

## Локализация и классификация аномалий в одномерных сигналах на основе вейвлет-анализа и математических методов оптимизации

Н.Д. Сакович<sup>1,2</sup>, Д.А. Аксенов<sup>1,2</sup>, Е.С. Плешакова<sup>3</sup>, С.Т. Гатауллин<sup>4</sup>

<sup>1</sup> Финансовый университет при Правительстве РФ, 125167, Москва, Ленинградский пр-кт, д. 49/2;

<sup>2</sup> НИИ – филиал АО "Гознак", 115162, Москва, Мытная ул., д. 19;

<sup>3</sup> МИРЭА – Российский технологический университет, 119454, Москва, пр. Вернадского, д. 78;

<sup>4</sup> Центральный экономико-математический институт Российской академии наук, 117418, Москва, Нахимовский пр., д. 47

### Аннотация

Метод классификации типов нестационарностей во временных рядах на основе вейвлет-анализа для локализации и детекции границ нестационарности; математических алгоритмов для сравнения типов нестационарностей на основе шаблонов эталонных сигналов и классификации аномалий при помощи неградиентных методов оптимизации; с возможностью дальнейшего применения для разработки автоматизированных информационных систем комплексного мониторинга состояния электросети и ее отдельных компонент представлен в данном исследовании. Кодовая база размещена в открытом репозитории проекта и доступна для воспроизведения вычислительных экспериментов. Предлагаемый подход имеет широкий спектр применения, но отдельного внимания заслуживает возможная интеграция в целях повышения надежности систем диагностики с технологиями Интернета вещей, построения цифровых двойников, агент-ориентированного моделирования сложных социально-экономических процессов, облачными высокопроизводительными вычислениями, позволяющими анализировать большие данные в реальном времени.

**Ключевые слова:** цифровая обработка сигналов, вейвлет-анализ, численные методы, алгоритмы.

**Цитирование:** Сакович, Н.Д. Локализация и классификация аномалий в одномерных сигналах на основе вейвлет-анализа и математических методов оптимизации / Н.Д. Сакович, Д.А. Аксенов, Е.С. Плешакова, С.Т. Гатауллин // Компьютерная оптика. – 2026. – Т. 50, № 1. – 1683. – DOI: 10.18287/COJ1683.

**Citation:** Sakovich ND, Aksenov DA, Pleshakova ES, Gataullin ST. Localization and classification of anomalies in one-dimensional signals based on wavelet analysis and mathematical optimization methods. Computer Optics 2026; 50(1): 1683. DOI: 10.18287/COJ1683.

### Введение

Проблематика обнаружения сбоев сигналов занимает ключевое место в сферах, связанных с анализом сейсмической активности [1 – 3], оценкой вибраций промышленного оборудования [4 – 6], контролем целостности конструкций и надежностью электрических сетей [7 – 9]. Эффективная диагностика сбоев позволяет оперативно выявлять и устранять проблемы, минимизируя риск аварий и повреждений оборудования. Однако существующие методы часто сталкиваются с ограничениями в точности и чувствительности, особенно при выявлении сложных и непосредственных сбоев. Традиционные методы, такие как оконное преобразование Фурье, не обеспечивают детального временно-частотного анализа сигналов, что затрудняет выявление незначительных изменений в сигналах. В связи с этим поиск новых и более эффективных методов диагностики становится одной из ключевых задач современной науки и инженерии. Одним из наиболее перспективных направлений в области обработки сигналов является использование вейвлет-преобразования. Вейвлет-анализ позволяет проводить детальный временно-частотный анализ сигналов, что делает его особенно привлекательным для диагностики сбоев в электрических цепях. Множество исследований по применению вейвлет-преобразования в различных областях, включая медицинскую диагностику [10 – 14], обработку аудиосигналов и анализ данных [15 – 18], подтверждает его эффективность и гибкость. В частности, вейвлет-преобразование позволяет выявлять мелкие и сложные изменения в сигналах, что является критически важным для диагностики сбоев в электрических цепях.

В данной работе представлена система классификации типов ошибок в одномерных сигналах на основе разработанного авторами алгоритма вейвлет-анализа, дополненного численными методами. Предложенный подход использует вейвлет-преобразование для локализации сбоев в сигнале для последующей классификации с использованием методов оптимизации через минимизацию целевой функции ошибки. Это позволяет не только улучшить точность идентификации сбоев, но и обеспечивает более глубокое понимание их природы. Предложенный метод может быть использован после адаптации в агентном моделировании сложных социально-экономических процессов в сочетании с высокопроизводительными вычислениями (HPC) [19 – 20].

Ранее авторами рукописи уже были получены значимые научные результаты в области приложений гибридных нейросетевых моделей цифровой обработки сигналов. Так, в [21] была разработана и протестирована

на наборе данных, собранном авторами в результате полевого исследования, нейросетевая модель вида Windowed Fourier Transform (WFT) – 2D-CapsNet мониторинга состояния промышленного оборудования для бурения. В работе [22] метод был модифицирован до вида Frequency Slice Wavelet Transform (FSWT) – 2D-CapsNet и произведено сравнение с базовым подходом. Несмотря на близкие количественные оценки производительности, модель FSWT – 2D-CapsNet оказалась более эффективной и была успешно использована авторами в задачах мониторинга физиологических показателей посредством анализа фотоплетизмограмм [23].

Современные подходы к задачам классификации данных всё чаще опираются на методы машинного обучения, в частности, на нейросетевые архитектуры. Однако использование нейронных сетей сопряжено с рядом ограничений, основное из которых заключается в необходимости формирования обширных обучающих выборок. Эти датасеты, как правило, собираются в рамках конкретной предметной области, что усложняет процесс адаптации модели к новым сферам применения и требует значительных вычислительных ресурсов для её переобучения.

В предлагаемом методе используется альтернативный подход, основанный на вейвлет-преобразовании и численных методах оптимизации. Вейвлет-преобразование [30, 31] выполняет задачу локализации значимых особенностей во входных данных, что позволяет эффективно выделять информативные признаки. В отличие от нейросетевых методов, требующих сложной процедуры обучения, классификация в нашем подходе осуществляется на основе математических моделей, использующих библиотеки базовых эталонных сигналов.

Ключевой аспект предлагаемого метода заключается в применении метрики среднеквадратичного отклонения (MSE) в качестве критерия сходства. Данный подход позволяет сравнивать входные данные с заранее сформированными эталонами и производить классификацию без необходимости предварительного обучения модели на специализированных выборках. Таким образом, обеспечивается гибкость метода и его переносимость в различные предметные области без значительных затрат на адаптацию.

Одним из важных преимуществ предлагаемого метода является его модульность. Данный подход можно рассматривать как сервисно-ориентированный, что даёт возможность заменять отдельные компоненты системы без нарушения её общей структуры. Например, этап классификации или локализации может быть заменён на нейросетевые модели, если это окажется целесообразным для конкретной задачи. Такой гибридный подход позволяет адаптировать алгоритм под различные сценарии, сочетая преимущества традиционных математических методов и машинного обучения.

Некоторые параметры алгоритма, такие как пороговые значения (threshold) и другие коэффициенты, зависят от предметной области и специфики решаемой задачи. Их выбор осуществляется с учётом характеристик данных и может требовать дополнительной калибровки для оптимального функционирования модели.

В целом, предложенный метод сочетает преимущества вейвлет-преобразования для локального анализа данных, численных методов оптимизации для эффективной классификации и гибкость модульного подхода, позволяя при необходимости интегрировать нейросетевые компоненты. Это делает его конкурентоспособной альтернативой чисто нейросетевым моделям, особенно в случаях, когда доступность больших обучающих выборок ограничена, а адаптация к новым данным должна быть выполнена с минимальными вычислительными затратами.

Организация настоящей рукописи выглядит следующим образом. В параграфе материалов и методов представлены базовые алгоритмы и подходы, схема многоуровневой архитектуры прикладного решения, схема и описание разработанного авторами метода локализации и классификации сбоев в сигнале; в экспериментальном параграфе приведены результаты численных экспериментов, описаны ограничения применимости метода и возможные пути улучшения эффективности работы системы, произведены тесты производительности; заключение финализирует исследование, обобщает полученные результаты и обозначает направления будущей работы.

## 1. Материалы и методы

Система [24] анализа сигналов (см. схему на рис. 1) на основе алгоритма вейвлет-анализа, дополненного численными методами (см. схему на рис. 2) использует следующие базовые методы и алгоритмы.

Данная архитектура представляет собой многоуровневую архитектуру, ориентированную на сбор, обработку, передачу, анализ и предоставление данных в рамках распределённой вычислительной среды. Она может быть рассмотрена как прикладной пример реализации алгоритмов обработки данных с использованием современных технологий распределённого хранения, распределённой обработки и анализа.

Алгоритм вейвлет-разложения, часто называемый алгоритмом Маллата [25] или быстрым вейвлет-преобразованием, представляет собой эффективный метод для разложения сигнала на несколько уровней детализации. Этот алгоритм является основой для многих приложений обработки сигналов, включая шумоподавление, сжатие данных и анализ особенностей. Псевдокод, представленный как ‘WaveletDecomposition’, описывает схему алгоритма Маллата для одномерных сигналов и соответствует реализации в библиотеке ‘pywt.wavdec’.

Метод Нелдера–Мида [26] является одним из популярных методов неградиентной оптимизации. Он основан на идее последовательного изменения и перемещения симплексов по пространству для поиска необходимого оптимума. В данном случае это актуально, так как не всегда получится взять градиент от целевой функции.

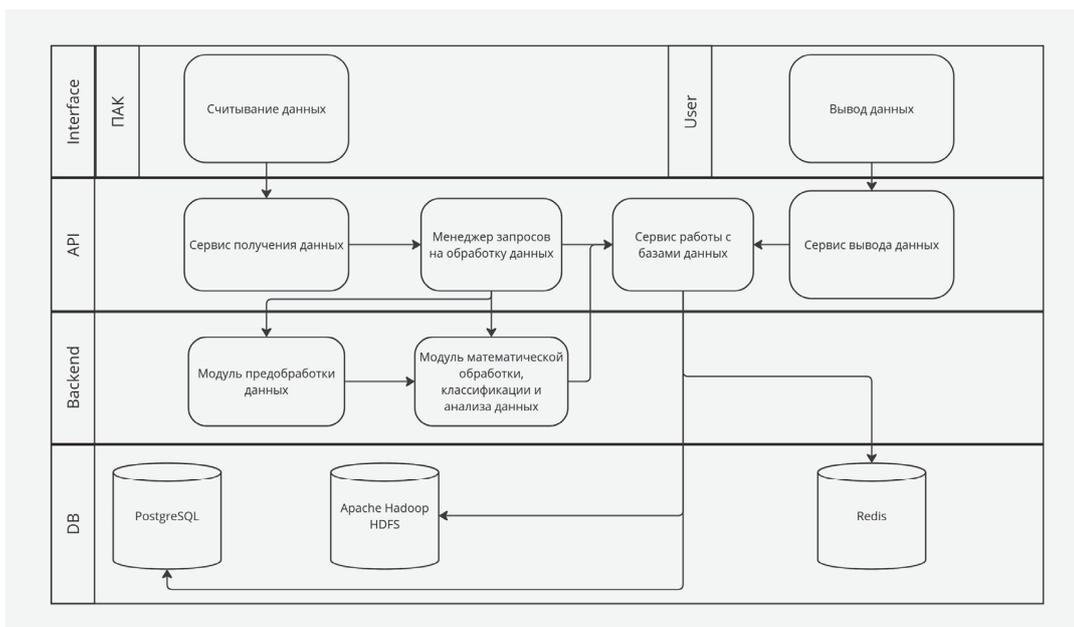


Рис. 1. Диаграмма многоуровневой архитектуры системы анализа сигналов

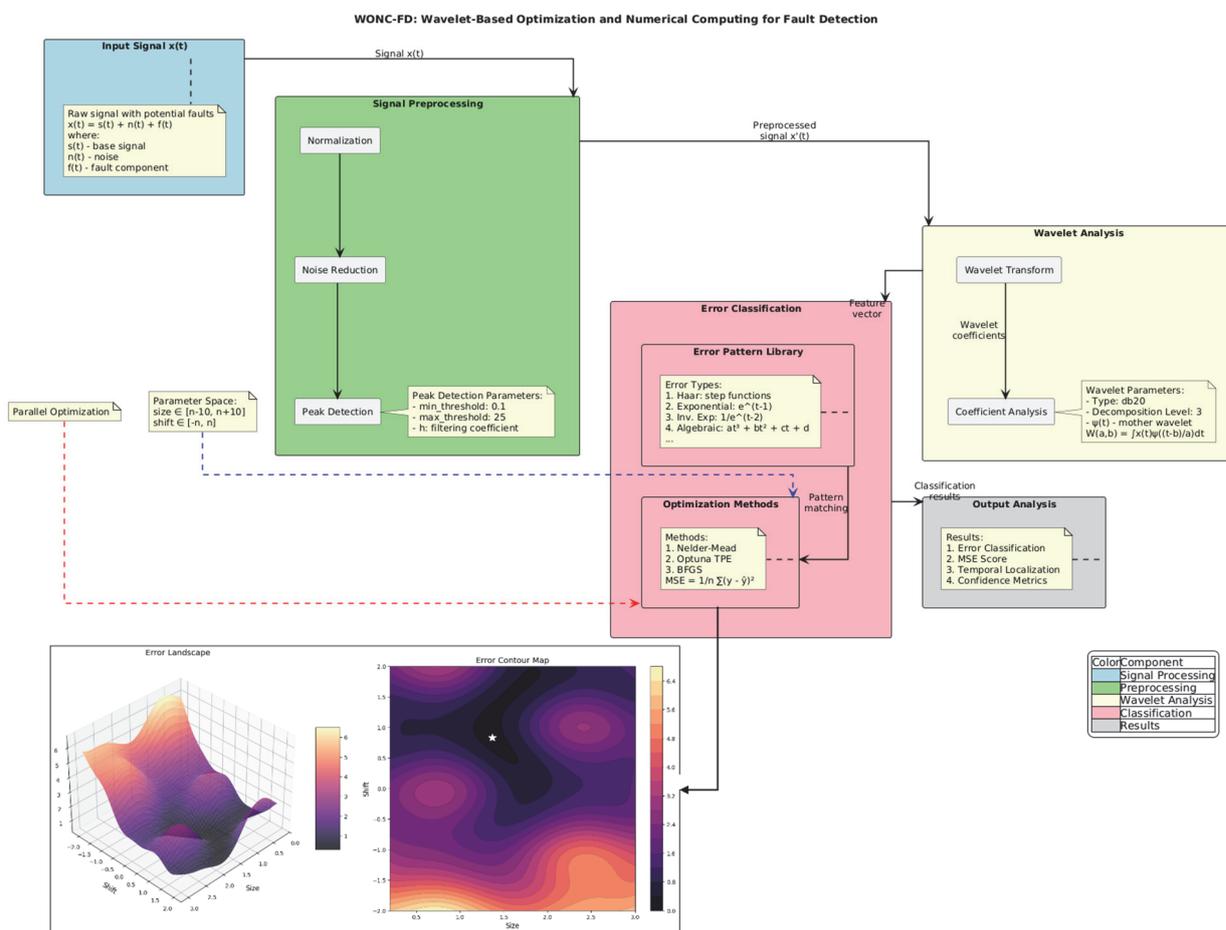


Рис. 2. Схема работы алгоритма локализации и классификации сбоев в сигнале

Алгоритм Бройдена–Флетчера–Гольдфарба–Шанно (BFGS) [27] – один из самых быстрых и точных алгоритмов оптимизации, использующий градиенты функций. В данном случае его эффективно использовать, если требуется быстрая скорость при необязательной высокой точности классификации.

Алгоритм байесовской оптимизации (Tree-structured Parzen Estimator, TPE) [28] используется в Optuna для эффективного поиска оптимальных параметров в задачах, где вычисление целевой функции является затратным. TPE является sampler-ом, то есть методом, который определяет, какие параметры следует исследовать на

следующей итерации оптимизации. В отличие от случайного поиска или grid search, TPE использует историю предыдущих испытаний, чтобы интеллектуально направлять поиск в наиболее перспективные области пространства параметров.

Алгоритм метода наименьших квадратов (Least Squares Method) [29].

### 1.1. Вейвлет-преобразование

Основополагающим методом для работы алгоритма является вейвлет-преобразование, которое позволяет разложить входной сигнал на детализирующие коэффициенты, которые можно использовать для последующего анализа сбоев и их локализации.

---

#### Алгоритм 1: WaveletDecomposition (схема Mallat's Algorithm, pywt.wavedec)

---

```

1: Input: signal — входной одномерный сигнал,
2: Input: wavelet — тип вейвлета (например, 'db20'),
3: Input: level — глубина разложения
4: Output: массив коэффициентов  $coeffs = [cA_{level}, cD_{level}, \dots, cD_1]$ 
5: Инициализировать список коэффициентов coeffs пустым
6: current_signal  $\leftarrow$  signal
7: for  $k \leftarrow 1$  до level do
8: Выполнить свёртку current_signal с фильтром масштабирования (low-pass) из выбранного вейвлета  $\rightarrow cA_k$ 
9: Выполнить свёртку current_signal с фильтром детализации (high-pass) из выбранного вейвлета  $\rightarrow cD_k$ 
10: Произвести операцию «даунсэмплинга» (выбрать каждый второй отсчёт) для  $cA_k$  и  $cD_k$ 
11: current_signal  $\leftarrow cA_k$   $\triangleright$  Сигнал на следующем шаге — это аппроксимирующие коэффициенты
12: Сохранить  $cD_k$  во внутренний буфер
13: end for
14: Сохранить current_signal (т.е.  $cA_{level}$ ) в coeffs
15: Добавить все детализирующие коэффициенты  $cD_k$  в coeffs в порядке от последнего к первому
16: return coeffs

```

---

### 1.2. Методы предобработки данных

Далее необходимо рассмотреть методы предобработки данных, то есть полученного сигнала для последующего анализа: локализации и классификации.

#### 1.2.1. Алгоритм intervals\_errors

---

#### Алгоритм 2: intervals\_errors

---

```

1: Input: signal — одномерный массив
2: Input: wavelet — тип вейвлета (например, 'db20'),
3: Input: level = 3 — уровень разложения
4: Input: epsilon = 2
5: Input: min_threshold = 0.1, max_threshold = 25
6: Input:  $h = 10^{-6}$ 
7: Output: список интервалов (пар), где предполагаются ошибки
8: coeffs  $\leftarrow$  WaveletDecomposition(signal, wavelet, level)  $\triangleright$  См. алгоритм 1.
9: cf_lvl  $\leftarrow$  coeffs[level]  $\triangleright$  Коэффициенты на заданном уровне (детализирующие или аппроксимирующие – зависит от библиотеки)
10: peaks  $\leftarrow$  detect_faults(cf_lvl, min_threshold, max_threshold)
11: Построить список eps_peaks путём peak –  $\epsilon$  для каждого  $peak \in peaks[0]$ 
12: intervals  $\leftarrow$  combinations(eps_peaks, 2)
13: return intervals

```

---

Алгоритм intervals\_errors предназначен для определения интервалов, в которых предположительно находятся ошибки или повреждения в сигнале. Он использует вейвлет-разложение, обнаружение пиков и генерацию пар для определения этих интервалов.

#### **Замечания:**

Алгоритм intervals\_errors комбинирует вейвлет-анализ с обнаружением пиков и комбинаторикой для определения интервалов ошибок. Использование вейвлет-разложения позволяет анализировать сигнал на разных частотных уровнях и, возможно, лучше выделять особенности, связанные с повреждениями. Параметр  $\epsilon$  и генерация пар могут быть связаны с попыткой определить интервалы вокруг пиков, в которых наиболее вероятно

наличие ошибки. Настройка параметров вейвлета, уровня разложения, порогов prominence и порога фильтрации, а также значения  $\varepsilon$  критически важна для эффективности алгоритма в конкретных приложениях.

### 1.2.2. Прочие алгоритмы

Алгоритм `find_peaks` предназначен для обнаружения локальных максимумов (пиков) в одномерном числовом массиве. Основная цель алгоритма – выделить индексы точек, которые соответствуют пику, удовлетворяющим заданным критериям, в частности, по 'prominence' (выделенности пика).

Алгоритм `detect_faults` предназначен для обнаружения потенциальных «повреждений» в одномерном сигнале. Он использует алгоритм поиска пиков `find_peaks` для идентификации пиков в абсолютном значении сигнала, которые интерпретируются как индикаторы повреждений.

Алгоритм `clean_signal` предназначен для «очистки» сигнала, представленного в виде структуры или массива с полями  $x$  и  $y$ . Алгоритм удаляет синусоидальную компоненту из сигнала  $y$  для исследования последующего исследования сбоя в сигнале.

Алгоритм `pad_array` предназначен для дополнения входного массива `arr` нулями до заданной длины `length`. Если входной массив уже имеет длину, большую или равную `length`, алгоритм возвращает исходный массив без изменений. В случае, если длина массива меньше требуемой, недостающие элементы добавляются симметрично с обеих сторон массива в виде нулей.

Алгоритм `resize_vector` предназначен для изменения размера одномерного массива (вектора) `vector` до новой заданной длины `new_size`. Алгоритм предусматривает два основных случая изменения размера: увеличение длины и уменьшение длины, и использует разные подходы для каждого случая.

### 1.3. Алгоритм

Теперь можно рассмотреть основные алгоритмы для локализации и классификации сбоя в сигнале, которые основываются на использовании предыдущих методов и алгоритмов.

Математически можно описать решаемую задачу следующим образом:

$$\operatorname{argmin}_{(s,t,a) \in S \times T \times A} \operatorname{MSE}(y, \operatorname{ErrorFunction}(s, t, a)),$$

где:

1.  $y$ : исходный сигнал, представленный в виде временного ряда. Этот сигнал является эталонным и используется для оценки точности приближения.
2.  $\hat{y}$ : приближенный сигнал, полученный в результате применения функции ошибки с определенными параметрами. Целью является максимально точное соответствие этого сигнала исходному сигналу  $y$ .
3. `ErrorFunction`: функция, которая принимает на вход параметры растяжения, сдвига и амплитуды и возвращает приближенный сигнал. Формально `ErrorFunction(s, t, a)` обозначает приближенный сигнал, полученный с учетом параметров  $s$ ,  $t$  и  $a$ .
4.  $S$ : множество возможных значений параметра растяжения сигнала по оси OX. Каждый элемент этого множества представляет собой коэффициент растяжения, который может быть применен к исходному сигналу.
5.  $T$ : множество возможных значений параметра сдвига сигнала. Каждый элемент этого множества представляет собой величину сдвига, которая может быть применена к исходному сигналу.
6.  $A$ : множество возможных значений параметра амплитуды сигнала. Каждый элемент этого множества представляет собой коэффициент амплитуды, который может быть применен к исходному сигналу.
7.  $S \times T \times A$ : декартово произведение множеств  $S$ ,  $T$  и  $A$ , которое представляет собой множество всех возможных комбинаций параметров растяжения, сдвига и амплитуды.

Целью работы метода является нахождение такой комбинации параметров  $(s, t, a)$ , которая минимизирует среднеквадратическую ошибку между исходным сигналом  $y$  и приближенным сигналом  $\hat{y}$ .

Таким образом, данная формула определяет процедуру поиска оптимальных параметров растяжения, сдвига и амплитуды, которые обеспечивают наилучшее приближение исходного сигнала в смысле минимизации среднеквадратической ошибки.

#### 1.3.1. Алгоритм `WONC-FD mse_classification`

---

##### Алгоритм 3: `WoncFD mse_classification`

---

- 1: **Input:**
- 2:  $y$  — одномерный массив точек
- 3: `errors` — список возможных типов ошибок (например, ["haar", "haar1", ...])
- 4: `leng` — базовая длина сигнала (например, 1000)
- 5: `epsilon` — маленький порог для проверки близости к 0
- 6:  $n$  — доля нулевых (по модулю) точек для отбраковки
- 7: **Output:** список: [название лучшего типа ошибки, MSE]

```

8: if (length(y) ≤ leng × 0,05) or ( $\sum_i 1(|y_i| < \varepsilon) > \text{length}(y) \times n$ ) then
9:   return ["Bad signal", "NaN"]
10: end if
11: Инициализировать пустой массив result
12: x ← массив равномерных значений от 0 до 1 размером length(y)
13: ampl ← 1
14: size_values ← {length(y), length(y) + 10, ..., 2 × length(y)}
15: shifts ← {−length(y), −length(y) + 1, ..., length(y)}
16: for каждый err из errors do
17:   mse_err ← 105   ▷Задаём начальное большое значение MSE
18:   for каждый s из size_values do
19:     for каждый shift из shifts do
20:       Error ← resize_vector(Error_signal(err, x, ampl, 1), s)
21:       if shift ≥ 0 then
22:         arr_error ← массив Error, дополненный нулями в начале (сдвиг вперёд)
23:       else
24:         arr_error ← массив Error, дополненный нулями в конце (сдвиг назад)
25:       end if
26:       features_matrix ← матрица (length(x), 2), где первая колонка — 1, вторая — arr_error
27:       Найти коэффициенты  $\beta$  методом наименьших квадратов:  $\min_{\beta} \| \text{features\_matrix} \cdot \beta - y \|^2$ 
28:       approximated_signal ← features_matrix ·  $\beta$ 
29:       mse ← MSE(y, approximated_signal)
30:       if mse < mse_err then
31:         mse_err ← mse
32:       end if
33:     end for
34:   end for
35:   Добавить mse_err в массив result
36: end for
37: Найти индекс i* = argminresult
38: return [errors[i*], result[i*]]

```

Алгоритм *mse\_classification* предназначен для классификации типа «ошибки» или «дефекта» в одномерном сигнале *y* путем сравнения его с набором predetermined «ошибочных сигналов» *Error\_signal* различных типов, перечисленных в массиве *errors*. Классификация основана на минимизации среднеквадратической ошибки (MSE) между входным сигналом и аппроксимацией, полученной с использованием различных типов ошибок, размеров и сдвигов.

#### Основные шаги алгоритма

- **Входные данные:** алгоритм принимает на вход одномерный массив точек *y* (которые являются интервалами, полученными после применения вейвлет-преобразования и деления на данные интервалы), список названий типов ошибок *errors*, базовую длину сигнала *leng*, малый порог  $\varepsilon$  для проверки близости к нулю и долю нулевых (по модулю) точек *n* для отбраковки сигнала.
- **Предварительная проверка сигнала:** алгоритм выполняет начальную проверку входного сигнала *y*, чтобы отсеять «плохие» сигналы, которые могут привести к некорректным результатам.
  - **Проверка длины сигнала:** Проверяется, не слишком ли короткий сигнал. Если длина *y* меньше или равна 5 % от базовой длины *leng* (т.е.,  $\text{length}(y) \leq 0,05 \times \text{length}(y)$ ), сигнал считается «плохим».
  - **Проверка на наличие почти нулевых точек:** подсчитывается количество точек в *y*, абсолютное значение которых меньше заданного малого порога  $\varepsilon$  ( $|y_i| < \varepsilon$ ). Если количество таких точек превышает долю *n* от общей длины сигнала (т.е. число точек  $|y_i| < \varepsilon > n \times \text{length}(y)$ ), сигнал также считается «плохим».
- **Инициализация:** для хороших сигналов алгоритм продолжает процесс классификации.
  - Инициализируется пустой массив *result* для хранения значений MSE для каждого типа ошибки.
  - Создается массив *x*, представляющий собой равномерно распределенные значения от 0 до 1, размером с длину сигнала *y*. Это может служить в качестве нормированной «временной» оси для генерации ошибок.
  - Определяется набор значений размеров *size\_values*. В псевдокоде это размеры от  $\text{length}(y)$  с шагом 10 до  $2 \times \text{length}(y)$ . Эти размеры будут использоваться для изменения размера генерируемых сигналов ошибок.

- Определяется набор значений сдвигов  $shifts$  от  $-length(y)$  до  $length(y)$  с шагом 1. Эти сдвиги будут использоваться для смещения генерируемых сигналов ошибок относительно начала сигнала  $y$ .
  - **Цикл по типам ошибок (внешний цикл FOR):** алгоритм итерируется по каждому типу ошибки  $err$  из списка  $errors$ . Для каждого типа ошибки выполняется следующий процесс:
    - **Инициализация минимальной MSE для типа ошибки:** устанавливается начальное значение минимальной MSE  $mse\_err$  для текущего типа ошибки на большое число (например,  $10^5$ ). Это значение будет обновляться, если будут найдены меньшие значения MSE.
    - **Цикл по размерам (средний цикл FOR):** для каждого размера  $s$  из набора  $size\_values$ :
      - \* **Цикл по сдвигам (внутренний цикл FOR):** для каждого сдвига  $shift$  из набора  $shifts$ :
        - 1) **Генерация и изменение размера сигнала ошибки:** генерируется «базовый» сигнал ошибки с использованием функции  $Error\_signal(err, x, ampl, 1)$  для текущего типа ошибки  $err$ , «временной оси»  $x$ , амплитуды  $ampl$  и масштаба 1. Затем с помощью функции  $resize\_vector$  размер сгенерированного сигнала ошибки изменяется до длины  $s$ .
        - 2) **Применение сдвига к сигналу ошибки:** в зависимости от знака сдвига  $shift$ , сигнал ошибки  $Error$  сдвигается либо вперед (если  $shift \geq 0$ ), либо назад (если  $shift < 0$ ) относительно сигнала  $y$ . Сдвиг реализуется путем добавления нулей в начало или конец сигнала ошибки и обрезания результата до длины  $length(x)$ .
        - 3) **Построение матрицы признаков:** создается матрица признаков  $features\_matrix$  размером  $(length(x), 2)$ . Первый столбец матрицы заполняется единицами, а второй столбец – сдвинутым сигналом ошибки  $err\_error$ . Первый столбец с единицами позволяет учесть смещение в модели.
        - 4) **Метод наименьших квадратов (МНК):** решается задача линейной регрессии методом наименьших квадратов для нахождения коэффициентов  $coefficients$  ( $\beta$ ) минимизирующих норму разности  $\|features\_matrix \cdot \beta - y\|^2$ . Это позволяет найти оптимальную линейную комбинацию (смещение и масштабирование) сигнала ошибки для аппроксимации сигнала  $y$ .
        - 5) **Вычисление аппроксимированного сигнала:** вычисляется аппроксимированный сигнал  $approximated\_signal$  как произведение матрицы признаков  $features\_matrix$  на найденные коэффициенты  $coefficients$ .
        - 6) **Вычисление MSE:** вычисляется среднеквадратическая ошибка (MSE) между исходным сигналом  $y$  и аппроксимированным сигналом  $approximated\_signal$ .
        - 7) **Обновление минимальной MSE:** если вычисленное значение MSE меньше текущей минимальной MSE  $mse\_err$  для данного типа ошибки, то  $mse\_err$  обновляется этим новым, меньшим значением MSE.
  - **Сохранение минимальной MSE для типа ошибки:** после завершения циклов по размерам и сдвигам минимальное найденное значение MSE  $mse\_err$  для текущего типа ошибки  $err$  добавляется в массив  $result$ .
- Определение лучшего типа ошибки:** После завершения цикла по всем типам ошибок, алгоритм находит индекс минимального значения в массиве  $result$ . Этот индекс соответствует типу ошибки который обеспечил наименьшую MSE при аппроксимации сигнала  $y$ .

**Выходные данные:** алгоритм возвращает список, содержащий два элемента: название лучшего типа ошибки (из списка  $errors$ ) и соответствующее минимальное значение MSE. Это указывает на тип ошибки, который лучше всего «объясняет» структуру сигнала  $y$  с точки зрения MSE.

**Замечания:** алгоритм  $mse\_classification$  представляет собой метод классификации, основанный на сопоставлении сигнала с набором эталонных «ошибочных» сигналов. Использование MSE в качестве метрики качества аппроксимации позволяет количественно оценить, насколько хорошо каждый тип ошибки соответствует входному сигналу. Перебор различных размеров и сдвигов сигналов ошибок призван учесть возможные различия в масштабе и фазе между ожидаемыми ошибками и реальными отклонениями в сигнале. Метод наименьших квадратов используется для оптимальной «подгонки» масштаба и смещения сигнала ошибки к входному сигналу. Эффективность алгоритма зависит от репрезентативности набора  $errors$ , адекватности функции  $Error\_signal$  для моделирования реальных ошибок, и правильного выбора параметров, таких как  $size\_values$ ,  $shifts$ ,  $\epsilon$  и  $n$ . Алгоритм вычислительно интенсивен из-за вложенных циклов и многократного решения задачи МНК. Для решения проблемы вычислительной сложности вложенных циклов используются неградиентные методы оптимизации, такие как метод Нелдера–Мида, BFGS и TPESampler, которые находят локальные минимумы целевой функции ошибки по параметрам сдвига и растяжения «эталонных» сигналов, это позволяет заменить грубый перебор на более эффективные методы.

#### 1.4. Эксперименты и анализ результатов

##### 1.4.1. Численный эксперимент

Для оценки эффективности предлагаемого метода был проведён эксперимент, в котором анализировался сигнал, содержащий несколько различных сбоев на основе синтетических данных. Входные данные подавались на вход алгоритма, последовательно проходя через этапы локализации и классификации.

В качестве тестового сигнала для данного эксперимента был выбран сигнал, содержащий два типа аномалий: haar и algb (рис. 3). Эти типы сбоя характеризуются различными особенностями: haar представляет собой резкие скачки или провалы в данных, в то время как algb отражает схожий с видом некоторого полинома сбой.

Описание некоторых других видов аномалий:

- 'haar' – ошибки типа вейвлета Хаара,
- 'haar1' – одноступенчатая ошибка Хаара,
- 'exp' – экспоненциальная ошибка,
- '-exp' – отрицательная экспоненциальная ошибка,
- 'Invexp' – обратная экспоненциальная ошибка,
- 'Invexp+' – обратная экспоненциальная ошибка со сдвигом вверх,
- 'algb' – полиномиальная ошибка 3-й степени.

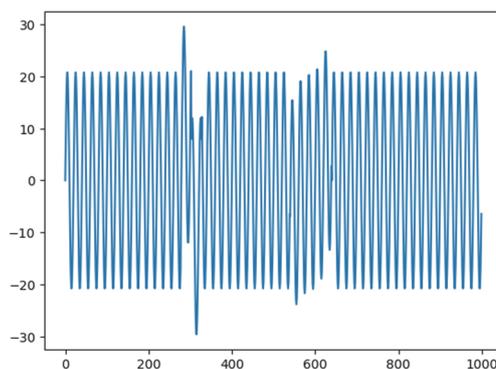


Рис. 3. Сигнал со сбоями

После предварительной обработки, включающей подавление шума и удаление синусоидальной составляющей, к сигналу были применены методы обнаружения аномалий. Целью данного этапа являлось разбиение сигнала на интервалы, с высокой вероятностью содержащие сбои (рис. 4).

Результаты работы алгоритмов представлены в табл. 1. Анализ таблицы показывает, что для всех интервалов, содержащих сбои типов haar и algb, алгоритмы продемонстрировали успешное распознавание. Количественная оценка качества, выраженная через среднеквадратичную ошибку (MSE), подтверждает этот вывод: для всех корректно идентифицированных сбоя haar и algb значение MSE не превышает 0,5.

Однако, помимо истинно положительных результатов, был зафиксирован один ложноположительный случай, связанный с алгоритмом Invexp. Интервал, ошибочно классифицированный как содержащий аномалию, на самом деле не имеет выраженных признаков сбоя.

Данную проблему можно решить несколькими путями:

- 1) Увеличение количества итераций: при оптимизации параметров алгоритма большее количество итераций работы методов оптимизации (BFGS, Nelder-Mead, TPESampler) может привести к более точной минимизации целевой функции и, как следствие, к снижению вероятности ложноположительных результатов.
- 2) Уменьшение порога допуска: в данный момент пороговое значение MSE для классификации интервала как аномального установлено на уровне 0,5. Снижение этого порога (например, до 0,4) повысит строгость критериев отбора и может уменьшить количество ложных срабатываний. Однако этот подход требует осторожности, так как слишком строгий порог может привести к пропуску реальных сбоя. Улучшение качества распознавания сигнала при увеличении количества итераций, описанное в пункте 1, позволит снизить порог без потери истинно положительных результатов.
- 3) Увеличение уровня детализации вейвлет-преобразования: Для вейвлет-анализа можно использовать более высокий уровень детализации (в рамках экспериментов использовался третий уровень детализации). Это позволит выявлять более тонкие особенности сигнала, что может способствовать более точному различению между нормальными участками и аномалиями.
- 4) Использование разных вейвлетов: разные вейвлеты обладают различной чувствительностью к различным типам аномалий. Экспериментирование с различными вейвлет-семействами (например, «Морле», «Мексиканская шляпа», «Добеши» и др.) может привести к улучшению качества распознавания для конкретного типа сигнала и сбоя.

Таким образом, несмотря на наличие ложноположительного срабатывания, результаты эксперимента в целом демонстрируют эффективность применения предложенного подхода для обнаружения сбоя типов haar и algb в зашумленном сигнале с синусоидальной составляющей.

#### 1.4.2. Тестирование производительности

Экспериментальная оценка производительности системы проводилась в условиях нарастающего числа сбоя в сигнале. На рис. 5 представлены графики зависимости времени выполнения от эпохи, где эпоха соответствует

одной итерации алгоритма, определяемой количеством сбоев в обрабатываемом сигнале. В экспериментах рассматривались пять эпох (от одного до пяти сбоев). Представленные результаты демонстрируют влияние количества сбоев на производительность алгоритма.

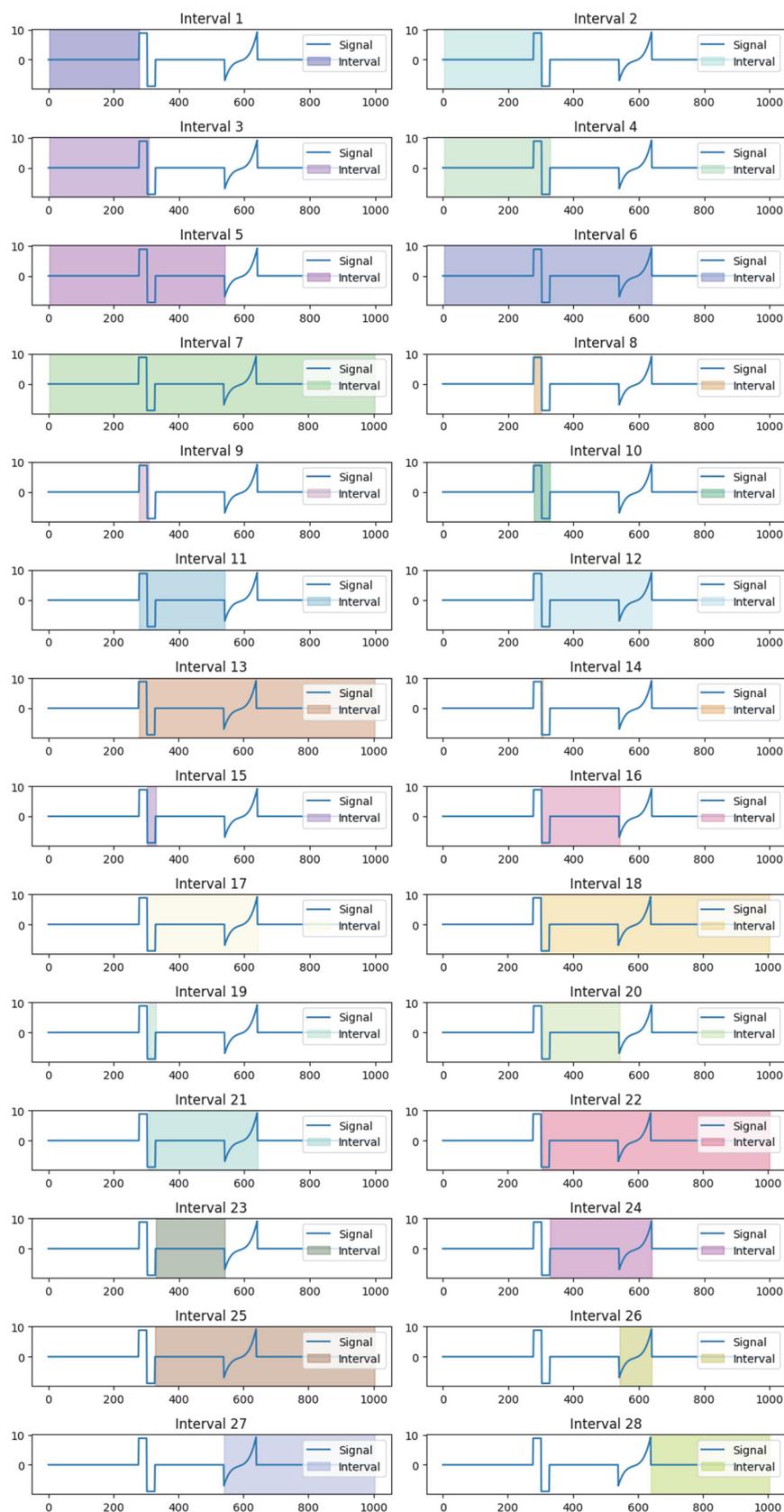


Рис. 4. Интервалы сигнала со сбоями

Табл. 1. Результаты

Interval	Fault Type	Error (MSE)
1	Bad signal	Nan
2	Invexp	0,739
3	Invexp	0,927
4	haar	0,448
5	haar	1,644
6	haar	1,662
7	haar	1,209
8	Bad signal	Nan
9	Bad signal	Nan
10	Bad signal	Nan
11	haar	0,383
12	haar	1,505
13	Invexp	3,785
14	Bad signal	Nan
15	Bad signal	Nan
16	Invexp	0,696
17	algb	2,142
18	Invexp	1,922
19	Bad signal	Nan
20	Invexp	0,481
21	-exp	1,902
22	Invexp	1,696
23	Bad signal	Nan
24	algb	0,209
25	algb	3,187e-13
26	algb	0,251
27	algb	2,174e-13
28	Bad signal	Nan

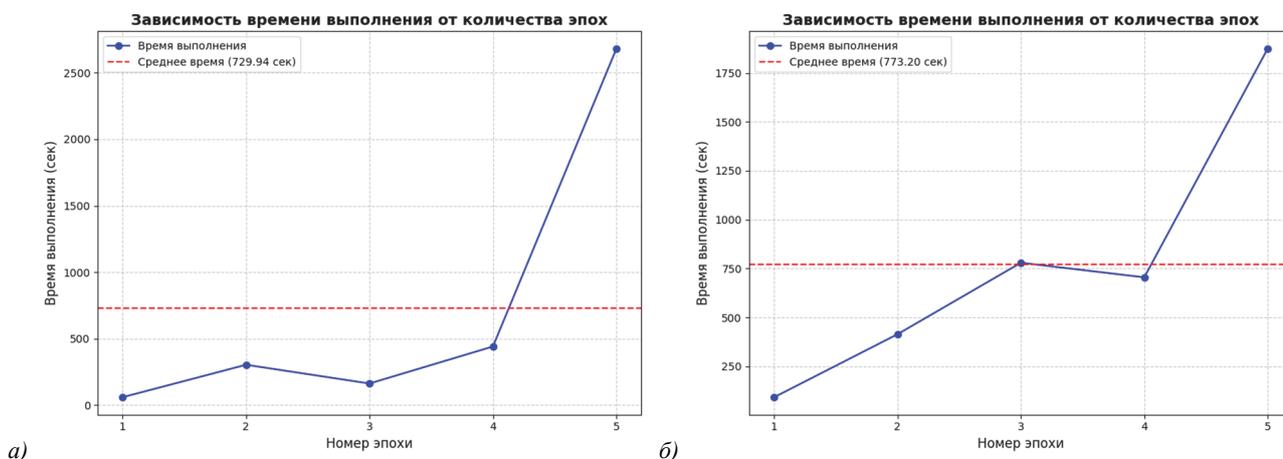


Рис. 5. Тест производительности: а) №1, б) №2

Результаты экспериментов демонстрируют тенденцию к увеличению времени выполнения алгоритма с ростом количества сбоев в сигнале. Это обусловлено особенностями вейвлет-анализа: большее число сбоев приводит к появлению большего количества значимых вейвлет-коэффициентов, что, в свою очередь, увеличивает число интервалов, требующих обработки алгоритмом. Однако следует отметить, что в отдельных случаях наблюдается отклонение от этой общей тенденции. Так, для некоторых сигналов с большим количеством сбоев время выполнения оказывалось меньше, чем для сигналов с меньшим числом сбоев. Это связано с тем, что вейвлет-преобразование не всегда приводит к равномерному увеличению количества интервалов при увеличении числа сбоев; в некоторых случаях структура вейвлет-коэффициентов может быть такова, что общее количество операций, необходимых для обработки, уменьшается. Таким образом, влияние сбоев на время выполнения алгоритма является нелинейным и зависит от конкретной структуры сигнала и его вейвлет-представления.

**Заключение**

В рукописи представлены метод, алгоритмы, пример многоуровневой архитектуры информационной системы и результаты тестирования системы анализа сигналов на основе разработанного авторами алгоритма вейвлет-анализа, дополненного численными методами. Несмотря на достигнутые результаты, дальнейшие исследования могут быть

направлены на повышение вычислительной эффективности, расширение библиотеки модельных ошибок, исследование робастности к параметрам и изучение альтернативных метрик качества классификации. Ключевыми преимуществами предложенной системы алгоритмов являются использование методов оптимизации и параллельных вычислений. Система демонстрирует потенциал для применения в различных областях обработки сигналов, где требуется автоматическое обнаружение и классификация аномалий или дефектов, таких как диагностика оборудования высокопроизводительных вычислительных кластеров или мониторинг сенсорных данных, являясь мощным и универсальным инструментом для анализа и интерпретации сложных сигнальных данных.

### *Благодарности*

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации в рамках проекта № 075-15-2024-525 от 23.04.2024.

### *References*

- [1] Bykov A, Grecheneva A, Kuzichkin O, Surzhik D, Vasilyev G. Mathematical description and laboratory study of electrophysical methods of localization of geodeformational changes during the control of the railway roadbed. *Mathematics*; 2021, 9 (24): 3164. DOI 10.3390/math9243164.
- [2] Daurenbayeva N, Atymtayeva L, Nurlanuly A, Bykov A, Akhmetov B, Shuitenov G, Umut A. A Machine Learning Approach to Microclimate Monitoring and Fault Detection. *Appl Math*; 2025, 19 (2): 327-334. DOI 10.18576/amis/190209.
- [3] Bykov AA, Surzhik DI, Kuzichkin OR. Assessing the Ground Structure in Drilling by Analysis of the Phase Shift in the Probing Electrical Signal. *Russian Engineering Research*; 2024, 44 (11): 1691-1696. DOI 10.3103/S1068798X24702800.
- [4] Abid A, Khan MT, Iqbal J. A review on fault detection and diagnosis techniques: basics and beyond. *Artificial Intelligence Review*; 2021, 54 (5): 3639-3664. DOI 10.1007/s10462-020-09934-2.
- [5] Hong YY, Pula RA. Methods of photovoltaic fault detection and classification: A review. *Energy Reports*; 2022, 8: 5898-5929. DOI 10.1016/j.egyr.2022.04.043.
- [6] Jua Y, Tian X, Liu H, Ma L. Fault detection of networked dynamical systems: A survey of trends and techniques. *International Journal of Systems Science*; 2021, 52 (16): 3390-3409. DOI 10.1080/00207721.2021.1998722.
- [7] Zhang L, Matt L, Yeonjin B, Borui C, Saptarshi B, Seungjae L, Piljae I, Veronica A, Draguna V, Teja K. Sensor impact evaluation and verification for fault detection and diagnostics in building energy systems: A review. *Advances in Applied Energy*; 2021, 3: 100055. DOI 10.1016/j.adapen.2021.100055.
- [8] Li G, Yue Z, Jiangyan L, Zhenxin Z, Chengliang X, Xi F, Qing Y. An improved stacking ensemble learning-based sensor fault detection method for building energy systems using fault-discrimination information. *Journal of Building Engineering*; 2021, 43: 102812. DOI 10.1016/j.jobe.2021.102812.
- [9] Wang Z, Wang L, Tan Y, Yuan J. Fault detection based on Bayesian network and missing data imputation for building energy systems. *Applied Thermal Engineering*; 2021, 182: 116051. DOI 10.1016/j.applthermaleng.2020.116051.
- [10] Aviña-Corral V, de Jesus Rangel-Magdaleno J, Peregrina-Barreto H, Ramirez-Cortes JM. Bearing fault detection in asd-powered induction machine using modwt and image edge detection. *IEEE Access*; 2022, 10: 24181-24193. DOI 10.1109/ACCESS.2022.3154410.
- [11] Gupta V. Wavelet transform and vector machines as emerging tools for computational medicine. *Journal of Ambient Intelligence and Humanized Computing*; 2023, 14 (4): 4595-4605. DOI 10.1007/s12652-023-04582-0.
- [12] Wu W et al. Wavelet-improved score-based generative model for medical imaging. *IEEE transactions on medical imaging*; 2023. DOI 10.1109/TMI.2023.3325824.
- [13] Ahmad MA, Elloumi M, Samak AH, Al-Sharafi AM, Alqazzaz A, Kaid MA, Iliopoulos C. Hiding patients' medical reports using an enhanced wavelet steganography algorithm in DICOM images. *Alexandria Engineering Journal*; 2022, 61 (12): 10577-10592. DOI 10.1016/j.aej.2022.03.056.
- [14] Karpov OE, Afinogenov S, Grubov VV, Maksimenko V, Korchagin S, Utyashev N, Hramov AE. Detecting epileptic seizures using machine learning and interpretable features of human EEG. *The European Physical Journal Special Topics*; 2023, 232 (5): 673-682. DOI 10.1140/epjs/s11734-022-00714-3.
- [15] Guo T, Zhang T, Lim EG, López-Benítez M, Ma F, Yu L. A review of wavelet analysis and its applications: Challenges and opportunities. *IEEE Access*; 2022, 10: 58869-58903. DOI 10.1109/ACCESS.2022.3179517.
- [16] Yan Z, Chao P, Ma J, Cheng D, Liu C. Discrete convolution wavelet transform of signal and its application on BEV accident data analysis. *Mechanical systems and signal processing*; 2021, 159: 107823. DOI 10.1016/j.ymsp.2021.107823.
- [17] Santhosh Kumar HS, Karibasappa K. An approach for brain tumour detection based on dual-tree complex Gabor wavelet transform and neural network using Hadoop big data analysis. *Multimedia Tools and Applications*; 2022, 81 (27): 39251-39274. DOI 10.1007/s11042-022-13016-6.
- [18] Zhao N, Wei J, Long Z, Yang C, Bi J, Wan Z, Dong S. An Integrated Method for Tunnel Health Monitoring Data Analysis and Early Warning: Savitzky-Golay Smoothing and Wavelet Transform Denoising Processing. *Sensors*; 2023, 23 (17): 7460. DOI 10.3390/s23177460.
- [19] Petushkov GV, Sigov AS. Analysis and selection of the structure of a multiprocessor computing system according to the performance criterion. *Russian Technological Journal*; 2024, 12 (6): 20-25. DOI 10.32362/2500-316X-2024-12-6-20-25.
- [20] Andriyanov N, Khasanshin I, Utkin D, Gataullin T, Ignar S, Shumaev V, Soloviev V. Intelligent system for estimation of the spatial position of apples based on YOLOv3 and real sense depth camera D415. *Symmetry*; 2022, 14 (1): 148. DOI 10.3390/sym14010148.

- [21] Osipov A, Pleshakova E, Bykov A, Kuzichkin O, Surzhik D, Suvorov S, Gataullin S. Machine learning methods based on geophysical monitoring data in low time delay mode for drilling optimization. *IEEE Access*; 2023, 11: 60349-60364. DOI 10.1109/ACCESS.2023.3284030.
- [22] Osipov AV, Pleshakova ES, Gataullin ST. Production processes optimization through machine learning methods based on geophysical monitoring data. *Computer Optics*; 2024, 48 (4): 633-642. DOI 10.18287/2412-6179-CO-1373.
- [23] Osipov A, Pleshakova E, Liu Y, Gataullin S. Machine learning methods for speech emotion recognition on telecommunication systems. *Journal of Computer Virology and Hacking Techniques*; 2024, 20 (3): 415-428. DOI 10.1007/s11416-023-00500-2.
- [24] <https://github.com/NekkittAY/WONC-FD-Method>; 2025.
- [25] Beylkin G, Coifman R, Rokhlin V. Fast wavelet transforms and numerical algorithms I. *Communications on pure and applied mathematics*; 1991, 44 (2): 141-183. DOI 10.1002/cpa.3160440202.
- [26] Nelder JA, Mead R. A simplex method for function minimization. *The computer journal*; 1965, 7 (4): 308-313. DOI 10.1093/comjnl/8.1.27.
- [27] Nocedal J, Wright SJ. *Numerical optimization*. Springer New York; 1999.
- [28] Bergstra J, Bardenet R, Bengio Y, Kégl B. Algorithms for hyper-parameter optimization. *Advances in neural information processing systems*; 2011, 24.
- [29] Björck A. *Numerical methods for least squares problems*. SIAM; 2024.
- [30] Ahmed N, Hashmani AA, Khokhar S, Tunio MA, Faheem M. Fault detection through discrete wavelet transform in overhead power transmission lines. *Energy Science Engineering*; 2023, 11: 4181-4197. DOI 10.1002/ese3.1573.
- [31] Hadad K, Pourahmadi M, Majidi-Maraghi H. Fault diagnosis and classification based on wavelet transform and neural network. *Progress in Nuclear Energy*; 2011, 53 (1): 41-47. DOI 10.1016/j.pnucene.2010.09.006.

### *Сведения об авторах*

**Сакович Никита Данилович**, специалист 2-й категории в лаборатории прикладного применения искусственного интеллекта НИИ – филиал АО «Гознак», обучающийся Финансового университета. Области научных интересов: алгоритмы, высокопроизводительные вычисления, вычислительная математика, информатика, логика, математические методы оптимизации, математическое моделирование, машинное обучение, прикладная математика, разработка программного обеспечения, численное моделирование, численные методы оптимизации. Профессиональные навыки и компетенции: C/C++, Cuda, Docker, Git, Google JAX, Google-colab, Julia, LaTeX, LightGBM, Linux, MS SQL Server, MS Visio, Matlab, Matplotlib, MongoDB, MySQL, NumPy, Numerical simulation, Octave, OriginLab, Pandas, PostgreSQL, Python, Redis, SQL, SQLite, SciLab, SciPy, Scikit-fem, Scikit-learn, TensorFlow, UML, XGBoost. E-mail: [217556@edu.fa.ru](mailto:217556@edu.fa.ru)

**Аксенов Дмитрий Андреевич**, заместитель заведующего лаборатории прикладного применения искусственного интеллекта НИИ – филиал АО «Гознак», руководитель Лаборатории цифровой трансформации бизнес-процессов в промышленности Финансового университета. Области научных интересов: физика элементарных частиц, практическая ядерная физика, физика высоких энергий, алгоритмы, высокопроизводительные вычисления, вычислительная математика, информатика, логика, математические методы оптимизации, математическое моделирование, машинное обучение, прикладная математика, разработка программного обеспечения, численное моделирование, численные методы оптимизации. E-mail: [Daaksenov@fa.ru](mailto:Daaksenov@fa.ru)

**Плешакова Екатерина Сергеевна**, специалист по технической защите информации, доцент кафедры индустриального программирования РТУ МИРЭА. Получены значимые научные результаты в технологиях искусственного интеллекта для обеспечения кибербезопасности, интеллектуальной отраслевой робототехнике, нейросетевых методах противодействия социальной инженерии. E-mail: [pleshakova@mirea.ru](mailto:pleshakova@mirea.ru)

**Гатауллин Сергей Тимурович**, специалист по технической защите информации, ведущий научный сотрудник Лаборатории социального моделирования ЦЭМИ РАН. Получены значимые результаты в математической экономике, компьютерных и инженерных науках, технологиях искусственного интеллекта для обеспечения кибербезопасности, интеллектуальной отраслевой робототехнике, нейросетевых методах противодействия социальной инженерии. E-mail: [sgataullin@cemi-ras.ru](mailto:sgataullin@cemi-ras.ru)

ГРНТИ: 27.41.17

Поступила в редакцию 13 февраля 2025 г. Окончательный вариант – 11 сентября 2025 г.

---

# Localization and classification of anomalies in one-dimensional signals based on wavelet analysis and mathematical optimization methods

*N.D. Sakovich<sup>1,2</sup>, D.A. Aksenov<sup>1,2</sup>, E.S. Pleshakova<sup>3</sup>, S.T. Gataullin<sup>4</sup>*

<sup>1</sup> *Financial University under the Government of the Russian Federation, 125167, Moscow, Leningradsky Prospekt, 49/2;*

<sup>2</sup> *The Scientific Research Institute of Goznak, 115162, Moscow, Mytnaya st., 19;*

<sup>3</sup> *MIREA – Russian Technological University, 119454, Moscow, Vernadsky Ave., 78;*

<sup>4</sup> *CEMI RAS, 117418, Moscow, Nakhimovsky pr., 47*

## **Abstract**

A method for classifying types of non-stationarities in time series based on wavelet analysis for localization and detection of non-stationarity boundaries; mathematical algorithms for comparing types of non-stationarities based on templates of reference signals and classifying anomalies using non-gradient optimization methods; with the possibility of further application for developing automated information systems for comprehensive monitoring of the state of the power grid and its individual components is presented in this study. The code base is located in the open project repository and available to reproduce computational experiments. The suggested approach has a wide range of applications, but special attention should be paid to the possible integration in order to improve the reliability of diagnostic systems with Internet of Things (IoT) technologies, the digital twins construction, agent-based modeling (ABM) for complex socio-economic processes, and cloud high-performance computing (HPC) that allows real time big data analytics.

**Keywords:** digital signal processing, wavelet analysis, numerical methods, algorithms.

**Citation:** Sakovich ND, Aksenov DA, Pleshakova ES, Gataullin ST. Localization and classification of anomalies in one-dimensional signals based on wavelet analysis and mathematical optimization methods. *Computer Optics* 2026; 50(1): 1683. DOI: 10.18287/COJ1683.

**Acknowledgements:** This research was funded by the Ministry of Science and Higher Education of the Russian Federation grant no. 075-15-2024-525.

---

## **Author's information**

**Nikita Sakovich**, Nikita Sakovich, 2nd Category Specialist in the Laboratory of AI Applications, Research Institute - branch of JSC «Goznak», student at the Financial University. Research interests: Algorithms, High-performance computing, Computational mathematics, Informatics, Logic, Mathematical methods of optimization, Mathematical modeling, Machine learning, Applied mathematics, Software development, Numerical modeling, Numerical methods of optimization. Professional skills and competencies: C/C++, Cuda, Docker, Git, Google JAX, Google-colab, Julia, LaTeX, LightGBM, Linux, MS SQL Server, MS Visio, Matlab, Matplotlib, MongoDB, MySQL, NumPy, Numerical simulation, Octave, OriginLab, Pandas, PostgreSQL, Python, Redis, SQL, SQLite, SciLab, SciPy, Scikit-fem, Scikit-learn, TensorFlow, UML, XGBoost. E-mail: [217556@edu.fa.ru](mailto:217556@edu.fa.ru)

**Dmitry Aksenov**, Deputy Head of the Laboratory of Artificial Intelligence Applications, Research Institute - branch of JSC «Goznak», Head of the Laboratory of Digital Transformation of Business Processes in Industry, Financial University. Research interests: Particle Physics, Practical Nuclear Physics, High Energy Physics, Algorithms, High Performance Computing, Computational Mathematics, Computer Science, Logic, Mathematical Methods of Optimization, Mathematical Modeling, Machine Learning, Applied Mathematics, Software Development, Numerical Modeling, Numerical Methods of Optimization. E-mail: [Daaksenov@fa.ru](mailto:Daaksenov@fa.ru)

**Ekaterina Pleshakova**, PhD in Technical Sciences, Information Security Specialist, Associate professor at the Industrial Programming Department, RTU MIREA. Significant scientific results have been obtained in AI technologies for ensuring cybersecurity, intelligent industrial robotics, neural network methods for countering social engineering. E-mail: [pleshakova@mirea.ru](mailto:pleshakova@mirea.ru)

**Sergey Gataullin**, Graduated from the State University of Management (RF); PhD in Economics (Mathematical and Instrumental Methods), Information Security Specialist, Lead Research Scientist at the Social Modeling Lab, CEMI RAS. Significant results have been obtained in mathematical economics, computer and engineering sciences, AI technologies for ensuring cybersecurity, intelligent industry robotics. E-mail: [sgataullin@cemi-ras.ru](mailto:sgataullin@cemi-ras.ru)

---

*Received February 13, 2025. The final version – September 11, 2025.*

---