

УДК 621.396.96

МУЛЬТИПЛИКАТИВНЫЕ МУЛЬТИФРАКТАЛЬНЫЕ ПРОЦЕССЫ В МОДЕЛИРОВАНИИ СЕТЕВОГО ТРАФИКА

А. Р. Бестугин,

канд. техн. наук, доцент

Санкт-Петербургский государственный университет аэрокосмического приборостроения

А. Ф. Богданова,

канд. техн. наук, доцент

Военно-космическая академия имени А. Ф. Можайского

Г. В. Стогов,

доктор техн. наук, профессор

Санкт-Петербургский государственный университет аэрокосмического приборостроения

Статистические свойства широкополосного сетевого трафика являются самоподобными в широком диапазоне временных шкал. Потоки с такими свойствами показывают сложную временную корреляцию, которая может быть охарактеризована долговременной зависимостью (LRD) или кратковременной зависимостью (SRD). Сетевое управление и администрирование зависят от характеристик таких процессов, которые моделируются как мультипликативные мультифрактальные процессы.

Statistical properties of broadband network traffic are self-similar across a wide range of time scales. Flows with such properties exhibit complex temporal correlation which can be characterized by the long-range dependence (LRD) or short-range dependence (SRD). Network control and management critically depends on the features of such processes that are modelled as multiplicative multifractal processes.

Внедрение новых информационных технологий, под которыми понимается совокупность систематических и массовых способов создания, накопления, обработки, хранения и распределения информации (данных, знаний) с помощью средств вычислительной техники и связи [1], требует разработки методов и алгоритмов управления информационными потоками, которые до сих пор традиционно проводились на основе пуассоновских (или, более обобщенно, марковских) предположений относительно структуры трафиковых прибытий и экспоненциальных предположений относительно требований удержания ресурса.

Высококачественные измерения, проводимые на высокоскоростных сетях связи, показали, что нагрузка современных высокоскоростных сетей обладает «фрактальной» природой, которая характеризуется такими свойствами, как самоподобие, пачечность, долговременная зависимость (Long-range dependence, LRD).

Процесс с LRD имеет автокорреляционную функцию $r(k) \sim k^{-\beta}$, когда $k \rightarrow \infty$, где $0 < \beta < 1$. Таким образом, автокорреляционная функция этого процесса следует степенному закону, в отличие от экспоненциального спада, показываемого традиционными моделями трафика (с кратковременной зависимостью – Short-range dependence, SRD). Спад по степенному закону медленнее, чем экспоненциальный спад, а так как $\beta < 1$, то ряд, образованный последовательными значениями коэффициента автокорреляции, расходится – $\sum_{k=1}^{\infty} r(k) = \infty$.

Исследования технического влияния самоподобия показали, что самоподобие может влиять на такие области, как сетевые измерения, качество обслуживания, управление перегрузкой, управление вызовом, мониторинг ошибок и другие, причем степень этого влияния изменяется в широком диапазоне.

Исследования технического влияния самоподобия показали, что самоподобие может влиять на такие области, как сетевые измерения, качество обслуживания, управление перегрузкой, управление вызовом, мониторинг ошибок и другие, причем степень этого влияния изменяется в широком диапазоне.

Учет этих свойств требует создания новых моделей сетевого трафика, которые в наибольшей степени отвечали бы реальным процессам, протекающим в реальных сетях. Такими моделями являются фрактальное броуновское движение (ФБД) и фрактальный гауссовский шум (ФГШ). Хотя ФБД и ФГШ относятся к мощным и легко обрабатываемым сигнальным моделям, они имеют значительные ограничения для моделирования некоторых типов процессов. Например, много процессов, включая сетевой трафик, обладают значительной долговременной зависимостью, но в то же время отображают кратковременные корреляции, а также масштабное поведение, несовместимое со строгим самоподобием. Многие процессы имеют инкременты, которые по своей природе положительные и, следовательно, негауссовские. Сигналы с такими свойствами, естественно, попадают в класс мультифрактальных процессов. Чтобы рассмотреть этот новый феномен, который связан с динамиками измеренного сетевого трафика на малых временных шкалах, рассматривается класс мультипликативных процессов, так называемых каскадов. Каскадом (или мультипликативным процессом) называется процесс, который дробит (фрагментирует) данный ряд на все меньшие и меньшие части в соответствии с некоторым геометрическим правилом и в то же самое время делит меру этих частей в соответствии с другим (возможно, случайным) правилом [3].

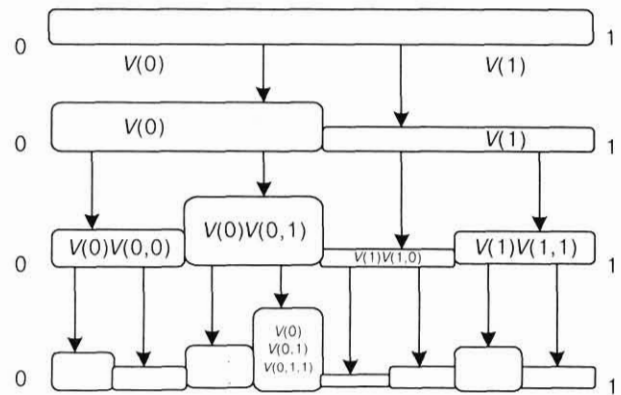
Обычно рассматриваются основная каскадная конструкция, называемая случайным каскадом, а также несколько других важных классов каскадов, включая хорошо изученные детерминированные и консервативные каскады [3, 4].

Конструкцию случайного каскада можно построить следующим образом. Воспользуемся механической интерпретацией, используемой в теории вероятности и математической статистике. Для этого рассмотрим единичный интервал $I = [0, 1]$. Свяжем его с некоей массой M , равномерно распределенной на единичном интервале. Единичный интервал делится на ряд $c, c^2, \dots, c^l, \dots$ малых под-интервалов. Каждый под-интервал l -й стадии делится на c подинтервалов, чтобы сформировать $(l + 1)$ -ю стадию. Обозначим интервалы, сгенерированные этим процессом построения:

$$I(j_1, \dots, j_l) = \left[\sum_{k=1}^l j_k c^{-k}, \sum_{k=1}^l j_k c^{-k} + c^{-l} \right],$$

для $j_k \in \{0, \dots, c - 1\}$ и $l = 1, 2, \dots$

Массу $MV(0), \dots, MV(c - 1)$ отнесем к подинтервалам первой стадии, где V – неотрицательная случайная переменная со средним $1/c$, а случайные переменные $V(0), \dots, V(c - 1)$ – независимые и имеют то же самое распределение, как V , генератор случайного каскада. Итерация этой процедуры генерирует последовательность случайных переменных $V(j_1), \dots, V(j_1, \dots, j_l)$, которые независимы и идентично распределенные, как генератор V . Пример конструкции со случайными каскадами при $M = 1$ и $c = 2$ показан на рис 1.



■ Рис. 1. Каскадная конструкция

Мера c -адического интервала $I(j_1, \dots, j_l)$, полученного такой конструкцией со случайными каскадами на стадии l , задается следующим образом:

$$\mu_l(I(j_1, \dots, j_l)) = MV(j_1)V(j_1, j_2) \dots V(j_1, \dots, j_l).$$

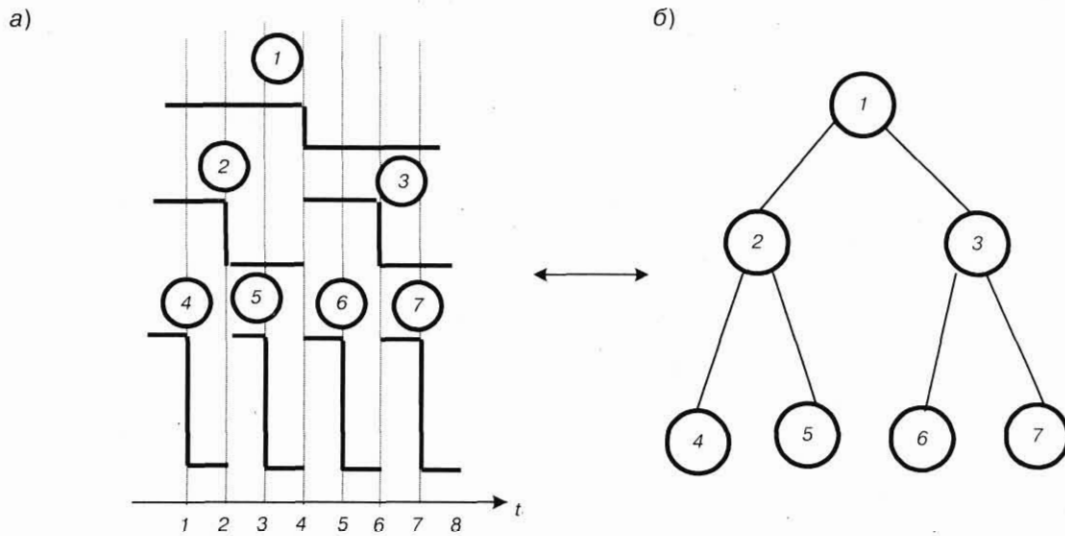
Заметим, что случайный каскад сохраняет массу только в смысле математического ожидания, т. е. для всех $l \geq 1$ мы имеем $E[\mu_l(I)] = M$.

Для иллюстрации детерминированного каскада выбираем для простоты $c = 2$. Детерминированный каскад – это специальный случай случайного каскада в том смысле, что в каждом подинтервале используется фиксированное умножение первичной массы (независимо от стадии каскадной конструкции). На первой стадии конструкции выбирается фиксированное $p \in (0, 1/2)$, и определяется масса M_p в левом интервале $I(0)$ и масса $M(1 - p)$ в правом интервале $I(1)$. При итерации этого процесса масса диадического интервала $I(j_1, \dots, j_l)$ при l -й стадии является l -множителем p и $(1 - p)s$, т. е.

$$\mu(I(j_1, \dots, j_l)) = M \left(\prod_{k=1}^l p_{j_k} \right) = Mp^{n_0}(1 - p)^{n_1},$$

где $j_k \in \{0, 1\}$, $p_0 = p$, $p_1 = 1 - p$ и где n_0 есть число нулей в (j_1, \dots, j_l) и $n_1 = l - n_0$ есть число единиц. Этот процесс сохраняет первоначальную массу M на каждой стадии.

Для удобства рассмотрения консервативных каскадов так же будем рассматривать случай для $c = 2$. В то время как случайные каскады могут быть подходящими физическими моделями для турбулентности, они не подходят для моделирования в контексте телекоммуникаций (также не подходят и детерминированные каскады). Поэтому необходим компромисс между высокогибким случайным каскадом и негибким детерминированным каскадом. Это требует свойства сохранения массы (детерминированного каскада) и полного случайного выбора (случайного каскада). Чтобы объединить эти две конкурирующие цели, было определено полуслучайное (или консервативное) правило, которое относит массу MW к интервалу $I(0)$ и массу $M(1 - W)$ к интервалу $I(1)$. Генератор W есть слу-



■ Рис. 2. Базисные функции с использованием вейвлета Хаара (а) и бинарное дерево, используемое для индексации случайного процесса при множестве разрешений (б)

чаянная переменная со средним 1/2, принимающая значения в (0, 1), и симметричная около своего среднего. Для итерации этого процесса рассмотрим последовательность случайных переменных $W(j_1, \dots, j_l)$, $l \geq 1$, со структурой, заданной как

$$W(j_1, \dots, j_{l-1}, 1) = 1 - W(j_1, \dots, j_{l-1}, 0),$$

где случайные переменные $W(j_1, \dots, j_{l-1}, 0)$ и $W(j_1, \dots, j_{l-1}, 1) = 1 - W(j_1, \dots, j_{l-1}, 0)$ идентично распределены как W .

Этот процесс строит консервативный каскад и последовательность мер μ_l . Для всех $l \geq 1$ мера диадического интервала $I(j_1, \dots, j_l)$ задаётся как

$$\mu_l(I(j_1, \dots, j_l)) = MW(j_1)W(j_1, j_2) \dots W(j_1, \dots, j_l)$$

и из-за своей мультипликативной структуры $\mu_l(I(j_1, \dots, j_l))$ – аппроксимативно логнормальна. Таким образом, для всех $l \geq 1$ имеем $\mu_l(I) = M$.

Главное различие между случайным и консервативным каскадами заключается в способе, которым при конструировании процесса происходит распределение массы на каждой стадии из первичного интервала к левому и правому подинтервалам.

Мультифрактальная природа сетевого трафика может быть более детально исследована с использованием анализа, основанного на вейвлетах.

Дискретное вейвлетное преобразование представляет собой одномерный реальный сигнал $X(t)$ в терминах сдвинутых и растянутых вариантов полосовой вейвлетной функции $\psi(t)$ и сдвинутого варианта низкочастотной масштабной функции $\phi(t)$. Используя специально выбранные вейвлетные и масштабные функции

$$\psi_{j,k}(t) := 2^{j/2} \psi(2^j t - k);$$

$$\phi_{j,k} := 2^{j/2} \phi(2^j t - k), \quad j, k \in \mathbb{Z},$$

можно сформировать ортогональный базис и получить представление сигнала в виде [5]

$$X(t) = \sum_k U_{J_0,k} \phi_{J_0,k}(t) + \sum_{j=J_0}^{\infty} \sum_k W_{j,k} \psi_{j,k}(t),$$

где $W_{j,k} = \int X(t) \psi_{j,k}(t) dt$; $U_{j,k} = \int X(t) \phi_{j,k}(t) dt$.

Для вейвлеты $\psi(t)$, сосредоточенной по времени и частоте f_0 , вейвлетный коэффициент $W_{j,k}$ измеряет содержание сигнала около времени $2^{-j}k$ и частоты $2^{-j}f_0$. В вейвлетном преобразовании j указывает масштаб анализа: J_0 – грубейшая шкала или наименьшее разрешение анализа, а наибольшее значение j соответствует наибольшему разрешению анализа. Хааровские масштабные и вейвлетные функции обеспечивают простейший пример ортогонального вейвлетного базиса. Материнская вейвлета Хаара имеет вид:

$$\phi(t) = \begin{cases} 1, & \text{если } 0 \leq t < 1/2, \\ -1, & \text{если } 1/2 \leq t < 1, \\ 0 & \text{– в остальных случаях.} \end{cases}$$

На рис. 2,а показан пример вейвлета Хаара для $K = 3$, на рис. 2,б – соответствующее дерево графа [5].

В области вейвлет моделирование положительных процессов осложняется тем, что ограничения на вейвлетные коэффициенты, необходимые, чтобы гарантировать положительный выход, нетривиальны. Однако при использовании вейвлета Хаара значения коэффициентов масштабного и вейвлетного преобразований могут быть рекурсивно вычислены при помощи следующих выражений:

$$U_{j,k} = 2^{-1/2}(U_{j+1,2k} + U_{j+1,2k+1}); \quad (1)$$

$$W_{j,k} = 2^{-1/2}(U_{j+1,2k} - U_{j+1,2k+1}). \quad (2)$$

Более того, известно, что в преобразовании Хаара положительных данных все $U_{j,k} \geq 0$, так как каждое $U_{j,k}$ равно масштабированному локальному среднему. Представим выражения (1) и (2) в следующем виде:

$$U_{j+1,2k} = 2^{-1/2}(U_{j,k} + W_{j,k});$$

$$U_{j+1,2k+1} = 2^{-1/2}(U_{j,k} - U_{j,k}).$$

Таким образом, можно найти простое ограничение для гарантии того, что процесс является положительным:

$$|W_{j,k}| \leq U_{j,k}. \quad (3)$$

Условие (3) является необходимым и достаточным.

Допустим, что мы хотим построить статистическую модель, для которой автоматически выполнялось бы условие (3). Это приводит нас к простой мультипликативной сигнальной модели. Пусть $A_{j,k}$ – случайная переменная на интервале $[-1, 1]$. Определим вейвлетные коэффициенты как

$$W_{j,k} = A_{j,k}U_{j,k}. \quad (4)$$

Мультифрактальная вейвлетная модель состоит из хааровского вейвлетного преобразования и структурного ограничения (4). Процедура синтеза мультифрактальной модели может быть представлена следующим образом [6].

1. Пусть $j = 0$. Зафиксируем или вычислим наигрубейший (корневой) масштабный коэффициент $U_{0,0}$.

2. На шкале j генерируем случайные множители $A_{j,k}$ и вычисляем каждые $W_{j,k}$ посредством (4) для $k = 0, \dots, 2^j - 1$.

3. На шкале j используем $U_{j,k}$ и $W_{j,k}$ из выражения (3), чтобы вычислить $U_{j+1,2k}$ и $W_{j+1,2k+1}$, масштабные коэффициенты на шкале $j + 1$ для $k = 0, \dots, 2^j - 1$.

4. Итерируем шаги 2) и 3), заменяя j на $j + 1$ до тех пор, пока не будет достигнута самая мелкая шкала $j = n$.

Так как масштабные коэффициенты генерировались одновременно с вейвлетными коэффициентами, нет необходимости инвертировать вейвлетное преобразование. Из-за простой структуры преобразования Хаара шаги 2) и 3) могут быть скомбинированы, полностью исключая вейвлетные коэффициенты:

$$U_{j+1,2k} = \left(\frac{1 + A_{j,k}}{\sqrt{2}} \right) U_{j,k}$$

и

$$U_{j+1,2k+1} = \left(\frac{1 - A_{j,k}}{\sqrt{2}} \right) U_{j,k}.$$

Литература

1. **Мичин И. А., Синицын И. Н., Доступов Б. Г.** и др. Развитие определений «информатика» и «информационные технологии» // Системы и средства информатики. Вып. 4. – М.: Наука, ИПИ АН, 1991. – 21с. (Препринт).
2. **Нейман В. И.** Самоподобные процессы и их применение в теории телетрафика // Тр. междунар. академии связи. – 1999. – № 1. – С. 11–15.
3. **Gilbert A. C., Willinger W., Feldman A.** Scaling analysis of conservative cascades, with application to network traffic // IEEE transaction on information theory. – 1999. – Vol.45 – No.3. – P. 971–991.
4. **Gao J, Rubin I.** Statistical properties of multiplicative multifractal processes in modelling telecommunications traffic streams // Electronics letters. – 2000. – Vol.36. – No.1, P. 101–102.
5. **Gao J, Rubin I.** Superposition of multiplicative multifractal traffic processes // Electronics letters. – 2000. – Vol.36. – No.8. – P. 761–762.
6. **Ma S., Ji C.** Modelling heterogeneous network traffic in wavelet domain // IEEE/ACM Transaction on networking. – 2001. – Vol.9. – No.5. – P. 634–649.