УДК 004.896

Д.А. Мурзагулов, А.В. Замятин

ОБНАРУЖЕНИЕ АНОМАЛИЙ ТЕХНОЛОГИЧЕСКИХ СИГНАЛОВ С ИСПОЛЬЗОВАНИЕМ АНСАМБЛЯ КЛАССИФИКАТОРОВ И ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ¹

Мурзагулов Дамир Альбертович, окончил Томский политехнический университет (ТПУ), главный аналитик Improvado.io, аспирант кафедры «Теоретические основы информатики» Томского государственного университета (ТГУ). Имеет статьи по обнаружению аномалий в технологических данных. [e-mail: murzagulov.damir@gmail.com].

Замятин Александр Владимирович, доктор технических наук, доцент, окончил ТПУ, заведующий кафедрой «Теоретических основы информатики» ТГУ. Имеет статьи и учебные пособия в области интеллектуального анализа данных. [e-mail: zamyatin@mail.tsu.ru].

Аннотация

Уровень развития современной ИТ-инфраструктуры промышленных предприятий позволяет осуществлять сбор и хранение технологической информации, тем самым открывая возможности применения интеллектуальных систем анализа данных. В работе рассматривается задача обнаружения аномалий в технологических сигналах в целях повышения качества мониторинга объектов управления. Для обнаружения аномалий предлагается ансамбль из базовых классификаторов на основе алгоритмов машинного обучения и вейвлет-преобразований. Рассмотрены специфика технологических сигналов и преимущества вейвлет-анализа для предварительной обработки сигналов. В работе разработан подход к обнаружению аномалий на основе ансамбля моделей и проведена его предварительная апробация на реальных технологических сигналах.

Ключевые слова: технологические сигналы, машинное обучение, вейвлет-преобразование, обнаружение аномалий, ансамбль.

doi: 10.35752/1991-2927-2021-1-63-20-26

THE PROCESS SIGNAL ANOMALY DETECTION USING CLASSIFIER ENSEMBLE AND WAVELET TRANSFORMS

Damir Albertovich Murzagulov, graduated from Tomsk Polytechnic University (TPU); Improvado.io Chief Analyst; Postgraduate Student of the Department of Fundamental Informatics and Information Technologies of Tomsk State University; an author of articles in the field of the process data anomaly detection. e-mail: murzagulov.damir@gmail.com.

Aleksandr Vladimirovich Zamyatin, Doctor of Sciences in Engineering, Associate Professor; graduated from Tomsk Polytechnic University (TPU); Head of the Department of Fundamental Informatics and Information Technologies of Tomsk State University; an author of articles and guides in the field of data mining. e-mail: zamyatin@mail.tsu.ru.

Abstract

The IT-infrastructure development level of manufacturing plants allows to collect and storage technological information, thereby creating the possibilities to adapt data-mining systems. The article deals with the problem of anomaly detection in process signals with a view to improving the quality of control object monitoring. The ensemble of base classifiers based on algorithms of machine learning and wavelet transforms is proposed to detect anomalies. Authors examine the process signal characteristics and wavelet analysis advantages for signal preprocessing. An approach to the anomaly detection was developed based on a model ensemble. This approach was previously tested on actual process signals.

Keywords: process signals, machine learning, wavelet transform, anomaly detection, ensemble.

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-37-90124 и при финансовой поддержке НИ ТГУ в рамках научного проекта № 8.1.62.2018.

ВВЕДЕНИЕ

Современная промышленная индустрия характеризуется активным переходом к этапу цифровой трансформации – помимо классических автоматизированных систем управления технологическими процессами (АСУ ТП), внедряются системы интернета вещей (IoT), больших данных (Big Data), машинного обучения [1–5]. Внедрение подобных автоматизированных систем поддержки принятия решения в основном направлено на две группы задач. К первой относятся вопросы повышения безопасности производства, включающие анализ и предупреждение аварий. Ко второй группе следует отнести экономические эффекты от оптимизации режимов потребления ресурсов, а также прогнозирование возможных поломок. Для решения обеих групп задач широко применяются методы интеллектуального анализа к массивам производственных данных [6]. При этом особое место в целях обеспечения надежности производства занимает задача обнаружения аномалий в данных, так как именно они зачастую являются предвестниками аварий, сбоев оборудования и нештатных ситуаций. Применительно к системам мониторинга техногенных объектов под аномалиями могут восприниматься любые события, нарушающие регламентное протекание контролируемого технологического процесса [7].

Основным источником производственных данных являются технологические сигналы, поступающие в АСУ ТП с различных измерительных датчиков (датчиков температуры, давления, вибрации и т. п.) [8]. Как правило, технологические сигналы имеют сложную структуру, они включают большое количество компонент, содержат локальные особенности различной формы и временной протяженности, а также шумовые факторы различной природы. Именно с этими факторами связана основная сложность автоматического обнаружения аномалий в технологических сигналах традиционными статистическими методами.

Широкое применение в задачах обнаружения аномалий нашли методы машинного обучения [9]. К достоинствам методов машинного обучения относятся их простая параметризуемость, возможность задавать различные метрики и количество степеней свободы модели. К алгоритмам машинного обучения, применяющимся для обнаружения аномалий, относятся алгоритм *k-means* [10], изолирующий лес [11], метод опорных векторов [12], фактор локальных выбросов [13] и др. Однако недостаточно только выбрать конкретный алгоритм машинного обучения, так как для эффективного выявления аномалий также необходимо построение информативного признакового пространства, в котором будет возможно разделение аномальных участков от участков со штатным протеканием процесса. Сложность построения признакового пространства для технологических сигналов заключается в том, что аномальные эффекты могут характеризоваться не только явными амплитудными изменениями, но и изменением ча-

стотно-временных свойств сигнала. Наибольшую сложность для частотно-временного (спектрального) анализа представляют нестационарные во времени сигналы, которые наиболее часто встречаются на практике. Как правило, такие сигналы состоят из близких по времени, коротких высокочастотных компонент и долговременных, близких по частоте низкочастотных компонент. Для анализа таких сигналов необходим метод, способный обеспечить хорошее разрешение и по частоте, и по времени. Первое необходимо для локализации низкочастотных составляющих, второе – для выделения компонент высокой частоты. Одним из эффективных методов спектрального представления сигналов является вейвлет-преобразование [14]. Основная область применения вейвлет-преобразований – анализ и обработка сигналов, нестационарных во времени или неоднородных в пространстве сигналов. На сегодняшний день применение вейвлет-анализа нашло широкое прикладное применение в распознавании человеческой речи, в задачах классификации биомедицинских данных, для фильтрации и сжатия данных [15–19].

Применение отдельных методов машинного обучения и вейвлет-преобразований дают в конкретных случаях удовлетворительные результаты в зависимости от типа аномалии и свойств самого сигнала [20]. Однако более высокий и стабильный результат обнаружения можно получить за счет создания ансамбля нескольких моделей обнаружения. В данной работе предлагается использовать ансамбль из методов машинного обучения для получения более стабильного и универсального обнаружения аномалий. Научная новизна данной работы заключается в комбинационном применении алгоритмов машинного обучения и вейвлет-анализа.

ПРЕДЛАГАЕМЫЙ ПОДХОД

Для построения ансамблей существуют ряд классических подходов:

а) «бэггинг» – последовательное обучение нескольких моделей на подмножестве обучающих данных и усреднение ответов по некоторому решающему правилу;

б) «стекинг» – параллельное обучение нескольких моделей и создание дополнительного метаалгоритма для обучения итогового решающего правила;

в) «бустинг» – последовательное обучение однородных моделей [21].

В данной работе предлагается использование «бэггинга» базовых классификаторов и функции моды для получения итоговой прогноза для сигнала. Допустим, имеется ряд классификаторов $b_1, b_2, \dots, b_n$, итоговым прогнозом будет являться значение во множестве прогнозов, которое встречается наиболее часто

$$B(X) = \text{mode}(b_1(X), b_2(X), \dots, b_n(X)),$$

где X – технологический сигнал,

mode – значение, которое встречается наиболее часто.

Заметим также, что в отличие от классического бэггинга, где предполагается что обучение базовых классификаторов производится на подмножестве данных, в данном случае предлагается каждый классификатор обучать на полном множестве обучающих данных. Выбор такого подхода определен следующим фактором: в ансамбле предполагается использовать разнородные классификаторы без учителя, за счет этого достигается независимость классификаторов, так как они не обучаются прогнозировать одну и ту же целевую переменную. Также при условии, что базовые алгоритмы имеют вероятность прогнозирования правильного ответа выше 0,5, то по неравенству Хёффдинга [22] ошибка ансамбля монотонно убывает с ростом количества базовых алгоритмов. В качестве метаалгоритма – решающего правила, по которому ансамбль формирует итоговый прогноз, выбрана мода Коши [23], которая часто используется в задачах бинарной классификации. Преимущество моды заключается в том, что она выбирает наиболее «типичный» прогноз, тем самым снижая вероятность ошибок второго рода, которые являются критичными для задачи обнаружения аномалий.

Обучение базовых классификаторов предлагается на основе амплитудных показаний сигнала и его вейвлет-преобразований совместно. Данный подход обоснован двумя факторами. Во-первых, вейвлет-преобразование предоставляет хорошее разрешение в частотно-временной области и способно локализовать аномалии в этой области. Во-вторых, исходные показания характеризуют аномалии, связанные с амплитудными изменениями. Другими словами, исходный анализируемый сигнал, представленный в виде дискретного ряда $S = \{s_1, s_2, \dots, s_t\}$, где s_t – показания технологического параметра, определенные в моменты времени t , описываются признаковым множеством:

$$X = \{S, \zeta(S)\}, \quad (1)$$

где $\zeta(S)$ – вейвлет-преобразование сигнала S .

Цель работы состоит в построении модели $B(X)$, которая разбивает множество X на подмножества X^m (кластеры) и для каждого кластера ставит в соответствие метку $[0,1]$, где значение 0 соответствует «нормальным», а 1 – аномальным значениям.

ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ В ЗАДАЧАХ АНАЛИЗА СИГНАЛОВ

Технология вейвлет-преобразования является обобщением спектрального анализа и представляет собой альтернативу широко используемому классическому преобразованию Фурье (ПФ) [24]. ПФ представляет сигнал, заданный во временной области, в виде разложения по ортогональным базисным функциям (синусоидам) с выделением частотных компонентов. Недостаток ПФ заключается в том, что частотные компоненты не могут быть локализованы во времени, обуславливая его применимость только к анализу стационарных сигналов, в то время как многие реальные сигналы

имеют сложные частотно-временные характеристики. Основное отличие вейвлет-преобразования от ПФ заключается в том, что для преобразования исходного сигнала используются в качестве базисных функций вейвлеты (от англ. «wavlet» – короткая волна). Вейвлет – это обобщенное название семейства математических функций определенной формы, которые локализованы по времени и частоте и в которых все функции получаются из одной базовой (порождающей) функции путем ее сдвигов и растяжения по оси времени. Существует два вида реализации вейвлет-преобразования – дискретное вейвлет-преобразование (ДВП) и непрерывное вейвлет-преобразование (НВП). В задачах фильтрации и сжатия данных преимущественно применяется ДВП, для проведения анализа сигналов интерес представляет НВП.

НВП определяется как сумма по всему времени сигнала, умноженного на масштабируемые, сдвинутые версии вейвлет-функции:

$$CWT_{\psi}(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{t-b}{a}\right) s(t) dt, \quad (2)$$

где $\psi(t)$ – вейвлет-функция, $s(t)$ – сигнал, параметр $b \in R$ отвечает за расположение во времени, а $a \in R$ – за масштаб. Большие значения параметра a соответствуют низким частотам, меньшие – высоким. Посредством уменьшения (увеличения) параметра a вейвлет-преобразование позволяет фокусироваться на локальных (глобальных) структурах сигнала, выявлять особенности и сегментировать сигнал по интенсивности.

В результате преобразования (2), вследствие зависимости вейвлетов от двух параметров и локализации вейвлет-функции, получается 2-мерный образ анализируемого сигнала в координатах плоскости (a, b) . На рисунке 1 изображен пример технологического сигнала и его НВП-спектрограмма. По оси абсцисс отложено время t , по оси ординат – частота F (является обратной a), в качестве базового вейвлета применён вейвлет типа «мексиканская шляпа». Градацией цвета показаны значения вейвлет-коэффициентов CWT_{ψ} . С ростом масштаба базисная вейвлет-функция увеличивается, захватывая все больший диапазон значений t , коэффициенты преобразования в соответствующей точке будут зависеть от значений части составляющих $s(t)$. Диапазон этот тем больше, чем больше масштаб, поэтому высокочастотная информация вычисляется на основе малых отрезков значений функции $s(t)$, а низкочастотная информация – на основе больших. Таким образом, локальные изменения сигнала, которые могут являться аномалиями, представляют собой высокочастотную информацию и вычисляются на малых масштабах преобразования.

С учётом соотношения (2) дополним определение признакового описания сигнала (1). Для сигнала $s(t)$ с

Рис. 1. Пример НВП-технологического сигнала: а) пример исходного технологического сигнала, б) вейвлет-спектрограмма технологического сигнала

N (количеством отсчётов) признаковое описание сигнала, принимает следующий вид:

$$X = \left\{ (s_1, CWT_{1,1}), \dots, (s_i, CWT_{j,i}) \right\}_{i=1}^N, \quad (3)$$

где $CWT_{j,i}$ – коэффициенты НВП, индекс j – масштаб разложения, i – номер отсчёта сигнала, s_i – показания сигнала при i -м отсчёте.

ОПИСАНИЕ ЧИСЛЕННОГО ЭКСПЕРИМЕНТА

В работе предлагается подход обнаружения аномалий технологических сигналов посредством базовых классификаторов в пространстве признаков (3). В качестве базовых классификаторов использовались следующие алгоритмы, широко применяемые в задачах обнаружения аномалий [25] (см. табл. 1).

Базовые классификаторы использовались с настройками по умолчанию. Для алгоритма GRB использовался статистический тест, применяемый для определения выбросов в одномерном наборе данных, подчиняющихся нормальному закону распределения. Критерий вычисляется следующим образом:

$$G = \frac{\max_{i=1, \dots, N} |X_i - \bar{X}|}{\sigma},$$

где σ – среднее квадратическое отклонение величины. При этом прогноз об аномальности принимается с введением некоторого эмпирического порога T_G , выше которого значение считается аномальным.

Для проведения численного эксперимента для каждого сигнала предварительно вычисляются ко-

эффициенты НВП в диапазоне масштабов $a = [0 : 10]$ согласно (3). Для каждого уровня разложения a производилось обучение классификаторов и тестирование данных. Апробация предложенного подхода осуществлялась на данных, описанных в работе [26]. Данные представляют собой реальные технологические сигналы, аугментированные различными типами аномалий. Реальные сигналы были получены с преобразователей давления, установленных на нефтепроводе, а аномалии были получены с помощью моделирования типовых отказов, встречающихся на практике. В наборе данных применены следующие типы аномалий:

- «изменение тренда» – наиболее простой и очевидный тип аномалии, характеризующий изменение амплитуды технологического сигнала;

Таблица 1

Алгоритмы, применяемые в задачах обнаружения аномалий

Обозначение алгоритма	Название алгоритма	Тип алгоритма
HBOS	Гистограммный определитель выброса	Линейный
KNN	К-ближайший соседей	Основан на оценке плотности данных
LOF	Локальный фактор выброса	Основан на оценке плотности данных
MCD	Определитель минимальной ковариации	Линейный
Isolation Forest	Изолирующий лес	Деревья решений
OCSVM	Одноклассовый классификатор опорных векторов	Линейный
GRB	Тест по критерию Граббса	Статистический

- «резкий скачок» – резкие изменения технологических параметров, как правило носят негативный характер;

- «шум» – такой тип аномалии может наблюдаться в механических узлах, где дефекты проявляются наличием повышенной вибрации и биений;

- «повышение дисперсии» – простыми словами, увеличение разброса значений сигнала, как правило, свидетельствует о наличии переходных процессов или нестабильной работы оборудования;

- «уменьшение дисперсии» – можно описать как вырождение колебательного процесса в гладкую прямую, такой тип аномалий можно наблюдать при отказах измерительных средств («замирание прибора»).

Использование синтезированных данных обусловлено двумя факторами:

а) возможность параметризации аномалий и получение более широкого и общего результата;

б) возможность достоверно измерить качество алгоритма на тестовых данных и определить область применения подхода.

Всего выборка состоит из 9 сигналов с ежеминутной дискретизацией продолжительностью 7 дней. Набор данных разбит на обучающую и тестовую выборки. Обучающая выборка состоит из реальных технологических сигналов в штатном режиме. На тестовую выборку случайным образом накладывались вышеуказанные аномалии.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В таблице 2 представлены результаты обнаружения аномалий по различным метрикам качества для каждого базового классификатора и для ансамбля классификаторов при масштабе разложения $j = 1$, при котором достигнуты максимальные результаты. Для алгоритма GRB определено субоптимальное значение порога $T_G = 4.8$ для данного набора данных.

Таблица 2
Результаты обнаружения аномалий

Алгоритм	F1 score	Precision	Accuracy
Ансамбль	0.78	0.98	0.84
HBOS	0.51	0.89	0.74
KNN	0.42	0.92	0.68
LOF	0.42	0.77	0.63
MCD	0.54	0.94	0.77
Isolation Forest	0.60	0.96	0.78
OCSVM	0.54	0.88	0.69
GRB	0.64	0.95	0.76

Рис. 2. Зависимость F-меры от уровня разложения вейвлет-преобразований

Как видно из результатов, по всем метрикам наибольшая точность достигается при использовании ансамбля базовых классификаторов. При этом стоит заметить, что значительный вклад в итоговый результат вносят модели Isolation Forest и GRB. Другие модели по отдельности демонстрируют низкий результат.

На рисунке 2 приведена зависимость F-меры от уровня разложения j для наиболее результативных моделей, из которой можно заключить, что наиболее информативной, с точки зрения аномалий, является высокочастотная составляющая сигнала. С переходом в низкочастотную область точность обнаружения уменьшается и при уровне $j > 5$ сводится практически к случайному результату.

ЗАКЛЮЧЕНИЕ

Предложен подход обнаружения аномалий в технологических сигналах построения ансамбля базовых классификаторов без учителя. В основе подхода лежит применение коэффициентов НВП для формирования информативного признакового пространства для обучения классификаторов и консолидации прогнозов каждого. Рассмотрен математический аппарат и особенности вейвлет-преобразования, позволяющие локализовать аномалии в частотно-временной области. В соответствии с предложенным подходом разработаны модели обнаружения аномалий с применением различных алгоритмов машинного обучения и ансамбль из этих моделей. Проведен вычислительный эксперимент с реальными технологическими сигналами и произведена оценка результатов. Анализ полученных результатов подтвердил теоретические выводы о вейвлет-преобразовании и эффективность ансамбля перед одиночными моделями. Полученный результат имеет прикладное значение в целях повышения безопасности и техникоэкономических показателей промышленных объектов. На основе предложенного подхода возможна разработка системы контроля и предиктивной диагностики технологического оборудования, идентификации стадий технологических процессов и производств.

СПИСОК ЛИТЕРАТУРЫ

1. Machine Learning for Predictive Maintenance: A Multiple Classifier Approach / Susto Gian A., Pampuri

Simone and other // *IEEE Transactions on Industrial Informatics*. 2015. V. 11, iss. 3. pp. 812–820.

2. Liu Y., Ding W. A KNNs based anomaly detection method applied for UAV flight data stream // *Prognostics and System Health Management Conference*, Beijing, 2015. DOI: 10.1109/PHM.2015.7380051.

3. Brad C., Radu S.N. Predictive maintenance applications for machine learning // *Reliability and Maintainability Symposium (RAMS)*, Saint Petersburg, 2017. DOI: 10.1109/RAM.2017.7889679.

4. Tibco Statistica Inc. URL: <http://statistica.io/wordpress/wp-content/uploads/Striim-Partner-Solution-Brief.pdf> (дата обращения: 24.03.2018).

5. Мурзагулов Д.А., Замятин А.В. Адаптивные алгоритмы машинного обучения в управлении технологическими процессами // *Автоматизация и современные технологии*. 2018. № 8. С. 354–361.

6. Хуссейн Х.Ш., Якунин А.Г. Методы выявления аномалий при контроле динамических процессов природных и техногенных объектов // *Вестник ИжГТУ им. М.Т. Калашникова*. 2015. № 1 (18). С. 79–83.

7. Селевцов Л.И. Автоматизация технологических процессов. М.: Академия, 2014. 345 с.

8. Заварзин Д.В. К вопросу поиска аномалий во временных рядах // *Инновации в науке*. 2014. № 29. С. 59–64.

9. Шкодырев В.П., Ягафаров К.И. Обзор методов обнаружения аномалий в потоках данных // *Proceedings of the Second Conference on Software Engineering and Information Management*. 2017. V. 1864. С. 330–343.

10. Kanungo T., Mount D.M. An efficient k-means clustering algorithm: analysis and implementation // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002. V. 24, iss. 7. pp. 881–892.

11. Liu Fei T., Ting Kai Ming, Zhou Zhi-Hua. Isolation Forest // *Data Mining*, 2008. Eighth IEEE International Conference on Project. ICDM '08. DOI: 10.1109/ICDM.2008.17.

12. Shengfeng T., Jian Y., Chuanhuan Y. Anomaly Detection Using Support Vector Machines // *Advances in Neural Networks*. 2004. Vol. 3173. pp 592–597.

13. LOF: Identifying Density-Based Local Outliers / Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, Joerg Sander // *Proc. ACM SIGMOD 2000 Int. Conf. on Management of Data*. Dallas, TX, 2000. DOI: 10.1145/342009.335388.

14. Мисриханов А.М. Применение методов вейвлет-преобразования в электроэнергетике // *Автоматика и телемеханика*. 2006. Вып. 5. С. 5–23.

15. Фильтрация речевого сигнала с помощью вейвлет-преобразования при решении задач распознавания речи / Н.Е. Балакирев, С.Ю. Гуснин, М.А. Малков, Л.М. Червяков // *Известия Юго-Западного государственного университета*. 2012. № 5-2 (44). С. 044–050.

16. Ручай А.Н. Модифицированный метод сегментации речевого сигнала на основе непрерывного вейвлет-преобразования // *Докл. Томского государ-*

ственного университета систем управления и радиоэлектроники. 2012. Т. 2, № 1. С. 189–192.

17. Классификация динамики локальных максимумов в вейвлет-спектрах электроэнцефалограмм / Я.А. Туровский, С.Д. Кургалин, А.Г. Семёнов, А.В. Максимов // *Вестник Тамбовского государственного технического университета*. 2013. Т. 19, № 1. С. 90–96.

18. Григорьев Д.С., Спицын В.Г. Применение нейронной сети и дискретного вейвлет-преобразования для анализа и классификации электрокардиограмм // *Известия Томского политехнического университета*. 2012. Т. 321, № 5. С. 57–61.

19. Влияние гипергомоцистемии на мозговой кровоток, по данным вейвлет-анализа / В.В. Александрин, Б.П. Лузянин, А.В. Иванов, А.А. Кубатиев // *Патологическая физиология и экспериментальная терапия*. 2011. № 2. С. 13–18.

20. Murzagulov Damir A., Zamyatin Alexander V., Ostrast Pavel M. An approach to the detection of anomalies of technological signals using classification and wavelet transforms // *IEEE International Russian Automation Conference*. Sochi, 2018, September 9–16. DOI: 10.1109/RUSAUTOCON.2018.8501786.

21. Замятин А.В. Интеллектуальный анализ данных: учеб. пособие. Томск: Издательский Дом государственного университета, 2020. 196 с.

22. Толстихин И.О. Неравенства концентрации вероятностной меры в трансдуктивном обучении и PAC-Байесовском анализе: автореф. дисс... канд. физ.-мат. наук: 05.13.17. М., 2014.

23. Обобщенное среднее. URL: <https://bit.ly/379TT5S> (дата обращения: 20.11.2020).

24. Малла С. Вейвлеты в обработке сигналов. М.: Мир, 2005.

25. Документация на библиотеку PYOD. URL: <https://pyod.readthedocs.io/en/latest/> (дата обращения: 20.11.2020).

26. Gavrin S., Murzagulov D., Zamyatin A. Anomaly Detection in Process Signals within Machine Learning and Data Augmentation Approach // *Machine Learning and Data Mining in Pattern Recognition. MLDM 2019: 15th International Conference on Machine Learning and Data Mining*, New York, 20-25 July 2019: proceedings. Vol. 2. Leipzig: Ibai-publishing, 2019. pp. 585–598.

REFERENCES

1. Susto Gian A., Pampuri Simone et al. Machine Learning for Predictive Maintenance: A Multiple Classifier Approach. *IEEE Transactions on Industrial Informatics*, 2015, vol. 11, iss. 3, pp. 812–820.

2. Liu Y., Ding W. A KNNs Based Anomaly Detection Method Applied for UAV Flight Data Stream. *Prognostics and System Health Management Conference*, Beijing, 2015. DOI: 10.1109/PHM.2015.7380051.

3. Brad C., Radu S.N. Predictive Maintenance Applications for Machine Learning. *Reliability and Maintainability Symposium (RAMS)*, Saint Petersburg, 2017. DOI: 10.1109/RAM.2017.7889679.

4. *Tibco Statistica Inc.* Available at: <http://statistica.io/wordpress/wp-content/uploads/Striim-Partner-Solution-Brief.pdf> (accessed 24.03.2018).
5. Murzagulov D.A., Zamiatin A.V. Adaptivnye algoritmy mashinnogo obucheniia v upravlenii tekhnologicheskimi protsessami [Adaptive Algorithms of Machine Learning in the Management of Technological Processes]. *Avtomatizatsiia i sovremennye tekhnologii* [Automation and Modern Technologies], 2018, no. 8, pp. 354–361.
6. Husseyn H.Sh., Yakunin A.G. Metody vyavleniia anomalii pri kontrole dinamicheskikh protsessov prirodnykh i tekhnogennykh obektov [Methods for Detection of Anomalies in Control of Dynamic Processes in Natural and Industrial Objects]. *Vestnik IzhGTU im. M.T. Kalashnikova* [Bulletin of Kalashnikov ISTU], 2015, no. 1 (18), pp. 79–83.
7. Selevtsov L.I. *Avtomatizatsiia tekhnologicheskikh protsessov* [Automation of Technological Processes]. Moscow, Akademiia Publ., 2014. 345 p.
8. Zavarzin D.V. K voprosu poiska anomalii vo vremennykh riadakh [About Anomalies Detection Techniques in Time Series]. *Innovatsii v nauke* [Innovations], 2014, no. 29, pp. 59–64.
9. Shkodyrev V.P., Iagafarov K.I. Obzor metodov obnaruzheniia anomalii v potokakh dannykh [Review of Anomaly Detection Techniques in Data Flows]. *Proceedings of the Second Conference on Software Engineering and Information Management*. 2017, vol. 1864, pp. 330–343.
10. Kanungo T., Mount D.M. An Efficient k-Means Clustering Algorithm: Analysis and Implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, vol. 24, iss. 7, pp. 881–892.
11. Liu Fei T., Ting Kai Ming, Zhou Zhi-Hua. Isolation Forest. *Eighth IEEE International Conference on Data Mining, ICDM '08. 2008*. DOI: 10.1109/ICDM.2008.17.
12. Shengfeng T., Jian Y., Chuanhuan Y. Anomaly Detection Using Support Vector Machines. *Advances in Neural Networks*, 2004, vol. 3173, pp. 592–597.
13. Markus M., Breunig, Hans-Peter Kriegel, Raymond T. Ng, Joerg Sander. LOF: Identifying Density-Based Local Outliers. *Proc. ACM SIGMOD 2000 Int. Conf. on Management of Data*. Dallas, TX, 2000. DOI:10.1145/342009.335388.
14. Misrikhanov A.M. Primenenie metodov veivlet-preobrazovaniia v elektroenergetike [Wavelet Transform Methods: Application in Electroenergetics]. *Avtomatika i telemekhanika* [Automation and Remote Control], 2006, iss. 5, pp. 5–23.
15. Balakirev N.E., Gusnin S.Iu., Malkov M.A., Chervakov L.M. Filtratsiia rechevogo signala s pomoshchiu veivlet-preobrazovaniia pri reshenii zadach raspoznavaniia rechi [Speech Signal Filtering Using Wavelet Method for Solving Speech Recognition Problems]. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta* [Proceedings of the Southwest State University], 2012, no. 5-2 (44), pp. 044–050.
16. Ruchai A.N. Modifitsirovannyi metod segmentatsii rechevogo signala na osnove nepreryvnogo veivlet-preobrazovaniia [Modified Method of Segmentation of Speech Signal by Continuous Wavelet Transform]. *Dokl. Tomskogo gosudarstvennogo universiteta sistem upravleniia i radioelektroniki* [Proc. of Tomsk State University of Control Systems and Radioelectronics], 2012, vol. 2, no. 1, pp. 189–192.
17. Turovskii I.A., Kurgalin S.D., Semenov A.G., Maksimov A.V. Klassifikatsiia dinamiki lokalnykh maksimumov v veivlet-spektrakh elektroentsefalogramm [Classification of Local Maxima Dynamics in Wavelet Spectrum of Electroencefalograms]. *Vestnik Tambovskogo gosudarstvennogo tekhnicheskogo universiteta* [Transactions of Tambov State Technical University], 2013, vol. 19, no. 1, pp. 90–96.
18. Grigorev D.S., Spitsyn V.G. Primenenie neironnoi seti i diskretnogo veivlet-preobrazovaniia dlia analiza i klassifikatsii elektrokardiogramm [Application of Neuronal Network and Discrete Wavelet Transform for Analyzing and Classifying Electrocardiograms]. *Izvestiya Tomskogo politekhnicheskogo universiteta* [Bulletin of the Tomsk Polytechnic University], 2012, vol. 321, no. 5, pp. 57–61.
19. Aleksandrin V.V., Luzianin B.P., Ivanov A.V., Kubatiev A.A. Vliianie gipergomotsisteinonii na mozgovoii krovotok, po dannym veivlet-analiza [Influence of Hyperhomocysteinemia on the Wavelet-Analysis of Cerebral Blood Flow]. *Patologicheskaiia fiziologiia i eksperimentalnaia terapiia* [Pathological Physiology and Experimental Therapy], 2011, no. 2, pp. 13–18.
20. Murzagulov Damir A., Zamiatin Aleksandr V., Ostrast Pavel M. An Approach to the Detection of Anomalies of Technological Signals Using Classification and Wavelet Transforms. *IEEE International Russian Automation Conference*. Sochi, 2018, September 9–16. DOI: 10.1109 / RUSAUTOCON.2018.8501786.
21. Zamiatin A.V. *Intellektualnyi analiz dannykh*. Ucheb. posobie [Data Mining. Tutorial]. Tomsk, State University Publ., 2020. 196 p.
22. Tolstikhin I.O. *Neravenstva kontsentratsii veroiatnostnoi mery v transduktivnom obuchenii i PAC-Baiesovskom analize. Avtoref. diss... kand. fiz.-mat. nauk. 05.13.17* [The Inequality of the Probability Measure Concentration in Transductive Learning and PAC-Bayesian Analysis. Author's Abstract of Cand. Phys. and Math. Sci. Diss. 05.13.17]. Moscow, 2014.
23. *Obobshchennoe srednee* [Generalized Mean Value]. Available at: <https://bit.ly/379TT5S> (accessed 20.11.2020).
24. Malla S. *Veivlety v obrabotke signalov* [Wavelets in Signal Processing]. Moscow, Mir Publ., 2005.
25. *Dokumentatsiia na biblioteku PYOD* [Documentation for PYOD Library]. Available at: <https://pyod.readthedocs.io/en/latest/> (accessed 20.11.2020).
26. Gavrin S., Murzagulov D., Zamiatin A. Anomaly Detection in Process Signals within Machine Learning and Data Augmentation Approach. *Proc. of 15th International Conference on Machine Learning and Data Mining. Machine Learning and Data Mining in Pattern Recognition. MLDM 2019*. Leipzig, Ibai-publishing, 2019, vol. 2, pp. 585–598.