

Вейвлет-обработка временных рядов для повышения точности представления информации

Ю.А. Кропотов¹, А.Ю. Проскуряков¹, А.А. Белов¹

¹Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», Орловская 23, Муром, Россия, 602264

Аннотация. В работе рассмотрен метод понижения погрешности представления информации в телекоммуникационных системах мониторинга путем фильтрации шумов и флуктуаций уровней отсчетов временных рядов. Для решения данной задачи используется метод вейвлет-обработки, в частности, уменьшение влияния флуктуаций временного ряда осуществляется путем вычисления аппроксимирующих коэффициентов n -го порядка, что соответствует выполнению многоуровневой статистической обработки значений отсчетов временных рядов и эквивалентно прохождению сигнала через фильтр нижних частот. Кроме того, точность представления информации повышается за счет подавления компонент шумового процесса методом выполнения операции трешолдинга над детализирующими коэффициентами до n -го уровня разложения. В этом случае предлагается реализовать алгоритм трешолдинга по критерию адаптивного штрафного порога.

1. Введение

Модель обработки сигналов в виде временных рядов, основанная на вейвлет-разложении, является эффективным методом для создания алгоритмов многоуровневой статистической обработки в системах мониторинга.

2. Модель обработки временных рядов данных систем мониторинга

В случае вейвлет-разложения сигналов, представленных в виде временных рядов, аппроксимирующие и детализирующие коэффициенты вычисляются по формулам [1]

$$C_{i,k} = \frac{1}{p} \sum_{i=1}^n C_{i-1} \varphi_{i,k}(2^i t - k), \quad d_{i,k} = \frac{1}{p} \sum_{i=1}^n C_{i-1} \psi_{i,k}(2^i t - k), \quad (1)$$

где i – уровень разложения; n – максимальный уровень разложения; $p = 2^{i/2}$ – коэффициент ортонормирования, обеспечивающий единичную норму скейлинг функции, $\varphi_{i,k}(2^i t - k)$ - скейлинг функция i -го уровня вейвлет-разложения, $\psi_{i,k}(2^i t - k)$ - вейвлет функция.

При $i=1$, C_0 и d_1 принимают вид:

$$C_{0,k} = x(k) \frac{1}{p} \varphi_{1,k}(t - k) \approx x(k), \quad d_{1,k}(t) = \frac{1}{p} C_{0,k} \psi_{1,k}(2t - k) \approx \frac{1}{p} x(k) \psi_{1,k}(2t - k), \quad (2)$$

где временной ряд $x(k)$ имеет вид

$$x(k) = s(k) + y(k),$$

$s(k)$ - отсчеты информационной функции, $y(k)$ – составляющая флуктуаций отсчетов, $C_{0,k}$ – локальное среднее значение сигнала $x(k)$ взвешенное со скейлинг функции $\varphi_{1,k}$.

Соответственно, при декомпозиции сигнала $x(k)$ до n -ого уровня и последующего восстановления его в виде [1,2]

$$s(k) = \sum_k C_{n,k} \varphi_{n,k} + \sum_k \sum_{i=1}^n d_{i,k} \psi_{i,k} = C_n + x(k) \psi_1(2t - k) + \sum_{i=2}^n C_{i-1} \varphi_i(2^i t - k), \quad s(k) = C_{n,k} + \sum_{i=1}^n d_i,$$

получают аппроксимирующие коэффициенты i -ого уровня путем локальных усреднений аппроксимирующих коэффициентов уровня $(i-1)$ с i -ой скейлинг функцией на каждом уровне разложения в виде (1).

Таким образом, при помощи вычисления $C_{n,k}$, осуществляется многоуровневая статистическая обработка входных сигналов, чем существенно подавляется составляющая $y(k)$ флуктуаций отсчетов информационной входной функции $x(k)$.

3. Исследование статистической обработки временных рядов методом вейвлет-преобразования

В ходе исследования был проведен экспериментальный сбор данных концентраций испарений ксилола в цехах лакокрасочной обработки деталей с датчиковой аппаратуры с частотой измерений равной один отсчет в 30 минут. Таким образом, в течение месяца были осуществлены 1488 измерения концентраций по каждому из анализируемых веществ. Данные с датчиков вместе со временем отбора записывались в файл и базу данных. После этого он преобразовывался в формат (.mat), доступный для вейвлет-анализа в системе Matlab.

На рисунке 1 приведен измеренный сигнал в виде экспериментального временного ряда данных о концентрациях ксилола (ПДК 50 мг/м³).

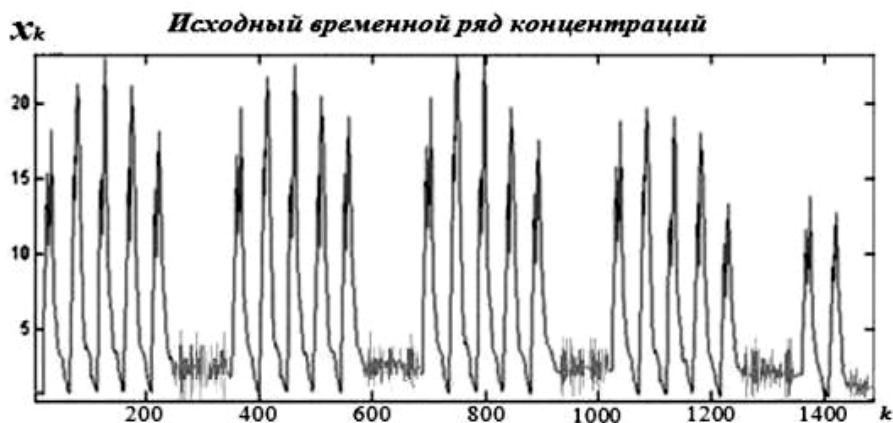


Рисунок 1. Представление временного ряда загрязняющих выбросов $x(k)$ с периодом $T=30$ мин.

На графике рисунок 1 отчетливо виден достаточно заметный уровень шумовых компонент, а также можно отметить периодическую ежесуточную и еженедельную структуры.

После вейвлет-обработки представленного временного ряда на рисунке 1 по формулам (1), (2), были исследованы АКФ для аппроксимирующих коэффициентов $C_i(k)$, для $i=1$ и $i=6$, по выражению

$$R_{a.k.}(\tau) = \frac{1}{N} \sum_{l=0}^{L-1} \sum_{k=0}^{N-1-L} C_i(k) C_i(k+l), \tag{3}$$

где $R(\tau)$ – АКФ последовательности $C_i(k)$, $0 \leq k \leq N-1$, lT – дискретное время сдвига, $L \leq \frac{N}{2}$ – максимальный сдвиг исследования, N – конечная длина исследуемой последовательности [3].

Графики АКФ аппроксимирующих коэффициентов вейвлет-разложения C_1 и C_6 приведены на рисунке 2.

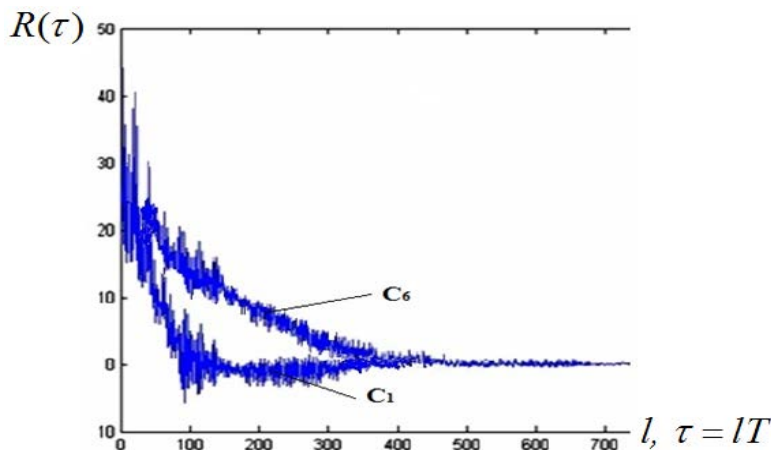


Рисунок 2. Графики АКФ аппроксимирующих функций.

Как видно из рисунка 2, графики АКФ аппроксимирующих коэффициентов $C_i(k)$ показывают увеличение времени корреляции при увеличении уровня вейвлет-разложения [4].

В работе были проведены исследования зависимости времени корреляции аппроксимирующих коэффициентов от уровня вейвлет-разложения. Также была исследована зависимость ослабления шумовых составляющих флуктуаций отсчетов временного ряда выбросов при обработке вейвлет-разложением и получением аппроксимирующих коэффициентов i -ого уровня.

На рисунке 3 показана зависимость ослабления $\beta(\tau)$ дисперсии компонент шума временного ряда от уровня вейвлет-обработки, где $\beta(\tau) = \frac{\sigma_{u,W}^2}{\sigma_{u,ex}^2}$, $\sigma_{u,W}^2$ – дисперсия шумовых составляющих на выходе системы вейвлет-обработки (W -фильтра).

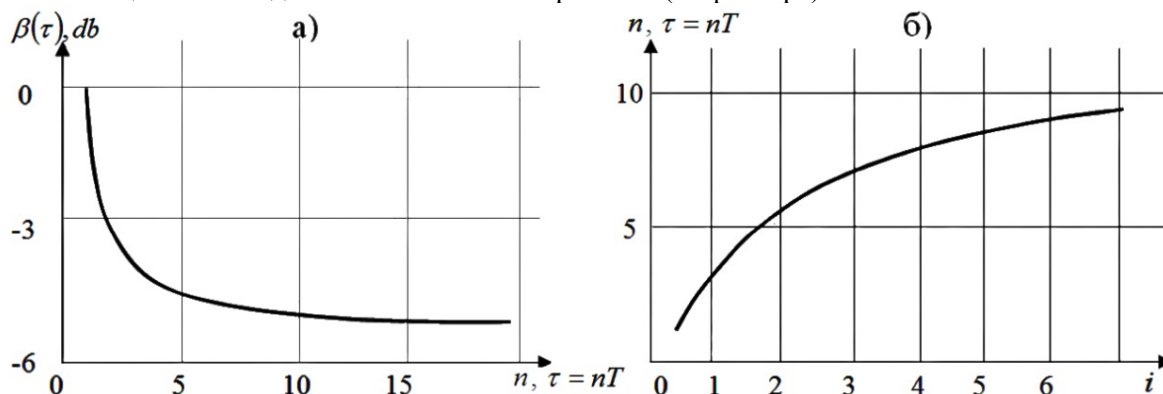


Рисунок 3. Зависимость ослабления шума от выбора уровня вейвлет-разложения i в W -фильтре ('а' – зависимость ослабления дисперсии компонент шума от времени корреляции шумовых компонент; 'б' – зависимость изменения времени корреляции от уровня разложения i аппроксимирующих коэффициентов).

По графику зависимости ослабления дисперсии компонент шума (рисунок 3“а”) видно, что наибольшее ослабление компонент шума $\beta = -(4 \div 5)dB$ может быть получено при увеличении времени корреляции C_i до значения $\tau = (5 \div 10)T$.

По графику на рисунке 3“б”, зависимости времени корреляции шумовых компонент временного ряда концентраций выбросов от уровня i вейвлет-разложения, от уровня i вычисленных аппроксимирующих коэффициентов C_i временного ряда можно отметить, что наибольшее увеличение времени корреляции шумовых компонент до $\tau = (6 \div 8)T$ можно получить при вейвлет-разложении аппроксимирующих коэффициентов от 3 до 6 уровня. При вычислении C_3 получено время корреляции $\tau = 6T$, при C_6 время корреляции $\tau = 8T$.

В этом случае после вейвлет-обработки до уровня n , получен временной ряд представления данных с более высокой точностью в виде

$$s(k) = \frac{1}{p} \left[\sum_k u_{0,k} \psi_{1,k}(2t - kT) + \sum_k \sum_{i=2}^n C_{(i-1),k} \psi_{i,k}(2^i t - k) + \sum_k C_{n,k} \right]. \quad (4)$$

Таким образом, была разработана и исследована математическая модель и её статистические параметры обработки вейвлет-преобразованием отсчетов временных рядов. Показано, что вейвлет-обработка временного ряда и применение аппроксимирующих коэффициентов до n -го уровня вейвлет-разложения повышает точность представления данных.

4. Алгоритм подавления шумовых составляющих методом вычисления детализирующих коэффициентов с последующей операцией трешолдинга

Детализирующие коэффициенты при вейвлет-разложении временных рядов в системах мониторинга содержат шумовые и флуктуационные составляющие, единичные выбросы и другие компоненты помех. Кратковременные единичные выбросы и флуктуации ряда данных могут быть вызваны, например, наводкой на сигнал, передаваемый с датчиков, программно-аппаратными флуктуациями параметров системы мониторинга. Возникновение шумовых, помеховых составляющих во временных рядах данных может стать следствием характерных для датчикового оборудования погрешностей сбора, искажением полезного сигнала, которое имеет место при передаче по каналам и интерфейсам системы мониторинга.

Шумы, флуктуации временного ряда вносят погрешность в оценку анализа его компонент, тем самым формируют ряд данных и соответственно восстановление с повышенными погрешностями.

Можно отметить, что использование в формуле восстановления аппроксимирующих коэффициентов n -ого уровня, полученных путем локальных усреднений с i -ой скейлинг-функцией на каждом уровне вейвлет-разложения $0 < i \leq n$, позволяет осуществить многоуровневую статистическую обработку входного временного ряда. Это существенно ослабляет компоненты шумовых флуктуаций. В тоже время, как показывают исследования, дополнительное ослабление шумовых компонент можно получить пороговой обработкой детализирующих коэффициентов с помощью алгоритма сглаживания. Так как полученные при вейвлет-разложении временного ряда детализирующие коэффициенты содержат высокочастотные (уточняющие) данные разных уровней вейвлет-разложения, в них также сосредоточены достаточно высокого уровня шумовые и помеховые составляющие. В связи с этим, пороговая обработка детализирующих коэффициентов является эффективным методом для дополнительного подавления шумовых и помеховых компонент. Алгоритм сглаживания детализирующих коэффициентов временного ряда данных или алгоритм их пороговой обработки путем обнуления значений детализирующих коэффициентов, которые не превышают заданный пороговый уровень, позволяют получить увеличенное отношение сигнал/шум (ОСШ). Это означает, что повышается точность представления данных в восстановленном временном ряде системы мониторинга.

Алгоритм пороговой обработки детализирующих коэффициентов вейвлет-разложения, алгоритм их сглаживания представлен на рисунке 4 [5, 6].

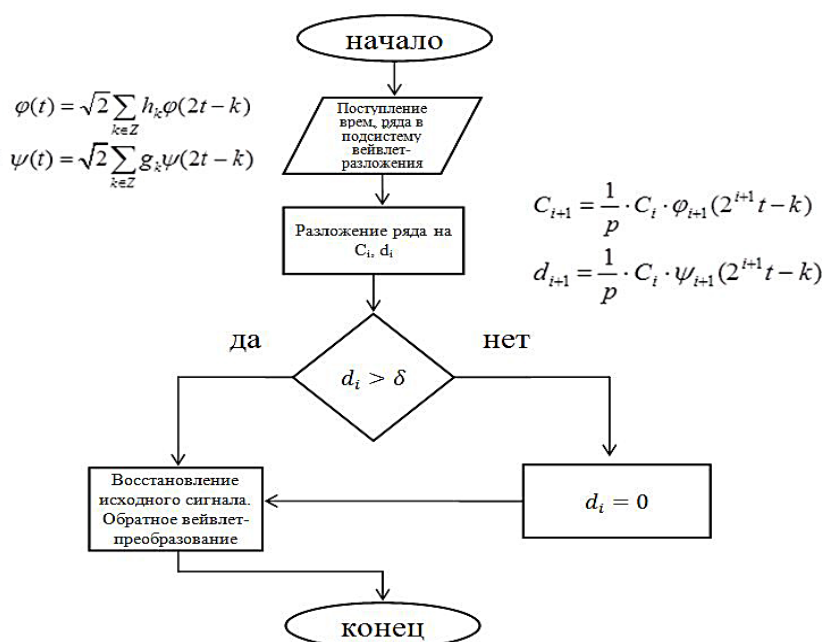


Рисунок 4. Алгоритм сглаживания, алгоритм пороговой обработки детализирующих коэффициентов.

Отношение сигнал/шум после вейвлет-обработки является критерием эффективности разработанного алгоритма сглаживания временных рядов. Отношение сигнал/шум на входе алгоритма сглаживания определяется выражением [9-11]

$$ОСШ_{вх} (\text{дБ}) = 10 \lg \frac{P_{с_вх}}{P_{ш_вх}}, \tag{5}$$

где $P_{с_вх} = \frac{1}{N} \sum_{k=0}^{N-1} x_k^2$, $P_{ш_вх} = \frac{1}{N} \sum_{k=0}^{N-1} (x_k - \bar{x})^2$, $\bar{x} = \sqrt{P_{с_вх}}$.

Выходное ОСШ обработанного и восстановленного временного ряда вычисляется по формулам

$$ОСШ_{вых} (\text{дБ}) = 10 \lg \frac{\frac{1}{N} \sum_{k=0}^{N-1} s_k^2}{P_{ш_вых}} = 10 \lg \frac{P_{с_вых}}{P_{ш_вых}}, \tag{6}$$

где $P_{ш_вых} = \frac{1}{N} \sum_{k=0}^{N-1} (s_k - \bar{s}_k)^2$, $\bar{s}_k = \sqrt{P_{с_вых}}$.

В программном пакете Wavelet Toolbox для обработки сглаживанием детализирующих коэффициентов вейвлет-разложения временных рядов может быть применен один из нескольких критериев по выбору оптимального порога ограничения детализирующих коэффициентов, несущих в себе шумовые составляющие [5, 7].

С использованием различных критериев по выбору пороговых уровней обработки детализирующих коэффициентов было проведено сглаживание тестового сигнала. В качестве него был применен экспериментальный временной ряд данных концентраций ксилола (1488 отсчета), полученных с датчиков системы мониторинга выбросов на участке лакокрасочных покрытий предприятия радиотехнического профиля (сбор данных осуществлялся с периодичностью в 30 минут). В таблице 1, представлены полученные расчетные данные результатов пороговой обработки при сглаживании временных рядов.

На рисунке 5 проиллюстрированы результаты пороговой обработки по критерию адаптивного штрафного порога.

Таблица 1. Результаты сглаживания с применением различных критериев.

	Исходный сигнал до фильтрации	Критерий Штейна	Эвристический критерий Штейна	Минимаксный критерий	Критерий адаптивного штрафного порога
Мощность $P_{вх.сигн.}$	5,608	5,608	5,608	5,608	5,608
Мощность $P_{вх.шума.}$	0,215	0,215	0,215	0,215	0,215
$ОСШ_{вх.}$	26,084	-	-	-	-
Мощность $P_{вых.сигн.}$	-	5,546	5,580	5,521	5,605
Мощность $P_{вых.шума.}$	-	0,153	0,187	0,128	0,119
$ОСШ_{вых.}$	26,084	36,248	29,839	43,133	46,747
$ОСШ_{вых.}, дБ$	14,16	15,59	14,48	16,48	16,69

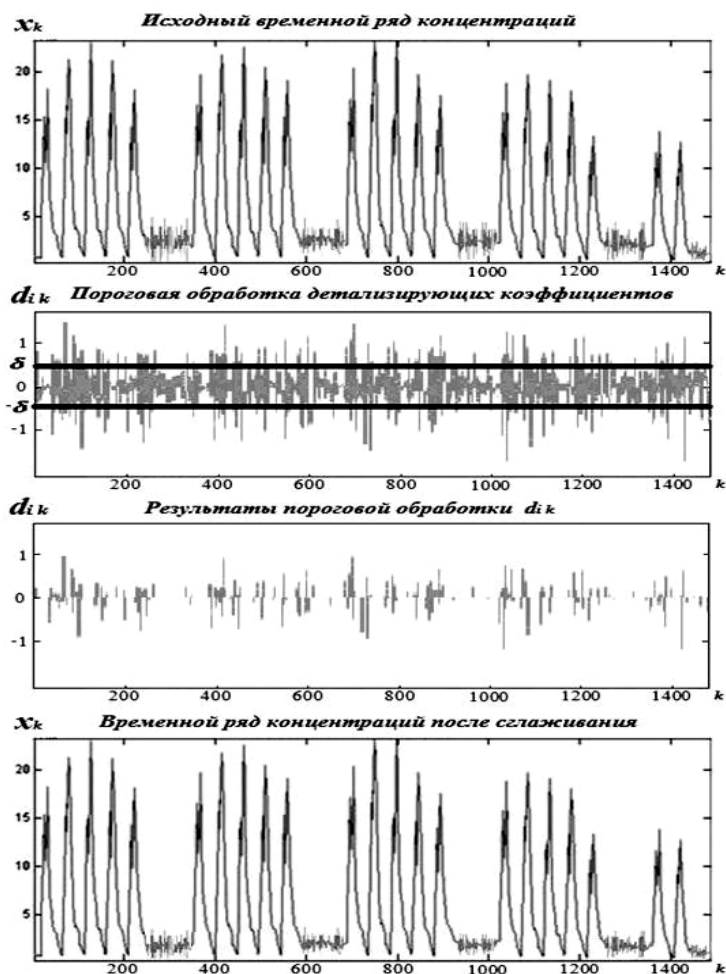


Рисунок 5. Результаты сглаживания с адаптивным штрафным порогом.

Проведенный анализ результатов применения различных критериев и подходов по сглаживанию на основе пороговой обработки детализирующих коэффициентов вейвлет-

разложения показал, что оптимальным при сглаживании временных рядов будет выбор адаптивного штрафного порогового уровня. При пороговой обработке по данному критерию максимально эффективно подавляются случайные шумовые и флуктуационные составляющие, кратковременные «выбросы» и «всплески» сигнала [8]. Отношение сигнал/шум при использовании данного критерия пороговой обработки составило 16,7 дБ, что на 2,53 дБ выше по сравнению с первоначальным значением отношения сигнал/шум [5].

5. Вейвлет обработка и восстановление временного ряда данных с пониженной погрешностью

Алгоритм вейвлет обработки временного ряда $x(k)$ для получения данных $s(k)$ с пониженной погрешностью представления данных представлен на рисунке 6.



Рисунок 6. Вейвлет обработка временного ряда.

Входные сигналы в виде временного ряда $x(k)$ подаются на W-фильтр обработки вейвлет-преобразованием. В W-фильтре формируются аппроксимирующие коэффициенты C_i и детализирующие коэффициенты d_i , n -ого уровня вейвлет-разложения. Детализирующие коэффициенты подаются на блок трешолдинга для их обработки по критерию адаптивного штрафного порога. После пороговой обработки алгоритмом сглаживания, полученные обработанные детализирующие коэффициенты $d_1^*, d_2^*, \dots, d_n^*$ поступают на сумматор - блок восстановления временного ряда. Также на сумматор подаются аппроксимирующие коэффициенты n -ого уровня разложения ($C_n(k)$). На выходе сумматора получаем восстановленный временной ряд данных с пониженной погрешностью представления данных.

6. Заключение

В работе разработан и исследован метод понижения влияния компонент шумового процесса и флуктуаций отсчетов временного ряда на точность представления информации. Предложенный алгоритм обработки временного ряда реализован с применением вейвлет-преобразования путем вычисления аппроксимирующих коэффициентов n -го уровня и детализирующих коэффициентов до n -го уровня с последующей обработкой трешолдингом по критерию адаптивного штрафного порога. Отношение сигнал/шум при использовании данной обработки увеличилось на более чем 2,53 дБ по сравнению с исходным.

7. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 17-48-330726.

8. Литература

- [1] Дремин, И.М. Вейвлеты и их использование / И.М. Дремин, О.В. Иванов, В.А. Нечитайло // *Успехи физических наук*. – 2001. – Т. 17, №5. – С. 465-501.
- [2] Яковлев, А.Н. Введение в вейвлет-преобразование: учебное пособие / А.Н. Яковлев. – Новосибирск: Издательство Новосибирского технического университета, 2003. – 102 с.
- [3] Рабинер, Л. Теория и применение цифровой обработки сигналов / Л. Рабинер, Б. Гоулд. – М.: «МИР», 1978. – 848 с.
- [4] Проскуряков, А.Ю. Модель прогнозирования временных рядов на трехслойном персептроне в информационной системе мониторинга загрязняющих выбросов / А.Ю. Проскуряков // *Радиотехнические и телекоммуникационные системы*. – 2014. – №2. – С. 57-63.
- [5] Белов, А.А. Сглаживание временных рядов на основе вейвлет-преобразования в системах автоматизированного экологического мониторинга / А.А. Белов, А.Ю. Проскуряков // *Методы и устройства передачи и обработки информации*. – 2010. – Т. 1, № 12. – С. 21-24.
- [6] Белов, А.А. Вопросы обработки экспериментальных временных рядов в электронной системе автоматизированного контроля / А.А. Белов, Ю.А. Кропотов, А.Ю. Проскуряков // *Вопросы радиоэлектроники*. – 2010. – Т. 1, № 1. – С. 95-101.
- [7] Белов, А.А. Автоматизированный анализ и обработка временных рядов данных о загрязняющих выбросах в системе экологического контроля / А.А. Белов, Ю.А. Кропотов, А.Ю. Проскуряков // *Информационные системы и технологии*. – 2010. – Т. 6, № 62. – С. 28-35.
- [8] Белов, А.А. Исследование вопросов сжатия и поиска картографической информации методом вейвлет-преобразований в экологической геоинформационной системе / А.А. Белов, Ю.А. Кропотов // *Вестник компьютерных и информационных технологий*. – 2008. – № 12. – С. 9-14.
- [9] Быков, А.А. Модель закона распределения вероятности амплитуд сигналов в базе экспоненциальных функций системы / А.А. Быков, Ю.А. Кропотов // *Проектирование и технология электронных средств*. – 2007. – № 2. – С. 30-34.
- [10] Кропотов, Ю.А. Алгоритм определения параметров экспоненциальной аппроксимации закона распределения вероятности амплитуд речевого сигнала / Ю.А. Кропотов // *Радиотехника*. – 2007. – № 6. – С. 44-47.
- [11] Кропотов, Ю.А. Алгоритм подавления акустических шумов и сосредоточенных помех с формантным распределением полос режекции / Ю.А. Кропотов, А.А. Быков // *Вопросы радиоэлектроники*. – 2010. – Т. 1, №1. – С. 60-65.

Wavelet processing of time series to improve the accuracy of information representation

Y.A. Kropotov¹, A.Y. Proskuryakov¹, A.A. Belov¹

¹Murom Institute (branch) Vladimir State University named after Alexander and Nickolay Stoletovs, Orlovskaya street 23, Murom, Russia, 602264

Abstract. The paper considers a method of reducing the error in the representation of information in telecommunication monitoring systems by filtering noise and fluctuations in the levels of samples of time series. To solve this problem, the wavelet-processing method is used, in particular, the effect of time-series fluctuations is reduced by calculating approximating coefficients of the n -th order, which corresponds to performing multilevel statistical processing of the values of the time series readings and is equivalent to passing the signal through the low-pass filter. In addition, the accuracy of information representation is enhanced by suppressing the components of the noise process by performing a tresholding operation on the detailing coefficients up to the n -th level of the decomposition. In this case, it is proposed to implement the tresholding algorithm according to the criterion of the adaptive penalty threshold.

Keywords: noise and fluctuations, penalty threshold, accuracy of information presentation, time series, wavelet processing, approximating coefficients, detailing coefficients.