

# ВЛИЯНИЕ ЭФФЕКТА САМОПОДОБНОСТИ НА РАБОТУ СЕТЕВЫХ СЕРВЕРОВ

**С. Ильницкий**

*Рижский технический университет*

*Рига, Латвия*

*Sergejs.Ilnickis@dati.lv*

В данной публикации проведен анализ реального трафика компьютерной сети, используя один из методов определения самоподобности. На основании данных, полученных в ходе экспериментов, доказана самоподобная природа реального сетевого трафика. Разработана модель, которая описывает самоподобный трафик с различными степенями самоподобности и проведено сравнение результатов с результатами, полученными для модели, обслуживающей системы с пуассоновским трафиком.

*Ключевые слова:* компьютерные сети, трафик, моделирование, методы расчета, самоподобность.

## 1. ВВЕДЕНИЕ

Исследования различных типов сетевого трафика доказывают, что сетевой трафик является самоподобным (self-similar) или фрактальным (fractal) по своей природе, т. е. в нем присутствуют так называемые всплески или пачки (burst) пакетов, наблюдаемые в различных временных интервалах (от миллисекунд до минут или даже часов). Из этого следует, что широко используемые в настоящее время методы моделирования и расчета сетевых систем, основанные на использовании пуассоновских потоков, не дают полной и точной картины происходящего в сети.

Целью данной публикации является анализ реального трафика компьютерной сети, используя один из методов определения самоподобности, разработка модели, которая описывает самоподобный трафик с различными степенями самоподобности и сравнение результатов с результатами традиционного моделирования.

## 2. АНАЛИЗ ТРАФИКА СЕТЕВОГО СЕРВЕРА

**2.1. Описание исследуемой сети.** Анализ производился в корпоративной сети компании "X", которая является компанией, разрабатывающей программное обеспечение, и которая широко использует для этого базы данных внутри и за пределами корпоративной сети. Максимальная скорость передачи данных с корпоративной сети равняется 10 Mbit/s. Из большого числа доступных серверов был выбран один, работавший большую часть времени с максимальной нагрузкой. Основными функциями данного сервера в сети являются: WEB-сервер, сервер баз данных, файловый сервер и сервер печати.

Результаты измерений трафика были получены с помощью Multi Router Traffic Grapher (MRTG). Данные собирались в течение месяца, 24 часа в сутки с интервалом 5 минут. Для дальнейшего анализа полученные данные были разделены на группы: трафик в течение суток, трафик в течение недели, трафик в течение месяца. Трафик в течение суток дополнительно был разделен на подгруппы, в зависимости от интенсивности нагрузки:

- с 02:00 до 08:00 – низкая интенсивность трафика нагрузки;
- с 08:00 до 14:00 – средняя интенсивность трафика нагрузки;
- с 14:00 до 20:00 – высокая интенсивность трафика нагрузки;
- с 20:00 до 02:00 – максимальная интенсивность трафика нагрузки.

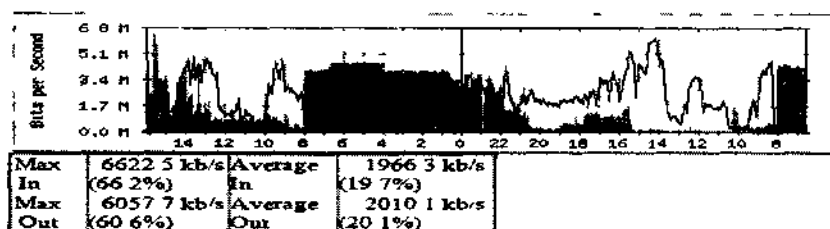


Рис 1 Трафик в течение суток

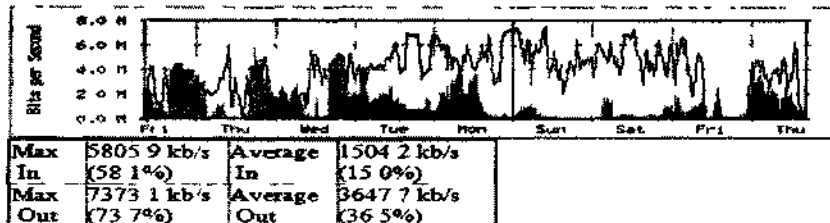


Рис 2 Трафик в течение недели

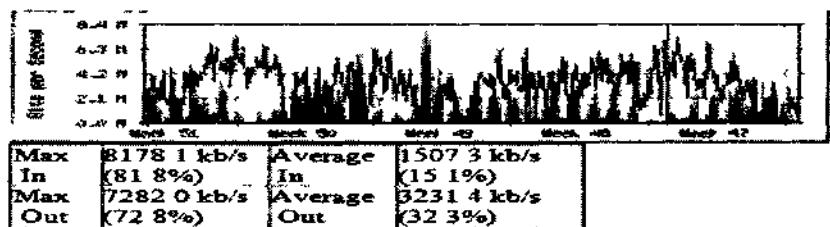


Рис 3 Трафик в течение месяца

**2.2. Результаты измерений.** Итак, предполагаем, что трафик является самоподобным. Наша цель доказать это утверждение. Для этого воспользуемся методом абсолютных моментов.

В данном методе исходная последовательность с длиной  $N$  разделяется на блоки с длиной  $m$ . На границах блока последовательность имеет среднее значение (дисперсия):

$$X^{(m)}(k) := \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i, k=1, 2, \dots, [N/m].$$

Для каждого блока рассчитываем дисперсию  $X^{(m)}$  и математическое ожидание  $\bar{X}$  для всей последовательности. После этого для каждого блока находим момент  $n$ :

$$AM_n^{(m)} = \frac{1}{N/m} \sum_{k=1}^{N/m} |X^{(m)}(k) - \bar{X}|^n.$$

В данном выражении  $n = 1$  (абсолютное среднее значение). Далее последовательность обозначим как  $m$  и построим график в логарифмическом масштабе – зависимость средних значений (дисперсий) абсолютных моментов для последовательности от  $m$ . С помощью полученных точек строим аппроксимирующую прямую по методу минимального среднеквадратического отклонения от экспериментальных данных. В результате получившийся наклон линии будет равен  $\beta$ . С помощью  $\beta$  можем найти коэффициент самоподобности  $H = 1 - |\beta|$ .

Необходимо, чтобы длина каждого блока и число блоков были большими. Если последовательность – это не процесс с медленно изменяющейся зависимостью, то  $H = 0.5$  и наклон аппроксимационной линии будет равен  $1/2$ . Если процесс самоподобен, тогда  $0.5 < H < 1.0$  и наклон линии будет меньше чем  $1/2$ .

Используя вышеописанную методику, были получены результаты, один из которых можно увидеть на рис. 4.

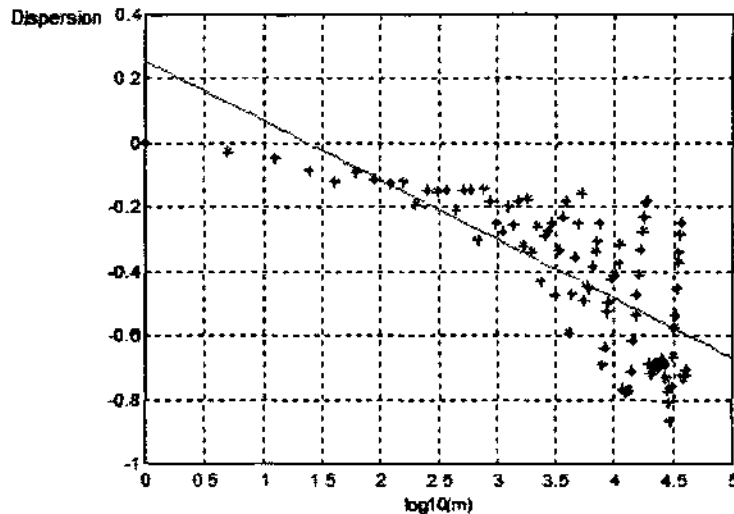


Рис. 4. Дисперсионно-временной график для данных, полученных в понедельник

Значения коэффициента Харста, рассчитанные с помощью метода абсолютных моментов для различных интервалов времени, показаны в табл. 1, 2.

Таблица 1

## Коэффициент Харста для трафика за неделю

День недели	$Abs(\beta)$	Коэф. Харста
Понедельник	0.1845	0.82
Вторник	0.2206	0.78
Среда	0.2537	0.75
Четверг	0.1620	0.84
Пятница	0.2120	0.79
Суббота	0.1518	0.85
Воскресенье	0.2441	0.76
Общий трафик за неделю	0.2131	0.79

Таблица 2

## Коэффициент Харста для трафика за сутки

Интервал времени	$Abs(\beta)$	Коэф. Харста
02:00 – 08:00	0.1845	0.71
Низкая интенсивность		
08:00 – 14:00	0.2206	0.73
Средняя интенсивность		
14:00 – 20:00	0.2537	0.76
Высокая интенсивность		
20:00 – 02:00	0.1620	0.81
Максимальная интенсивность		
Общий трафик за сутки	0.2118	0.75

Трафик за сутки был специально разделен на интервалы. Существует мнение, что коэффициент самоподобности зависит от интенсивности трафика, а также от типа передаваемой информации. На самом деле, интенсивность трафика очень мала в интервале времени 02:00 – 08:00, и это связано с особенностями поведения пользователей. В данном интервале времени параметр Харста минимален. Рабочий день начинается в 08:00 и интенсивность трафика невелика, но, несмотря на это, параметр Харста уже увеличивается. В интервале с 20:00 до 02:00, когда программисты начинают тестировать написанный программный код в автоматическом режиме, интенсивность потока заявок к серверу вырастает очень значительно и параметр Харста также возрастает до значения 0.81. Сетевой трафик в случае максимальной интенсивности (20:00 – 02:00) имеет очень высокую степень самоподобности.

После анализа результатов можем сделать некоторые выводы:

- Трафик в локальной сети – это самоподобный процесс.
- Эффект самоподобности проявляется в широком диапазоне времени: от нескольких часов до нескольких месяцев.
- Коэффициент самоподобности трафика меняется в пределах  $\approx 0.7-0.85$ .
- В случае если возрастает интенсивность трафика, также возрастает коэффициент самоподобности.
- Для трафика в течение суток значение коэффициента Харста одинаково для любой интенсивности.
- Для трафика в течение месяца коэффициент Харста равен  $\approx 0.81$ .

**2.3. Моделирование самоподобного трафика в среде GPSS.** Для того чтобы смоделировать самоподобный входной трафик, для начала необходимо задать все главные параметры, характерные для трафика, использовавшегося в ходе экспериментов.

Из топологии исследуемой сети известно, что у нас имеется 350 рабочих станций, посылающих запросы к серверу, подключенному к маршрутизатору. Поэтому для моделирования нам потребуется 350 ON-OFF источников запросов, один маршрутизатор (для упрощения будем рассматривать его как концентратор) и один прибор массового обслуживания (сервер) (рис. 5).

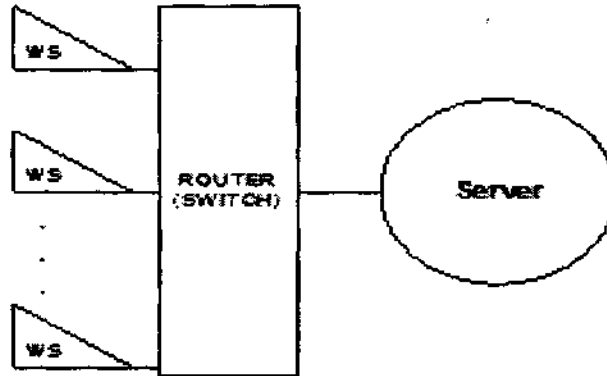


Рис. 5. Упрощенная схема анализируемой сети

Для примера возьмем трафик за сутки, полученный в ходе экспериментов (рис. 6).

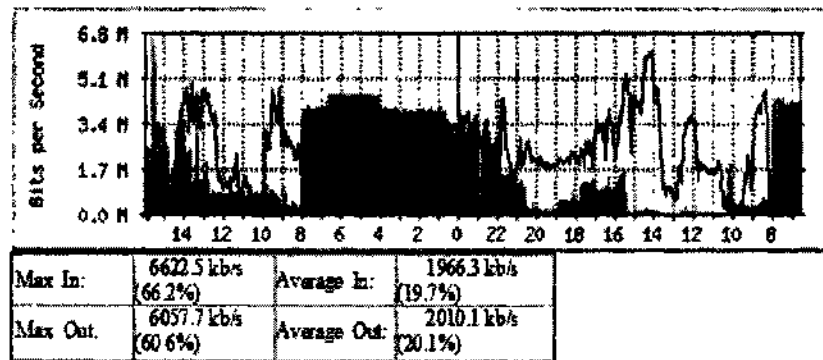


Рис. 6. Трафик за сутки

Как видим, средний выходной трафик достигает значения 2 Mbit/s. Это означает, что канал загружен на 20 %. Будем использовать это значение в качестве средней интенсивности работы ON-OFF источников. Следовательно, ON-период в 5 раз меньше, чем OFF-период. Функция GPSS позволяет генерировать распределение Парето с параметром  $\alpha$ . В нашем эксперименте параметр равен  $\alpha = 1.4$ , что соответствует значению коэффициента Харста  $H = 0.8$ . Усредняющий коэффициент  $m$  будем изменять от 1 до 10 000.

Проверим, является ли полученный таким образом поток самоподобным с параметром Харста  $H = 0.8$  (значение получено ранее с помощью метода абсолютных моментов).

Таблица 3  
Дисперсионно-временная  
зависимость

m	log(m)	Var(x)	log[Var]
1	0	2.371	0.374932
2	0.30103	2.334	0.368101
3	0.477121	2.305	0.362671
5	0.69897	2.51	0.352375
10	1	2.115	0.32531
20	1.30103	1.843	0.65525
50	1.69897	1.297	0.11294
100	2	1.018	0.007748
200	2.30103	0.826	0.08302
500	2.69897	0.632	0.19928
1000	3	0.5	0.30103
2000	3.30103	0.396	0.4023
5000	3.69897	0.281	0.55129
10000	4	0.222	0.65365

Используя данные из табл. 3, строим дисперсионно-временной график, который показан на рис. 7.

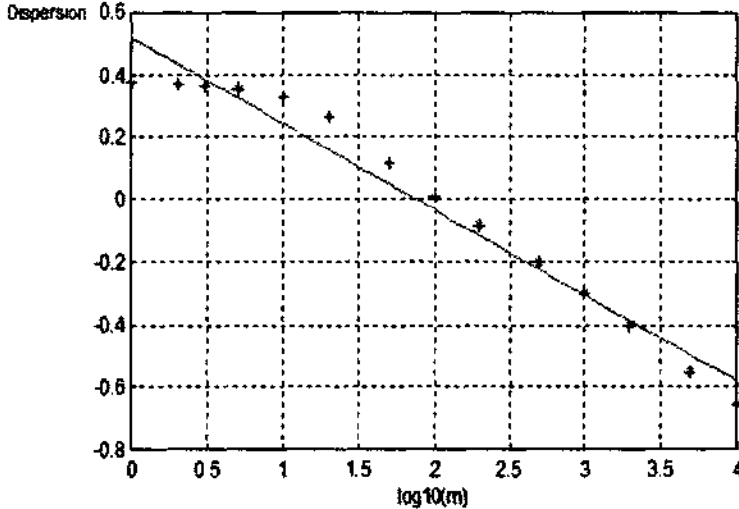


Рис. 7. Дисперсионный график для сгенерированного трафика,  
 $\beta = -0.2744$ ,  $H = 0.8628$

Угол наклона прямой составляет  $-0.2744$ . Это означает, что значение параметра Харста равно  $0.8628$ , что приблизительно соответствует значению параметра Харста, полученного в ходе измерения реального трафика сервера в локальной сети.

Это дает нам возможность использовать разработанную модель для исследования реальных сетей и сравнивать рабочие параметры потока в случаях, когда входной поток имеет самоподобную структуру, или в случае пуассоновского входного потока.

## ЛИТЕРАТУРА

1. *Leland W., Taqqu M., Willinger W., Wilson D.* On The Self-Similar Nature Of Ethernet Traffic // Proc. ACM SIGCOMM'93. San-Fransisco. 1993. P. 183-193.
2. *Willinger W., Taqqu M., Sherman R., Wilson D.* Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level // IEEE/ACM Transactions on Networking, 1997. V. 5. № 1.
3. *Popescu A.* Traffic Self-Similarity. University of Karlskrona/Ronneby: Bucharest, 2001.
4. *Paxson V.* Fast Approximation of Self-Similar Network Traffic, preprint Lawrence Berkely Laboratory and EECS Disvision. University of California, 1996.
5. *Ulanovs P., Petersons E.* Modeling of Self-similar traffic in high-performance Computer Networks, RTU, report collection for 42 International Conference, Riga. 2001.
6. *Цыбаков Б. С.* Модель телеграфика на основе самоподобного случайного процесса // Радиотехника. 1999. № 5.
7. *Willinger W., Taqqu M. S., Erramilli A.* A Bibliographical Guide to Self-Similar Traffic and Performance Modeling for Modern High-Speed Networks, Stochastic Networks: Theory and Applications, Oxford University Press, 1996.
8. *Stalling W.* High-Speed Networks TCP/IP and ATM Design Principles, Prentice-Hall Inc., 1998.
9. *Mandelbrot B. B.* Long-Run Linearity, Locally Gaussian Processes, H-Spectra and Infinite Variances // International Economic Review. 1969. V. 10. P. 82-113.