КОМПЛЕКСНЫЙ ПОДХОД К ИССЛЕДОВАНИЮ ФРАКТАЛЬНЫХ ВРЕМЕННЫХ РЯДОВ

Людмила Кириченко, Лариса Чалая

Аннотация: *В работе предложен комплексный поход к анализу фрактальных свойств самоподобных случайных процессов по временным рядам небольшой длины. Приведена последовательность этапов проведения фрактального анализа. Этими этапами являются: предварительный анализ, включающий удаление краткосрочной зависимости и выявление истинной долгосрочной зависимости; проверка гипотезы о наличии свойства самоподобия; несмещенное интервальное оценивание показателя Херста в случаях стационарных и нестационарных временных рядов несколькими методами; уточнение полученной оценки показателя Херста.*

Ключевые слова: *самоподобный стохастический процесс, временной ряд, показатель Херста, методы оценивания показателя Херста.*

ACM Classification Keywords: *G.3 Probability and statistics - Time series analysis, Stochastic processes, G.1 Numerical analysis, G.1.2 Approximation - Wavelets and fractals.*

Введение

Задачи современной нелинейной физики, радиоэлектроники, теории управления, обработки изображений, требуют разработки и применения новых математических моделей, методов и алгоритмического обеспечения анализа данных. В настоящее время стало общепризнанным, что многие стохастические процессы в природе и технике обладают долгосрочной зависимостью и фрактальной структурой. Наиболее адекватным математическим аппаратом для исследования динамики и структуры таких процессов является фрактальный анализ.

Самоподобие случайных процессов заключается в сохранении вероятностных характеристик при изменении масштаба времени. Стохастический процесс $X(t)$ является самоподобным с параметром H, если процесс $a^{-H}X(at)$ описывается теми же законами конечномерных распределений, что и $X(t)$:

$$\text{Law}\{a^{-H}X(at)\} = \text{Law}\{X(t)\},\ \forall a>0,\ t>0\ . \tag{1}$$

Параметр H, $0<H<1$, называемый показателем Херста, представляет собой степень самоподобия процесса. Наряду с этим свойством, показатель $H>0.5$ характеризует меру долгосрочной зависимости стохастического процесса, т.е. убывание автокорреляционной функции $r(k)$ по степенному закону:

$$r(k) \sim k^{-\beta},\ k\to\infty,\ 0<\beta<1,\ H=1-(\beta/2).$$

Одними из первых реальных стохастических процессов, у которых были обнаружены самоподобные свойства, являются информационные потоки данных в телекоммуникационных сетях. Для самоподобного трафика методы расчета характеристик компьютерной сети (пропускной способности каналов, емкости буферов и пр.), основанные на классических моделях, не соответствуют необходимым требованиям и не позволяют адекватно оценивать нагрузку в сети. Существует большое количество публикаций, посвященных анализу фрактальных свойств трафика и их влияния на функционирование и качество обслуживания телекоммукационной сети [Willinger, 1996; Stollings, 2003; Sheluhin, 2007].

Другим примером фрактальных стохастических структур являются современные финансовые рынки. Гипотеза фрактальности финансовых рядов предполагает, что рынок представляет собой саморегулируемую макроэкономическую систему с обратной связью, использующую информацию о прошлых событиях, влияющих на решения в настоящем, и содержащую долговременные корреляции и тренды. Рынок остается стабильным, пока он сохраняет свою фрактальную структуру. Анализируя динамику возникновения участков с различной фрактальной структурой, можно диагностировать и прогнозировать нестабильные состояния (кризисы) рынка. [Peters, 1996; Ширяев, 1998]

В последние годы многочисленные исследования показали, что многие биоэлектрические сигналы обладают фрактальной структурой. Отчетливые изменения фрактальных характеристик кардио- и энцефалограмм проявляются при различных заболеваниях, при изменении умственной и физической нагрузки на организм. Фрактальный анализ биоэлектрических сигналов может являться основой для проведения статистических исследований, что позволит сформулировать методики, которые будут значимы и для клинической практики. [Peng, 1994; Kantelhardt, 2003].

Очевидно, что оценивание показателя Херста по экспериментальным данным играет важнейшую роль в изучении процессов, обладающими свойствами самоподобия. Существует множество методов оценивания параметра самоподобия, каждый из которых несет отпечаток той области научных приложений, где он первоначально разрабатывался [Willinger, 1996; Clegg, 2005; Sheluhin, 2007; Kantelhardt, 2008]. При оценивании показателя Херста на практике наиболее часто используются методы нормированного размаха, изменения дисперсии ряда, флуктуационного анализа. Особое значение среди методов исследования фрактальных нестационарных процессов имеют методы, основанные на вейвлет-преобразованиях. Основные идеи вейвлет-фрактальных методов анализа сформулированы в работах [Mallat, 1998; Abry, 1998, 2002; Flandrin, 2009].

С каждым днем число публикаций, связанных с практическим применением фрактального анализа растет. Однако в настоящее время не существует универсального подхода к оцениванию фрактальных характеристик, основанного на предварительном исследовании корреляционной структуры процесса. Основными недостатками в применении методов фрактального анализа является: отсутствие предварительного исследования корреляционной структуры процесса, применение только одного метода анализа, слабое исследование статистических свойств оценок фрактальных характеристик, полученных по временным рядам малой длины.

Целью работы является разработка комплексного использования методов фрактального анализа для исследования временных рядов небольшой длины, с применением специальных методов предварительного исследования данных.

Основные методы оценивания показателя Херста

Моменты q-го порядка самоподобного случайного процесса (1) можно выразить следующим образом:

$$\mathrm{M}\left[\left|X(t)\right|^{q}\right] = C(q)\cdot t^{qH}\ ,\ C(q) = \mathrm{M}\left[\left|X(1)\right|^{q}\right]. \qquad (2)$$

Фактически все методы оценивания параметра самоподобия по временному ряду $x(t)$, $t = 1,\dots,N$, базируются на выполнении соотношения (2) при значении $q = 2$. Метод нормированного размаха, предложенный Г.Херстом и до сих пор являющийся одним из наиболее популярных в исследованиях фрактальных рядов самой различной природы [Feder, 1988], основан на скейлинговом соотношении $M\left[R(\tau)/S(\tau)\right] \propto \tau^H$, где $R(\tau)$ – размах кумулятивного ряда $x^{cum}(t,\tau)$, $S(\tau)$ – среднее квадратическое отклонение исходного ряда. Метод изменения дисперсии агрегированного ряда наиболее часто используется при исследовании процессов в телекоммуникационных сетях [Stollings, 2003] и базируется на том, что дисперсия агрегированных временных серий $x_k^{(m)} = \frac{1}{m}\sum_{t=km-m+1}^{km} x(t)$, $k = 1,\ldots,N/m$, подчиняется зависимости $Var(x^{(m)}) \propto \frac{Var(x)}{m^\beta}$. Метод детрендированного флуктуационного анализа (ДФА), первоначально предложенный в работе [Peng, 1994], в настоящее время является основным методом определения самоподобия для нестационарных временных рядов [Kantelhardt, 2001, 2008]. В методе ДФА вычисляется функция функция $F^2(\tau) = \frac{1}{\tau}\sum_{t=1}^{\tau}(y(t) - Y_m(t))^2$, где $Y_m(t)$ – локальный m-полиномиальный тренд. Функция $F(\tau)$, усредненная по всему ряду $y(t)$, обладает скейлинговой зависимостью $F(\tau) \propto \tau^H$. Вейвлет-оценивание показателя Херста базируется на свойствах детализирующих вейвлет-коэффициентов, которые на каждом уровне разложения j также обладают самоподобием. Метод вейвлет-оценивания [Abry, 1998] основан на том, что изменение значений вейвлет-энергии E подчиняется скейлинговому отношению $E_j \propto 2^{(2H+1)j}$.

Комплексный подход к оцениванию показателя Херста

В работах [Kirichenko 2010; 2011; Кириченко, 2006; 2010] выполнен сравнительный анализ статистических характеристик оценок параметра Херста, полученных вышеперечисленными методами по временным рядам малой длины. Подводя итоги исследований, можно предложить следующую схему проведения фрактального анализа некоторого случайного процесса, представленного временным рядом длины N. В основных этапах фрактального анализа задействованы методы нормированного размаха, ДФА и вейвлет-оценивания. Поскольку для применения аппарата вейвлет-преобразований необходимо соответствующее программное обеспечение и опыт работы, описание алгоритма построено таким образом, что использование методов вейвлет-оценивания является желательным, но не обязательным элементом. Однако применение метода ДФА является необходимым по двум причинам: этот метод обладает достаточной точностью и предназначен для работы с нестационарными временными рядами. Рассмотрим поэтапную реализацию комплексного подхода к оцениванию фрактальных свойств самоподобных временных рядов.

Этап 1. Предварительное исследование структуры временного ряда.

1. Прежде чем приступать к фрактальному анализу временного ряда, необходимо выяснить из априорно известной информации, является ли ряд кумулятивным (например, курс валюты) или представляет собой ряд приращений (например, телекоммуникационный трафик данных). Если по своей природе ряд является кумулятивным, то нижеуказанные этапы фрактального анализа относятся к соответствующему ряду приращений.

2. Определение интервалов различных скейлингов.

Если самоподобный процесс обладает несколькими скейлингами, зависящими от временных интервалов (например, дневные и часовые данные для валютных рядов), то на каждом таком интервале динамика временного ряда определяется соответствующим показателем Херста. Для определения таких интервалов надо построить показатель Херста, как функцию отсчетов времени. Такой подход возможен при применении метода нормированного размаха, когда интервалы времени τ изменяются малыми приращениями $H(\tau) = f[\log\frac{R}{S}(\tau)]$ [Peters, 1996] .

Кроме метода нормированного размаха для выявления интервалов разных скейлингов можно использовать построение флуктуационной функции $F_{DFA}(\tau)$ методом ДФА. Если существует несколько скейлингов, функция $F_{DFA}(\tau)$ изменит угол наклона. Однако, в отличие от метода нормированного размаха, который исследует долгосрочную зависимость временного ряда в диапазоне всей его длины, флуктуационная функция может быть корректно построена только на интервале до значений $N/4$ [Kantelhardt, 2001].

3. Выявление и удаление краткосрочной авторегрессионной зависимости. Проверка гипотезы о наличии самоподобия.

R/S-анализ позволяет обнаружить и устранить краткосрочную зависимость, характерную для авторегрессионных процессов. Наличие авторегрессионной зависимости смещает значения показателя Херста и демонстрирует ложную долговременную память [Peters, 1996; Ширяев, 1998; Кириченко, 2006]. Поэтому, при выяснении фрактальной структуры временного ряда необходимо сначала выяснить наличие краткосрочной зависимости. Для этого надо значения временного ряда $x(t)$ регрессировать как зависимую переменную против $x(t-1)$ и найти линейную зависимость между ними: $x(t) = a + b \cdot x(t-1)$. Значимость коэффициента b свидетельствует о наличии краткосрочной зависимости. Для ее устранения определяется остаток: $S(t) = x(t) - (a + b \cdot x(t-1))$.

После этого проводится R/S-анализ остаточного ряда $S(t)$. Если исходный ряд $x(t)$ имел долгосрочную зависимость, то она сохраняется, в то время, как краткосрочная зависимость устраняется. Если авторегрессионная зависимость является значимой, то все ниже указанные этапы фрактального анализа относятся к остаточному временному ряду.

Если значение параметра Херста H близко к 0.5, необходимо проверить гипотезу о наличии самоподобия. В качестве нулевой гипотезы обычно постулируется, что приращения случайного процесса имеют независимый характер. В работе [Peters, 1996] представлены критерии и области принятия данной гипотезы.

Качественной проверкой наличия свойств статистического самоподобия является построение агрегированных временных серий $x_k^{(m)} = \frac{1}{m}\sum_{t=km-m+1}^{km} x(t)$, $k = 1,\ldots,N/m$, и построение для них выборочных функций распределения. В случае самоподобности временного ряда $x(t)$ агрегированные серии имеют одинаковое распределение, подтвержденное статистическими критериями.

Этап 2. Оценивание показателя Херста по стационарному ряду.

Для оценивания показателя Херста необходимо определить, является ли ряд $x(t)$ стационарным или нет известными статистическими методами. Если ряд стационарен, то степень самоподобия H и интервальную оценку показателя Херста можно определить вышеуказанными или какими-либо другими методами.

Результаты исследований показали, что оценки показателя Херста H, которые получены рассмотренными методами по реализациям небольшой длины, является смещенными нормально распределенными случайными величинами. Для каждого метода смещение зависит от истинного значения степени самоподобия процесса и длины временного ряда. Минимальное смещение имеют оценки, полученные методом ДФА и с помощью вейвлет-преобразования. Средние квадратические отклонения оценок зависят от метода оценивания и уменьшаются с ростом длины ряда. Минимальные средние квадратические отклонения имеют оценки, полученные с помощью вейвлет-анализа.

Были исследованы выборочные законы распределения оценок параметра Херста и показано, что они имеют нормальное распределение. В этом случае оценка показателя Херста может быть представлена интервалом значений, внутри которого с заданной вероятностью находится истинное значение H:

$$\widehat{H} + \Delta - t_{\alpha} S < H < \widehat{H} + \Delta + t_{\alpha} S, \tag{3}$$

где N – длина исследуемого временного ряда; $method$ – выбранный метод оценивания; $\widehat{H} = \widehat{H}(N, method)$ – полученное значение оценки показателя Херста по реализации длины N; $\Delta = \Delta(N, method)$ – величина систематического смещения оценки, рассчитанная по модельным реализациям длины N; $S = S(N, method)$ – среднее квадратическое отклонение, рассчитанное по модельным реализациям длины N; α – требуемый уровень значимости; t_{α} – квантиль простейшего нормального распределения.

Этап 3. Оценивание показателя Херста по нестационарному ряду

Если ряд нестационарен, то степень самоподобия и интервальную оценку показателя Херста нужно определять методом ДФА или с помощью вейвлет-оценивания.

1. Если исследуемый временной ряд нестационарен, необходимо исследовать его структуру с помощью построения корреляционной функции (спектральной плотности) и спектра вейвлет-энергии, которые позволяют выявить трендовые и циклические составляющие ряда.

2. При оценивании показателя Херста методом ДФА, необходимо сначала провести прикидочное оценивание, используя локальные полиномы увеличивающейся степени и определить наименьшую степень полинома, начиная с которой оценка показателя Херста перестает изменяться. После этого для оценивания самоподобия временного ряда необходимо удалять локальный полиномиальный тренд найденной степени.

3. Вейвлет–оценивание существенно нестационарного временного ряда можно проводить согласно методам, представленным в работах [Кириченко 2009; 2010]. В этом случае оценка показателя H существенно зависит от выбранного материнского вейвлета.

Этап 4. Уточнение полученной оценки показателя Херста.

Анализ корреляционной зависимости между оценками параметра Херста, полученными разными методами, показал, что выборочные коэффициенты корреляции оказались лежащими в диапазоне абсолютных значений в основном меньшем 0.5. Корреляция вейвлет-оценок с оценками, полученными другими методами, является незначимой. Поэтому, для увеличения точности оценивания можно использовать среднее арифметическое несмещенных оценок, полученных с помощью нескольких методов оценивания.

Для повышения точности вейвлет-оценивания оценок был проведен сравнительный анализ статистических характеристик оценок, полученных с помощью разных вейвлет-функций [Кириченко 2009]. Корреляционный анализ вейвлет-оценок показателя Херста, полученными с помощью разных вейвлетов, показал, что более эффективной оценкой показателя Херста является среднее арифметическое оценок, полученных с помощью нескольких разных материнских вейвлет-функций.

Заключение

Таким образом, в работе предложен комплексный подход к анализу фрактальных свойств временных рядов. Предложенный метод предусматривает предварительное исследование структуры временного ряда, несмещенное интервальное оценивание параметра самоподобия и совместное использование нескольких методов фрактального анализа, что позволяет повысить достоверность получаемых оценок. Данный подход применим для исследования самоподобных временных рядов различной природы: телекоммуникационных трафиков, финансовых показателей, биомедицинских сигналов и др.

Литература

[Abry, 1998] P. Abry. Wavelet analysis of long-range dependent traffic. P. Abry, D. Veitch. IEEE/ACM Transactions Information Theory № 1(44), 1998.

[Abry, 2003] P. Abry. Self-similarity and long-range dependence through the wavelet lens. P. Abry, P. Flandrin, M.S. Taqqu, D. Veitch. Theory and applications of long-range dependence, Birkhäuser, 2003.

[Chen, 2002] Z. Chen. Effect of non-stationaritieson detrended fluctuation analysis. Z. Chen, P.Ch. Ivanov, K. Hu, H.E. Stanley. Phys. Rev. E 65, 041107, 2002.

[Clegg, 2005] R.G. Clegg. A practical guide to measuring the Hurst parameter. R. G. Clegg. Computing science technical report, № CS–TR–916, 2005.

[Feder, 1988] J. Feder. Fractals. J. Feder. Plenum, New York, 1988.

[Flandrin, 2009] Patrick Flandrin. Scale Invariance and Wavelets. Patrick Flandrin, Paulo Gonзalves and Patrice Abry in Scaling, Fractals and Wavelets. Ed. by P. Abry, P. Gonçalves, J. Lévy Véhel. John Wiley & Sons, London, 2009.

[Kantelhardt, 2001] J.W. Kantelhardt. Detecting long-range correlations with detrended fluctuation analysis. J.W. Kantelhardt, E. Koscielny-Bunde, H.H.A. Rego, S. Havlin, A. Bunde. Phys (A 295, 441), 2001.

[Kantelhardt, 2003] J.W. Kantelhardt Breathing during REM and non-REM sleep:correlated versus uncorrelated behavior. J.W. Kantelhardt, T. Penzel, S. Rostig, H. F. Becker, S. Havlin, A. Bunde. Physica A 319 (2003) P.447 – 457.

[Kantelhardt, 2008] J. W. Kantelhardt. Fractal and Multifractal Time Series. J. W. Kantelhardt. http://arxiv.org/abs/0804.0747, 2008.

[Kirichenko, 2011] L. Kirichenko. Comparative Analysis for Estimating of the Hurst Exponent for Stationary and Nonstationary Time Series L. Kirichenko,T.Radivilova, Zh. Deineko . Information Technologies & Knowledge. – 2011. – Vol.5. – № 4. – P. 371–388.

[Kirichenko, 2011] L. Kirichenko. Comparative analysis of statistical properties of the Hurst exponent estimates obtained by different methods L. Kirichenko, T.Radivilova. Information Models of Knowledg / ed. K. Markov, V. Velychko, O. Voloshin. – Kiev–Sofia: ITHEA. – 2010. – P. 451–459.

[Mallat, 1998] S. Mallat. A wavelet tour of signal processing. S. Mallat. Academic Press, San Diego, London, Boston, N.Y., Sydney, Tokyo, Toronto, 1998.

[Peng, 1994] C.-K. Peng. Mosaic organization of DNA nucleotides. C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger. Phys. Rev. (E 49, 1685),1994.

[Peters, 1996] Edgar E. Fractal Market Analysis: applying chaos theory to investment and economics . Edgar E. Peters. Wiley, 2 edition, 2003.

[Sheluhin, 2007] Oleg I. Sheluhin. Self-similar processes in telecommunications. Oleg I. Sheluhin, Sergey M. Smolskiy, Andrey V. Osin. JohnWiley & Sons Ltd, Chichester, 2007.

[Stollings, 2003] W. Stollings. High-speed networks and Internets. Performance and quality of service. W. Stollings. New Jersey, 2002.

[Willinger, 1996] W. Willinger. Bibliographical guide to self-similar traffic and performance modeling for modern high-speed network in «Stohastic networks: theory and applications». W. Willinger, M. S. Taqqu, A. A. Erramilli. Claredon Press (Oxford University Press), Oxford, 1996.

[Кириченко, 2006] Л. О Кириченко. Исследование долгосрочной зависимости сетевого трафика методом R/S-анализа Л. О. Кириченко, Т. А. Радивилова. Автоматизированные системы управления и приборы автоматики. – 2006. – Вып. 135. – С. 51–55.

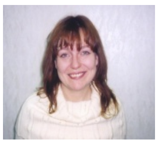
[Кириченко, 2010] Л. О Кириченко. Оценивание параметра Хёрста для временных рядов с трендом методом вейвлет-преобразования. Л. О Кириченко, Ж.В. Дейнеко. Системи управління навігації та зв'язку. Вип 4 (16), Київ, 2010.

[Ширяев, 1998] А. Н. Ширяев. Основы стохастической финансовой математики. А. Н. Ширяев. – М. : Фазис, 1998. – Т. 1: Факты. Модели.

Информация об авторах

Людмила Кириченко *– к.т.н., доцент Харьковского национального университета радиоэлектроники; пр. Ленина 14, 61166, Харьков, Украина; e-mail: ludmila.kirichenko@gmail.com*

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, фрактальный анализ, вейвлет-анализ, детерминированные хаотические системы.



Лариса Чалая *– к.т.н., доцент Харьковского национального университета радиоэлектроники; пр. Ленина 14, 61166, Харьков, Украина; e-mail: kovalivnich@yahoo.com*

Основные области научных исследований: искусственный интеллект, обработка естественно-языковой информации, фракталы.

Integrated Approach to the Study of Fractal Time Series

Lyudmyla Kirichenko, Larysa Chala

Abstract: *In this works we propose an integrated approach to the analysis of self-similar properties of stochastic processes for time series of short length. The sequence of steps of the fractal analysis was given. These steps are: preliminary analysis, including the removal of short-term dependence and revealing the true long-term dependency; hypothesis testing of a self-similarity; unbiased interval estimation of the Hurst exponent in cases of stationary and non-stationary time series by several methods; correction of the resulting estimate of the Hurst exponent.*

Keywords: *self-similar stochastic process, time series, Hurst exponent, methods for estimating the Hurst exponent*