

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ТРАФИКА ПРОВОДНЫХ И БЕСПРОВОДНЫХ СЕТЕЙ

Косиор А.В.

Донецкий Национальный Технический Университет

Кафедра автоматики и телекоммуникаций

E-mail: alenakosior@gmail.com

Аннотация

Косиор А.В. Сравнительный анализ трафика проводных и беспроводных сетей.

Рассмотрено понятие фракталов и самоподобия, математическое описание дискретного фрактального процесса. Проанализирован трафик беспроводной и проводной сетей, доказана масштабная инвариантность. Проведена проверка независимости АКФ от уровня агрегирования, рассчитан параметр Хэрста различными методами.

Общая постановка проблемы

При проектировании, запуске и эксплуатации информационных телекоммуникационных сетей одной из основных проблем является задача обеспечения качества обслуживания (заданных уровней задержек, потерь и пр.) при обработке потока данных - трафика, являющегося следствием информационного обмена между системами.

До недавнего времени теоретическую базу для проектирования систем распределения информации обеспечивала теория телетрафика, которая является одной из ветвей теории массового обслуживания и появилась в результате работ А.К. Эрланга, Т. Энгсета, Г. О.Делла, К. Пальма, А.Я. Хинчина и др. [6]

Проблема изучения моделирования трафика является актуальной в наши дни и изучается многими исследователями, такими как Петров В.В., Платов В., Цыбаков Б.С., Воробьев О.В., Жуков И.А., Пустовойтов П.Е.

Наиболее распространенной моделью потока вызовов (данных) в теории телетрафика является простейший поток (стационарный ординарный поток без последствия), также называемый стационарным пуассоновским потоком.

Однако статистические наблюдения за сетевым трафиком в реальных компьютерных сетях показали, что входящий поток пакетов данных обладает свойством самоподобия. В большей степени свойство самоподобия выражается в наличии последствия в случайном характере трафика. Это значит, что рост нагрузки в сети может привести к еще большему увеличению нагрузки в дальнейшем. В результате исследований W.Leland, M.Taqqu, W.Willinger и D.Wilson был сделан вывод о том, что потоки в современных сетях нельзя аппроксимировать простейшими, поскольку они имеют иную структуру, чем принято в классической теории телетрафика.[4]

Анализ фрактальных свойств трафика позволит выделить важные числовые характеристики, на основе которых могут быть построены адаптивные алгоритмы статистического управления и прогнозирования. В результате использование свойств самоподобия автокорреляционной функции трафика можно обеспечить достижение высокой степени масштабируемости прогноза, что в свою очередь позволяет получать оценки для широкого диапазона временных интервалов на основе результатов измерения ограниченного набора данных.

Целью данной статьи является анализ реализаций трафика беспроводной и проводной сетей, выявление их особенностей, определение требований к дальнейшей реализации модели имитации трафика в магистерской работе.

Математическое описание дискретного фрактального процесса

Примерами самоподобных объектов являются фракталы (от лат. fractus – дробный, изрезанный) – структуры, состоящие из частей, которые в каком-то смысле подобны целому (согласно определению Мандельброта). Фракталами являются геометрические объекты: линии, поверхности, пространственные тела, имеющие значительно изрезанную форму и обладающие свойством самоподобия. [6]

Самоподобный характер трафика, наблюдаемый в современных высокоскоростных сетях связи, обусловлен тем, что сети являются интегральными и используются для передачи речи, данных, изображений (в том числе и по факсу), файлов и других видов информации, представляемых в форме стандартизованных пакетов. Такой трафик оказывает значительное влияние на характеристики систем связи. В частности, при увеличении размера буфера на входе канала вероятность потерь падает значительно медленнее, чем по экспоненциальному закону, свойственному широко используемым классическим моделям телетрафика.

Фрактальная размерность D временного ряда связана с показателем степени его фрактальности (показателем Хэрста) H формулой $H = 2 - D$. Параметры самоподобия H и D представляют собой меры устойчивости статистического явления или меры длительности долгосрочной зависимости стохастического процесса. Значения $H=0,5$ или $D=1,5$ указывают 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. Он более изменчив, чем случайный ряд, поскольку состоит из частых реверсов спад-подъем. При $0,5 < H \leq 1$ ряд трендонеустойчив. Тенденция его изменения может быть спрогнозирована.

Таким образом, важным параметром, характеризующим степень самоподобия, является параметр Хэрста H , который определяется для временного ряда из соотношения

$$R/S = (aN)^H$$

где $R = \max(X_k) - \min(X_k)$ – размах отклонения,

$$S = \sqrt{\frac{1}{N-1} \sum_{k=1}^N (x_k - \bar{x})^2}$$
 – стандартное отклонение,

N – число членов временного ряда X_k , a – константа.

Используя значение показателя Хэрста H , выделяют три типа случайных процессов:

$0 \leq H \leq 0,5$ – случайным процесс является антиперсистентным, или эргодическим, рядом, который не обладает самоподобием;

$H = 0,5$ – полностью случайный ряд, аналогичный случайным смещениям частицы при классическом броуновском движении;

$H > 0,5$ – персистентный (самоподдерживающийся) процесс, который обладает длительной памятью и является самоподобным.

Следует заметить, что в случае $0,5 < H < 1$ говорят о персистентном (поддерживаемом) поведении процесса, либо о том, что процесс обладает длительной памятью [7]. Другими словами, если в течение некоторого времени в прошлом наблюдались положительные приращения процесса, то есть происходило увеличение, то и впредь в среднем будет происходить увеличение. Иначе говоря, вероятность того, что процесс на $i+1$ шаге отклоняется от среднего в том же направлении, что и на i шаге настолько велика, насколько параметр H близок к 1. Таким образом, персистентные стохастические процессы обнаруживают четко выраженные тенденции изменения при относительно малом „шуме”.

Существует два класса фрактальных процессов, так называемые точно самоподобные и асимптотически самоподобные процессы. Процесс X называется точно самоподобным с

параметром β ($0 < \beta < 1$), если выполняются следующие условия: $D_m = \frac{D}{m^\beta}$, D – дисперсия

процесса X , D_m – дисперсия агрегированного процесса $X^{(m)}$, полученного уменьшением размера шкалы наблюдений X в m раз. Автокорреляционная функция (АКФ) сохраняется на всех масштабах: $R(k, X^{(m)}) = R(k, X)$.

Процесс X называется асимптотически самоподобным если для больших k выполняются условия: $D_m = \frac{D}{m^\beta}$, D – дисперсия процесса X , D_m – дисперсия агрегированного процесса $X^{(m)}$, полученного уменьшением размера шкалы наблюдений X в m раз. Параметр β связан с параметром Хэрста H соотношением $\beta = 2(1 - H)$.

АКФ сохраняется при $m \rightarrow \infty$: $R(k, X^{(m)}) = R(k, X)$.

Наиболее точным свойством самоподобных процессов является то, что АКФ не вырождается при $m \rightarrow \infty$, в отличие от стохастических процессов, где $R(k, X) \rightarrow 0$ при $m \rightarrow \infty$.

Анализ трафика сетей

Методика анализа трафика представляет собой процедуру агрегирования реализации трафика, доказан масштабная инвариантность процесса, проведение проверки независимости АКФ от уровня агрегирования, расчет параметра Хэрста различными методами.

За основу для исследования взята одна из реализаций трафика беспроводной сети, полученных в процессе мониторинга магистрального канала Интернет-провайдера.

Weth04.src – трафик уровня протокола Fast Ethernet, зафиксированный в точке, с направлением от беспроводных клиентов в сеть Интернет. Содержит информацию об Ethernet-кадрах и их размере. [2]

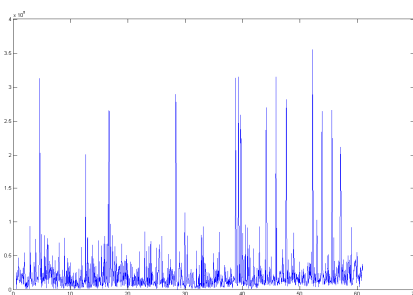
Данная реализация не соответствует Пуассоновскому распределению, т.к. не выполняется условие равенства математического ожидания и дисперсии для распределения количества моментов поступления пакетов, попавших в последовательность временных интервалов заданной длины, причем МО отличается от Д больше чем на порядок.

Mean.....: 141,9

Variance...: 69391

Std.....: 263,4

Рассмотрен анализ трафика на различных временных интервалах, а именно: на промежутке 60 сек, 2 часа, 7 часов. Была проведена процедура агрегирования с уровнем агрегирования $dT=0,05$ сек, которая заключается в разделении всего временного интервала на промежутки длиной dT , каждому из которых ставится в соответствие отношение



прошедшей за этот промежуток информации (байт) к его длине (сек). На рисунках 1-3 представлены графики агрегированной реализации трафика на заданных интервалах (по оси ОХ – время в сек, по оси ОУ – интенсивность передачи данных в байт/сек)

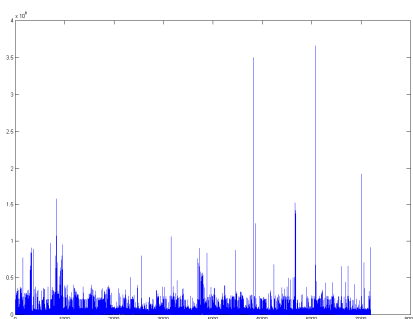
Рис 1. График реализации трафика, агрегированной по

уровню 0,05 (60 сек)

Характеристики:

Длительность реализации	1210 отсчетов по 0,05 сек, 60сек
Среднее	2,53e+4 байт/сек
СКО	4,167e+4 байт/сек
СКО/среднее	1,647

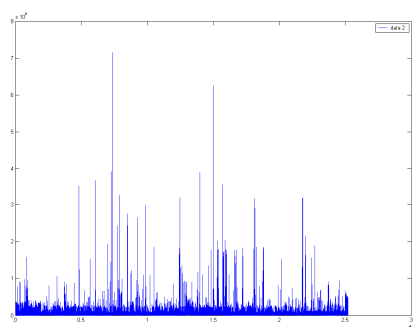
Максимум	3,553e+5 байт/сек
----------	-------------------



Длительность реализации	144010 отсчетов по 0,05 сек, 2 часа
Среднее	2,566e+4 байт/сек
СКО	5,123e+4 байт/сек
СКО/среднее	1,966
Максимум	3,656e+6 байт/сек

Характеристики:

Рис.2 – график реализации трафика, агрегированной по уровню 0,05 (7200 сек)



Характеристики:

Длительность реализации	504010 отсчетов по 0,05 сек, 7 часов
Среднее	2,741e+4 байт/сек
СКО	6,396e+4 байт/сек
СКО/среднее	2,33
Максимум	7,148e+6 байт/сек

Рис.3– график реализации трафика, агрегированной по уровню 0,05 (25200 сек)

Из приведенных графиков видно, что наблюдаемый процесс довольно неравномерный, и соответствует описанным выше признакам самоподобных процессов – наблюдается масштабная инвариантность, наличие всплесков.

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).

Выполним проверку независимости АКФ от уровня агрегирования. На рис. 4-5 – графики АКФ для реализации на интервале 60 сек с различным уровнем агрегирования.

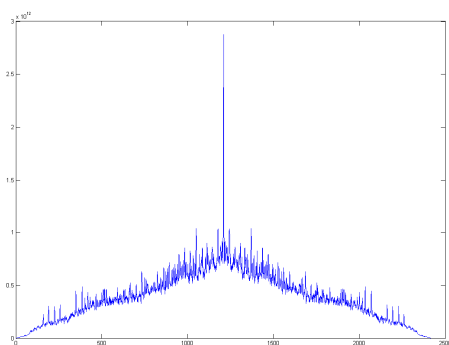


Рис.4– АКФ реализации трафика, агрегированной по уровню 0,05 (60 сек)

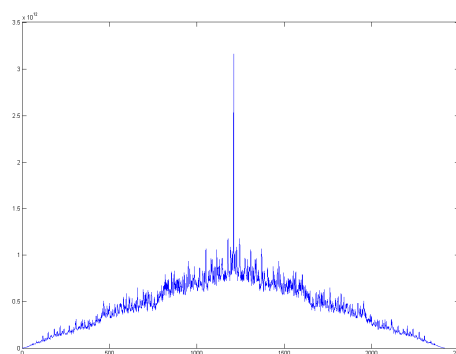


Рис.5 – АКФ реализации трафика, агрегированной по уровню 0,1 (60 сек)

Прослеживается достаточно медленное убывание функции. В то же время на графике АКФ отчетливо прослеживаются периодические пульсации. Также видно, что графики практически идентичны, что является еще одним признаком самоподобия рассматриваемого

процесса. При расчете параметра Хэрста данного трафика, видно что он лежит в пределах $H > 0,5$ –т.е. процесс является самоподобным.

Расчет параметра Хэрста проведен сначала для реализации трафика беспроводной сети.[5]

- 1)R/S Estimator Hurst Exponent Estimate :0.729 Correlation Coefficient: 97.73%
- 2)Aggregate Variance Estimator Hurst Exponent Estimate :0.825 Correlation Coefficient: 97.05%
- 3)Absolute Moments Estimator Hurst Exponent Estimate :0.705 Correlation Coefficient: 68.47%

Расчет параметра Хэрста для реализации трафика проводной сети.

Одна из изучаемых реализаций была получена Will Leland и Dan Wilson. Она названа авторами BC-Oct89Ext.TL и представляет собой зависимость от времени размеров сетевых кадров протокола второго уровня модели OSI (Ethernet-кадров в данном случае), проходящих через вычислительную сеть Bellcore Morristown Research and Engineering facility в глобальную сеть Internet. Реализация исследуется в работах [3] и [4]. Первый столбец представляет собой время в секундах. Второй длину поля данных Ethernet-кадра в байтах (не включая преамбулу, заголовка и контрольной суммы CRC).

BC-Oct89Ext4-пакеты между Bellcore и Интернет.[1]

- 1)R/S Estimator Hurst Exponent Estimate :0.867 Correlation Coefficient: 97.18%
- 2)Aggregate Variance Estimator Hurst Exponent Estimate :0.901 Correlation Coefficient: 65.21%
- 3)Absolute Moments Estimator Hurst Exponent Estimate :0.765 Correlation Coefficient: 57.59%

Выводы

Исследован поток, взятый из реализации трафика Интернет-провайдера. Данный поток не является простейшим (математическое ожидание не равно дисперсии), а является самоподобным процессом. В пакете Matlab построены графики реализации трафика на различных временных интервалах. Они отвечают свойствам самоподобных процессов-наблюдается масштабная инвариантность, наличие всплесков. Построены также автокорреляционные функции для разных уровней агрегирования, наблюдается инвариантность АКФ при изменении уровня агрегирования при условии медленно убывающей зависимости. Изучая результаты измерений показателя H методами Aggregate Variance, R/S, Absolute Moments Estimator можно отметить, что для всех реализаций сетевого трафика $H > 0.5$, то есть трафик относится к классу персистентных процессов.

Кроме исследования трафика беспроводной сети, также был рассмотрен трафик проводной сети компании Bellcore. АКФ второго имеет большие пульсации, однако также прослеживается достаточно медленное убывание функции. Сравнивая показатели Хэрста двух трафиков разных сетей, можно сделать вывод, что в беспроводной сети параметр Хэрста меньше, чем в проводной.

Проанализировав трафик сетей, можно выдвинуть требования к созданию модели имитации трафика сети. Использование традиционных методов для его оценки, основанных на пуассоновских моделях и формулах Эрланга, приводит к сильной недооценке реальной нагрузки компьютерной сети. Потому модель должна будет имитировать самоподобный трафик. Для представления его свойств наиболее подходят модели распределениями: Вейбулла и Парето.

Литература:

1. <http://ita.ee.lbl.gov/html/traces.html> Traces available in the Internet Traffic Archive
2. Петров В.В. Статистический анализ сетевого трафика,- М,2003
3. Н. J. Fowler, W. E. Leland. [Local Area Network Traffic Characteristics, with Implications for Broadband Network Congestion Management](#), *IEEE JSAC*, 9(7), September 1991, pp. 1139-1149.
4. W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the Self-Similar Nature of Ethernet Traffic (Extended Version), *IEEE/ACM Transactions on Networking*, 2(1), February 1994,pp. 1-15.
5. Selfis v0.1b. программа для анализа экспоненты Хэрста разработки Thomas Karagiannis: <http://www.cs.ucr.edu/~tkarag/>