>
Уровень агрегирования	0,05 сек
Длительность реализации	504 010 отсчетов по 0,05 сек, 7 часов
Среднее	1.348998e+05 байт/сек ~ 1,079 Мбит/с
СКО	2.558526e+05 байт/сек ~ 2,046 Мбит/с
СКО/Среднее	1,89

В частности, наибольший выброс в данной реализации превышает 58 Мбит/с (см. рис.3) при потенциально достижимой скорости 100 Мбит/с в данном канале согласно спецификации *Fast Ethernet*, однако среднее значение трафика соответствует ~1 Мбит/с (см. табл. 1).

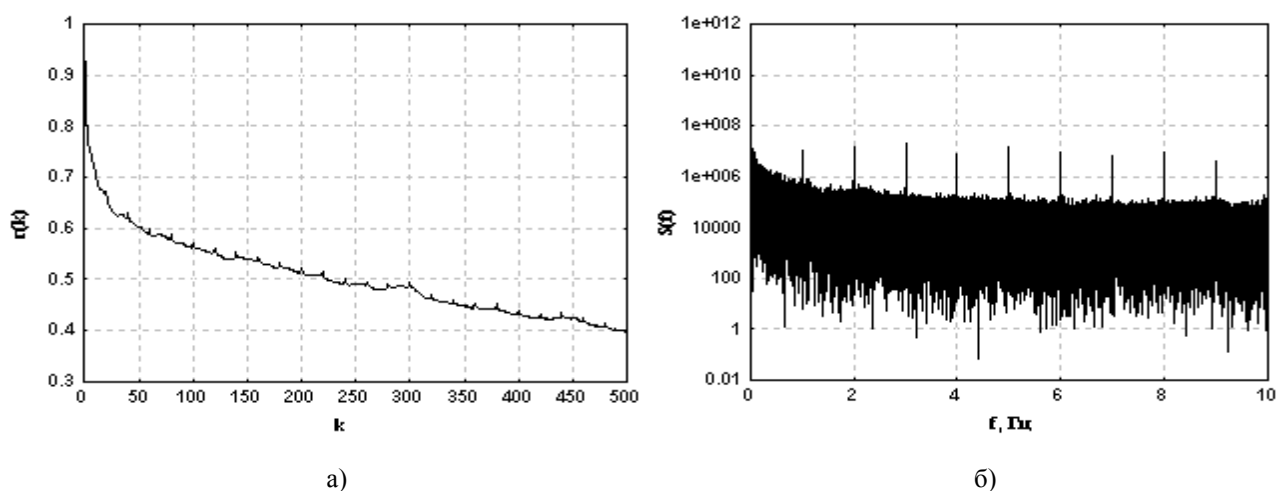


Рис. 4. а) Автокорреляционная функция (АКФ) временного ряда *Weth04_0.05.dat*
 б) Энергетический спектр временного ряда *Weth04_0.05.dat*

На рис. 4 можно видеть автокорреляционную функцию (а) и энергетический спектр (б) анализируемого процесса (в логарифмическом масштабе по оси ординат). Изучая АКФ, легко заметить достаточное слабое ее убывание: даже при сдвиге 500 значение АКФ остается выше 0,4. В то же время на графике АКФ отчетливо прослеживаются периодические пульсации.

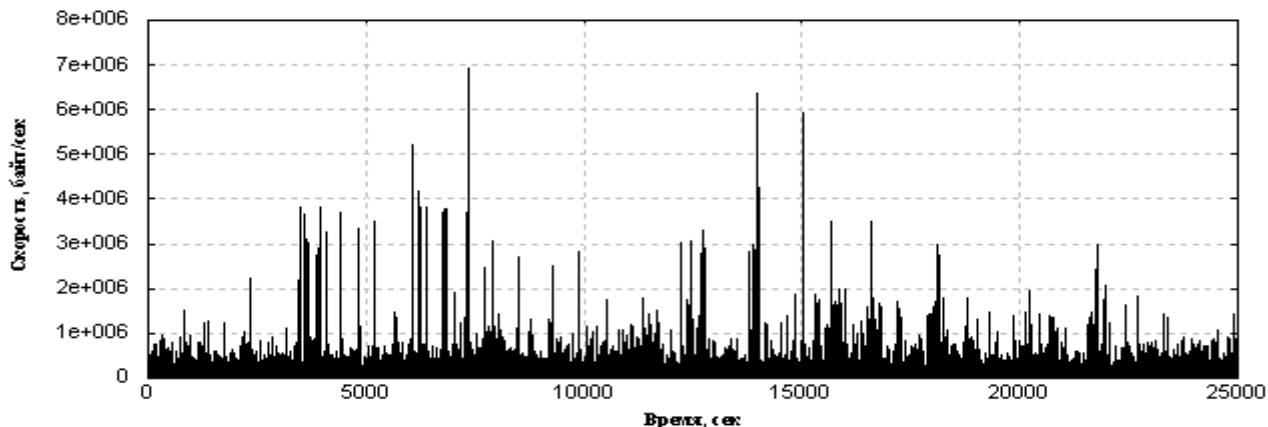


Рис. 5. Временной ряд *Wtcp04_0.05.dat*

Присутствие значительных гармонических компонент также можно отследить по графику энергетического спектра. Очевидно наличие гармоник с частотами 1, 2, 3... Гц, которые

не исчезают при изменении уровня агрегирования. Благодаря каким именно сетевым механизмам и протоколам возникают данные гармоники сказать с уверенностью на данном этапе пока затруднительно. Отметим, однако, что аналогичные явления уже отмечались в [8], [9] (~11,9 Гц и 1,25 Гц), а также в [10] (~5 Гц).

На рис.5 представлен график реализации трафика *Wtcp04_0.05.dat*, представляющий собой агрегированный по уровню 0,05 сек временной ряд, соответствующий трафику протокола *TCP*, который был зафиксирован в точке “2” сети. Как можно видеть, профиль данного временного ряда практически полностью повторяет профиль временного ряда трафика *Weth04_0.05.dat*. Изучая автокорреляционные функции и энергетические спектры (см. рис. 6), также можно отметить присутствие выделяющихся гармонических компонент и слабое убывание АКФ данного вида трафика.

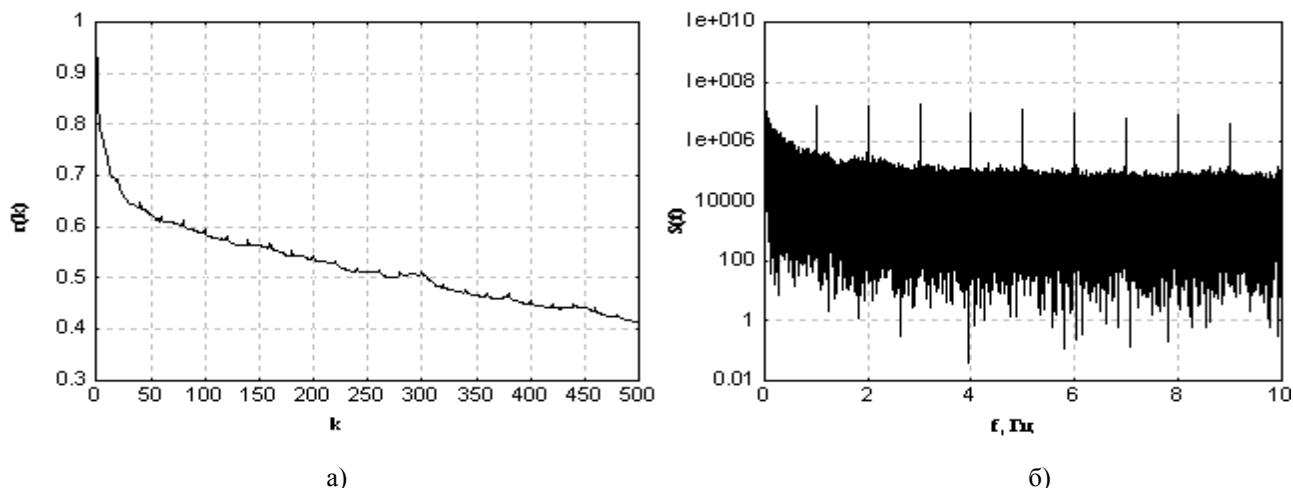


Рис. 6. а) Автокорреляционная функция временного ряда *Wtcp04_0.05.dat*.
 б) Энергетический спектр временного ряда *Wtcp04_0.05.dat*.

Таким образом, гармонические компоненты, профиль и самоподобная структура сетевого трафика проявляют себя уже на транспортном уровне, и, по-видимому, определяются либо спецификой работы непосредственно самого протокола *TCP*, либо алгоритмами работы протоколов и механизмов верхних уровней модели *OSI*.

Табл.2. Участки временных реализаций

Номер участка	Положение на оси времени, секунд с начала реализации	Длительность
1	0,52 – 2257,58	45 143 отсчетов по 0,05 сек; 0,63 ч
2	3954,87 – 6142,53	43 753 отсчетов по 0,05 сек; 0,61 ч
3	8070,16 – 10108,64	40 770 отсчетов по 0,05 сек; 0,57 ч
4	12546,54 – 13933,84	27 746 отсчетов по 0,05 сек; 0,38 ч

Более детальный анализ полученных реализаций в настоящей работе производился с помощью изучения и сравнения характеристик временных рядов *Weth04_0.05.dat*, *Weth04_0.05.src*, *Weth04_0.05.dst*, *Wtcp04_0.05.dat* на одинаковых временных интервалах. Для этого во всех изучаемых реализациях были выделены участки 1, 2, 3 и 4, положение на оси времени и длительность которых отражены в табл. 2.

На каждом из участков для изучаемых реализаций трафика были оценены средние, среднеквадратические отклонения (СКО), параметра Хэрста *H* (методами Виттла и Эбри-Вейча). Основные характеристики изучаемых реализаций приведены в табл.3,4,5,6.

Табл.3. Характеристики реализации Weth04_0.05.dat

Параметр	Участок 1	Участок 2	Участок 3	Участок 4
Среднее, байт/сек	8,981711e+04	1,227245e+05	1,182806e+05	2,878128e+05
СКО, байт/сек	9,694183e+04	1,898454e+05	1,245987e+05	6,336444e+05
СКО/Среднее	1,07	1,54	1,05	2,2
Параметр Херста (Whittle Estimator)	0,780 (95% Confidence Intervals [0,773-0,787])	0,964 (95% Confidence Intervals [0,957-0,972])	0,772 (95% Confidence Intervals [0,765-0,780])	0,978 (95% Confidence Intervals [0,967-0,988])
Параметр Херста (Abyr-Veitch Estimator)	0,859 (95% Confidence Intervals [0,851-0,867])	1,045 (95% Confidence Intervals [1,037-1,054])	0,811 (95% Confidence Intervals [0,803-0,819])	0,953 (95% Confidence Intervals [0,941-0,965])

Табл.4. Характеристики реализации Weth04_0.05.dst

Параметр	Участок 1	Участок 2	Участок 3	Участок 4
Среднее, байт/сек	4,313164e+04	6,309383e+04	6,765121e+04	5,603362e+04
СКО, байт/сек	7,803364e+04	8,726733e+04	1,096809e+05	7,574341e+04
СКО/Среднее	1,8	1,3	1,6	1,35
Параметр Херста (Whittle Estimator)	0,817 (95% Confidence Intervals [0,809-0,824])	0,721 (95% Confidence Intervals [0,713-0,728])	0,796 (95% Confidence Intervals [0,788-0,803])	0,751 (95% Confidence Intervals [0,741-0,761])
Параметр Херста (Abyr-Veitch Estimator)	0,884 (95% Confidence Intervals [0,876-0,892])	0,762 (95% Confidence Intervals [0,754-0,770])	0,835 (95% Confidence Intervals [0,827-0,843])	0,820 (95% Confidence Intervals [0,809-0,832])

Табл.5. Характеристики реализации Weth04_0.05.src

Параметр	Участок 1	Участок 2	Участок 3	Участок 4
Среднее, байт/сек	2,424995e+04	2,370111e+04	2,845095e+04	3,321249e+04
СКО, байт/сек	4,263671e+04	5,473894e+04	7,595552e+04	5,369175e+04
СКО/Среднее	1,7	2,3	2,6	1,6
Параметр Херста (Whittle Estimator)	0,801 (95% Confidence Intervals [0,794-0,808])	0,716 (95% Confidence Intervals [0,709-0,723])	0,858 (95% Confidence Intervals [0,851-0,865])	0,716 (95% Confidence Intervals [0,706-0,726])
Параметр Херста (Abyr-Veitch Estimator)	0,856 (95% Confidence Intervals [0,848-0,864])	0,742 (95% Confidence Intervals [0,734-0,751])	0,925 (95% Confidence Intervals [0,917-0,933])	0,770 (95% Confidence Intervals [0,758-0,781])

Табл.6. Характеристики реализации Wtcp04_0.05.dat

Параметр	Участок 1	Участок 2	Участок 3	Участок 4
Среднее, байт/сек	5.227452e+04	8.656856e+04	8.085651e+04	2.287812e+05
СКО, байт/сек	7.978429e+04	1.761661e+05	1.092429e+05	5.932242e+05
СКО/Среднее	1,53	2,03	1,35	2,59
Параметр Херста (Whittle Estimator)	0,770 (95% Confidence Intervals [0,763-0,777])	0,973 (95% Confidence Intervals [0,966-0,980])	0,755 (95% Confidence Intervals [0,747-0,762])	0,983 (95% Confidence Intervals [0,973-0,993])
Параметр Херста (Abyr-Veitch Estimator)	0,834 (95% Confidence Intervals [0,826-0,842])	1,057 (95% Confidence Intervals [1,049-1,065])	0,784 (95% Confidence Intervals [0,776-0,792])	0,967 (95% Confidence Intervals [0,955-0,978])

Анализируя результаты оценок основных параметров, можно заметить, что трафик в канале *Internet* асимметричен: уровень среднего исходящего трафика (реализация Weth04_0.05.src) беспроводных клиентов меньше, чем уровень среднего входящего (реализация Weth04_0.05.dst). Отмеченное явление достаточно типично для данного случая.

Значение параметра Херста для всех рассмотренных случаев стабильно больше 0,5, что говорит о проявлении эффекта самоподобия в изучаемых реализациях. Однако с течением времени значение параметра Херста трафика несколько изменяется. Это особенно отчетливо видно для реализации *Wtcp04_0.05.dat*: оно лежит в интервале от 0,770 до 0,967.

Выводы

Исследования, проведенные в настоящей работе, подтверждают наличие самоподобных свойств в трафике современных телекоммуникационных сетей, использующих в том числе технологии беспроводного доступа *IEEE 802.11b*.

Самоподобные свойства проявляют себя в трафике как на канальном (*Fast Ethernet*), так и на транспортном (*TCP*) уровнях. Как видно из сравнения рис. 3 и 5 структура трафика канального уровня практически полностью определяется трафиком транспортного уровня.

С течением времени уровень самоподобия (значение параметра Хэрста для трафика) изменяется.

В трафике канального и транспортного уровней обнаружены значительные гармонические составляющие с частотами 1, 2, 3... Гц. В этой связи при разработке адекватных математических моделей телетрафика следует обращать внимание на присутствие периодических компонент.

Сравнивая полученные в настоящей работе результаты с аналогичными результатами для трафика проводных сетей можно заключить, что, несмотря на различные принципы функционирования канального и физического уровней, с точки зрения самоподобной структуры принципиальных отличий между данными видами трафика не обнаружено.

Реализации трафика, полученные и изученные в настоящей работе можно найти на информационном портале "*Self-Similarity h.u.*", посвященном проблеме самоподобного телетрафика по адресу www.teletraffic.ru.

С авторами можно связаться по электронной почте mail@teletraffic.ru, vplatov@dlink.ru.

Литература

1. W.E.Leland, M.S.Taqqu, W.Willinger, and D.V.Wilson. On the self-similar nature of Ethernet traffic (extended version).IEEE/ACM Transactions of Networking, 2(1):1-15,1994.
2. Шелухин О.И., Тенякишев А.М., Осин А.В. Фрактальные процессы в телекоммуникациях. Монография./ Под ред. О.И. Шелухина. – М.: Радиотехника, 2003.-480с.
3. Криштофович А.Ю. Самоподобие трафика сети ОКС №7. – МКИССиТ, Санкт-Петербург, 2002 г.
4. T. D. Dang, B. Sonkoly, S. Molnár, Fractal Analysis and Modelling of VoIP Traffic, NETWORKS2004, Vienna, Austria, 2004.
5. Feng W., Tinnakornrisuphap P. The Failure of TCP in High-Performance Computational Grids //SC2000: High-Performance Network and Computing Conference, Dallas, TX , November 2000.
6. Internet traffic archive <http://ita.ee.lbl.gov/>
7. Pablo Brenner, A Technical Tutorial on the IEEE 802.11 Protocol, 1997.
8. G. Olowoyeye, B. Kim, K. Chandra. Modelling Spectral Features in TCP Traffic.
9. K. Chandra, C. You, G. Olowoyeye and C. Thompson, Non-linear Time-Series Models of Ethernet Traffic. Submitted to INFOCOM'99, July 1998.
10. V. Petroff, Self-Similar Network Traffic: From Chaos and Fractals to Forecasting and QoS. NEW2AN, 2004.