

УДК 004.7

АНАЛИЗ СЕТЕВОГО ТРАФИКА МЕТОДОМ АГРЕГИРОВАНИЯ

Бельков Д.В., Едемская Е.Н.

Донецкий национальный технический университет
кафедра вычислительной математики и программирования

E-mail: belkov@telenet.dn.ua

Аннотация

Бельков Д.В., Едемская Е.Н. Анализ сетевого трафика методом агрегирования. В работе проведен анализ потока Ethernet-кадров методом агрегирования и определены фрактальные характеристики. Показано, что исследуемый трафик является масштабно-инвариантным.

Общая постановка проблемы

До недавнего времени теоретическую базу для проектирования систем распределения информации составляла теория массового обслуживания. Моделью потока вызовов (данных) в этой теории является простейший поток (стационарный ординарный поток без последствия). В 1993 году группа американских исследователей: W.Leland, M.Taqqu, W.Willinger и D.Wilson опубликовали результаты работы, которая в корне изменила представления о процессах, происходящих в телекоммуникационных сетях с коммутацией пакетов. Оказалось, что потоки в современных сетях нельзя аппроксимировать простейшими, поскольку они имеют иную структуру, чем принято в классической теории телетрафика. Было установлено, что трафик сети обладает свойством самоподобия (масштабной инвариантности), имеет память (последствие), а также обладает высокой пульсацией. По этой причине 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 с самоподобием. Статистические характеристики (среднее значение, спектральная плотность, автокорреляционная функция и др.) самоподобного трафика имеют характер спада сильно отличающийся от экспоненциального. Поэтому требуют корректировки исходные предпосылки, которые делались ранее при разработке многих сетевых устройств.

Несмотря на продолжительный период изучения проблемы самоподобия телетрафика, остается ряд нерешенных задач:

- фактически отсутствует строгая теоретическая база, которая пришла бы на смену классической теории массового обслуживания при проектировании современных систем распределения информации с самоподобным трафиком;
- нет единой общепризнанной модели самоподобного трафика;
- не существует достоверной и признанной методики расчета параметров и показателей качества систем распределения информации при влиянии эффекта самоподобия;
- отсутствуют алгоритмы и механизмы, обеспечивающие качество обслуживания в условиях самоподобного трафика [1].

Решение указанных задач имеет не только теоретическое, но и практическое значение. Большинство современных приложений являются синхронными и предъявляет высокие требования к качеству соединения. Сократить задержку передачи данных по сравнению с протоколом TCP, позволяет протокол без гарантированной доставки UDP. Однако, обеспечить повышенные требования к качеству соединения только с помощью

транспортного протокола (UDP или TCP) затруднительно, поскольку причины, приводящие к большим задержкам, большей частью находятся на сетевом уровне [2].

Свойство масштабной инвариантности сетевого трафика позволяет разработать алгоритмы прогнозирования, которые смогут посредством анализа трафика на относительно небольшом отрезке времени предсказать его поведение на более длительных интервалах. Используя такие прогнозы, можно будет создавать более эффективные методы управления пропускной способностью, что позволит сократить задержки передачи данных по сети и потери пакетов.

Цель и задачи работы

Ситуация, сложившаяся в современных глобальных компьютерных сетях, наличие большого количества сетевых маршрутов на которых периодически возникают резкие колебания задержки в передаче данных и большой процент потерь пакетов, появление новых свойств сетевого трафика, необходимость обеспечения высокого качества обслуживания различных категорий приложений, делают актуальной задачу исследования фрактального трафика.

Целью данной работы является определение фрактальных характеристик (H и D) временного ряда методом агрегирования, который может быть использован при исследовании сетевого трафика. Его применение показано на конкретном примере.

Задачи работы:

1. Сформулировать свойства фрактальных процессов;
2. Определить степень фрактальности трафика методом агрегирования;
3. Определить фрактальную размерность трафика методом агрегирования;

Анализ фрактальных свойств трафика позволяет выделить важные числовые характеристики, на основе которых могут быть построены адаптивные алгоритмы статистического управления и прогнозирования. Использование свойств самоподобия трафика может обеспечить достижение высокой степени масштабируемости прогноза сетевой нагрузки и получить оценки для широкого диапазона временных интервалов на основе результатов измерения ограниченного набора данных. Предсказание во временной области изменения значения потоков, проходящих через канал позволяет:

- 1) Динамически перераспределять размеры буферов, компенсируя недостаток пропускной способности, выделенный для низкоприоритетных данных.
- 2) Обеспечивать оптимальные временные характеристики, например задержку при буферизации.
- 3) Формировать локальные контуры управления перегрузкой буферов и каналов связи АТМ сети.
- 4) Дать возможность источникам, передающим данные с помощью транспортного протокола с обратной связью (TCP), получать прогнозы доступной пропускной способности сети и соответственно модифицировать стратегии предотвращения перегрузок.

Для этих задач можно использовать реализацию алгоритмов адаптации и управления с помощью самонастраивающихся сетевых компонент, получивших название интеллектуальных агентов. В этом случае передача данных и выработка управляющих воздействий не формируется в одном программном модуле, а локализуется в том месте сетевой инфраструктуры, где соответствующие задачи могут быть решены с привлечением минимальных сетевых ресурсов.

Математическое описание дискретного фрактального процесса

Фракталы - это структуры, которые, несмотря на свою крайнюю нерегулярность, на разных масштабах выглядят примерно одинаково. Мультифракталы - неоднородные фрактальные объекты для полного описания которых, в отличие от регулярных фракталов, недостаточно введения всего лишь одной фрактальной размерности, а необходим спектр

таких размерностей. Причина этого заключается в том, что наряду с чисто геометрическими характеристиками такие фракталы обладают и некоторыми статистическими свойствами.

Фрактальная размерность D временного ряда связана с показателем степени его фрактальности (показателем Херста) 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.

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: $D_m = D / m^\beta$, 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: $D_m = D / m^\beta$, 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 Херста H соотношением $\beta = 2(1 - H)$. АКФ сохраняется при $m \rightarrow \infty$: $R(k, X^{(m)}) \rightarrow R(k, X)$.

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 $m \rightarrow \infty$.

Метод агрегирования

Исходный ряд (1000 значений) является частью ряда BC-Oct89Ext4.TL [3]. Он показан на рисунке 1 и представляет собой зависимость от времени объемов Ethernet-кадров, проходящих через сеть Bellcore Morristown Research and Engineering facility в Internet. Минимальный и максимальный объемы кадров равны соответственно 64 и 594 байтам.

Для исходного ряда осуществлен следующий агрегационный процесс. Выполнено уменьшение размера шкалы наблюдений в 2 раза. Для этого сформирован новый ряд, полученный при помощи операции нахождения среднего каждых двух последовательных исходных наблюдений. Полученный ряд состоит из 500 событий. Произошло уменьшение рассматриваемой шкалы в 2 раза: каждое единичное деление новой шкалы содержит 2 единицы исходной. Затем аналогично выполнено уменьшение размера исходной шкалы наблюдений в m раз, для $m=4, 5, 10, 20, 25, 40, 50, 100, 200, 250$. Каждое деление новой шкалы содержит m единиц исходной. Структура полученных рядов подобна структуре исходного ряда.

Согласно определению самоподобного процесса, имеет место следующее соотношение дисперсий временных рядов:

$$D_{X^m} = \frac{D_X}{m^\beta} \quad (1)$$

Логарифмируя выражение (1), получим:

$$\ln(D_{X^m}) = \ln(D_X) - \beta \cdot \ln(m) \quad (2)$$

Поскольку $\ln(D_X)$ является константой, не зависящей от m , то график зависимости $\ln(D_{X^m})$ от $\ln(m)$ представляет собой прямую с наклоном, равным $(-\beta)$. Построив график

зависимости (2) и линию тренда, как показано на рисунке 2, определим аппроксимированное значение β : $\beta = 0,5138 \approx 0,514$. Учитывая, что параметр β связан с показателем Херста H как $H = 1 - \frac{\beta}{2}$, получим значение H : $H = 0,743$. Фрактальная размерность D временного ряда в таком случае: $D = 2 - H = 2 - 0,743 = 1,257$. Поскольку $H > 0,5$, то степень устойчивости долгосрочной зависимости исследуемого временного ряда выше среднего и ряд является самоподобным (фрактальным). На рисунке 3 видна крайне слабая зависимость параметра β от масштаба. Исследуемый ряд является масштабно-инвариантным.

Выводы

Многочисленные исследования, проведенные за последнее десятилетие учеными разных стран, свидетельствуют, что трафик современных компьютерных сетей обладает особой самоподобной структурой, не позволяющей использовать при проектировании традиционные методы, основанные на марковских моделях. Игнорирование особенностей трафика приводит к недооценке сетевой нагрузки. В самоподобном трафике присутствует много сильных выбросов на фоне низкого среднего уровня, что значительно увеличивает задержки, даже в случаях, когда средняя интенсивность трафика намного ниже максимальной скорости передачи в данном канале. Самоподобные процессы имеют длинную память, что позволяет предсказать их будущее, зная недавнее прошлое. Прогнозирование трафика важно для обеспечения качества обслуживания (QoS) [4].

В работе проведен анализ потока Ethernet-кадров [3] методом агрегирования и определены фрактальные характеристики. Показано, что исследуемый трафик является масштабно-инвариантным.

Литература

1. Петров В.В. Структура телетрафика и алгоритм обеспечения качества обслуживания при влиянии эффекта самоподобия. Автореферат диссертации. Москва. – 2005. – 20 с.
2. Иванов А.В. Разработка и исследование алгоритмов прогнозирования и управления очередями в компьютерных сетях. Автореферат диссертации. Санкт-Петербург.–2001. – 18 с.
3. Internet traffic archive. <http://ita.ee.lbl.gov/>
4. Величко В.В, Субботин Е.А., Шувалов В.П., Ярославцев А.Ф. Телекоммуникационные системы и сети. Том 3. Мультисервисные сети. Москва: Горячая линия-Телеком,2005.– 592 с.

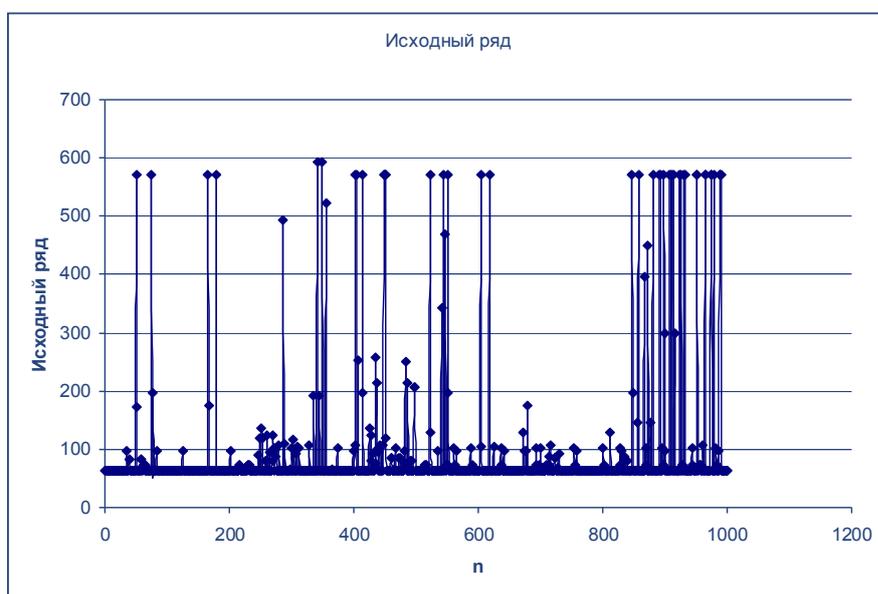


Рисунок 1.- Исходный ряд

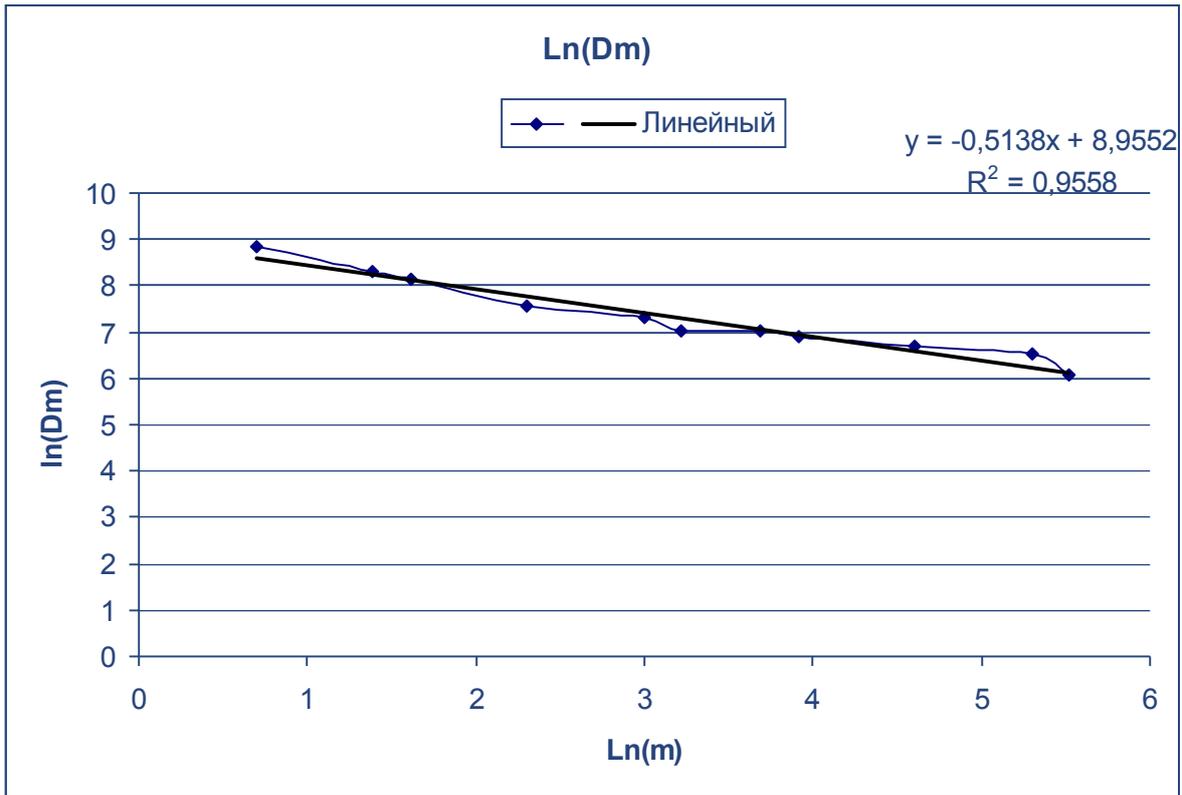


Рисунок 2.- Линия тренда для определения параметра β

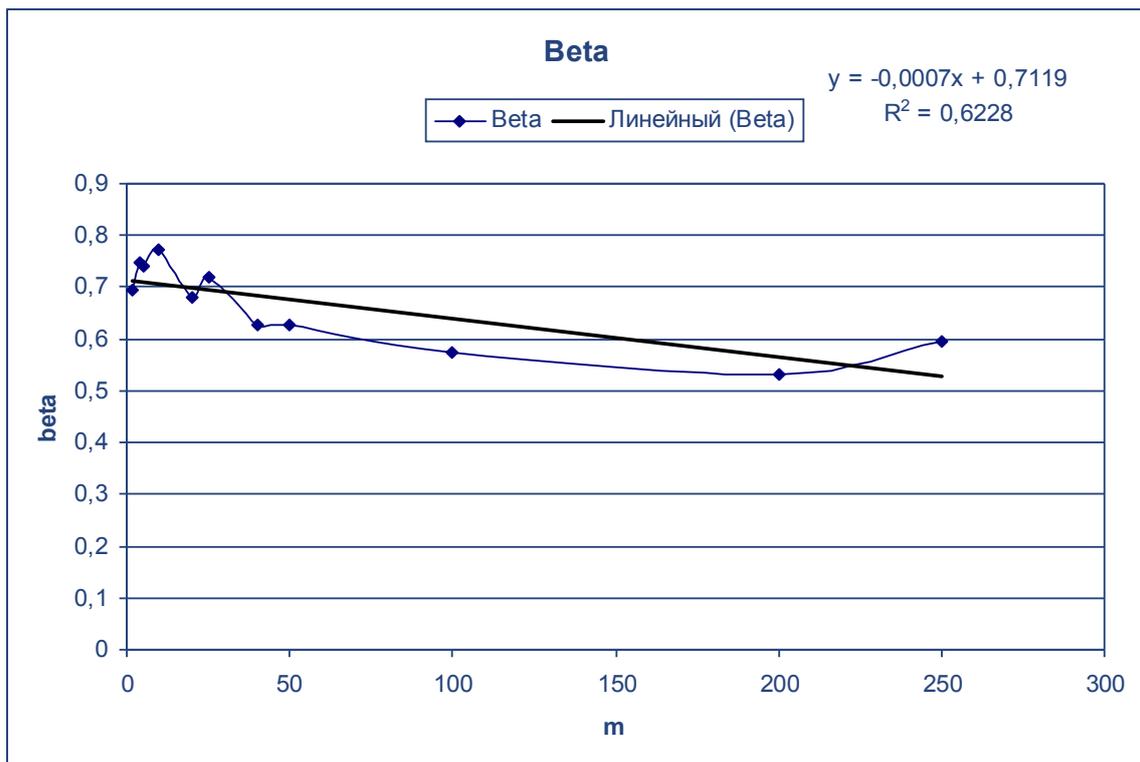


Рисунок 3.- Зависимость параметра β от масштаба