

УДК 621.317

ИСПОЛЬЗОВАНИЕ Д-КАНАЛЬНОГО ВЕЙВЛЕТ-РАЗЛОЖЕНИЯ ДЛЯ АНАЛИЗА ЭКОНОМИЧЕСКИХ ДАННЫХ

П. Н. Подкур

В данной работе показаны возможности N -кратномасштабного вейвлет-анализа для изучения временных рядов, возникающих в экономике. В качестве примера рассмотрены 15-минутные данные об объемах продаж акций компании Лукойл на ММВБ за 2006 год. Найдено вейвлет-разложение временного ряда объемов продаж сначала с коэффициентом $N=4$, затем - с коэффициентом $N=8$ и наконец, с коэффициентом $N=5$. Разделение сигнала на серию высокочастотных компонент позволяет выделить основные периоды колебаний объемов продаж, что невозможно при обычном спектральном анализе. Предложенные методы применимы и для анализа цен акций.

1. Введение. Как известно [1], [2] вейвлет-анализ успешно применяется при анализе сигналов: для выявления локализованных изменений сигнала; для сжатия сигнала и удаления шума; для обработки изображений. Обычно используются вейвлеты с параметром масштабирования $N = 2$. Однако для изучения массивов данных экономического происхождения более правильно использовать кратномасштабный вейвлет-анализ с параметром масштабирования $N > 2$. Обычными примерами являются массивы данных об объемах или ценах продаж товарного или фондового рынка. Такие временные ряды имеют периодичность по времени. Например, при анализе 15-минутных данных о цене акции естественно использовать вейвлеты с параметром масштабирования $N = 4$ (поскольку в часе таких

периодов – 4). Далее, при анализе часовых данных о цене ценной бумаги фондового рынка естественно использовать вейвлеты с параметром масштабирования $N = 8$ (поскольку день работы рынка содержит 8 таких отсчетов). При анализе дневных данных естественно использовать вейвлеты с параметром масштабирования $N = 5$ (поскольку неделя работы рынка содержит 5 рабочих дней). На товарном рынке при анализе дневных показателей продаж товара естественно использовать вейвлеты с параметром масштабирования $N = 7$ (недельный цикл) с отдельной выборкой данных за выходные дни.

В данной работе мы покажем на примере возможности N -кратномасштабного вейвлет-анализа для изучения объемов продаж акций на ММВБ. В качестве примера рассмотрим 15-минутные данные об объемах продаж акций компании Лукойл на ММВБ за 2006 год. Этот выбор объясняется тем, что объемы продаж акции за период (15 минут) ведут себя менее регулярно, чем цены акции. Найдем вейвлет-разложение временного ряда объемов продаж сначала с коэффициентом $N = 4$, затем – с коэффициентом $N = 8$ и, наконец, с коэффициентом $N = 5$. Будем использовать вейвлеты Хаара с параметром масштабирования N , построенные в работе [4]. Для численной обработки данных используем систему MATLAB. Анализ цены акции можно провести совершенно аналогично.

Перед использованием массивы данных были предварительно обработаны. Дело в том, что работа ММВБ начинается после 10.30 и заканчивается после 18.30. В первый час работы имеются два отсчета данных: в 10.30 и 10.45, в последний час работы – три отсчета: 18.00, 18.15 и 18.30. Таким образом,

первый и последний часы работы не содержат полных четырех периодов времени. Эту проблему можно решить разными способами. Например, можно продолжить массивы, повторив в 18.45 данные за 18.30, а в первый час – добавив два отсчета в 10.00 и 10.15 данными за 10.30. Тогда рабочий день будет состоять из 9-и часов работы. В работе принят следующий способ обработки. Недостающие данные в 18.45 заполнены данными в 10.45 следующего дня, а данные в 10.30 – удалены, поскольку продаж на этот момент не было. В этом случае рабочий день ММВБ будет состоять из 8-и часов работы. Кроме того, 6 неполных недель (которые образовались за счет праздничных дней) искусственно дополнены днями с данными за предыдущий рабочий день.

2. Вейвлет-анализ объемов продаж. Рассмотрим 15-минутные данные об объемах продаж акций компании Лукойл на ММВБ за 2006 год, предварительно обработанные указанным выше способом. Длина выбранного сигнала $L = 8160$. График сигнала, вместе с данными на закрытие Close, изображен на рис. 1. Обычный частотный анализ этих данных результатов не дает – низкие частоты подавляют все остальные. Для разделения сигнала на определенные частотные составляющие применим вейвлет-разложение и выделим из сигнала VOL высокочастотные составляющие – детализирующие коэффициенты разложения (будем в дальнейшем обозначать их символами cSign41, cSign42, cSign43) и низкочастотную компоненту, которую представляет массив cSign40 коэффициентов аппроксимации.

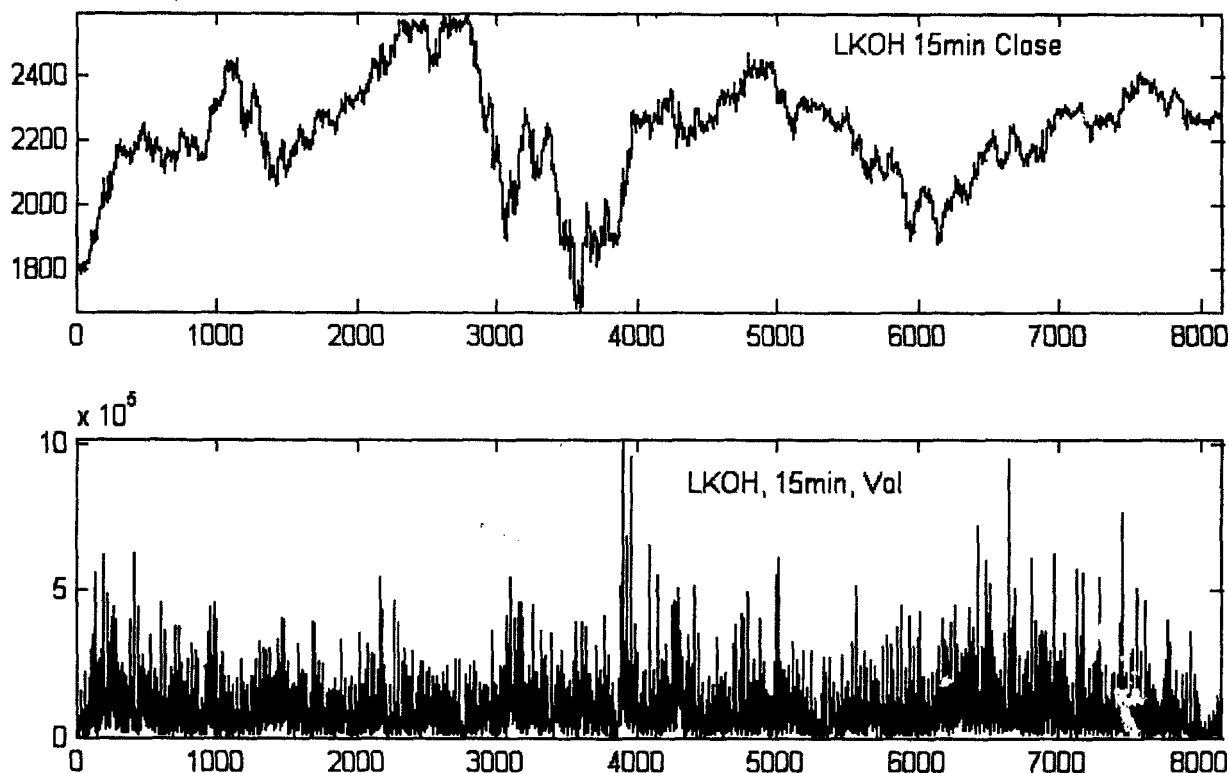
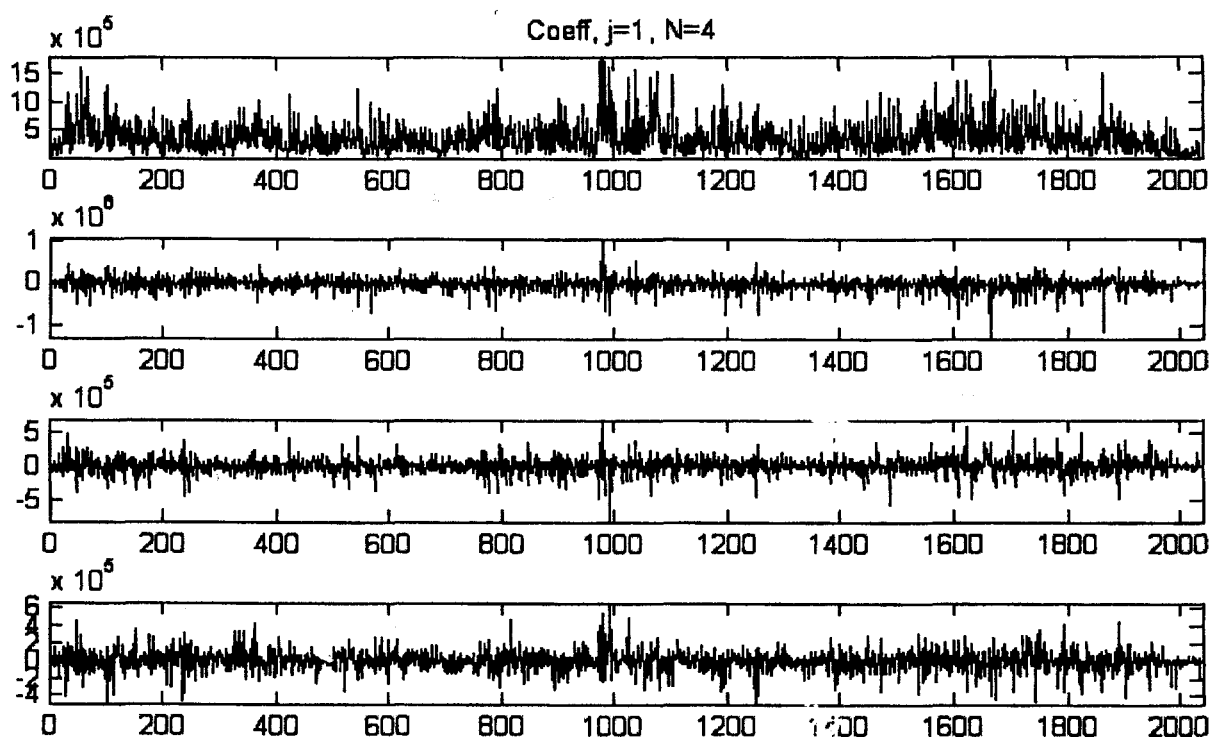


Рис. 1. Графики массива Vol вместе с данными цен на закрытие Close

Рис. 2. Графики коэффициентов разложения, $N = 4$

2.1. Первый уровень вейвлет-разложения, $N = 4$. Выбираем следующие фильтры $H40$, $H41$, $H42$, $H43$ вейвлетов Хаара:

$H40 = [1, 1, 1, 1]/2$; % Фильтр вейвлета ϕ .
 $H41 = [1, 1, -1, -1]/2$; % Фильтр вейвлета ψ^1 .
 $H42 = [1, -1, 1, -1]/2$; % Фильтр вейвлета ψ^2 .
 $H43 = [1, -1, -1, 1]/2$; % Фильтр вейвлета ψ^3 .

При получении коэффициентов разложения учетом специфику анализируемого массива. Когда мы переходим от 15-минутных данных на уровень часовых данных, то здесь следует брать не средние значения 15-минутных данных за час, а накапливать эти данные, чтобы получить объем продаж за один час. Поэтому мы будем умножать коэффициенты разложения на 2 и тогда смысл вейвлет-коэффициентов будет следующий:

- $cSign40$ – коэффициенты аппроксимации получаются суммированием четырех 15-минутных объемов, это объем продаж за час;
- $cSign41$ – первые детализирующие коэффициенты, это разность получасовых объемов продаж;
- $cSign42$ – вторые детализирующие коэффициенты, это есть сумма разностей объемов 15-минутных продаж в получасах;
- $cSign43$ – третьи детализирующие коэффициенты, это разность разностей 15-минутных объемов в получасах.

Вейвлет-коэффициенты отражают колебания объема продаж акции относительно объема продаж за период 1 час. Длина массивов коэффициентов равна 2040.

Вычислим статистические характеристики: стандартное отклонение Std , среднее абсолютное отклонение $Mean$ и максимум отклонения Max детализирующих коэффициентов $cSign41$, $cSign42$ и $cSign43$ с учетом поправочных коэффициентов с тем, чтобы данные имели денежный смысл. Данные (в тысячах рублей) приведены в табл. 1.

Таблица 1

Статистические характеристики вейвлет-коэффициентов

Коэффициенты	Std	$Mean$	Max
$cSign41$	147.030	97.705	1.324.202
$cSign42$	105.420	71.848	791.325
$cSign43$	108.750	75.748	666.511

Таким образом, колебание объема продаж акции в пределах 100 тыс. руб. в течение часа можно считать нормальным.

Сделаем дискретное преобразование Фурье вейвлет-коэффициентов, найдем интенсивности частотного спектра и изобразим их на графиках. Результаты отражены на рис. 3.

Частотные спектры вейвлет-коэффициентов показывают наличие выделяющихся частот в первом наборе детализирующих коэффициентов. Частота $n = 256$ соответствует периоду $T = 2040/256 = 8$ часов, частота $n = 511$, период $T = 2040/511 = 4$ часа и частота $n = 764$, период $T = 2040/764 = 2,67 = 8/3$ часа – третья часть рабочего дня.

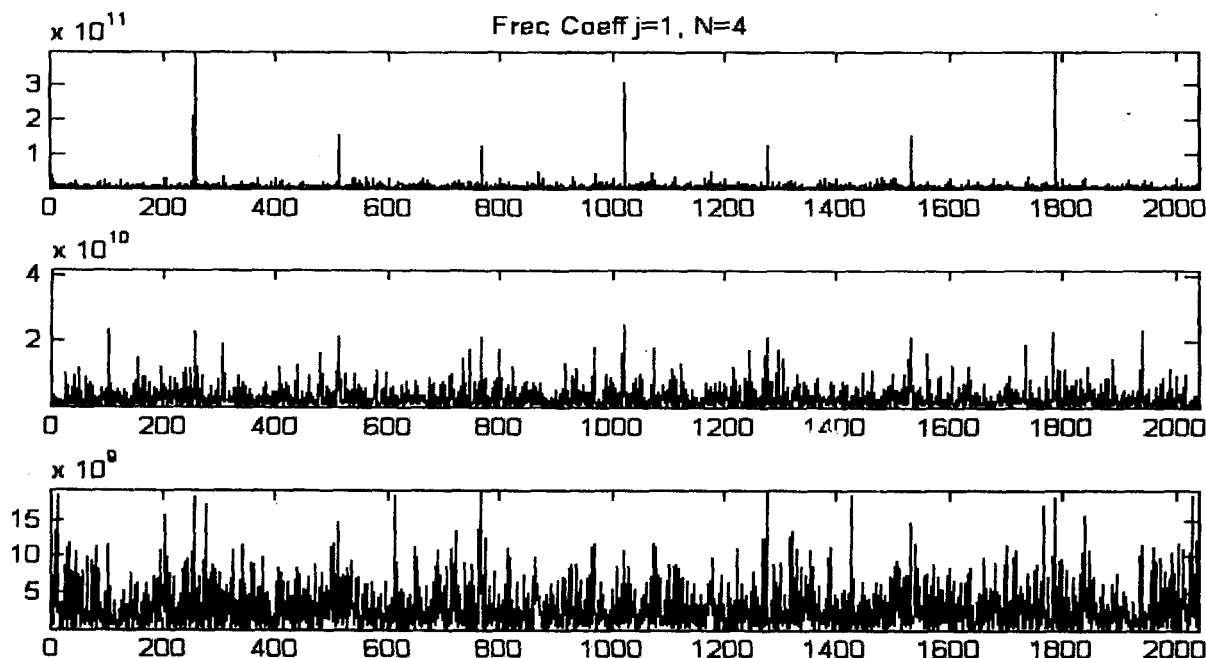


Рис. 3. Частотные спектры коэффициентов детализации

2.2. Второй уровень разложения, $N = 8$. Средние часовые данные. Будем раскладывать полученные при первом уровне разложения коэффициенты аппроксимации cSign40, длины $L_4=2040$, которые отражают средние за час объемы продаж акции. Поскольку рабочий день составляет 8 часов, то при последующем разложении естественно взять вейвлеты с параметром масштабирования $N=8$.

Выберем фильтры вейвлетов Хаара $H80, H81, H82, H83, H84, H85, H86, H87$ следующим образом:
 $H80 = [1, 1, 1, 1, 1, 1, 1, 1] / (8^{1/2})$; % Фильтр вейвлета ψ^0 .

$H81 = [1, -1, 0, 0, 0, 0, 0, 0] / (2^{1/2})$; % Фильтр вейвлета ψ^1 .

$H82 = [1, 1, -2, 0, 0, 0, 0, 0] / (6^{1/2})$; % Фильтр вейвлета ψ^2 .

$H83 = [1, 1, 1, -3, 0, 0, 0, 0] / (12^{1/2})$; % Фильтр вейвлета ψ^3 .

$H84 = [1, 1, 1, 1, -4, 0, 0, 0] / (20^{1/2})$; % Фильтр вейвлета ψ^4 .

$H85 = [1, 1, 1, 1, 1, -5, 0, 0] / (30^{1/2})$; % Фильтр вейвлета ψ^5 .

$H86 = [1, 1, 1, 1, 1, 1, -6, 0] / (42^{1/2})$; % Фильтр вейвлета ψ^6 .

$H87 = [1, 1, 1, 1, 1, 1, 1, -7] / (56^{1/2})$; % Фильтр вейвлета ψ^7 .

При вейвлет-разложении весь сигнал разбивается на восьмерки подряд идущих элементов, и коэффициенты получаются различными комбинациями элементов восьмерки. Количество вейвлет-коэффициентов $L_8 = 255$. Вейвлет-коэффициенты отражают следующие характеристики (с точностью до коэффициентов):

- коэффициенты аппроксимации cSign80 – сумма восьми часовых объемов продаж, это объем за день;

- детализирующие коэффициенты cSign81 – разность объемов за первый и второй час работы;

- детализирующие коэффициенты cSign82 – разность между средним значением объемов за два первых часа работы и объемом продаж третьего часа;

- детализирующие коэффициенты cSign83 – разность между средним значением объемов за три первых часа работы и объемом продаж четвертого часа;

- детализирующие коэффициенты cSign84 – разность между средним значением объемов за четыре первых часа работы и объемом продаж пятого часа;

- детализирующие коэффициенты cSign85 – разность между средним значением объемов за пять первых часов работы и объемом продаж шестого часа;

- детализирующие коэффициенты cSign86 – разность между средним значением объемов за шесть первых часов работы и объемом продаж седьмого часа;

- детализирующие коэффициенты cSign87 – разность между средним значением объемов за семь первых часов работы и объемом продаж восьмого часа.

Выбор вейвлетов Хаара объясняется именно тем, что можно понять смысл коэффициентов разложения. При других вейвлетах, возможно, получатся более интересные результаты.

Вычислим статистические характеристики высокочастотных компонент сигнала: стандартное отклонение, среднее абсолютное значение и максимальное значение. Высокочастотная компонента сигнала получается при восстановлении сигнала только по выбранным детализирующим коэффициентам при нулевых других коэффициентах. Данные (в тысячах рублей) приведены в табл. 2.

Таблица 2

Статистические характеристики высокочастотных компонент

	Sign81	Sign82	Sign83	Sign84	Sign85	Sign86	Sign87
Std	54.340	65.750	59.960	57.440	65.160	73.240	136.820
Mean	20.114	29.449	30.364	28.317	30.993	33.983	70.292
Max	481.700	697.000	409.500	722.500	626.900	1.066.200	1.278.000

Выводы

Из приведенных в таблице результатов следует, что колебание объемов продаж акции в пределах от 50 до 100 тысяч рублей в час, можно считать нормальным. Наибольшие отклонения у компоненты cSign87 – она отражает разность между средним значением за семь предыдущих часов работы и объемами в последний час работы. Наибольшие колебания объемов продаж акции происходят в последний час работы. В целом колебания объемов стали меньше, чем на первом уровне разложения, потому, что теперь за основу взяты усредненные часовые данные.

Интерес представляют частотные характеристики компонент сигнала, т. е. компонент, восстановленных только по одному набору детализирующих коэффициентов. Сделаем дискретное преобразование Фурье, найдем интенсивности частотного спектра и изобразим их на графиках (рис. 4).

Частотный спектр вейвлет-коэффициентов показывает наличие тех же основных частот в большинстве компонент сигнала. Частота $n = 256$ соответствует периоду $T = 2040/256 = 8$ часов, частота $n = 511$, период $T = 2040/511 = 4$ часа и частота $n = 766$, период $T = 2040/766 = 2,66 = 8/3$ часа – третья часть рабочего дня.

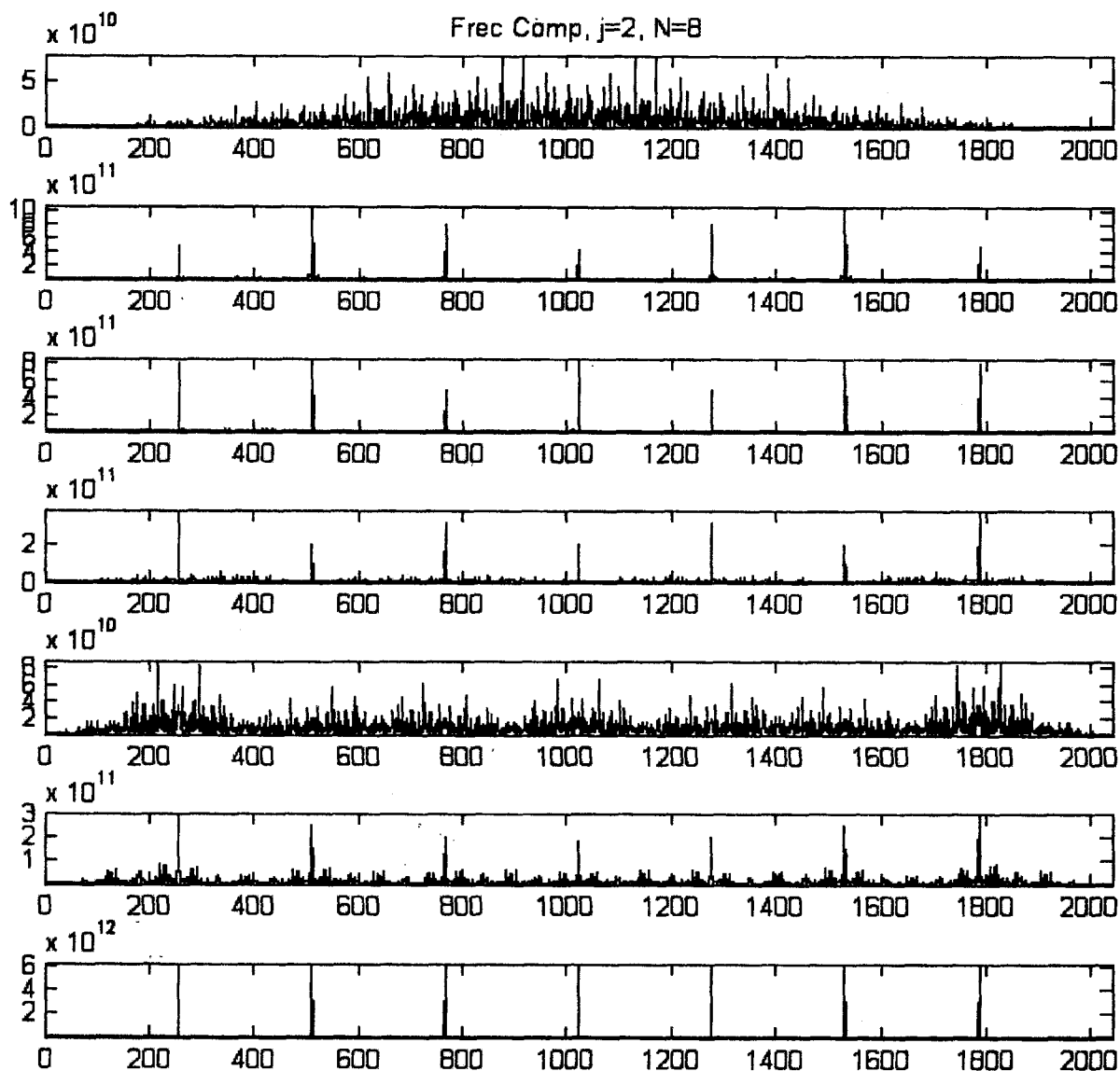


Рис. 4. Частотный спектр компонент детализации

Вейвлеты Хаара выбираются неоднозначно [4]. При другом выборе фильтров $H81$, $H82$, $H83$, $H84$, $H85$, $H86$, $H87$ мы получим другой смысл детализирующих коэффициентов. Например, можно создать фильтры для $N = 8 = 2^3$ из фильтров обычных вейвлетов Хаара,

$H80 = [1, 1, 1, 1, 1, 1, 1, 1]/(8^{1/2});$ % Фильтр вейвлета ψ^0 .

$H81 = [1, -1, 1, -1, 1, -1, 1, -1]/(8^{1/2});$ % Фильтр вейвлета ψ^1 .

$H82 = [1, 1, -1, -1, 1, 1, -1, -1]/(8^{1/2});$ % Фильтр вейвлета ψ^2 .

$H83 = [1, 1, 1, 1, -1, -1, -1, -1]/(8^{1/2});$ % Фильтр вейвлета ψ^3 .

$H84 = [1, 1, -1, -1, -1, 1, 1, 1]/(8^{1/2});$ % Фильтр вейвлета ψ^4 .

$H85 = [1, -1, 1, -1, -1, 1, 1, 1]/(8^{1/2});$ % Фильтр вейвлета ψ^5 .

$H86 = [1, -1, -1, 1, 1, -1, -1, 1]/(8^{1/2});$ % Фильтр вейвлета ψ^6 .

$H87 = [1, -1, -1, 1, -1, 1, 1, -1]/(8^{1/2});$ % Фильтр вейвлета ψ^7 .

В этом случае получается частотный спектр компонент сигнала, изображенный на рис. 5.

Отметим более четкое разложение сигнала $s\text{Sign}40$ по частотам. Частотный спектр вейлет-коэффициентов показывает наличие тех же выделенных частот в большинстве компонент сигнала. Частота $n = 256$ соответствует периоду $T = 2040/256 = 8$ часов, частота $n = 511$, период $T = 2040/511 = 4$ часа и частота $n = 766$, период $T = 2040/766 = 2,66 = 8/3$ часа – третья часть рабочего дня.

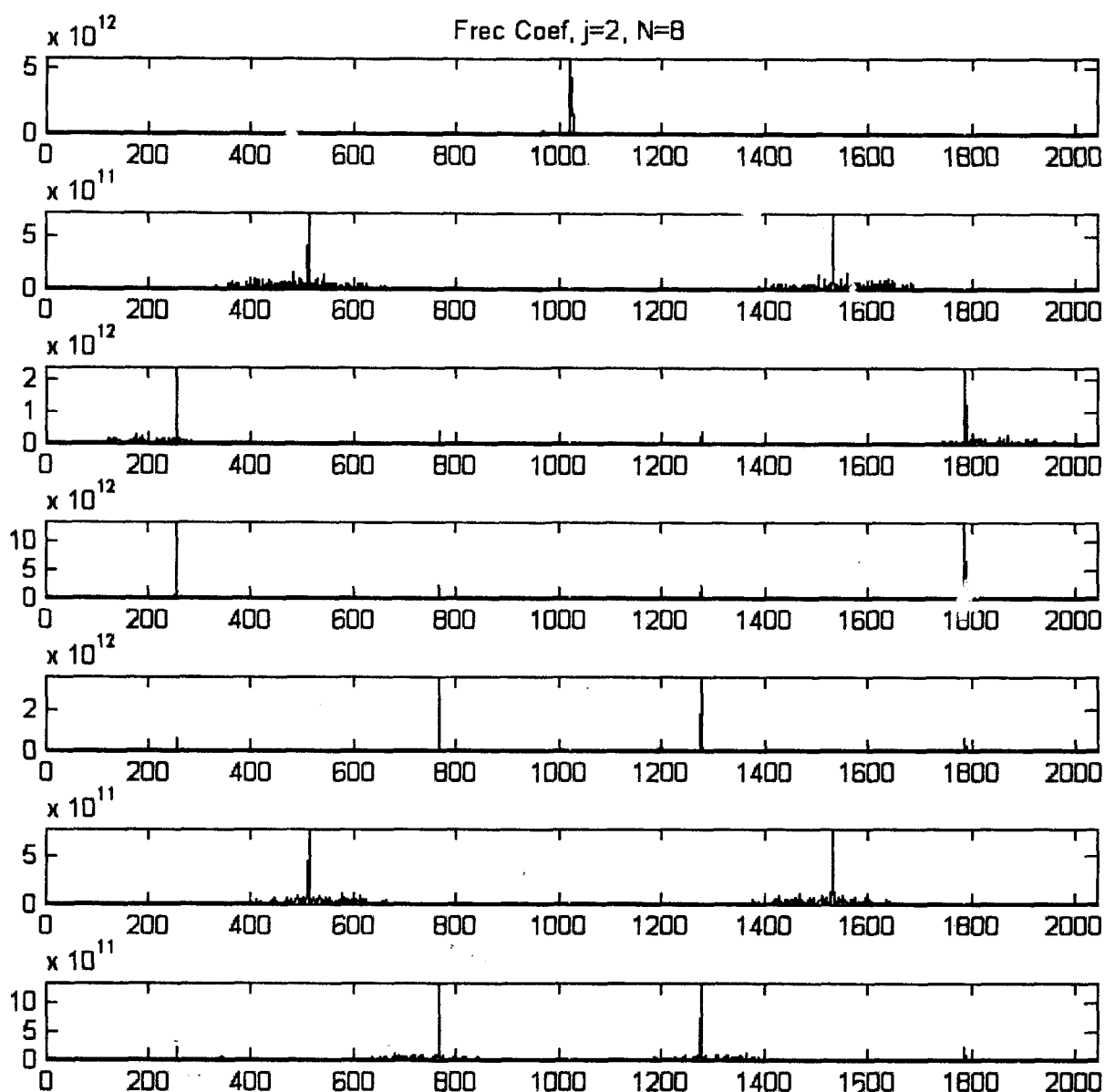


Рис. 5. Частотный спектр компонент детализации при другом выборе фильтров

2.3. Третий уровень разложения, $N = 5$. Средние дневные данные. Рассматриваем коэффициенты аппроксимации cSign80, длины $L5 = 255$, которые отражают средние за день объемы продаж акции. Поскольку неделя составляет 5 рабочих дней, то при разложении естественно взять вейвлеты с параметром масштабирования $N = 5$.

Выбираем фильтры $H50, H51, H52, H53, H54$ вейвлетов Хаара,

$H50 = [1, 1, 1, 1, 1] / (N5^{1/2});$ % Фильтр вейвлета ϕ .

$H51 = [1, -1, 0, 0, 0] / (2^{1/2});$ % Фильтр вейвлета ψ^1 .

$H52 = [1, 1, -2, 0, 0] / (6^{1/2});$ % Фильтр вейвлета ψ^2 .

$H53 = [1, 1, 1, -3, 0] / (12^{1/2});$ % Фильтр вейвлета ψ^3 .

$H54 = [1, 1, 1, 1, -4] / (20^{1/2});$ % Фильтр вейвлета ψ^4 .

В данном случае весь сигнал разбивается на пятерки подряд идущих элементов и коэффициенты получаются различными комбинациями элементов. Вейвлет-коэффициенты имеют аналогичный смысл, что и при разложении часовых данных.

На рис. 6 изображены графики частотных спектров компонент разложения.

Частотный спектр компонент разложения показывает наличие выделенных частот в большинстве компонент сигнала. Частота $n = 52$ соответствует периоду $T = 255/52 = 4,9$ дней (неделя), частота $n = 103$, период $T = 255/103 = 2,475$ дней (половина рабочей недели).

Литература

1. Добеши, И. Десять лекций по вейвлетам / И. Добеши. – М.; Ижевск: РХД, 2001. – 494 с.
2. Смоленцев, Н. К. Основы теории вейвлетов. Вейвлеты в MATLAB / Н. К. Смоленцев. – М.: ДМК Пресс, 2005. – 303 с.
3. Подкур, И. Н. О некоторых типах вейвлетов с параметром масштабирования 3 // Вестник КемГУ. – Вып. 3(27). – 2006. – С. 21 – 25.
4. Smolentsev, N. K. Construction of some types wavelets with coefficient of scaling N / N. K. Smolentsev, P. N. Podkur // <http://arXiv.org/math.FA/0612573v1>, 2006. – 19 p.

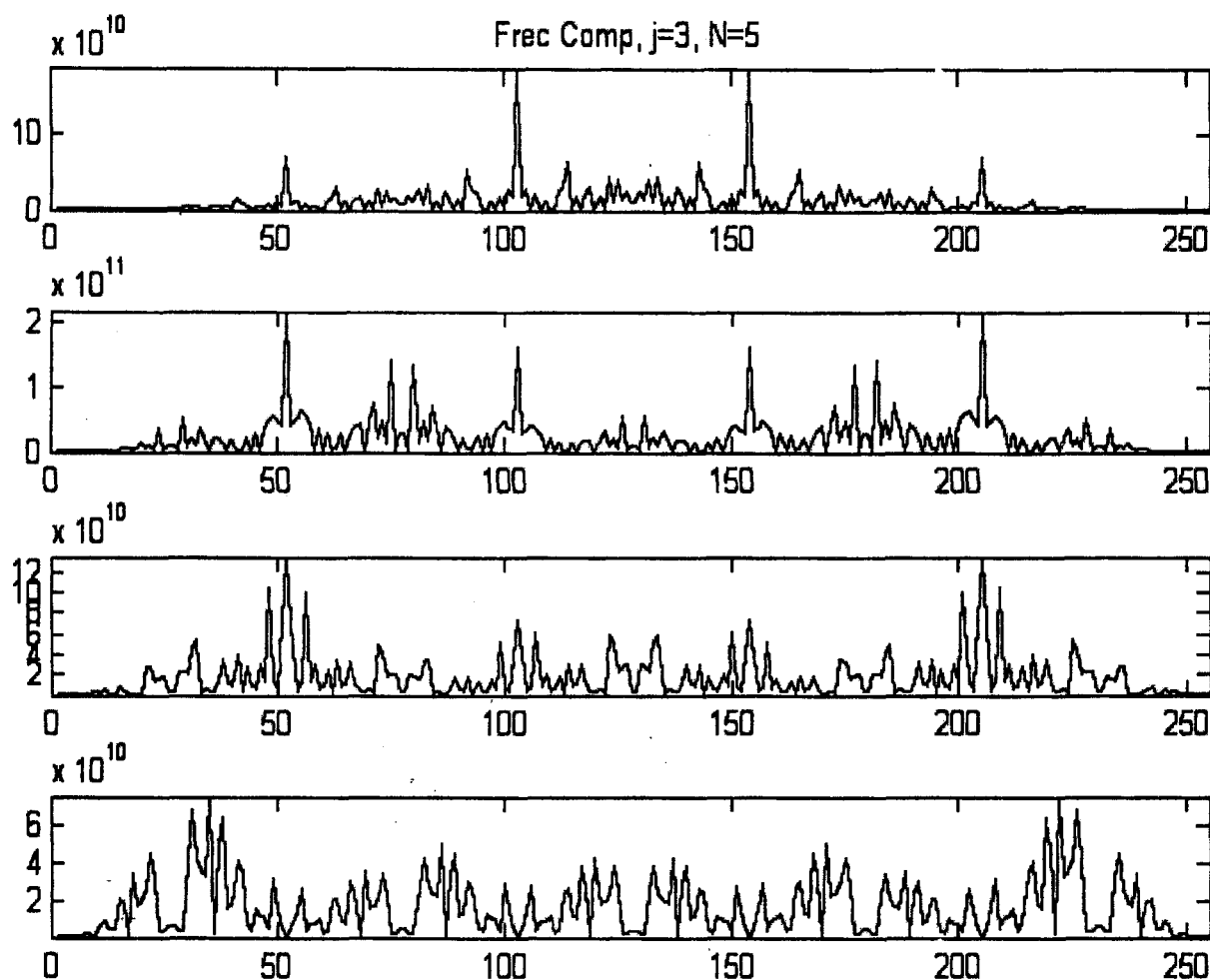


Рис. 6. Частотный спектр компонент разложения