

Комплексный алгоритм обнаружения аномальных особенностей в природных временных рядах

А.Р. Лисс¹, Б.С. Мандрикова², О.В. Мандрикова²

¹ Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), 197376, Россия, г. Санкт-Петербург, ул. Профессора Попова, д. 5;

² Институт космических исследований и распространения радиоволн ДВО РАН, 684034, Россия, Камчатский край, с. Паратунка, ул. Мирная, д. 7

Аннотация

Предложен комплексный автоматизированный алгоритм анализа природных временных рядов и обнаружения аномальных особенностей. Алгоритм включает в себя алгоритм определения информационных компонент сигнала и алгоритм адаптивной вейвлет-фильтрации сигнала. Алгоритм определения информационных компонент сигнала выполняет подавление коррелированного шума и определение информационных составляющих сигнала. Алгоритм адаптивной вейвлет-фильтрации сигнала выполняет обнаружение аномальных особенностей и оценку их интенсивности. Данные алгоритмы могут применяться как совместно, так и независимо друг от друга. Основу алгоритмов составляют разработанные авторами правила. На основе правил оцениваются параметры пороговой функции и определяется наилучший аппроксимирующий вейвлет. В статье описаны операции комплексного алгоритма и представлена блок-схема его реализации. Также приведены результаты применения комплексного алгоритма с использованием данных вторичных космических лучей и модельных данных, построенных по их подобию. Результаты подтвердили эффективность разработанных правил и предлагаемого комплексного алгоритма.

Ключевые слова: вейвлет-преобразование, теория рисков, природные аномалии, космические лучи.

Цитирование: Лисс, А.Р. Комплексный алгоритм обнаружения аномальных особенностей в природных временных рядах / А.Р. Лисс, Б.С. Мандрикова, О.В. Мандрикова // Компьютерная оптика. – 2025. – Т. 49, № 6. – С. 1030-1036. – DOI: 10.18287/2412-6179-CO-1652.

Citation: Liss AR, Mandrikova BS, Mandrikova OV. Complex algorithm for detecting anomalous features in natural data. Computer Optics 2025; 49(6): 1030-1036. DOI: 10.18287/2412-6179-CO-1652.

Введение

Известной проблемой извлечения информации из данных в большинстве прикладных областей (физика, биология, медицина и др.) является высокая доля неполных априорных знаний об информационном сигнале и шуме. Процедуру выделения полезной информации из природных данных также существенно затрудняет наличие коррелированных шумов. Применение современных нейросетевых методов в некоторых случаях не обеспечивает требуемой точности решения данной задачи. Это связано с существенной нестационарностью природных данных вследствие постоянной изменчивости природных процессов и отсутствием представительной выборки для обучения нейронных сетей. Указанные проблемы требуют усовершенствования методов обработки и анализа природных данных и создания новых более эффективных подходов и методов, в том числе с использованием нейросетевых технологий. Например, в работе [1] для решения задачи обнаружения аномалий в технологических сигналах авторами предлагается подход, основанный на комбинации ансамбля из базовых классификаторов на основе алгоритмов машинного обучения и вейвлет-

преобразования. Применение вейвлет-коэффициентов на этапе формирования информативного признакового пространства для обучения классификаторов и консолидации прогнозов позволило повысить точность распознавания и повысить качество мониторинга объектов управления [1]. В работе [2] для задачи обнаружения ранних признаков отказов и поломок оборудования представлен подход к обнаружению аномалий в нестационарных технологических сигналах с использованием преобразования Гильберта–Хуанга совместно со статистической моделью классификации. Способность адаптации к нестационарным данным и высокая детализация в частотно-временной области подхода [2] позволила повысить точность обнаружения моделированных аномалий до 94 %.

Задачи обработки и анализа природных данных, ввиду пассивности проводимых экспериментов, составляют еще большую сложность. Например, в работе [3], в области медицины, для оптимизации исследований головного мозга по данным функциональной магнитно-резонансной томографии (МРТ) предложена нейросетевая технология обнаружения ступенчатых аномалий с обучением на частично синтезированных данных с адаптацией на основе метаобучения.

Разработанная процедура формирования синтетического набора данных позволила существенно оптимизировать процедуру нейросетевого обучения [3]. Примером из области космофизики является работа [4], описывающая новый комбинированный метод выявления скрытых аномалий в вариациях галактических космических лучей (КЛ). Метод основан на совмещении спектрально-сингулярного разложения сигналов КЛ и вейвлет-преобразования. Как показано в работе [4], спектрально-сингулярное разложение сигналов КЛ позволяет изучить динамику вариаций, а вейвлет-преобразование позволяет вычислять энергию короткопериодных вариаций произвольной формы (скрытых аномалий в КЛ) на фоне шума, превышающего по амплитуде полезный сигнал. В работах [5, 6] рассмотрены методы анализа природных данных для задач космической погоды. В частности, в работе [6] предложен автоматизированный метод анализа параметров ионосферы и обнаружения ионосферных аномалий на основе комплексного подхода, объединяющего методы вейвлет-преобразования с моделями авторегрессии – проинтегрированного скользящего среднего. На основе метода [6] исследователями обнаружены короткопериодные аномальные изменения, предшествующие магнитным бурям и характеризующие возникновение колебательных процессов в ионосфере на фоне повышенной солнечной активности, что позволило авторам получить важный прикладной результат.

Объектом исследования являются алгоритмы анализа данных нейтронных мониторов (НМ), отражающих интенсивность вторичных космических лучей в околоземном пространстве [7]. В периоды повышенной солнечной активности в данных НМ наблюдаются аномальные особенности, обнаружение которых важно в задачах космической погоды [7]. Аномальные особенности в данных НМ возникают в случайные моменты времени и имеют разнообразную форму, что усложняет процедуру их обнаружения. Высокая доля неопределенности знаний о полезной составляющей данных НМ и их нестационарность делает неэффективным применение классических методов анализа временных рядов (моделей ARIMA, методов декомпозиции временных рядов и т.п.). Применение методов машинного обучения, в частности нейронных сетей, требует периодической адаптации алгоритмов ввиду существенной изменчивости временного хода данных НМ. Это снижает их эффективность и не обеспечивает возможность получения результата в оперативном режиме.

Учитывая структуру анализируемых данных, в работе предложен комплексный алгоритм обнаружения аномальных особенностей, основанный на синтезе вейвлет-преобразования с адаптивными пороговыми функциями. Методы вейвлет-преобразования имеют обширный набор базисов разной

формы и эффективны для анализа данных нестационарной структуры. Использование вейвлетов в работе позволило реализовать операции обнаружения и оценки параметров аномальных особенностей разной формы и длительности, в том числе резких всплесков, особенностей пикообразной формы и др. Компактный носитель вейвлетов обеспечил высокую степень локализации получаемой информации о моментах возникновения аномальных особенностей. Построенный комплексный алгоритм основан на разработанных авторами правилах, применение которых позволяет минимизировать погрешность оценки информационного сигнала. Предложенный алгоритм основан на теории, описанной в работах [6, 8], и включает в себя алгоритм определения информационных компонент сигнала и алгоритм адаптивной вейвлет-фильтрации сигнала. Алгоритм определения информационных компонент сигнала (алгоритм 1) подавляет шум, в том числе коррелированный, и определяет информационные составляющие сигнала. Алгоритм адаптивной вейвлет-фильтрации сигнала (алгоритм 2) путем комбинации дискретного вейвлет-преобразования и пороговых функций обнаруживает аномальные особенности и оценивает их интенсивность. Алгоритм 1 и алгоритм 2 являются законченными и могут быть применены независимо. В статье описаны операции комплексного алгоритма, представлена блок-схема его реализации. Показаны результаты применения комплексного алгоритма с использованием природных данных вторичных космических лучей и модельных данных, построенных по их подобию.

1. Описание комплексного алгоритма

Регистрируемый сигнал рассматривается как комбинация

$$X[t] = F[t] + V[t], \quad (1)$$

где $F[t]$ – полезный сигнал, F принадлежит пространству Гильберта H [9], $V[t]$ – шум.

Выполняя разложение сигнала X в пространстве H по базису $\mathfrak{B} = \{g_m\}_{m \in \mathbb{N}}$ (\mathbb{N} – натуральные числа) и применяя пороговые функции, получаем:

$$\tilde{F} = \sum_{m \in \mathbb{N}} \Pi(X, g_m) g_m, \quad (2)$$

где

$$\Pi(x) = \begin{cases} x, & \text{если } |x| \geq T, \\ 0, & \text{если } |x| < T, \end{cases}$$

пороговая функция, $\langle \cdot \rangle$ – скалярное произведение.

На основе операции (2) стоит задача подавить шум и выделить информационные компоненты сигнала. Погрешность оценки (2) для дискретного сигнала $F[t_n]$, $t_n \in \{1, \dots, N\}$, N – длина сигнала, есть

$$\begin{aligned}
 r(\tilde{F}[t_n], F[t_n]) &= E\{\tilde{F}[t_n] - X[t_n]\}^2 = \\
 &= E\left\{\sum_{t_n=1}^N |\tilde{F}[t_n] - X[t_n]|\right\}^2 = \\
 &= E\left\{\sum_{t_n \in \{t_n | |X[t_n], g_m[t_n]| < T\}} |X[t_n], g_m[t_n]|\right\}^2,
 \end{aligned} \tag{3}$$

E – математическое ожидание, $\| \cdot \|$ – норма в пространстве H .

Очевидно, погрешность (3) определяется абсолютными значениями коэффициентов $|\langle X, g_m \rangle|$ разложения по базису и зависит от величины порога T .

Определение величины порога.

Применяя критерий отношения [10, 11] в окрестности $O_\delta[t_i] = \{t_n: |t_n - t_i| < \delta\}$ некоторой точки $t_i, t_i \in \{1, \dots, N\}$, будем иметь:

$$L(X[t_i]) = \frac{W(X[t_i]/s_1)}{W(X[t_i]/s_0)} \geq T_\delta, \tag{4}$$

где $L(X[t_i])$ – функция правдоподобия, $W(X[t_i]/s_0), W(X[t_i]/s_1)$ – плотности вероятностей при отсутствии (справедливости гипотезы G_0) и наличии информационной составляющей сигнала (справедливости гипотезы G_1) соответственно, s_0, s_1 – возможные состояния сигнала, T_δ – величина порога в окрестности O_δ – некоторой точки t_i .

Тогда, используя критерий Неймана–Пирсона [10, 11], для оценки порога $T_{\alpha, \delta}$ в окрестности O_δ точки t_i с заданной доверительной вероятностью α будем решать следующую оптимизационную задачу:

$$\begin{aligned}
 \int_A^0 W(X[t_i]/s_0) dX &= \alpha = const, \\
 \int_A^0 W(X[t_i]/s_1) df &= \gamma \rightarrow \max,
 \end{aligned} \tag{5}$$

где α – заданная ошибка первого рода, $\gamma = 1 - \beta$ – мощность критерия, β – ошибка второго рода, A – критическая область.

Поскольку гипотеза G_1 в рассматриваемой постановке задачи является простой альтернативой, можно рассматривать только гипотезу G_0 , так как отклонение G_0 означает принятие гипотезы G_1 .

Тогда из соотношений (3) – (5) получаем следующее *Правило определения величины порога (Правило 1)*: порог $T_{\alpha, \delta}$ в окрестности O_δ точки t_0 с заданной доверительной вероятностью α будем определять, как

$$T_{\alpha, \delta} = \tau_\alpha \cdot \hat{\sigma}_\delta, \tag{6}$$

где τ_α – α -квантили распредел. Стьюдента [11],

$$\hat{\sigma}_\delta = \sqrt{\frac{1}{2\delta} \sum_{n=1}^{2\delta} \left(|X[t_n], g_m[t_n]| - |X[t_n], g_m[t_n]| \right)^2} -$$

выборочное стандартное отклонение величины $|\langle X[t_n], g_m[t_n] \rangle|$ в окрестности O_δ точки t_i ,

$$\overline{|X[t_n], g_m[t_n]|} -$$

среднее значение.

В случае разложения функции $X[t_n]$ по вейвлет-базису [12, 13], формула (2) примет вид [6, 8]:

$$\tilde{F}[t_n] = \sum_k \sum_{n=1}^N \Pi(X[t_n], \Psi_{k,n}[t_n]) \Psi_{k,n}[t_n], \tag{7}$$

где $\Psi_{k,n}[t_n] = 2^{k/2} \Psi(2^k t - n)$ – вейвлеты,

$$\begin{aligned}
 \Pi(X[t_n], \Psi_{k,n}[t_n]) &= \\
 &= \begin{cases} X[t_n], \Psi_{k,n}[t_n], & \text{если } |X[t_n], \Psi_{k,n}[t_n]| \geq T_{\alpha, k, n}, \\ 0, & \text{если } |X[t_n], \Psi_{k,n}[t_n]| < T_{\alpha, k, n}, \end{cases} -
 \end{aligned}$$

пороговые функции, $T_{\alpha, k, n} = \tau_\alpha \cdot \hat{\sigma}_{k, n}$ – пороги на масштабе k в окрестности O_δ точки $t = n$.

Замечание 1. В операции (7) пороги $T_{\alpha, k, n}$, следуя соотношению (6), оцениваются на масштабе k в окрестности O_δ точки t_n . Адаптация порогов $T_{\alpha, k, n}$ под изменяющиеся свойства сигнала позволяет подавить коррелированный шум.

Замечание 2. Попадание вейвлет-коэффициентов в критическую область A (см. соотн. (5)) означает принятие гипотезы G_1 и, как показано в [8], в вейвлет-пространстве обеспечивает обнаружение аномальных особенностей сигнала.

Погрешность $r(\tilde{F}[t_n], F[t_n])$ оценки \tilde{F} (см. (3)), очевидно, зависит от аппроксимирующего вейвлета (см. (6)). Следуя результатам [9], в вейвлет-базисе погрешность может быть оценена как

$$\begin{aligned}
 r_\Psi(\tilde{F}[t_n], F[t_n]) &= \\
 &= \sum_{k, n} \min \left(|\tilde{F}[t_n], \Psi_{k, n}[t_n]|^2, (\sigma_{k, n}^\Psi)^2 \right).
 \end{aligned} \tag{8}$$

Тогда, ограничив погрешность (8) величиной

$$r_\Psi(\tilde{F}[t_n], F[t_n]) \leq \sum_{k, n} (\sigma_{k, n}^\Psi)^2,$$

введем следующее правило определения наилучшего аппроксимирующего вейвлета (Правило 2).

Правило 2. Наилучшим аппроксимирующим вейвлетом Ψ_{min} для сигнала F будем считать вейвлет, удовлетворяющий условию

$$r_{\Psi_{min}}(\tilde{F}[t_n], F[t_n]) = \min_{\Psi \in \Omega} \sum_{k, n} (\sigma_{k, n}^\Psi)^2, \tag{9}$$

где Ω – множество рассматриваемых вейвлетов,

$$\sigma_{k, n}^\Psi = \sqrt{\frac{1}{2\delta} \sum_{n=1}^{2\delta} \left(|X[t_n], \Psi_{k, n}[t_n]| - \overline{|X[t_n], \Psi_{k, n}[t_n]|} \right)^2}.$$

Табл. 1. Погрешности для разных вейвлетов

$\frac{\sum_{k,n} (\sigma_{k,n}^\Psi)^2}{X}$	Вейвлет-функция					
	coif_1	coif_2	coif_3	db_1	db_2	db_3
ст. Инувик	0,0175	0,0170	0,0176	0,0178	0,0174	0,0176
	0,0177	0,0175	0,0174	0,0178	0,0177	0,0173
ст. Оулу	0,0189	0,0187	0,0188	0,0190	0,0189	0,0188
	0,0202	0,0202	0,0201	0,0203	0,0200	0,0201
РН	6	12	18	2	4	6
ЧНМ	2	4	6	1	2	4

Замечание 3. В данной работе наилучший аппроксимирующий вейвлет определялся в рамках семейств Добеши и Койфлеты [12, 13]. Выбор семейств основывался на таких характеристиках вейвлетов, как гладкость, число нулевых моментов и размер носителя, обоснован в работе [14].

Результаты применения Правила 2 для вейвлетов семейств Добеши и Койфлеты представлены в

$$B_j^p = \begin{cases} \text{если } \sum_{m \in I^p} |\langle X, \Psi_{j,m}^p \rangle|^2 \geq \sum_{m \in I^p} |\langle X, \Psi_{j+1,m}^{2p} \rangle|^2 + \sum_{m \in I^{2p+1}} |\langle X, \Psi_{j+1,m}^{2p+1} \rangle|^2, \\ \text{если } \sum_{m \in I^p} |\langle X, \Psi_{j,m}^p \rangle|^2 < \sum_{m \in I^p} |\langle X, \Psi_{j+1,m}^{2p} \rangle|^2 + \sum_{m \in I^{2p+1}} |\langle X, \Psi_{j+1,m}^{2p+1} \rangle|^2, \end{cases}$$

где множество индексов

$$I^l, l = p, 2p, 2p+1, m \in I^l,$$

если $|X[t_n], \Psi_{j,m}^p| \geq T_{\alpha,j,m}^p, T_{\alpha,j,m}^p = \tau_\alpha \cdot \delta_{j,m}^p$.

Шаг 2. Выполняем вейвлет-восстановление преобразованного на основе шага 1 сигнала

$$\tilde{F}[t_n] = \sum_{j=0}^J \sum_{m=1}^N \Pi(X[t_n], \Psi_{j,m}^p) \Psi_{j,m}^p[t_n],$$

где N – длина сигнала, J – наибольший масштаб.

Шаг 3. По правилу 2 определяем наилучший аппроксимирующий вейвлет Ψ_{min} . На выходе Алгоритма имеем вектор $\tilde{F}_{\Psi_{min}}[t_n]$.

Алгоритм адаптивной вейвлет-фильтрации сигнала (Алгоритм 2):

Шаг 1. Выполняем дискретное вейвлет-преобразование сигнала $\tilde{F}_{\Psi_{min}}[t_n]$ и применяем пороговые функции (величины порогов определяем по правилу 1):

$$\tilde{F}[t_n] = \sum_k \sum_{n=1}^N \Pi(\tilde{F}_{\Psi_{min}}[t_n], \Psi_{k,n}) \Psi_{k,n}[t_n],$$

$$\Pi(\tilde{F}_{\Psi_{min}}[t_n], \Psi_{k,n}[t_n]) = \begin{cases} \tilde{F}_{\Psi_{min}}[t_n], \Psi_{k,n}[t_n], & \text{если } |\tilde{F}_{\Psi_{min}}[t_n], \Psi_{k,n}[t_n]| \geq T_{\alpha,k,n}, \\ 0, & \text{если } |\tilde{F}_{\Psi_{min}}[t_n], \Psi_{k,n}[t_n]| < T_{\alpha,k,n} \end{cases},$$

$T_{\alpha,k,n} = \tau_\alpha * \delta_{k,n}$ – пороги.

Полученный преобразованный сигнал $\tilde{F}[t_n]$ содержит аномальные особенности.

Шаг 2. Оцениваем суммарную интенсивность выделенных аномальных особенностей:

табл. 1. Оценки выполнялись для вейвлетов порядков 1 – 3, с учетом допустимых размеров их носителей. Погрешности аппроксимации для разных вейвлетов рассчитывались по данным станций Инувик и Оулу в периоды высокой и низкой активности Солнца. В работе учитывались такие характеристики вейвлетов, как размер носителя (РН) и число нулевых моментов (ЧНМ).

Комплексный алгоритм анализа данных и обнаружения аномалий.

Алгоритм определения информационных компонент сигнала (Алгоритм 1):

Шаг 1. Используя базисы вейвлет-пакетов: $W_j^p = \bigoplus_{l=0}^j W_j^l, \{\Psi_j^p(2^j t - m)\}_{m \in \mathbb{N}}$, выполняем разложение сигнала и применяем пороговые функции, величины порогов определяем по правилу 1:

Базис $B_j^p = \{\Psi_j^p(2^j t - m)\}_{m \in \mathbb{N}}$ пространства W_j^p есть базис

$$E[t_n] = \sum_{k=0}^K \Pi(\tilde{F}_{\Psi_{min}}[t_n], \Psi_{k,n}[t_n]).$$

Замечание 4. Для обеспечения оперативности получения результата алгоритма и сокращения времени его выполнения в шаге 1 Алгоритма 1 использовался предварительно определённый по Правилу 2 набор вейвлет-функций (см. табл. 1).

На рис. 1 представлена блок-схема предлагаемого комплексного алгоритма.

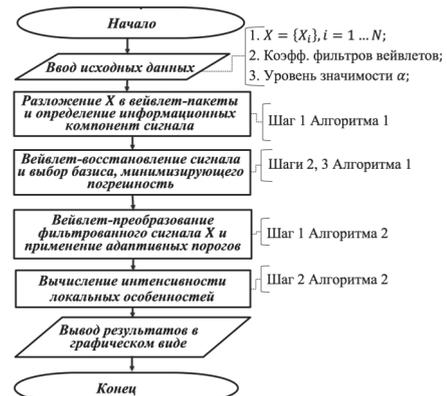


Рис. 1. Схема реализации комплексного алгоритма

2. Результаты применения алгоритма

В работе использовались природные данные вторичных космических лучей [15] и модельные данные, построенные по их подобию. Вторичные космические лучи – это частицы высоких энергий, регистрируемые наземными станциями нейтронных мониторов [15].

Аномальные изменения в потоке вторичных космических лучей свидетельствуют о возмущениях в околоземном космическом пространстве, которые способны нарушить работу систем спутниковой, навигационной и радиосвязи, а также повысить риск обострения сердечно-сосудистых заболеваний людей [7, 16]. Аномальные изменения в космических лучах могут проявляться в виде так называемых Форбуш-эффектов (внезапное локальное снижение интенсивности космических лучей) [17] и GLE-событий (сильное локальное возрастание интенсивности космических лучей, регистрируемых на поверхности Земли) [18].

На рис. 2 представлен пример дерева вейвлет-пакетов, построенного по алгоритму определения информационных компонент сигнала. Красным цветом отмечены выбранные информационные узлы. Ниже, на рис. 2, синим цветом показан исходный сигнал, оранжевым – результат данного алгоритма. Результат подтверждает нестационарную структуру регистрируемых данных вторичных космических лучей и наличие шума.

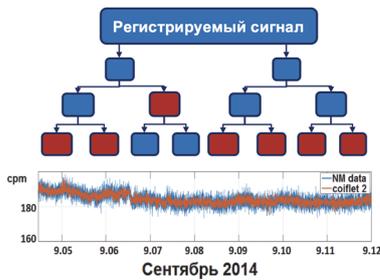


Рис. 2. Результат алгоритма 1

На рис. 3 показан результат обнаружения малоамплитудной продолжительной аномальной особенности (отношение сигнал/шум 1.1) в построенном модельном сигнале на фоне коррелированного шума (добавлен аддитивный розовый шум). Обнаружение выполнено на основе применения операции непрерывного вейвлет-преобразования (НВП) (рис. 3в, г) и на основе предложенного комплексного алгоритма (рис. 3 д, е). Следует отметить, что аномальная особенность в зашумленном модельном сигнале не видна (рис. 3б). Полученный результат показывает, что ввиду наличия коррелированного шума её обнаружить на основе непрерывного вейвлет-преобразования невозможно. Предложенный алгоритм позволил подавить шум и обнаружить аномальную особенность.

На рис. 4 представлен пример применения комплексного алгоритма к данным нейтронного монитора ст. Оулу за 15–18 июля (слева) и модельным данным (справа), содержащим аномальную особенность. В модельный сигнал добавлялись аномальные особенности вида треугольный импульс с отношением сигнал/шум = 1,5. Момент регистрации Форбуш-эффекта в природных данных [19] и добавленная аномальная особенность в модельные данные отмечены красными пунктирными вертикальными линиями. Результаты

комплексного алгоритма представлены по мере поступления данных в систему обработки. Результат алгоритма подтверждает его эффективность как для модельных, так и для природных данных. Алгоритм позволил обнаружить аномальную особенность своевременно.

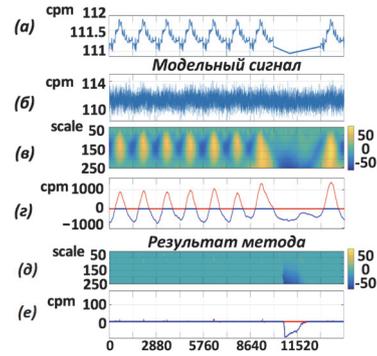


Рис. 3. Результаты обработки: (а) тренд+аномалия; (б) тренд+аномалия+шум; (в, г) результат применения НВП; (д, е) результат применения комплексного алгоритма

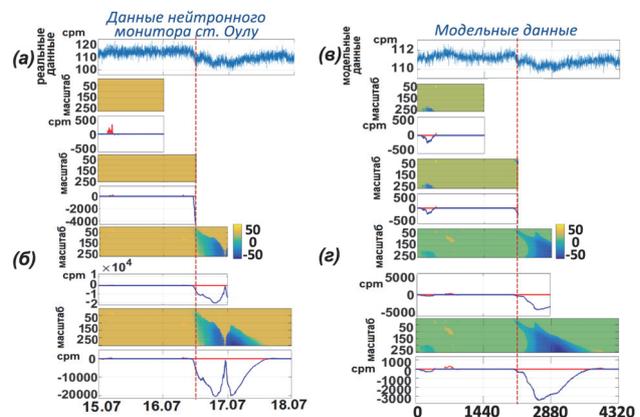


Рис. 4. Результат обработки: (а) данные нейтронного монитора ст. Оулу; (б) результаты применения комплексного алгоритма к данным нейтронного монитора (обработка выполнена по мере поступления данных в систему обработки); (в) модельные данные, построенные по подобию природных; (г) результаты применения комплексного алгоритма к модельным данным (обработка выполнена по мере поступления данных в систему обработки)

В табл. 2 представлен расчет временных затрат на работу комплексного алгоритма. Расчет выполнен в зависимости от объема входных данных. Реализация комплексного алгоритма выполнена в программе, написанной на языке Java в среде разработки IntelliJ IDEA 2020.3.4 [20]. Временные затраты позволяют применять алгоритм в оперативном анализе данных нейтронных мониторов.

Табл. 2. Временные затраты на работу комплексного алгоритма (мкс)

Операции	Объем входных данных (отсчеты)		
	N = 2880	N = 4320	N = 5760
Вывод результата программы	75,45	109,78	127,05

Заключение

Результаты исследования показали эффективность предложенного комплексного алгоритма для анализа природных данных и обнаружения аномальных особенностей разной формы и продолжительности. Применение алгоритма позволило обнаружить аномальную особенность малой амплитуды (отношение сигнал / шум 1.1) на фоне коррелированного шума. Данный результат свидетельствует о высокой эффективности разработанного алгоритма. Выполненный расчет временных затрат на выполнение алгоритма показал возможность его применения в режиме, близком к реальному времени, что важно для задач космической погоды.

Благодарности

Работа выполнена за счет Государственного задания ИКИР ДВО РАН (рег. № темы 124012300245-2).

References

- [1] Murzagulov DA, Zamyatin AV. The process signal anomaly detection using classifier ensemble and wavelet transforms [In Russian]. *Automated Control Systems* 2021; 1(63): 20-26. DOI: 10.35752/1991-2927-2021-1-63-20-26.
- [2] Murzagulov DA, Zamyatin AV, Romanovich OV. Approach to the detection of anomalies in process signals by using the Hilbert–Huang transform. *Optoelectron Instrum Data Process* 2021; 57(1): 27-36. DOI: 10.3103/S8756699021010076.
- [3] Davydov NS, Evdokimova VV, Serafimovich PG, Protsenko VI, Khramov AG, Nikonov AV. Neural network for step anomaly detection in head motion during fMRI using meta-learning adaptation. *Computer Optics* 2023; 47(6): 991-1001. DOI: 10.18287/2412-6179-CO-1337.
- [4] Borog VV, Kryanev AV, Udumyan DK. Combined method for detecting hidden anomalies in galactic cosmic ray variations. *Geomagn Aeron* 2011; 51(4): 475-482. DOI: 10.1134/S0016793211040086.
- [5] Vorobev AV, Vorobeva GR. Geoinformation system for amplitude-frequency analysis of geomagnetic variations and space weather observation data. *Computer Optics*

- 2017; 41(6): 963-972. DOI: 10.18287/2412-6179-2017-41-6-963-972.
- [6] Mandrikova OV. Intelligent methods for natural data analysis: application to space weather. *Computer Optics* 2024; 48(1): 139-148. DOI: 10.18287/2412-6179-CO-1367.
- [7] Kuznetsov VD. Space weather and risks of space activity [In Russian]. *Space Engineering and Technology* 2014; 3(6): 3-13.
- [8] Mandrikova O, Mandrikova B. Hybrid model of natural time series with neural network component and adaptive nonlinear scheme: Application for anomaly detection. *Mathematics* 2024; 12(7): 1079. DOI: 10.3390/math12071079.
- [9] Mallat SG. *A wavelet tour of signal processing*. San Diego, CA: Academic Press; 1999.
- [10] Levin BR. *Theoretical foundations of statistical radio engineering* [In Russian]. Moscow: "Radio i Svyaz" Publisher; 1989.
- [11] Polyakova VV, Shabrova NV. *Fundamentals of the theory of statistics* [In Russian]. Ekaterinburg: Publishing House of the Ural University; 2015.
- [12] Daubechies I. *Ten lectures on wavelets*. Philadelphia, Pennsylvania: Society for Industrial and Applied Mathematics; 1992. ISBN: 978-0-898712-74-2.
- [13] Chui CK. *An introduction in wavelets*. Vol 1. New York, USA: Academic Press; 1992. ISBN: 978-0121745844.
- [14] Mandrikova OV, Geppener VV, Mandrikova BS. Method of cosmic ray data analysis based on vector quantization neural networks. *J Phys: Conf Ser* 2019; 1368(5): 052026. DOI: 10.1088/1742-6596/1368/5/052026.
- [15] Neutron Monitor Network Database. 2024. Source: <www.nmdb.eu>.
- [16] Samsonov SN, Manykina VI, Parshina SS. Space weather influence on cardiovascular system of healthy people and people with weakened adaptive capability [In Russian]. *Psychosomatic and Integrative Research* 2016; 2(1): 0102.
- [17] Dorman LI. Cosmic ray variations and space weather. *Phys-Usp* 2010; 53(5): 496-503. DOI: 10.3367/UFNe.0180.201005g.0519.
- [18] Firoz KA, Gan WQ, Li YP, Rodríguez-Pacheco J, Kudela K. On the possible mechanism of GLE initiation. *Astrophys J* 2019; 872(178): 178. DOI: 10.3847/1538-4357/ab0381.
- [19] Space Weather Forecasting Center IZMIRAN. 2024. Source: <<http://spaceweather.izmiran.ru/rus/>>.
- [20] Java and Kotlin IDE. 2024. Source: <<https://www.jetbrains.com/idea/>>.

Сведения об авторах

Лисс Александр Рудольфович родился в 1946 году. В 1969 году окончил Ленинградский государственный университет, в 1988 году присуждена ученая степень доктора технических наук, в 1991 году присуждено ученое звание профессора. Профессор кафедры математического обеспечения и применения ЭВМ Санкт-Петербургского государственного электротехнического университета ЛЭТИ имени В.И. Ульянова (Ленина). Научные интересы включают методы цифровой обработки сигналов, статистические методы принятия решений, разработку систем обработки сигналов реального времени. Автор более 80 публикаций. E-mail: alexander.liss@mail.ru

Мандрикова Богдана Сергеевна родилась в 1996 году. В 2023 году окончила аспирантуру Санкт-Петербургского государственного электротехнического университета ЛЭТИ имени В.И. Ульянова (Ленина). Научный сотрудник лаборатории системного анализа Института космофизических исследований и распространения радиоволн Дальневосточного отделения Российской академии наук. Научные интересы включают интеллектуальные

методы анализа геофизических данных, вейвлеты, нейронные сети. Автор более 50 публикаций (книг, статей). E-mail: 555bs5@mail.ru

Мандрикова Оксана Викторовна родилась в 1972 году. В 1995 году окончила Киевский национальный университет имени Т. Шевченко, в 2009 году присвоено звание доктора технических наук, в 2021 году присуждено ученое звание профессора. Заведующая лабораторией системного анализа Института космофизических исследований и распространения радиоволн Дальневосточного отделения Российской академии наук. Научные интересы включают интеллектуальные методы анализа геофизических данных, вейвлеты, нейронные сети, ионосферу, магнитосферу и аномалии сигналов. Автор более 100 публикаций (книг, статей). E-mail: oksanam1@mail.ru

ГРНТИ: 27.41.17

Поступила в редакцию 26 ноября 2024 г. Окончательный вариант – 10 февраля 2025 г.

Complex algorithm for detecting anomalous features in natural data

A.R. Liss¹, B.S. Mandrikova², O.V. Mandrikova²

¹ Saint Petersburg Electrotechnical University, St Petersburg, Russia,
197376, Russia, St. Petersburg, st. Professora Popova, 5;

² Institute of Cosmophysical Research and Radio Wave Propagation,
Far Eastern Branch of the Russian Academy of Sciences,
684034, Kamchatskiy Kray, Paratunka, Russia, Mirnaya st. 7

Abstract

A complex automated algorithm for analyzing natural data and detecting anomalous features is proposed. The algorithm includes an algorithm for determining the information components of a signal and an algorithm for adaptive wavelet filtering of a signal. The algorithm for determining the information components of a signal suppresses correlated noise and determines the information components of a signal. The algorithm for adaptive wavelet filtering of a signal detects anomalous features and estimates their intensity. The algorithms are based on the rules developed by the authors. Based on the rules, the parameters of the threshold function are estimated and the best approximating wavelet is determined. The article describes the operations of the complex algorithm and presents a block diagram of its implementation. Also presented are the results of applying a complex algorithm using data from secondary cosmic rays and model data constructed in their likeness. The results confirmed the effectiveness of the developed rules and the proposed complex algorithm.

Keywords: wavelet transform, risk theory, natural anomalies, cosmic rays.

Citation: Liss AR, Mandrikova BS, Mandrikova OV. Complex algorithm for detecting anomalous features in natural data. *Computer Optics* 2025; 49(6): 1030-1036. DOI: 10.18287/2412-6179-CO-1652.

Acknowledgements: The work was supported by IKIR FEB RAS State Task (subject registration No. 124012300245-2).

Author's information

Alexander Rudolfovich Liss was born in 1946. In 1969 he graduated from Leningrad State University, in 1988 he was awarded the academic degree of Doctor of Technical Sciences, in 1991 he was awarded the academic title of professor. Professor of the Department of Mathematical Support and Application of Computers of the Saint Petersburg State Electrotechnical University LETI named after V.I. Ulyanov-Lenin. Research interests include methods of digital signal processing, statistical methods of decision making, development of real-time signal processing systems. Author of more than 80 publications. E-mail: alexander.liss@mail.ru

Bogdana Sergeevna Mandrikova was born in 1996. In 2023 she graduated from the postgraduate program of the St. Petersburg State Electrotechnical University LETI named after V.I. Researcher at the Laboratory of Systems Analysis of the Institute of Cosmophysical Research and Radio Wave Propagation of the Far Eastern Branch of the Russian Academy of Sciences. Research interests include intelligent methods for analyzing geophysical data, wavelets, neural networks. Author of more than 50 publications (books, articles). E-mail: 555bs5@mail.ru

Oksana Viktorovna Mandrikova was born in 1972. Graduated from Shevchenko Kyiv National University in 1995 and was awarded the title of Doctor of Technical Science in 2009, in 2021 he was awarded the academic title of professor. Head of the Laboratory of System Analysis at the Institute of Cosmophysical Research and Radio Wave Propagation, Far East Branch, Russian Academy of Sciences. Her scientific interests are intelligent techniques for geophysical data analysis, wavelets, neural networks, the ionosphere, the magnetosphere, and signal anomalies. Author of more than 150 publications (books, papers). E-mail: oksanaml@mail.ru

Received November 26, 2024. The final version – February 10, 2025.
