

Методы и средства обработки сигналов и изображений

УДК 681.325

А.В. Волошко, канд. техн. наук, Д.О. Иванько, К.Ю. Гура, Д.К. Мищенко

К вопросу применения кратномасштабного анализа для сжатия и восстановления графика электрической нагрузки

Проведен анализ применения кратномасштабного анализа для сжатия и восстановления графика электрической нагрузки предприятия. Показано, что график нагрузки может быть сжат и восстановлен с необходимой точностью в зависимости от уровня его декомпозиции.

The application of multiresolution analysis for compression and reconstruction of the enterprise electric loading graph is considered. It is shown that the graph of loading can be compressed and reconstructed with the necessary accuracy depending on the level of its decomposition.

Введение

Масштабное внедрение автоматизированных систем учета, контроля и управления электропотреблением приводит к существенному возрастанию объема информации, подлежащей хранению, обработке и исследованию. Это, в свою очередь, обуславливает необходимость увеличения вычислительных, коммуникационных ресурсов и разработки действенных алгоритмов сжатия информации с последующей возможностью ее восстановления [1,7].

График электрических нагрузок является одним из основных исходных данных при проектировании, использовании и управлении энергохозяйством. В связи с этим представляется необходимым разработка алгоритма его сжатия и восстановления с необходимой точностью в зависимости от решаемых задач.

Все алгоритмы сжатия данных делятся на две группы [6]: кодирование без потери (когда исходные данные восстанавливаются полностью в результате процедуры декодирования) и с потерей информации. Алгоритмы первой группы часто называют архивацией данных. Для второй группы алгоритмов сжатия данных вводится понятие «качество сжатия», которое задается параметром с возможными значениями от 0 до 100, где 100 обозначает минимальное сжатие (наилучшее качество) и восстановленное значение данных практически неотличимо от исходного, тогда как 0 – максимальное сжатие, при котором

в восстановленных данных все еще можно различить основные характерные значения.

Кратномасштабный анализ

Кратномасштабный анализ (multiresolution analysis), или КМА основан на представлении пространства сигналов U в виде системы вложенных подпространств U_j , отличающихся друг от друга только перемасштабированием независимой переменной. Данный вид анализа базируется на следующих предпосылках [6]:

- Пространство сигналов U может быть разбито на иерархически вложенные подпространства U_j , которые не пересекаются и объединение которых дает в пределе $L^2(R)$;
- Для любой функции $s(t) \in U_j$ ее сжатая версия принадлежит пространству U_{j-1} ;
- Существует такая функция $\phi(x) \in U_0$, для которой ее сдвиги $\phi_{0,k}(t) = \phi(t-k)$ при $k \in Z$ образуют ортонормированный базис пространства U_0 .

Так как функции $\phi_{0,k}(t)$ образуют ортонормированный базис пространства U_0 , то следующие функции образуют ортонормированный базис пространства U_j :

$$\phi_{j,k}(t) = 2^{-j} \phi(2^{-j}t - k) \quad (1)$$

Эти функции (ϕ – функции) называются масштабирующими именно потому, что они создают свои масштабированные вершины в пространстве сигнала. При этом сигнал $s(t)$ может быть представлен множеством последовательных приближений $s_j(t)$ в субпространствах U_j . Переменная j при этом называется масштабным коэффициентом. Если дерево декомпозиции сигнала отсчитывать вниз, то можно сказать, что сигнал $s(t)$ есть предел аппроксимации $s_j(t) \in U_j$ при $j \rightarrow -\infty$, т.е. можно записать:

$$s(t) = \lim_{j \rightarrow -\infty} S_j(t) \quad (2)$$

В соответствии с (2) при больших j мы получаем грубые приближения сигнала, а при малых – точные. Приближение (аппроксимация) сигнала соответствует итерационной формуле:

$$s_j(t) = \sum_j c(j,k) \cdot \phi_{j,k}(t) \quad (3)$$

причем

$$\phi_{o,o}(t) = 2 \sum_k h_k \cdot \phi(2t - k), \quad (4)$$

где h_k – некоторая последовательность. Сумма приближенной и детализирующей составляющих в конечном счете и дает исходный сигнал с тем или иным приближением.

Рассмотрим применение КМА для анализа графика электрических нагрузок с применением функции Хаара (Haar) [4]. Функция Хаара соответствует прямоугольному весовому окну и записывается в следующем виде:

$$\phi_{j,k}(t) = p \cdot \phi(2^j \cdot t - k), \quad (5)$$

где $p=2^{j/2}$, коэффициент ортонормирования, который обеспечивает единичную норму т.н. скейлинг-функции.

Пусть имеем почасовой суточный (24 часа) график электрической нагрузки предприятия $\{p(t)\}$ (рис. 1). Данный график представлен 24 значениями, с равномерным интервалом дискретизации (1 час). Число ступеней графика нагрузки 24 не соответствует целому числу степени 2. Поэтому недостающее значение графика $(25 \div 31)$ дополним нулевыми значениями (при этом масштабный коэффициент $j=5$) [9].

Математически данную функцию (график нагрузки) можно представить в виде ряда с разложением по системе ортонормированных функций $\phi_{j,k}(t)$, образованных прямоугольным весовым окном ϕ шириной $\Delta t = 1/2^j$, ортогональность которых обеспечивается последовательным сдвигом функций друг относительно друга на величину Δt . При сдвиговой ортогональности прямоугольных базисных функций разложения прямое преобразование (проекции сигнала на базис) выполняется по формуле:

$$C_{j,k} = \frac{1}{P} \int_{k \cdot \Delta t}^{(k+1) \cdot \Delta t} p(t) \cdot \phi(2^j t - k) dt. \quad (6)$$

В нашем случае для дискретного задания графика по интервалам $\Delta t = \frac{1}{2^j}$, максимальная детализация соответствует $j=5$. То есть:

$$C_{j,k} = \frac{p(k \Delta t)}{\sqrt{2^j}}. \quad (7)$$

Восстановление графика нагрузки с пятого уровня декомпозиции выполняется по формуле:

$$P(5, t) = \sum_{k=0}^{31} C_{5,k} \cdot \phi_{5,k}(t). \quad (8)$$

При этом функция $\phi_{5,k}$ исходя из формулы (1) примет значение:

$$\phi_{5,k} = \sqrt{2^5} \cdot \phi(2^5 \cdot k \cdot \Delta t - k) = \sqrt{2^5}, \quad (9)$$

для $k=0..31$.

В связи с тем, что график нагрузки задан дискретно, его восстановление скейлинг функцией Хаара (умножением константы шириной Δt на соответствующее значение $C_{5,k}$) производится без погрешности.

На следующем уровне декомпозиции графика нагрузки, при $j=4$, скейлинг-функция расширяется по t до $1/16$ путем усреднения отсчетов по двум соседним интервалам исходного графика. Количество коэффициентов соответственно уменьшается в два раза. Расчет коэффициентов $C_{4,k}$ выполняется либо по формуле (6), либо по известным значениям коэффициентов предыдущего уровня декомпозиции $C_{5,k}$, с учетом изменения нормировочного множителя $1/p$ в формуле (6):

$$C_{j-1,k} = \frac{1}{\sqrt{2}} (C_{j,2k} + C_{j,2k+1}), \quad (10)$$

$$t.e. \quad C_{4,k} = \frac{1}{\sqrt{2}} (C_{5,2k} + C_{5,2k+1}). \quad (11)$$

Результат расчета аппроксимирующих коэффициентов данного уровня декомпозиции, представлен на рис. 2 ($j=4$). Подобным образом проведена декомпозиция последующих уровней при $j=3, 2, 1, 0$. Результаты вычислений представлены соответствующими графиками на рис. 2.

Как известно, произвольный информационный сигнал в теории цифровых временных рядов обычно рассматривается в виде суммы разнотипных составляющих [3]:

- изменение функции тренда средних значений по большему интервалу усреднения;
- циклические компоненты с определенным периодом повторения;
- локальные особенности разного порядка, вызванные резкими изменениями, флуктуациями, разрывами).

Выделение из исходного графика нагрузки коэффициентов его изменения в пределах нового интервала усреднения (локальных особенностей), называемых детализирующими, производится по формуле:

$$d_{j-1,k} = \frac{1}{\sqrt{2}} (C_{j,2k} - C_{j,2k+1}). \quad (12)$$

То есть, детализирующие коэффициенты определяются как разность значений аппроксимирующих коэффициентов первой и второй половины интервала, и представляют собой разность между исходным сигналом $p(k \cdot \Delta t)$ и аппроксимированным графиком $p(4; 2 \cdot k \cdot \Delta t)$.

На рис.2 представлены графики детализирующих коэффициентов ($j=4, 3, 2, 1, 0$), т.е. реконструкция информационных частей графика нагрузки (флуктуаций относительно аппроксимационной части) по формуле (12).

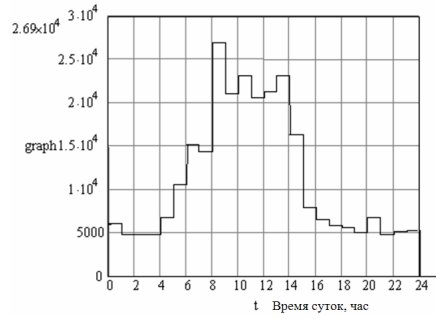


Рис. 1. Почасовой график электрической нагрузки предприятия

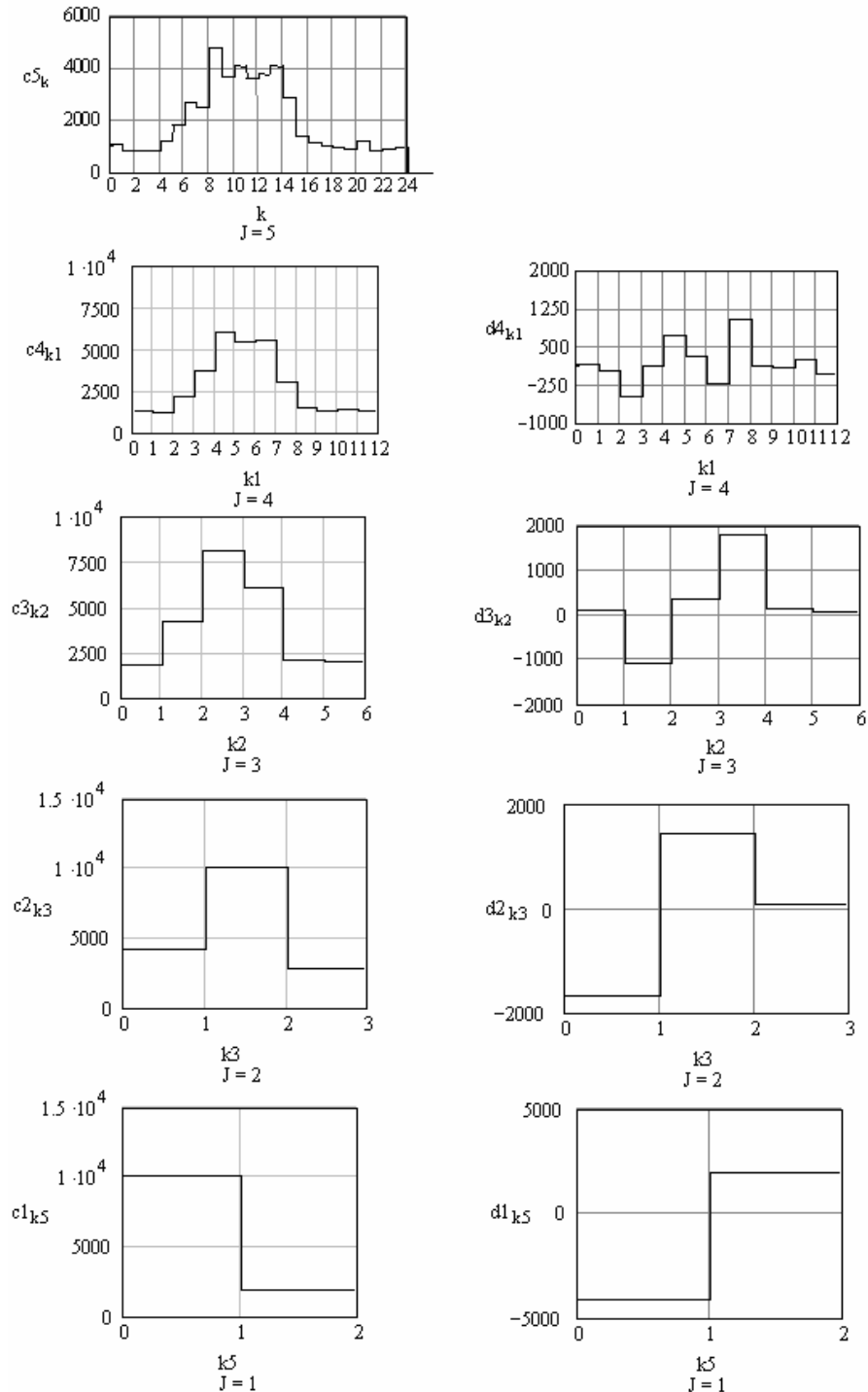


Рис. 2. Графики аппроксимирующих c_j и детализирующих d_j коэффициентов

Анализ графиков, представленных на рис.2, позволяет определить локальные особенности графика нагрузки по интервалу его задания. Значение коэффициентов более высокого уровня j восстанавливаются парой коэффициентов $C_{j-1,k}$ и $d_{j-1,k}$:

$$C_{j,2k} = \frac{1}{\sqrt{2}}(C_{j-1,k} + d_{j-1,k}), \quad (13)$$

$$C_{j,2k+1} = \frac{1}{\sqrt{2}}(C_{j-1,k} - d_{j-1,k}),$$

и, соответственно, по ним, возможно восстановление исходного графика нагрузки. Необходимо обратить внимание на тот факт, что для восстановления значений в исходных интервалах значение аппроксимирующего коэффициента на первой половине интервала $C_{j-1,k}$ складывается с детализирующим коэффициентом $d_{j-1,k}$, а на второй половине – вычитается.

Для математического отображения данной операции введена функция ψ , вид которой определяется следующей формулой [2]:

$$\Psi_{j,k}(t) = p\Psi(2^j t - k), \quad p = 2^{j/2}. \quad (14)$$

Данная функция является ортонормированным базисом разложения детализирующих коэффициентов, и ее использование позволяет проводить восстановление графика нагрузки с любого уровня детализации по формуле:

$$P(j, t) = \sum_{k=0}^{2^j-1} c_{j,k} \cdot \phi_{j,k}(t) + \sum_{k=0}^{2^j-1} d_{j,k} \cdot \Psi_{j,k}(t),$$

где $\phi_{j,k}(t)$ и $\psi_{j,k}(t)$ – ортонормированные базисы разложения аппроксимирующих и детализирующих коэффициентов соответственно.

В общем виде, начиная с последнего уровня разложения ($j=0$), полное восстановление графика электрической нагрузки производится по формуле:

$$P(t) = C_0 \cdot \phi_0(t) + d_0 \cdot \psi_0(t) + \sum_{k=0}^1 d_{1,k} \cdot \Psi_{1,k}(t) + \sum_{k=0}^3 d_{2,k} \cdot \Psi_{2,k}(t) + \sum_{k=0}^7 d_{3,k} \cdot \Psi_{3,k}(t) + \sum_{k=0}^{15} d_{4,k} \cdot \Psi_{4,k}(t) \quad (15)$$

То есть, если выполняется разложение до масштаба n , то для полного восстановления графика нагрузки (любой функции) необходимо сохранить только вектор аппроксимирующих коэффициентов на масштабе n и детализирующих коэффициентов на остальных масштабах. Восстановление графика производится по схеме [4] и представляет собой выполнение свертки аппроксимирующего и детализирующего набора

коэффициентов с фильтрами низкой и высокой частоты.

Анализ формулы (15) показывает, что:

1) Значение аппроксимирующего коэффициента C_0 представляет собой среднее значение графика электрической нагрузки P_{cp} на интервале его задания.

2) Аппроксимирующие коэффициенты соответствуют «низкочастотной» части восстанавливаемого графика нагрузки, в которой, как правило, сосредоточено большее количество полезной информации.

3) Детализирующие коэффициенты соответствуют высокочастотной части графика нагрузки и, имеющие, как правило, относительно небольшие значения.

4) Точность и количество информации, необходимые для хранения и восстановления графика электрических нагрузок зависит от степени детализации, поэтому, обнуляя часть детализирующих коэффициентов, можно с необходимой точностью хранить и восстанавливать значения графика нагрузок.

Погрешность восстановления графика электрических нагрузок

Анализ проведем в зависимости от метода фильтрации детализирующих коэффициентов и уровня детализации.

Для восстановления графика нагрузки с помощью пороговой фильтрации коэффициентов детализации используем методы так называемой «мягкой» и «жесткой» пороговой фильтрации. При «мягком» методе коэффициенты, абсолютные значения которых меньше порогового, обнуляются, а остальные – «подтягиваются» к нулевому значению на величину порога θ :

$$y = \begin{cases} x + \theta, & \text{если } x < 0 \text{ и } |x| > \theta \\ x - \theta, & \text{если } x > 0 \text{ и } |x| > \theta \\ 0, & \text{если } x > 0 \text{ и } |x| \leq \theta \end{cases} \quad (16)$$

где θ - пороговое значение, x – значение детализирующего коэффициента до фильтрации, y - значение детализирующего коэффициента после фильтрации.

При «жесткой» фильтрации значения детализирующих коэффициентов, которые по абсолютной величине меньше порогового значения отбрасываются.

Процесс пороговой фильтрации производится в несколько этапов. Вначале определяется среднее значение детализирующих коэффициентов на каждом уровне детализации. Затем определяются модифицированные детализирующие коэффициенты, отбрасываются приравненные к нулю и производится восстановление

графика электрических нагрузок. Расчет погрешности восстановления графика нагрузки (рис. 3) проведем в соответствии с чебышевской метрикой по формуле:

$$dif = \frac{\sum_{k=0}^{23} |graph_k - g1_k|}{\sum_{k=0}^{23} |graph_k|}, \quad (17)$$

где $graph_k$ – значение ординаты исходного графика, $g1_k$ – значение ординаты восстановленного графика.

В таблице 1 приведены результаты расчетов восстановления графика электрической нагрузки при пороговой фильтрации детализирующих коэффициентов.

Коэффициент сжатия определяется по следующей формуле [5]:

$$K_{сжат} = \frac{W_{исх}}{W_{сжат}}$$

где $W_{исх}$ – количество исходной информации; $W_{сжат}$ – количество сжатой информации.

Экономия памяти определяется исходя из того, что для хранения одного значения ординаты графика электрических нагрузок резервировалось 2 байта.

Относительная погрешность значения мощности и энергии определялось по формулам [8].

Анализируя данные результаты можно отметить следующее. Применение «мягкой» пороговой фильтрации более эффективно по сравнению с «жесткой» фильтрацией. Так, при относительной погрешности восстановления 0,02 экономия памяти для хранения графика электрической нагрузки составляет 30%. Следует также отметить

тот факт, что скорость передачи данных о графике нагрузки увеличивается в 1,6 раза.

Таблица 1. Восстановление графика электрической нагрузки с помощью «мягкой» и «жесткой» пороговой фильтрации

Параметр	Вид фильтрации	
	«мягкая»	«жесткая»
Относительная погрешность восстановления значения мощности, %	0,02	0,479
Экономия памяти, %	30	46
Коэффициент сжатия, раз	1,6	1,85
Относительная погрешность восстановления энергии, %	$2,6 \cdot 10^{-3}$	8,26

На рис.4. отчетливо видно, что на «гладких» участках графика нагрузок погрешность восстановления равномерная. Это явно указывает на зависимость точности восстановления (ко всему прочему) и от формы графика нагрузки.

Проанализируем погрешность восстановления графика нагрузки с различных уровней детализации. Восстановление с третьего уровня детализации ($j=3$) производится обнулением детализирующих коэффициентов предыдущего ($j=4$) уровня детализации и так до первого уровня. На рис. 5. представлены графики относительных погрешностей восстановления графика электрических нагрузок с различных уровней детализации.

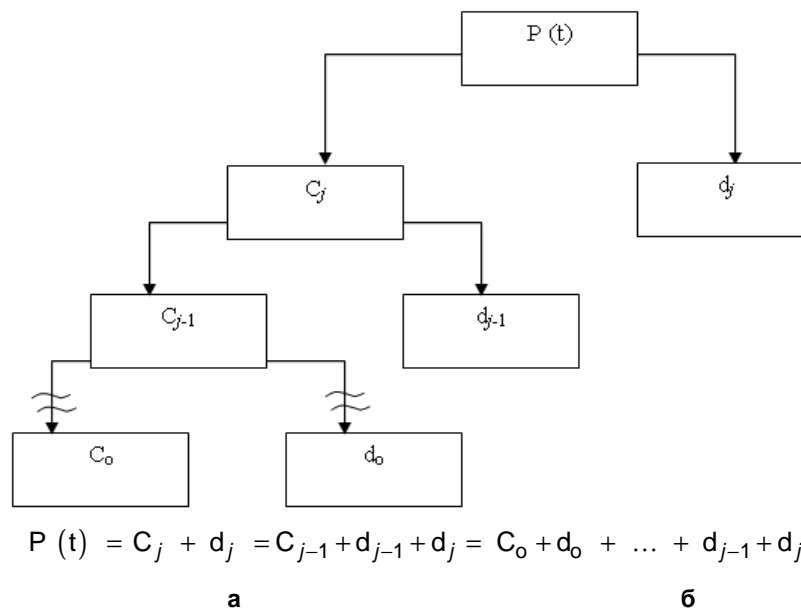


Рис. 3. Структурная схема кратномасштабного разложения (а) и восстановления (б) графика нагрузки. j – уровень разложения (детализации)

Анализ результатов расчета (таблица 2) показывает, что использование методики обнуления детализирующих коэффициентов позволяет эффективно реализовывать сжатие с потерями (т.е. восстановленный график нагрузки отличается от исходного в заданных пределах).

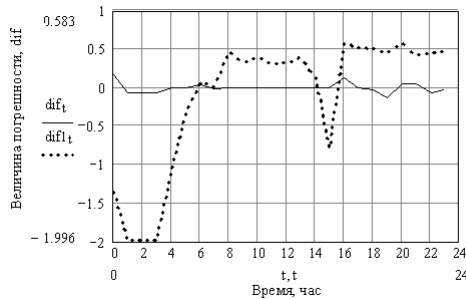


Рис. 4. Графики относительных погрешностей восстановления графика электрических нагрузок

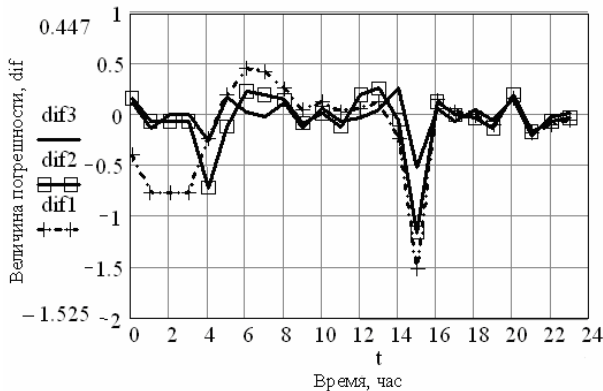


Рис. 5. Графики относительных погрешностей восстановления графика электрических нагрузок с различных уровней детализации

Таблица 2. Восстановление графика электрической нагрузки с различных уровней детализации

Параметр	Уровень детализации			
	3	2	1	0
Относительная погрешность восстановления значения мощности, %	0,107	0,172	0,243	0,479
Экономия памяти, %	59,0	71,0	84,0	91,0
Коэффициент сжатия, раз	2,4	3,4	6,0	12,0
Погрешность восстановления энергии, %	$2,6 \cdot 10^{-3}$	$2,8 \cdot 10^{-3}$	$2,9 \cdot 10^{-3}$	8,26

Выводы

1) Величина погрешности восстановления графика электрических нагрузок зависит от уровня декомпозиции, вида пороговой фильтрации, величины детализирующих коэффициентов и формы графика нагрузки.

2) Анализ и передачу данных об электропотреблении можно осуществлять последовательно, начиная с самых грубых уровней, несущих в предельно компактной форме основной объем информации, с дальнейшим последовательным (при необходимости) уточнением, вплоть до полного восстановления. На нужном уровне детализации (точности) передачу данных можно прекратить и перейти к следующему пакету передачи данных.

3) Свойство ортогональности кратномасштабного анализа позволяет получать независимую информацию на разных масштабах, а нормируемость обеспечивает сохранение величины информации на различных этапах детализации.

Литература

1. Волошко А.В. Метод формирования признаков классификации графиков электрических нагрузок на основе вейвлет преобразования // Промэлектро. – 2009. – №1. – С.39-43.
2. Горошко И.О. Метод спектральной реконструкции и коррекции в задачах восстановления сигналов и данных // Электронное моделирование. – 2005. – Т. 27. – №3 – С. 67-73.
3. Дремин И.М., Иванов О.В., Нечитайло В.А. Вейвлеты и их использование // Успехи физических наук. – 2001. – Т.17. – №5. – С. 465-501.
4. Засядько А.А., Олецкий А.В. Использование метода собственных функций при решении задачи восстановления сигналов // Электронное моделирование. – 2003. – Т. 25. – №1. – С. 107-113.
5. Королев А.В., Баранник В.В. Оценка степени сжатия изображений // Электронное моделирование. – 2002. – Т. 24. – №4. – С. 33-42.
6. Малла С. Вейвлеты в обработке сигналов: Пер. с англ. – М.: Мир. 2005. – 671 с.
7. Петухов А.П. Введение в теорию базисов всплесков. – СПб.: Изд. СПбГТУ. 1999. – 132 с.
8. П.П. Орнажский. Автоматические измерения и приборы (аналоговые и цифровые). – К.: Вища шк., 1986 – 504 с.
9. Воробьев В.И., Грибунин В.Г. Теория и практика вейвлет преобразование. СПб.: ВУС, 1999-2008 с.