

ПРОТЯЖЕННЫЕ СТОХАСТИЧЕСКИЕ И ДИНАМИЧЕСКИЕ ПРОЦЕССЫ В КОМПЬЮТЕРНЫХ СЕТЯХ: МОДЕЛИ, МЕТОДЫ АНАЛИЗА ДЛЯ СИСТЕМ ЗАЩИТЫ ИНФОРМАЦИИ

Заборовский В.С.

ЦНИИ РТК

Введение. Компьютеры и компьютерные сети являются неотъемлемыми компонентами современных технологий управления и производства. В настоящее время их основные ресурсы - производительность процессоров и пропускная способность линий связи - даже для настольных компьютерных систем сравнима с характеристиками супервычислителей начала 90-х годов. Возросшие вычислительные и коммуникационные возможности стали предпосылкой для создания нового класса сетевых приложений: видеоконференций, IP-телефонии, анимации в реальном времени, распознавания голоса, распределенных вычислений, интерактивной графики, виртуальной реальности и др. Все из перечисленных выше приложений функционируют в сетевой среде и используют методы пакетной коммутации. Совокупность сетевых устройств, протоколов и средств управления составляют основу интегрированной среды обработки, хранения и передачи данных - распределенной инфраструктуры современных компьютерных телекоммуникаций. В результате парадигма современных компьютерных технологий, развиваясь из концепции "компьютер - это сеть", получила инверсную, хотя и менее лаконичную формулировку "сеть - это распределенное вычислительное устройство для обработки и передачи пакетного трафика". В такой системе взаимодействие между устройствами или приложениями осуществляется с помощью создания виртуальных соединений, на управление которыми заметное влияние оказывают особенности стохастической динамики процессов пакетной коммутации. [1].

1. Новые технологии управления в компьютерных сетях. Анализ результатов многочисленных экспериментов по измерению характеристик трафика показывает, что переход к интегрированным сетевым приложениям создает предпосылки для рассмотрения задач управления отдельными виртуальными соединениями с позиций теории динамических систем. И хотя уже более десяти лет технология динамической маршрутизации широко используется для передачи пакетного трафика, но до сих пор этот механизм управления рассматривается в отрыве от динамических процессов, возникающих при установлении транспортных соединений или управлении перегрузками. В результате сеть как объект управления испытывает влияние многочисленных регулирующих и возмущающих воздействий. В определенных режимах работы объекта отсутствие координации таких воздействий на транспортном и сетевом уровнях может приводить к существенному ухудшению характеристик функционирования информационных приложений.

В связи с этим особую актуальность приобретают вопросы разработки конструктивных методов исследования процессов в распределенных системах с целью анализа безопасности с учетом особенности динамики трафика в компьютерных сетях. В последнее время для этой цели стала широко применяться методы анализа фрактальных процессов и концепция активной сетевой среды. Ключевым звеном этой концепции является структура системы управления, построенная на прогнозировании состояния виртуальных соединений и использующая для этой цели специальные программные модули или интеллектуальные сетевые агенты, которые позволяют контролировать поведение пользователей. Таким образом, значительная часть задач

интеграции телекоммуникационных и информационных приложений в единую информационную инфраструктуру может быть решена на базе создания интеллектуальных компьютерных систем управления сетевыми ресурсами.

2. Статистическая динамика сетевых процессов. При разработки методов управления распределенными сетевыми приложениями необходимо учитывать то, что воздействия различных возмущающих факторов могут приводить к возникновению сложных стохастических или бифуркационных процессов. В качестве примера стохастической динамики процессов, возникающих при передаче пакетного трафика, на Рис.1 показан график изменения величин задержек в виртуальном канале, в том числе и для контроля их безопасности.

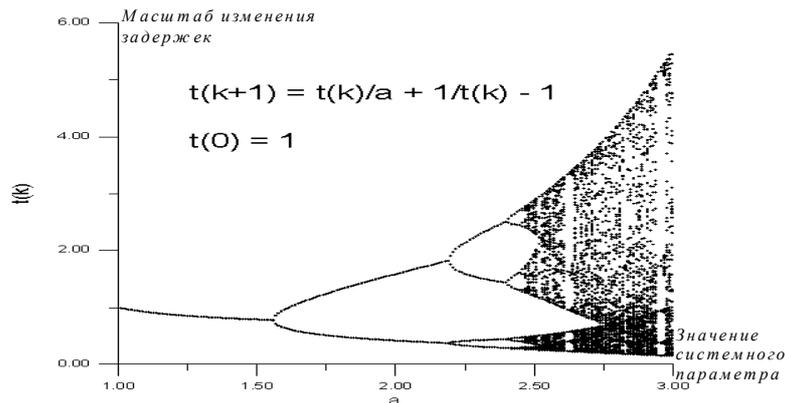


Рис.1. Бифуркации величин временных интервалов $t(k)$ между пакетами в виртуальном канале TCP на фазе медленного старта.

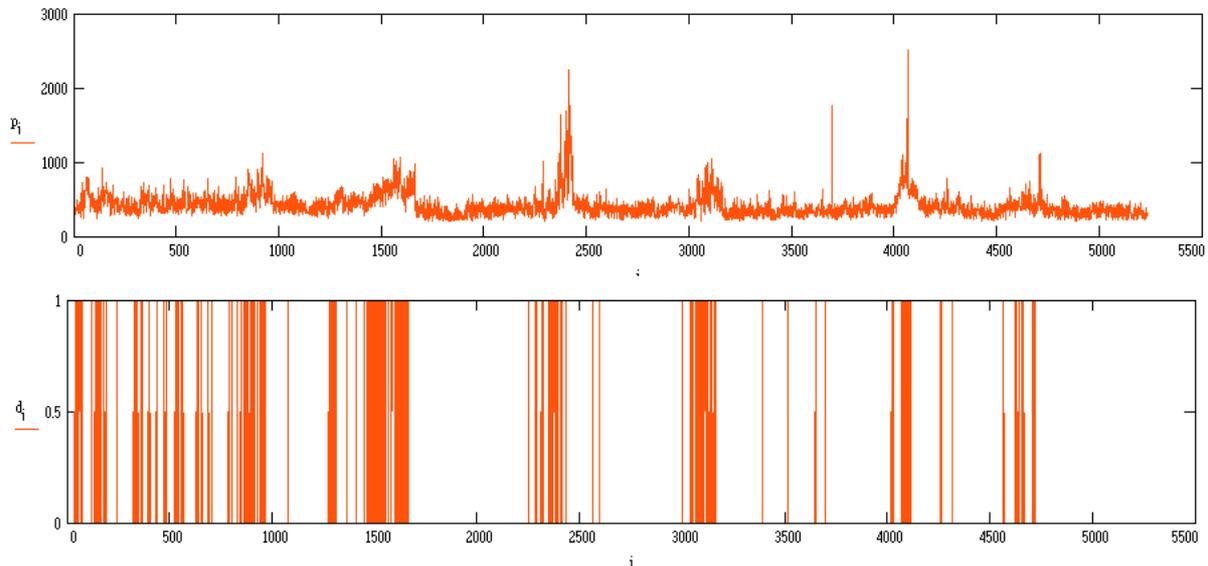


Рис 2. Характер изменения задержек при передаче данных в виртуальном канале Россия (С.-Петербург)-Канада. Апрель 1999 г:

- а) абсолютные значения интервалов времени между отправкой пакета и получением подтверждения в мсек;
- в) структура потока, превышающего пороговое значение 600 мсек в.

График изменения величин задержек в виртуальном канале при передаче трафика в одном из сегментов глобальной сети Интернет показан на Рис.2. Подобная

стохастическая динамика сетевых процессов обычно не принимается во внимание при разработке информационных приложений в локальных сетевых сегментах, где вариации параметров могут быть строго ограничены. Однако при создании глобальной инфраструктуры компьютерных телекоммуникаций методы прогнозирования состояния виртуальных соединений становятся весьма актуальными.

Широкие перспективы применения адаптивных транспортных протоколов для мультимедиа приложений стимулировали проведение большого комплекса научных исследований. С этими исследованиями связано появление новых представлений о процессах передачи данных в компьютерных сетях.

2. Влияние характеристик трафика на сетевую производительность. Построение модели функционирования сети Интернет представляет актуальную научную и практическую задачу. При этом интересно заметить, что такая сеть как технический объект характеризуется рядом свойств, которые не были итогом целенаправленного проектирования, а стали проявляться вследствие ее быстрого роста и территориальной распределенности. Многие особенности процессов в сети Интернет в настоящее время еще слабо изучены. При этом статистические характеристики трафика обладают свойством временной масштабной инвариантности, которое в современной научной литературе принято называть самоподобием. Данное свойство является следствием протяженных статистических зависимостей (ПЗ), проявляющихся в степенном характере затухания автокорреляционных функций временных отсчетов. Поэтому степень влияния ПЗ на сетевую производительность определяется тем, как успешно система управления виртуальным соединением может влиять на интенсивность генерации данных в зависимости от прогнозируемых изменений сетевых ресурсов. Поэтому уменьшение объема буферной памяти выделяемой для соединения требует значительно более точного прогнозирования состояния сетевых ресурсов. В противном случае возникновение локальных перегрузок приводит к фактическому разрыву виртуальных соединений.

На транспортном уровне существенное влияние на ПЗ сетевого трафика оказывает характер распределения времен прихода пакетов подтверждения. На рис.3 показана экспериментальная функция распределения времен подтверждения передачи пакетов t_{rtt} для трафика состоящего из 5000 пакетов.

$$\ln P(t_{rtt} = x)$$

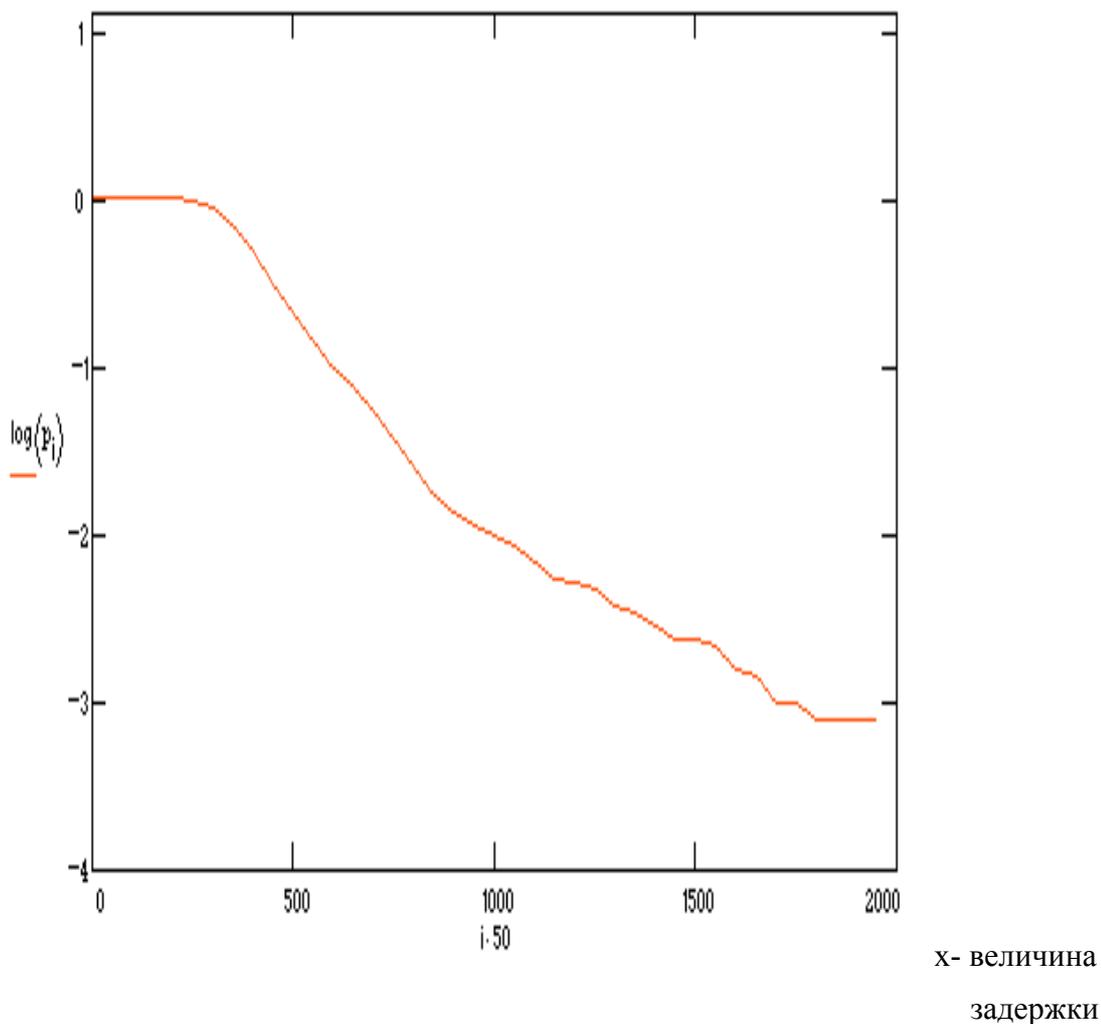


Рис.3 Экспериментально полученное распределение времен прихода пакетов подтверждения

Учет ПСЗ позволят построить модель статистической связанности процессов и, таким образом, обеспечить возможность адаптации характеристик транспортных протоколов к изменениям состояния сети. Поэтому, рассматривая транспортные протоколы как основной механизм перераспределения сетевых ресурсов между отдельными виртуальными соединениями, можно сделать вывод, что характер межсетевого взаимодействия в сети Интернет определяется ПСЗ, которые необходимо учитывать при синтезе нового поколения алгоритмов маршрутизации и методов статистического мультиплексирования.

3. Влияние динамической маршрутизации на характеристики соединений.

Особенности процессов ПСЗ могут быть использованы для изучения свойств транспортных протоколов, функционирующих в сетях с динамически изменяющейся топологией. В таких сетях разрыв логических соединений может быть вызван как переполнением промежуточных буферов, так и неисправности физического соединения. В силу того, что протоколы маршрутизации используют только локальную информацию о связанности узлов пакетной коммутации, то для них индикатором разрыва соединения является увеличение объема выходной очереди и ее возможное переполнение, если время перекоммутации виртуального соединения превышает определенное пороговое значение. В силу случайного характера рассматриваемых

событий их вероятностное моделирование должно отражать возможность появления подобных резких отклонений. Согласно выводам теории робастных статистик в этом случае вероятностное распределение анализируемых сигналов, например времен подтверждения, может иметь вид аддитивной смеси двух распределений $H(x,a)$ с различными наборами параметров a_1 и a_2

$$F(x) = (1-e) \cdot H(x,a_1) + e \cdot H(x,a_2), \quad (1)$$

где величина $(1-e)$ характеризует вероятность появления сигналов с параметрами a_1 , причем $0 < e \ll 1$. Если эти отклонения в k -ый момент обозначить T_k , то можно рассмотреть следующую модель подобных событий

$$T_k = T_0 + \Delta T_k, \quad (2)$$

где T_0 – время, затрачиваемое пакетом на прохождение по каналам связи; ΔT_k – случайная составляющая вариации задержек, асимптотическое распределение которой имеет вид степенной зависимости

$$1 - H(x,a) = P(\Delta T_k > x) \sim x^{-a} \text{ при } x \rightarrow \infty, \quad (3)$$

где $P(\cdot)$ символ вычисления вероятности события, $0 < a < 2$. Для ТСП протокола, поддерживающего упорядоченную доставку потока байт, индикаторами потерь наряду с задержками являются нарушения порядка прихода пакетов подтверждений, что непосредственно связано с функционированием протоколов динамической маршрутизации. Влияние подобных событий на сетевую производительность может быть существенно снижено с помощью введения робастных алгоритмов прогнозирования пропускной способности виртуальных каналов, когда оптимальные оценки вычисляются с использованием моделей (1) и (3). Разработка таких алгоритмов требует построения моделей ПСЗ процессов, параметры которых могут быть оценены в процессе функционирования сетевого приложения.

3. Математические модели ПСЗ процессов. В ряде исследований, выявивших наличие ПСЗ в трафике, были предложены различные модели основанные на обобщении свойств классического Броуновского движения. С помощью концепции обобщенного Броуновского движения $V_H(t)$ удастся объединить различные аспекты динамического и стохастического поведения сложных объектов с помощью весьма простых и наглядных математических моделей. Обычно $V_H(t)$ рассматривается как функция одной переменной t , приращение которой

$$\Delta V_H(t) = V_H(t_2) - V_H(t_1)$$

имеет Гауссовское распределение с нулевым математическим ожиданием и асимптотической дисперсией

$$M\{\Delta V_H^2(t)\} \sim \Delta t^{2H}, \quad (4)$$

где $M(\cdot)$ – символ математического ожидания, $\Delta t = |t_2 - t_1|$ – временной интервал; H – параметр, принимающий значение $0 < H < 1$. При $H = 0.5$ $V_H(t)$ является классическим Броуновским движением, приращение которого соответствует некоррелированному Гауссовскому белому шуму. При $H > 0.5$ приращение имеет положительную, а при $H < 0.5$ – отрицательную корреляцию. Для моделирования сложных сетевых процессов существенно то, что такая корреляционная связанность значений процессов сохраняется на произвольно длинном интервале времени. Действительно, рассмотрим два неперекрывающихся временных интервала $(-t, t_0), (t_0, t)$. Нормированная корреляционная функция при $V_H(t_0) = 0$ имеет вид

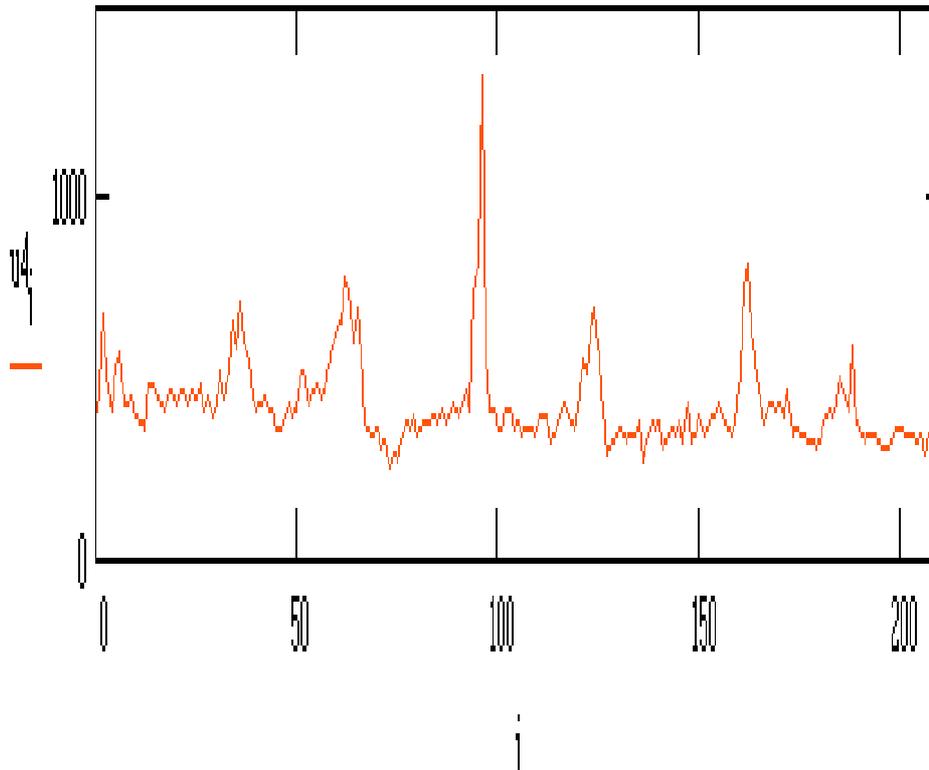
$$r_H(t) = \frac{M\{-B_H(-t)B_H(t)\}}{M\{B_H^2(t)\}} = \frac{M\{[-B_H(-t) + B_H(t) - B_H(t)][B_H(t) + B_H(-t) - B_H(-t)]\}}{M\{B_H^2(t)\}}. \quad (5)$$

С учетом (4), получим

$$r_H(t) = 2^{2H-1} - 1. \quad (6)$$

Выражение для $r_H(t)$ не зависит явно от времени, поэтому рассматриваемый процесс $B_H(t)$ является ПСЗ процессом. Экспериментальным подтверждением статистически протяженного характера сетевых процессов может служить график (Рис.5) из 200 отсчетов усредненного в масштабе М 25:1 потока данных, снятого на сегменте глобальной сети Интернет.

Задержка (мсек)



отчеты значений

Рис.4 Агрегированный трафик М 25:1, состоящий из 5000 отсчетов времен подтверждения прихода пакетов.

В агрегированном трафике отчетливо выделяется случайная составляющая резких отклонений, корреляционная связанность которых сохраняется на весьма продолжительном интервале наблюдения.

4. Модели сетевых процессов на основе операции дробного интегрирования.

Вопрос о построении моделей динамических объектов и технических систем с помощью операций дробного интегрирования (дифференцирования) является предметом широких дискуссий. Поиск убедительных примеров эффективности использования таких моделей для описания протяженных зависимостей непосредственно связан с возможностью применения методов фрактального анализа физических процессов для построения нового класса технических систем [2]. С формальной точки зрения речь идет о возможности создания специального типа

линейных фильтров, использующих операцию дробного интегрирования для вычисления операции свертки входных сигналов.

Для случая сетевых приложений естественным способом введение операции дробного интегрирования является рассмотрение процесса прохождения пакетного трафика через каскад промежуточных сетевых устройств с буферными накопителями в форме интегрирующих звеньев (рис.5).

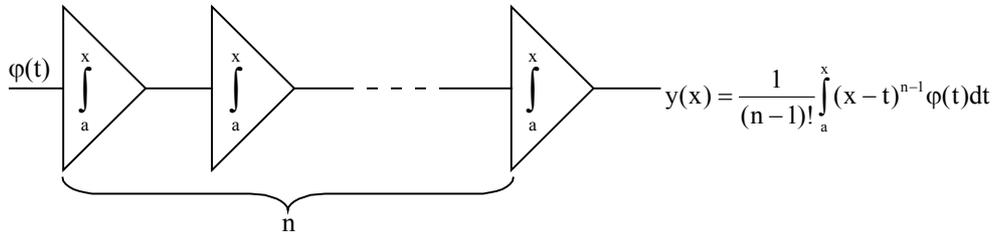


Рис.5. Преобразование трафика при прохождении через каскад интегрирующих звеньев.

Для \$n\$-кратного интеграла известна формула Коши

$$y(x) = \int_a^x dx \int_a^x dx \dots \int_a^x \varphi(x) dx = \frac{1}{(n-1)!} \int_a^x (x-t)^{n-1} \varphi(t) dt, \quad \text{где } n-1 > 0. \quad (7)$$

Так как \$(n-1)! = \Gamma(n)\$, то правой части (4) можно придать смысл и при нецелых значениях \$n\$. Тогда левосторонним (это означает, что \$x > a\$) интегралом дробного порядка \$\alpha > 0\$ называется выражение

$$(I_{a+}^{\alpha} \varphi)(x) = \frac{1}{\Gamma(\alpha)} \int_a^x (x-t)^{\alpha-1} \varphi(t) dt. \quad (8)$$

Операцию дробного дифференцирования порядка \$\alpha < 1\$ естественно ввести как операцию, обратную дробному интегрированию:

$$(D_{a+}^{\alpha} f)(x) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_a^x \frac{f(t) dt}{(x-t)^{\alpha}}. \quad (9)$$

Можно показать, что справедливо обобщенное представление операций дробного дифференцирования

$$(D_{a+}^{\alpha} f)(x) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dx} \right)^{n-x} \int_a^x \frac{f(t) dt}{(x-t)^{\alpha-n+1}},$$

которое может применяться для случая \$\alpha > 1\$, где \$n = [\alpha] + 1\$.

Решение уравнения

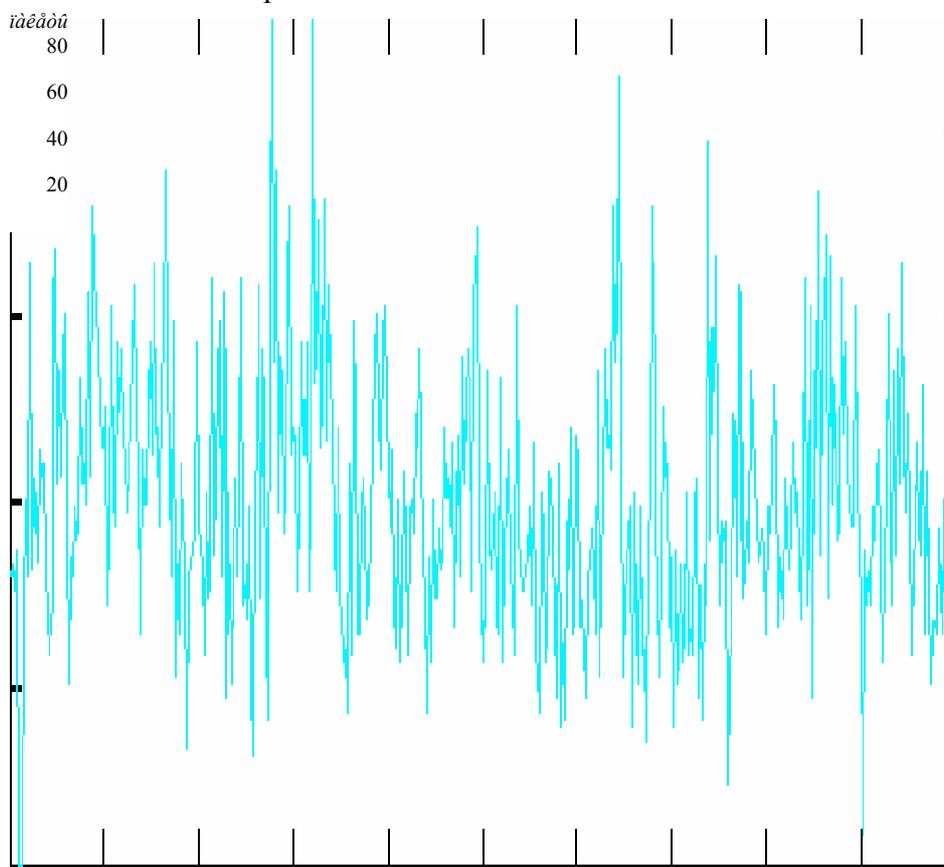
$$(I_a^{\alpha} \varphi)(x) = f(x), \quad x > a, \quad \alpha \in [0, 1] \quad (10)$$

то есть нахождение \$\varphi(x)\$ по заданной функции \$f(x)\$, используется в ряде прикладных задач, связанных с переносом процесса \$\varphi(x)\$ с помощью процесса \$f(x)\$ []. В этом случае решение (5) может быть записано в форме дифференциального уравнения с дробным показателем.

$$\varphi(x) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_a^x \frac{f(t) dt}{(x-t)^{\alpha}}. \quad (11)$$

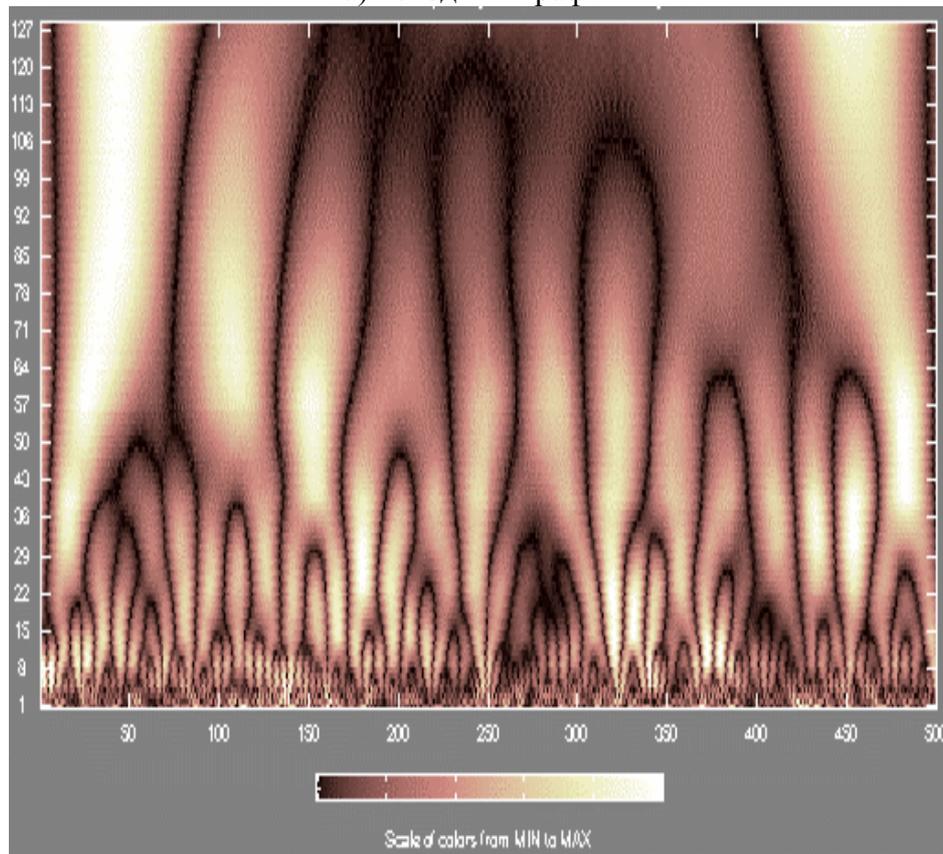
Такое представление решения позволяет учесть изменение скорости передачи трафика, связанные с вариациями задержек в буферах промежуточных сетевых устройств. Интенсивность этих вариаций может быть оценена с помощью асимптотической характеристики затухания функции распределения (3).

5. Применение теории всплесков для анализа свойств трафика. Преобразование всплесков вейвлет-анализ [3] для одномерного сигнала состоит из его разложения по иерархическому базису, сконструированному из солитоноподобных функций с помощью масштабных преобразований и переносов. Каждая из функций базиса характеризует как определенную частотную составляющую сигнала, так и локализацию этой составляющей в пространстве (времени), где этот сигнал рассматривается. Таким образом, в отличие от традиционного преобразования Фурье обеспечивается двумерная развертка исследуемого одномерного сигнала, а частота и координата рассматриваются как независимые переменные. Такой подход дает возможность анализировать свойства сигнала одновременно во временном и частотном пространствах. Полученный пространственно-временной спектр является важной характеристикой процессов в сложных нелинейных динамических системах, для которых имеет место взаимодействие различных возмущающих факторов в широких диапазонах пространственно-временных частот. Это позволяет выявить пространственно распределенные свойства сигналов, получить локальную высокочастотную и глобальную крупномасштабную информацию об исследуемом объекте, что особенно важно для оптимизации процессов управления в сети Интернет. В этом случае преобразование всплесков обнаруживает внутреннюю структуру существенно неоднородных потоков данных (трафика) и позволяет обнаружить масштабно инвариантные свойства применительно к их интегральным или статистическим характеристикам. Применительно к сетевым процессам для целей прогнозирования важное значение имеет свойство масштабной инвариантности трафика. В результате по оценкам, полученным на весьма коротких интервалах времени и с учетом параметров самоподобия можно осуществить прогноз состояния сетевого соединения на большие временные интервалы. Структура коэффициентов преобразования всплесков представлена на Рис.6



0 50 100 150 200 250 300 350
400 450 500 сек

а) исходный трафик



б) уровни изменения значений коэффициентов преобразования.

Рис.6 Преобразование всплесков сетевого:

а) исходный процесс; б) коэффициенты преобразования.

Заключение. Важной особенностью современных компьютерных систем является распределенность функций управления и недетерминированный характер связей между отдельными компонентами. В результате сеть приобретает черты сложного динамического объекта, характеристики которого подвержены случайным, но сильно коррелированным возмущениям, для прогнозирования которых требуется привлечение методов фрактального анализа.

ЛИТЕРАТУРА:

- 1.Zaborovsky "High-Speed Network Traffic Management: Automatic Control Approach", Proc. of INET'97, June, 1997, pp.107-108 (<http://www.neva.ru/~conference>).
- 2.Чукбар К.В. Стохастический перенос и дробные производные. ЖЭТФ. 108. Вып.5(11). с.185-1884.
- 3.Астафьева Н.М. Вейвлет-анализ: основы теории и примеры применения. УФН.166, 11 с.1145-1170.