

ВЕЙВЛЕТ-ОЦЕНИВАНИЕ ПОКАЗАТЕЛЯ ХЕРСТА ПО ВРЕМЕННЫМ РЯДАМ СО ЗНАЧИТЕЛЬНЫМИ ЦИКЛИЧЕСКИМИ КОМПОНЕНТАМИ С ПРИМЕНЕНИЕМ ПАКЕТНОГО ВЕЙВЛЕТ- ПРЕОБРАЗОВАНИЯ

Предлагается метод вейвлет-оценивания показателя Херста, основанный на использовании пакетного преобразования. Применение разработанного метода позволяет определять степень самоподобия для рядов небольшой длины, обладающих существенными трендовыми и циклическими компонентами.

1. Введение

В настоящее время стало общепризнанным, что многие информационные, биологические, физические, технологические процессы имеют сложную фрактальную структуру. Фрактальный анализ используется для моделирования, анализа и контроля сложных систем в различных областях науки и техники. В частности, в геологии – для прогнозирования сейсмической активности и цунами, определения возраста геологических пород; в биологии и медицине – для диагностики заболеваний и физиологического состояния по записям ЭКГ и ЭЭГ, при исследовании мутаций и изменений на генетическом уровне; в экономике – для прогнозирования кризисных ситуаций и оценивания риска по финансовым рядам; в физике – для исследования турбулентности и термодинамических процессов [1-4].

Оценка степени самоподобия (показателя Херста) по экспериментальным данным играет важнейшую роль в изучении процессов, имеющих фрактальные свойства. Степень самоподобия стохастического процесса, который определяется показателем Херста, является показателем сложности динамики и корреляционной структуры временного ряда.

Для оценки показателя Херста по временному ряду существует ряд методов [5-6]. Однако большинство методов применимы только к стационарным рядам, в то время как множество природных, технических и информационных процессов являются нестационарными. Типичным нарушением стационарных свойств, которые встречаются на практике, является наличие трендовой и циклической составляющей ряда. Эффективным инструментом исследования структуры нестационарных временных рядов является аппарат вейвлет-анализа, который позволяет выделять структурные компоненты ряда. В частности, методы, основанные на дискретном вейвлет-преобразовании, позволяют оценивать степень самоподобия для нестационарных рядов [7-9].

Однако стандартный метод вейвлет-оценивания не позволяет получить адекватные оценки показателя Херста при наличии значительных циклических компонент временного ряда [10, 11]. Целью представленного исследования является разработка метода оценивания показателя Херста для рядов с существенными трендовыми и циклическими компонентами, основанного на применении пакетного вейвлет-преобразования.

2. Декомпозиция временных рядов с помощью дискретного вейвлет-преобразования [7-9,12]

Вейвлет-преобразование временного ряда – это его представление в виде обобщенного ряда или интеграла по системе базисных функций

$$\Psi_{ab}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right),$$

полученных из материнской вейвлет-функции $\psi(t)$ за счет операций сдвига во времени b и изменения временного масштаба a . В дискретном вейвлет-преобразовании (ДВП) параметры масштаба и сдвига задаются обычно в виде степенных функций: $a = 2^{-m}$, $b = k_0 \cdot 2^{-m}$, $m, k_0 \in Z$. Дискретные вейвлеты используются, как правило, в паре со связанными с ними дискретными скейлинг-функциями. Скейлинг-функции имеют с вейвлетами общую об-

ласть задания и определенное соотношение между значениями. При заданных материнском вейвлете $\psi(t)$ и соответствующей скейлинг-функции $\varphi(t)$ аппроксимирующие коэффициенты $\text{apr}(j, k)$ и детализирующие коэффициенты $\text{det}(j, k)$ ДВП для процесса $X(t)$ определяются следующим образом:

$$\text{apr}(j, k) = \int_{-\infty}^{\infty} X(t)\varphi_{j,k}(t)dt, \quad \text{det}(j, k) = \int_{-\infty}^{\infty} X(t)\psi_{j,k}(t)dt,$$

где j – параметр масштаба; k – параметр сдвига ($j, k \in Z$) и базисные функции равны:

$$\varphi_{j,k}(t) = 2^{-j/2}\varphi(2^{-j}t - k),$$

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

Декомпозиция временного ряда, выполняемая с помощью ДВП, заключается в разбиении исследуемого ряда на две составляющие: аппроксимирующую и детализирующую, с последующим аналогичным дроблением аппроксимирующей до заданного уровня разложения. Временной ряд $X(t)$ представляется в виде суммы аппроксимирующей компоненты $\text{approx}_N(t)$ и детализирующих компонент $\text{detail}_j(t)$:

$$X(t) = \text{approx}_N(t) + \sum_{j=1}^N \text{detail}_j(t) =$$

$$= \sum_{k=1}^{N_a} \text{apr}(N, k)\varphi_{j,k}(t) + \sum_{j=1}^N \sum_{k=1}^{N_j} \text{det}(j, k)\psi_{j,k}(t),$$

где N – выбранный максимальный уровень разложения; N_j – количество детализирующих коэффициентов на уровне j ; N_a – количество аппроксимирующих коэффициентов на уровне N .

Выбор типа вейвлет-функции и количества уровней разложения является важным вопросом при выполнении ДВП. Обычно вейвлет-функция подбирается в зависимости от временных и частотных характеристик каждого анализируемого сигнала. Максимальный уровень разложения зависит от того, какие частотные диапазоны необходимо исследовать.

В обычном алгоритме ДВП при переходе с масштабного уровня m на уровень $m+1$ функция аппроксимирующих коэффициентов $\text{apr}(m, k)$ разделяется на низкочастотную $\text{apr}(m+1, k)$ и высокочастотную $\text{det}(m+1, k)$ части спектрального диапазона. При дальнейшем увеличении масштабных уровней разложению подвергаются только низкочастотные аппроксимирующие функции, детализирующие коэффициенты далее не анализируются (рис. 1 слева)).

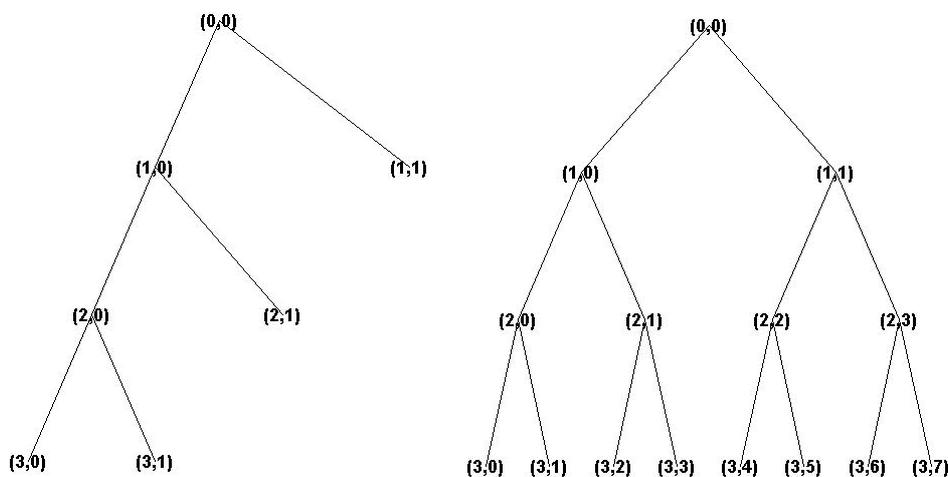


Рис. 1. Дерево обычного ДВП (слева) и пакетного ДВП (справа)

Идея пакетного вейвлет-анализа состоит в том, чтобы раскладывать детализирующие функции, используя тот же самый метод разложения. В пакетном алгоритме ДВП операция последовательного частотного расщепления применяется как для низкочастотных аппроксимирующих, так и для высокочастотных детализирующих коэффициентов. Таким образом пакетное ДВП способствует лучшей частотной локализации. В результате возникает дерево разложения, пример которого показан на рис. 1 справа. При разложении данных некоторого узла ниже слева находятся аппроксимирующие коэффициенты, а справа – детализирующие.

Каждый узел дерева пакетного ДВП содержит набор вейвлет-коэффициентов, соответствующих определенному частотному диапазону. Временной ряд может быть однозначно восстановлен по пакетным вейвлет-коэффициентам, находящимся в терминальных узлах дерева.

3. Оценивание самоподобия стохастических процессов с помощью ДВП [7-10]

Стохастический процесс $X(t)$ является самоподобным с параметром самоподобия H , если процесс $a^{-H}X(at)$ описывается теми же конечномерными законами распределений, что и $X(t)$: $\text{Law}\{X(t)\} = \text{Law}\{a^{-H}X(at)\}$, $\forall a > 0, t > 0$.

Параметр $H, 0 < H < 1$, называемый показателем Херста, представляет собой степень самоподобия и меру долгосрочной зависимости стохастического процесса.

Метод оценивания степени самоподобия базируется на свойствах детализирующих вейвлет-коэффициентов, полученных в результате декомпозиции процесса. Если случайный процесс $X(t)$ является самоподобным, то детализирующие коэффициенты на каждом уровне разложения также обладают свойством самоподобия:

$$\text{Law}\{\text{det}(j, k)\} = \text{Law}\{2^{j(H+1/2)} \text{det}(0, k)\}, \quad (1)$$

где $\text{det}(j, k)$ – k -й детализирующий коэффициент уровня разложения j , $k = 1, 2, \dots, N_j$; N_j – количество вейвлет-коэффициентов на уровне разложения j ; H – показатель Херста. Это свойство вытекает из свойств вейвлет-базиса, полученного путем масштабирования материнских вейвлетов.

Если для самоподобного процесса $X(t)$ существуют моменты 2-го порядка, то для вейвлет-коэффициентов, полученных в результате декомпозиции данного процесса, выполняется следующее равенство:

$$M|\text{det}(j, k)|^2 = M|\text{det}(0, k)|^2 2^{-j(2H+1)}. \quad (2)$$

Величина вейвлет-энергии $M|\text{det}(j, k)|^2$ на уровне j вычисляется следующим образом:

$$E_j = \frac{1}{N_j} \sum_{k=1}^{N_j} \text{det}^2(j, k). \quad (3)$$

Метод вейвлет-оценивания степени самоподобия H базируется на свойствах детализирующих вейвлет-коэффициентов, полученных на разных уровнях декомпозиции исследуемого временного ряда. Основу данного метода составляет положение, следующее из (1)-(2), о том, что изменение значений вейвлет-энергии на масштабном уровне j подчиняется скейлинговому отношению:

$$E_j \propto 2^{(2H+1)j}. \quad (4)$$

График зависимости $\log_2(E_j)$ от j представляет собой прямую линию, аппроксимированную методом наименьших квадратов. Значение параметра H можно найти, оценив тангенс угла наклона прямой. На рис. 2 представлен типичный спектр вейвлет-энергии самоподобного процесса (вверху) и зависимость $\log_2(E_j)$ от номера уровня разложения j (внизу).

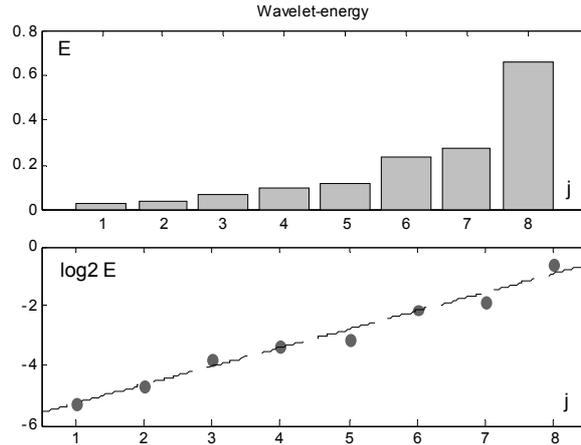


Рис. 2. Спектр вейвлет-энергии самоподобного процесса (вверху) и зависимость $\log_2(E_j)$ (внизу)

Одной из важных характеристик вейвлет-функции $\psi(t)$ является наличие нулевых моментов высокого порядка: $\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0, \quad k = \overline{0, p-1}$. Число нулевых моментов p характеризует способность вейвлет-преобразования анализировать мелкомасштабные флуктуации и особенности высокого порядка, игнорируя регулярные полиномиальные составляющие степени $m < p$.

4. Моделирование реализаций фрактального процесса

Одной из наиболее известных и простых моделей стохастической динамики, обладающих фрактальными свойствами, является фрактальное броуновское движение (ФБД), которое широко применяется в физике, химии, биологии, экономике и теории сетевого трафика [1-4].

Гауссовский процесс $X(t)$ называется фрактальным броуновским движением с параметром $H, 0 < H < 1$, если приращения случайного процесса $\Delta X(\tau) = X(t + \tau) - X(t)$ имеют распределение вида:

$$P(\Delta X < x) = \frac{1}{\sqrt{2\pi\sigma_0\tau^H}} \cdot \int_{-\infty}^x \text{Exp} \left[-\frac{z^2}{2\sigma_0^2\tau^{2H}} \right] dz,$$

где σ_0 – коэффициент диффузии.

ФБД с параметром $H = 0,5$ совпадает с классическим броуновским движением. Приращения ФБД называются фрактальным гауссовским шумом (ФГС), дисперсия которого подчиняется соотношению $D[X(t + \tau) - X(t)] = \sigma_0^2 \tau^{2H}$.

Существует несколько методов построения ФБД для случая дискретного времени. Эти модели имеют определенные недостатки, одним из которых является занижение (завышение) степени самоподобия процесса при малых или больших теоретических значениях параметра Херста и небольшой длине модельной реализации [13].

Одним из методов, направленных на преодоление указанных недостатков, является построение ФБД с помощью биортогональных вейвлетов [14-15]. В этом случае построение реализации ФБД осуществляется с помощью дискретного вейвлет-преобразования, где детализирующие вейвлет-коэффициенты на каждом уровне являются независимыми нормально распределенными случайными величинами, а аппроксимирующие вейвлет-коэффициенты получены с помощью фрактального процесса авторегрессии и скользящего среднего FARIMA:

$$B_H(t) = \sum_{k=-\infty}^{\infty} \Phi_H(t-k) S_k^{(H)} + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} 2^{-jH} \Psi_H(2^j t - k) \varepsilon_{j,k} - b_0,$$

где Ψ_H – биортогональная базовая вейвлет-функция; Φ_H – соответствующая Ψ_H скейлинг-функция; $S_k^{(H)}$ – стационарный гауссовский процесс FARIMA с параметром дробного дифференцирования $d = H - 0,5$; $\varepsilon_{j,k}$ – независимые стандартные нормальные случайные величины; b_0 – константа, такая что $B_H(0) = 0$.

На рис.3 приведены реализация ФБД (вверху) и соответствующий ряд ФГШ (внизу), полученные описанным выше методом. Длина реализации равна 1024 значениям, теоретический показатель Херста $H=0.8$.

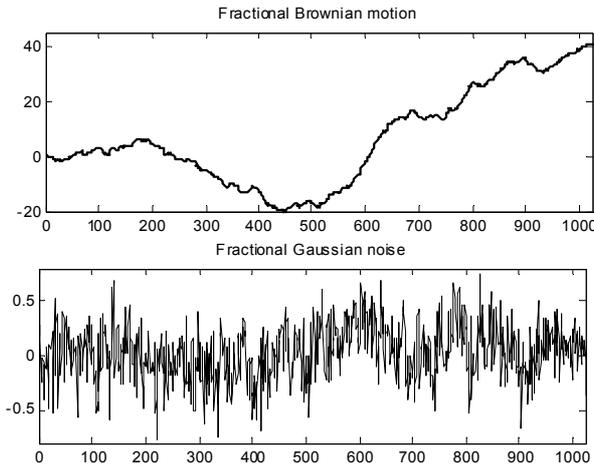


Рис.3. Реализация ФБД (вверху) и ФГШ (внизу)

5. Оценивание самоподобия временных рядов с циклическими составляющими

В работе использовались различные модельные сигналы – временные реализации, представленные суммой ФГШ с заданным показателем Херста и периодической составляющей. Суммарный сигнал $X(t)$ можно представить как

$$X(t) = P(t) + fgn(t),$$

где $P(t)$ – циклический тренд; $fgn(t)$ – фрактальный шум. Соотношение циклического тренда и модельного фрактального шума можно определить как

$$\text{Ratio} = \frac{S_{\text{trend}}}{S_{\text{noise}}},$$

где S_{trend} – среднее квадратическое отклонение тренда; S_{noise} – среднее квадратическое отклонение фрактального процесса.

Для оценивания показателя H были выбраны вейвлет-функции семейства Добеши (db2-db20), которые обладают необходимым числом нулевых моментов (соответственно 2-20). При небольших значениях отношения тренда к ФГШ ($\text{Ratio} \leq 2$) метод вейвлет-оценивания (4) позволяет получить адекватные оценки показателя Херста при использовании вейвлет-функций с большим числом нулевых моментов. На рис. 4 приведены периодическая составляющая $P(t)$, суммарный сигнал $X(t)$ и спектр вейвлет-энергии суммарного сигнала. В данном случае теоретический показатель $H=0,8$, величина $\text{Ratio} = 0.7$. Применена вейвлет-функция db20 и получена оценка показателя $\hat{H} = 0.808$.

Однако с увеличением значения Ratio и числа периодов во временной реализации требуются вейвлет-функции с все большим числом нулевых моментов, что сложно реализуемо на практике. Кроме того, использование вейвлет-функций, обладающих большим числом нулевых моментов, приводит к менее точному оцениванию показателя H [16].

В работах [11,17] предложен метод корректного оценивания параметра H , если соотношение тренда и фрактального шума является значительным, т.е. значение $Ratio=1$. В этом случае необходимо предварительно провести анализ спектра вейвлет-энергии сигнала. Если сигнал содержит значительную трендовую или циклическую составляющую, спектр вейвлет-коэффициентов демонстрирует большие значения E_k на соответствующих уровнях разложения k , и при оценивании показателя H преобладающие в сигнале трендовые составляющие подавляют все остальные. Для этого случая предложено при оценке наклона графика функции $\log_2(E_j)$ не учитывать значения трендовой составляющей E_k , тогда оценка показателя Херста будет корректной. На рис. 5 представлены модельный сигнал ($Ratio = 4$), спектр вейвлет-энергии, полученный с помощью вейвлет-функции db4, и зависимость $\log_2(E_j)$ с аппроксимирующими прямыми, построенными с учетом и без учета значений E_k при значении $k = 8$.

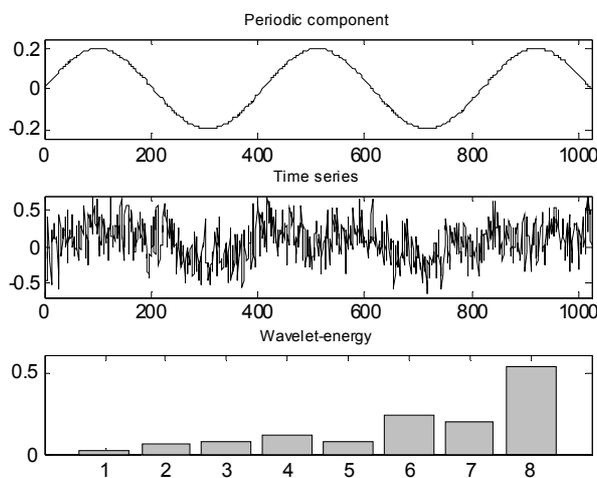


Рис. 4. Периодическая компонента, суммарный сигнал и спектр вейвлет-энергии суммарного сигнала

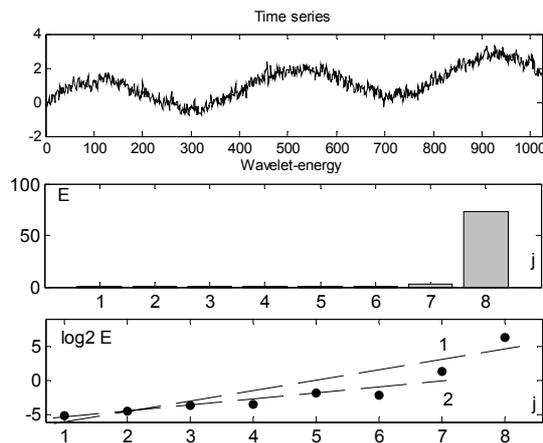


Рис. 5. Сигнал, спектр вейвлет-энергии, зависимость $\log_2(E_j)$ и аппроксимирующие прямые с учетом (1) и без учета (2) значений E_8

Однако такой подход не приемлем для временных рядов небольшой длины или в случае, когда циклические компоненты попадают сразу в несколько уровней вейвлет-разложения и удаление нескольких значений вейвлет-спектра E_k может привести к неправильному оцениванию показателя Херста. В данной работе предложен метод вейвлет-оценивания пока-

зателя Херста для временных рядов со значительными трендовыми или циклическими составляющими, основанный на использовании пакетного вейвлет-преобразования. Рассмотрим его на примере сигнала, показанного на рис. 6. В данном случае теоретический показатель Херста $H=0,8$, величина $Ratio = 3$. Спектр вейвлет-энергии получен с использованием вейвлет-функции $db10$. Непосредственное вычисление значения H приводит к некорректной завышенной оценке показателя $\hat{H} = 1.04$.

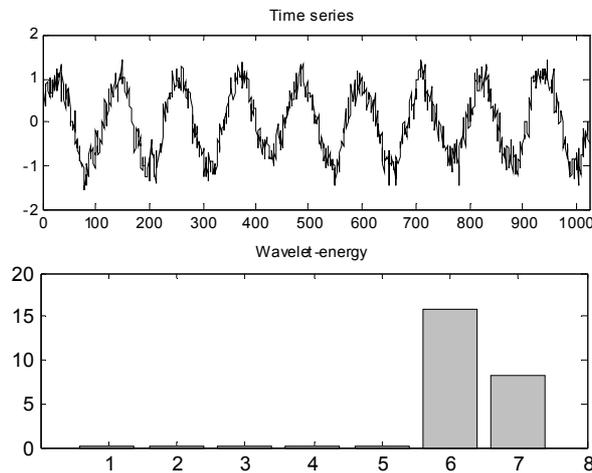


Рис. 6. Сигнал и спектр вейвлет-энергии

Очевидно, что основные циклические составляющие попадают в 6-й и 7-й уровни вейвлет-разложения. Детализируем эти уровни, применяя пакетное вейвлет-преобразование с той же вейвлет-функцией $db10$. Часть дерева пакетного вейвлет-разложения, соответствующая узлам 6-го и 7-го уровней, представлена на рис.7.

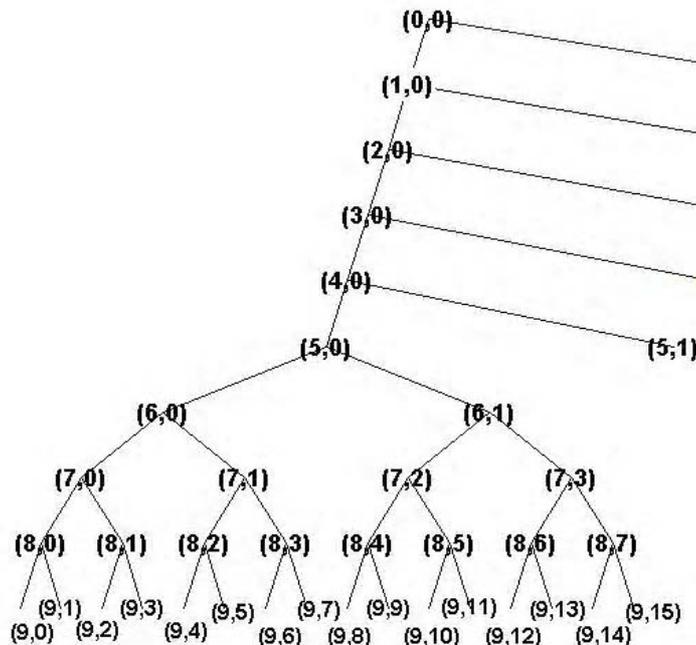


Рис. 7. Дерево пакетного вейвлет-разложения сигнала, показанного на рис. 6

Узел 6.1 дерева пакетного ДВП содержит набор вейвлет-коэффициентов, соответствующих частотному диапазону 6-го уровня разложения обычного ДВП. Узлы 8,4-8,7 пред-

ставляют частотные поддиапазоны с интервалом частот в четыре раза уже диапазона узла 6,1. По формуле (3) рассчитаем значения вейвлет-энергии в узлах 8,4-8,7. Среднее значение этого спектра соответствует значению вейвлет-энергии в узле 6,1, т.е. 6-му уровню разложения ДВП:

$$E_6 = \frac{1}{2}(E_{7,2} + E_{7,3}) = \frac{1}{4}(E_{8,3} + E_{8,4} + E_{8,5} + E_{8,6}).$$

Аналогичные расчеты выполнены для узла 7,1, т.е. 7-го уровня разложения ДВП. Значения спектра вейвлет-энергии представлены на рис.8. Очевидно, что основная энергия 6-го уровня сосредоточена в узлах 8,5-8,6, а энергия 7-го уровня - в узлах 9,3-9,4. Для наглядности восстановленные компоненты сигнала, которые соответствуют узлам 8,5-8,6, представлены в нижней части рис.8.

Теперь для оценивания показателя Херста в соответствии с отношением (4) в качестве пересчитанных значений энергий 6-го и 7-го уровней возьмем значения

$$E_6 = \frac{1}{2}(E_{8,3} + E_{8,4}), \quad E_7 = \frac{1}{2}(E_{9,5} + E_{9,6}).$$

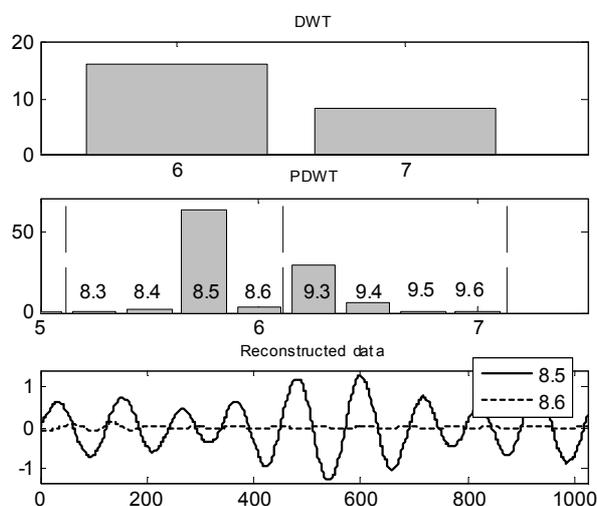


Рис.8. Часть спектра вейвлет-энергии ДВП, вейвлет-энергия узлов 6,1 и 7,1 пакетного ДВП, восстановленные сигналы узлов 8,5-8,6

На рис. 9 представлен спектр вейвлет-энергии исследуемого сигнала (см. рис. 6) с пересчитанными значениями 6-го и 7-го уровней и полученная зависимость $\log_2(E_j)$ от номера уровня разложения j . Тангенс угла наклона аппроксимирующей прямой позволяет оценить параметр H как $\hat{H} = 0.784$, что согласуется со статистическими характеристиками вейвлет-оценок показателя Херста по временным рядам аналогичной длины [16,17].

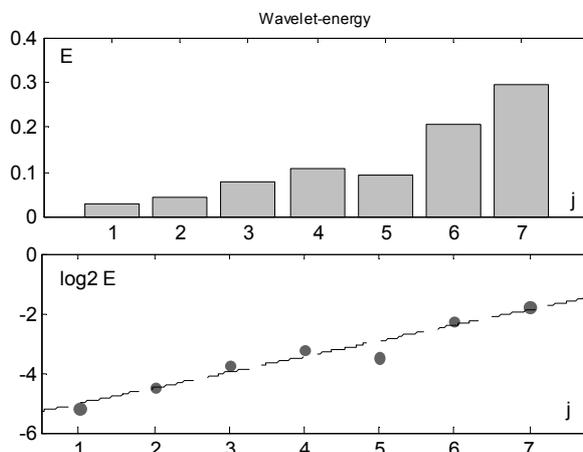


Рис. 9. Пересчитанный спектр вейвлет-энергии и соответствующая зависимость $\log_2(E_j)$

Предложенный метод можно применять к любому узлу дерева пакетного ДВП с разной степенью детализации. Таким образом можно получать вейвлет-оценки параметра Херста для самоподобных временных рядов со значительными трендовыми и циклическими составляющими, не теряя значений спектра вейвлет-энергии. Однако надо отметить, что не существует универсального подхода к определению уровня детализации пакетного вейвлет-дерева. Выбор подходящей вейвлет-функции и количества уровней разложения зависит от временных и частотных характеристик каждого анализируемого сигнала, а также от параметров циклических компонент.

Выводы

Предложен новый метод оценивания показателя Херста, основанный на использовании пакетного вейвлет-преобразования. Применение разработанного метода позволяет определять степень самоподобия для рядов небольшой длины, обладающих существенными трендовыми и циклическими компонентами.

Практическая значимость метода состоит в том, что он основан на алгоритме быстрого вейвлет-преобразования, а это позволяет с минимальными затратами времени и объема информации определять степень самоподобия сложных сигналов, что, в частности, важно для работы в режиме реального времени.

Литература: **1.** Федер Е. Фракталы. М.: Мир, 1991. 254 с. **2.** Мандельброт Б. Фрактальная геометрия природы. М.: Институт компьютерных исследований, 2002. 656 с. **3.** Шелухин О. И. Самоподобие и фракталы. Телекоммуникационные приложения / О. И. Шелухин, А. В. Осин, С. М. Смольский. М.: Физматлит, 2008. 368 с. **4.** Шредер М. Фракталы, хаос и степенные ряды. Миниатюры из бесконечного рая / М. Шредер. Ижевск: НИЦ «Регулярная и хаотическая динамика», 2005. 528 с. **5.** Clegg R.G. A practical guide to measuring the Hurst parameter / R.G. Clegg // Computing science technical report. 2005. – № CS–TR–916. **6.** Kantelhardt J.W. Fractal and Multifractal Time Series. 2008 [Электронный ресурс]: <http://arxiv.org/abs/0804.0747>. **7.** Малла С. Вэйвлеты в обработке сигналов. М.: Мир, 2005. 671 с. **8.** Abry P. Scaling, Fractals and Wavelets / P. Abry, P. Goncalves, J. L. Veheil. – London: John Wiley & Sons, Inc., 2009. 464 p. **9.** Abry P. Self-similarity and long-range dependence through the wavelet lens / P. Abry, P. Flandrin, M. S. Taqqu // Theory and applications of long-range dependence. Birkhuser, 2003. P. 34–38. **10.** Abry P. Wavelet analysis of long-range dependent traffic / P. Abry, D. Veitch // IEEE/ACM Transactions Information Theory. 1998. № 1 (44). P. 2–15. **11.** Дейнеко Ж. В. Оценивание параметра Хёрста для временных рядов с трендом методом вейвлет-преобразования / Ж. В. Дейнеко, Л. О. Кириченко // Системи управління, навігації та зв'язку. 2010. Вип. 4 (16). С. 85–89. **12.** Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в MATLAB // Н.К. Смоленцев. М.: ДМК Пресс, 2005. 304 с. **13.** Jeongy H.-D. J. A Comparative Study of Generators of Synthetic Self-Similar Teletraffic / H.-D. J. Jeongy, D. McNickle, K. Pawlikowski // Department of Computer Science and Management, University of Canterbury, 1998. **14.** Abry P. The wavelet-based synthesis for the fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation / Abry P., Sellan F. // Appl. and Comp. Harmonic Anal. V. 3(4). 1996. P. 377-383. **15.** Misiti M. Wavelet Toolbox for use with MATLAB / M. Misiti, Y. Misiti, G. Oppenheim, J.-M. Poggi // The MathWorks Inc. 941 p. 2000. **16.** Кириченко Л. О. Оценивание самоподобия стохастического временного ряда методом вейвлет-анализа / Л. О. Кириченко, Ж. В. Дейнеко // Радіоелектронні і комп'ютерні системи. 2009. № 4 (38). С. 99–105. **17.** Kirichenko L. 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.

Поступила в редколлегию 12.03.2012

Кириченко Людмила Олеговна, канд. техн. наук, доцент каф. ПМ ХНУРЕ. Научные интересы: вейвлеты, самоподобные и мультифрактальные случайные процессы, хаотическая динамика. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, ludmila@kture.kharkov.ua