

ОБНАРУЖЕНИЕ АНОМАЛИЙ В ТЕХНОЛОГИЧЕСКИХ СИГНАЛАХ С ПРИМЕНЕНИЕМ НЕЙРОСЕТЕВОЙ АРХИТЕКТУРЫ ТИПА АВТОЭНКODER

В.А. Мальцев, Д.А. Мурзагулов, А.В. Замятин (Национальный исследовательский Томский государственный университет)

Исследована задача обнаружения аномалий в технологических сигналах. Представлен подход к обнаружению аномалий, основанный на нейронной сети с архитектурой автоэнкодера, позволяющей учитывать контекст изменений сигнала. Показано, что применение архитектуры автоэнкодера для LSTM сетей за счет механизма долгой-краткосрочной памяти решает проблему долговременных зависимостей в сигнале и позволяет выявлять аномалии с учетом временного контекста изменений сигнала. Представлено формальное описание подхода на основе автоэнкодера и проведен численный эксперимент на модельных данных. На основе анализа экспериментальных результатов приведена методика подбора обучающего порога, что позволяет работать в адаптивном режиме без априорной информации об аномалиях и самом сигнале^{1,2}.

Ключевые слова: аномалия, технологический сигнал, автоэнкодер, нейронная сеть, АСУТП, предиктивная аналитика, мониторинг оборудования.

Введение

К настоящему времени развитие АСУТП позволяет получать и накапливать большой объем технологических данных, связанный с состоянием технологического оборудования во время его работы [1]. Опираясь на эти данные, развиваются технологии предиктивной аналитики, позволяющие распознавать нежелательное (нештатное) поведение оборудования, соответствующее его скорому выходу из строя, либо уже произошедшей неисправности [2]. Знания, полученные из подобного рода данных при помощи средств предиктивной аналитики, позволяют оптимизировать ремонт и обслуживание сложного технологического оборудования, оптимизировать протекание процессов и выявлять истинные причины отказов, сбоев и аварий на производстве. Для построения систем предиктивного анализа на сегодняшний момент активно применяются методы интеллектуального анализа и алгоритмы машинного обучения [3].

Одна из ключевых задач анализа технологических данных — это определение штатных (нормальных) и нештатных (ненормальных) состояний технологического оборудования. Штатное состояние оборудования или процесса обычно описывается в технических регламентах или в документации от производителя. Нештатное поведение носит непредсказуемый характер, что зачастую негативно сказывается на показателях эффективности работы. Другими словами, в терминологии анализа данных такая задача сводится к обнаружению аномалий в данных.

Как правило, нештатные ситуации являются редкими событиями на производстве, поэтому во всем объеме технологических данных преобладает штатное поведение. В таких условиях подходы к задаче, такие как управляемая классификация на размеченных данных, могут показать себя неэффективно на небольших объемах данных [3]. Большую популярность имеют подходы обучения модели без учителя — обучение

только на штатном поведении. Таким образом, основная задача обнаружения аномалий в технологических сигналах сводится к определению момента времени перехода из штатного режима в нештатный. Поскольку характеристики штатного режима могут варьироваться в широких пределах, а также возможна и обратная ситуация, когда характеристики сигнала остаются в пределах нормы, но участок является аномальным с точки зрения паттерна сигнала, то задача выявления аномалий становится нетривиальной. При этом одной из проблем является учет краткосрочного и долгосрочного контекста изменений сигнала. Это связано с тем, что состояние сигнала зависит от различных внешних факторов, таких как режим эксплуатации, цикл производства, смена измеряемой среды и т.д. Для возможности учитывать информацию такого рода необходимо модель, способная обобщать долгосрочные зависимости и запоминать краткосрочный контекст. Такой подход реализуется в LSTM сетях [4, 5] — рекуррентных нейронных сетях с долгой краткосрочной памятью. Однако сама по себе LSTM сеть не эффективна применительно к задаче обнаружения аномалий, если соотношение нормальных и аномальных данных не сбалансировано. Для преодоления этого ограничения предлагается использовать архитектуру сети типа автоэнкодера для LSTM. Автоэнкодер состоит из двух нейронных сетей, которые отвечают за разную функциональность, из Энкодера и Декодера. В такой схеме входные данные переводятся кодировщиком (энкодер) в некое векторное представление, которое затем будет декодировано при помощи де-кодировщика (декодер) обратно в исходную последовательность. Типовой алгоритм применения автоэнкодера в задаче обнаружения аномалий состоит в следующем [6].

1. Модель обучается только на штатных сигналах (сигналы, которые соответствуют безаварийным состояниям объекта), стремясь минимизировать ошибку их восстановления.

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

² Исследование выполнено при финансовой поддержке НИ ТГУ в рамках научного проекта №8.1.62.2018.

2. Далее модель тестируется на сигналах с нештатными фрагментами (сигнал содержит аномалии). При этом предполагается, что ошибка восстановления на тестовых сигналах будет выше из-за присутствия аномальных фрагментов, которые ранее не наблюдались для данного сигнала.

3. На основе выбора некоторого уровня ошибки в качестве критического уровня происходит выявление аномальных данных.

При этом применение архитектуры автоэнкодера предоставляет ряд преимуществ.

- Подход не требует наличия аномальных значений в обучающих данных, так как модель обучается восстанавливать только нормальные данные, и если ошибка восстановления выше, чем при обучении, то можно заключить, что модель никогда не работала с такими данными, эти данные значительно отличаются от обучающих. На этом основании строится предположение об аномальности данных [7].

- Обобщающая способность — энкодер учится выделять главные признаки входной последовательности, а декодер по этим признакам восстанавливает исходную последовательность, то есть происходит обобщение входных данных [8].

- Возможность предварительного обучения по отдельности энкодера и декодера.

Отметим, что модель автоэнкодер может быть реализована на разных типах нейронных сетей. Например, на полносвязных, LSTM, сверточных нейронных сетях или их комбинациях. В данной работе выбрана сеть LSTM в силу вышеупомянутых соображений, а именно:

- за счет рекуррентной составляющей LSTM [9] могут учитывать предыдущие значения входной последовательности;

- механизм долгой-краткосрочной памяти решает проблему долговременных зависимостей, которая возникает, когда текущие значения последовательности значительны в контексте относительно далеких в прошлом;

Таким образом основная идея предлагаемого подхода к обнаружению аномалий заключается в применении архитектуры автоэнкодер для LSTM сети. Входные данные предлагается формировать в оконном режиме — рассматривать значения сигнала в окне с фиксированной шириной и единичным сдвигом по времени. Это позволит рассматривать значение сигнала в различном временном контексте. Цель работы состоит в разработке алгоритма обнаружения аномалий в сигналах на основе архитектуры автоэнкодер для LSTM сети.

Научная новизна работы заключается в предложенном подходе детектирования аномалий, основанном на схеме автоэнкодер LSTM сетей и в оконном режиме анализа сигнала, что позволит преодолеть преобладание штатных («нормальных») экземпляров данных над аномальными и учитывать временной контекст сигнала.

В связи с тем, что технологическое оборудование большую часть времени находится в штатном состоянии, необходимо применить подход, который учтет такого рода дисбаланс штатных и аномальных экземпляров данных. Иными словами, простая бинарная классификация на размеченных данных в этом случае покажет себя не самым лучшим образом. Следовательно, необходима архитектура нейронной сети, которая позволила бы, обучившись только на штатных фрагментах сигнала, определять нештатные фрагменты. Для этих целей применима схема сети автоэнкодер.

Рассмотрим каждую точку в контексте скользящего окна размера n , перемещаемого на одно значение. При таком подходе для каждой точки, за исключением первых $n-1$, будем иметь статистику в контексте каждого окна, в которое она попала.

Рассмотрим подробнее такой подход:

1. Рассмотрим сигнал $S(t) = \{s_1, s_2, \dots, s_N\}$, где N — число измерений, $s_N \in R$. Предполагается, что сигнал не содержит аномальные фрагменты.

2. Установим размер окна анализа n и сформируем матрицу U размерностью $[n \times n]$, полученной при движении по ряду окном n , так что $U_1 = [s_1, s_2, \dots, s_n]$, $U_2 = [s_2, s_3, \dots, s_{1+n}]$ и т. д.

3. Построим модель $A(U)$ на основе архитектуры автоэнкодер. В качестве метрики точности восстановления будем использовать *среднюю абсолютную ошибку* (Mean Absolute Error, MAE).

4. Модель $A(U)$ обучается на данных U , минимизируя MAE. В качестве метода минимизации функции потерь использовался алгоритм ADAM [10].

5. Далее, применим обученную модель к тестовому сигналу $P(t)$ (предварительно применив аналогичное преобразование с окном размера n), в котором присутствуют нормальные и аномальные фрагменты, и посчитаем ошибку γ восстановления для каждого окна. Такая ошибка будет представлять собой ошибку восстановления каждой точки в рамках текущего окна. Таким образом, будем получать n значений для каждого момента времени кроме первых $n-1$, которые будут соответствовать ошибке восстановления значения сигнала в контексте скользящего окна.

6. Исходя из распределения ошибки γ , определим пороговое значение как:

$$P = \tilde{\gamma} + \sqrt{\frac{1}{n} \sum_{i=1}^n (\gamma_i - \tilde{\gamma})^2},$$

где $\tilde{\gamma}$ — среднее значение ошибки. Функция определения порога выбрана исходя из нормального распределения ошибки.

7. Далее, возьмем среднее значение \bar{p} для $P(t)$ и применим к ней следующую дельта-функцию:

$$\delta(p) = \begin{cases} 1, & \bar{p} \geq \gamma \\ 0, & \bar{p} < \gamma \end{cases},$$

которая сравнивает полученное значение с пороговым и принимает значение 1, когда ошибка восста-

новления больше, чем выбранное пороговое значение, или 0 в противном случае. Таким образом значение функции 1 соответствует аномалии, а 0 нормальному состоянию.

Отметим, что число скрытых слоев и их размерность определялись экспериментальным путем и имеют следующие локально оптимальные значения: энкодер и декодер имеют по три скрытых слоя. Самый большой по размерности слой состоит из 256 нейронов, средний — из 128 нейронов, а слой с самой маленькой размерностью из 32 нейронов. Зависимость ширины окна от числа слоев и их размерности в данной работе не исследовались.

Постановка эксперимента и описание используемых данных

Тестирование предложенного подхода предлагается в следующей последовательности: 1) конфигурирование архитектуры сети и инициализация модел³; 2) обучение модели на сигналах, которые не содержат аномалий; 3) тестирование модели на сигналах, содержащих как аномалии, так и нормальные фрагменты; 4) оценка точности модели и проверка результатов.

В работе использовались данные полученные в [11]. Краткая характеристика данных.

1. Реальные сигналы с измерительных преобразователей давления, аугментированные различными типами аномалий. На реальный сигнал были наложены аномалии нескольких типов: мелкодисперсный шум, сдвиг, разброс амплитуды.

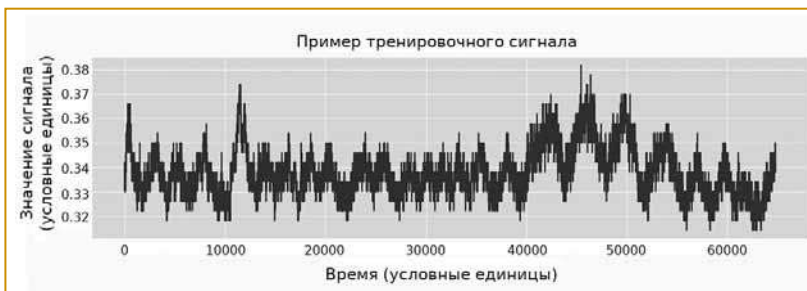


Рис. 1. Пример обучающего сигнала без аномалий



Рис. 2. Тестовый сигнал с разметкой аномалий

2. На сигналах имеется разметка — аномалия помечено единицей, штатное поведение нулем.

3. Всего в наборе представлено четыре типа аномалий: шум, изменение тренда, повышение дисперсии, снижение дисперсии;

4. Всего имеется по девять обучающих и тестовых сигналов, полученных с различных измерительных средств и представляющих временной интервал 6 месяцев.

На рис. 1 приведен пример тренировочного сигнала, а на рис. 2 проверочного с разметкой.

Результаты

Эксперимент проводился для различных размеров скользящего окна n : 120, 1200 и 3000 значений. Для оценки точности использовались метрики $F1$. Предлагаемый подход сравнивался с методом, основанным на обучении с учителем (бинарная классификация), и методом, который предсказывает по окну следующее значение последовательности (регрессия). Методы для сравнения реализованы с использованием LSTM для достоверности эксперимента.

Результат по размерам скользящего окна в зависимости от подхода представлен на рис. 3.

Как и предполагалось, точность бинарной классификации значительно выше предложенного подхода и регрессионной модели. Однако на практике такой подход мало реализуем из-за отсутствия разметки в реальных сигналах и дисбаланса классов, поэтому для дальнейшего анализа бинарную классификацию рассматривать не будем. Для предложенного подхода точность по метрике классификации $F1$ в среднем составляет 55% с достижением точности до 70% для отдельных сигналов. Заметим закономерность: чем больше размер окна, тем более точно модель определяет аномалию, однако в некоторых случаях, окно размером $w = 1200$ показывает лучшую точность, чем $w = 3000$. Рассмотрим метрику классификации $F1$ в контексте этих двух подходов (рис. 4).

Как видно из рисунка, в большинстве случаев лучшие результаты демонстрирует предлагаемый подход с автоэнкодером, однако точность в целом остается довольно на низком уровне — в среднем 50% для обоих подходов. Вопрос точности этих двух подходов остается открытым, так как в целом результат варьируется от одного тестового сигнала к другому. В первую очередь это обусловлено тем, что в этих тестовых сигналах присутствуют некоторые локальные выбросы, которые модель ложно отмечают аномалиями. Этот

³ В работе используя фреймворк для глубокого обучения Keras на языке программирования Python.

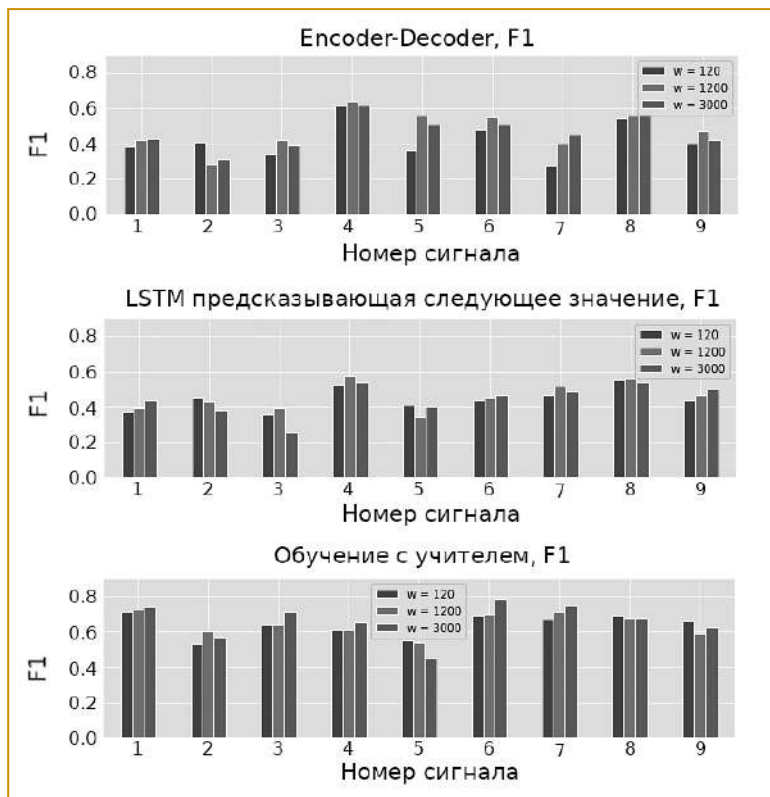


Рис. 3. Зависимость F1 от размеров окна для разных моделей

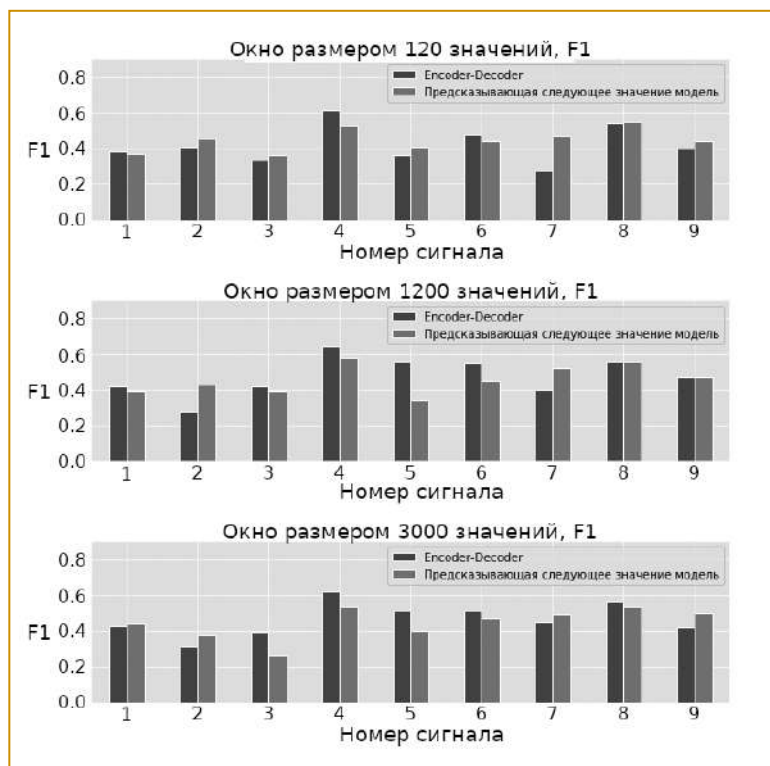


Рис. 4. Сравнение метода, предсказывающего следующее значение, и предлагаемого подхода

фактор можно устранить за счет гиперпараметров модели, таких как пороговое значение ошибки для каждого сигнала. Для некоторых сигналов этот порог должен быть выше, для других он будет совпадать с расчетным, для некоторых потребуется снизить по-

Таблица. Данные по детектированию конкретных типов аномалий схемой Encoder-Decoder

Тип аномалии	Размер окна	F1
Шум	120	0,86
Шум	1200	0,88
Шум	3000	0,90
Изменение тренда	120	0,88
Изменение тренда	1200	0,90
Изменение тренда	3000	0,90

рог. Такое же поведение можно встретить в других методах обучения без учителя [12]. Дополнительно эксперименты показали, что подход демонстрирует более низкую точность для аномалий, связанных со значительным изменением дисперсии сигнала. Исследование модели на основе автоэнкодера на способность определять аномалии разных видов позволило сделать вывод: модель не способна выявить аномалии вида “разброс амплитуды”, но эффективно обнаруживает аномалии типа “шум” и “изменение тренда”. Полученные точности по метрике классификации F1 представлены в таблице.

Таким образом, подход демонстрирует высокую точность — до 90% на отдельных типах аномалий. Этот фактор определяет область применения подхода на практике. Изменение тренда сигнала и наличие шума охватывает большой объем реальных отказов для измерений разного рода. Первое больше характерно в случае нарушения технологического процесса и связанных с ним увеличением/уменьшением технологического параметра. Второй тип аномалии более характерен для механических узлов оборудования, которые подвержены физическим дефектам, например, образование «задигов» на подшипниках насосов, которое ведет к появлению шумовых компонентов в сигнале. Примеры успешного обнаружения аномалий предложенным подходом представлены на рис. 5.

Закключение

В работе исследована задача обнаружения аномалий в технологических сигналах, актуальная в рамках всеобщей цифровой трансформации, направленная на оптимизацию технического обслуживания и ремонта оборудования.

Предложен оригинальный подход, базирующийся на схеме нейронной сети автоэнкодера, для создания модели обнаружения аномалий. Основными преимуществами такого подхода являются:

- а) анализ значений сигнала с учетом контекста изменений — на результат анализа влияет не только

текущее состояние сигнала, но и исторические;

б) подход не требует наличия аномалий в обучающем сигнале.

В результате выполнения задач были достигнуты следующие результаты:

- 1) разработана архитектура нейронной сети, реализующая предложенный подход;
- 2) проведен вычислительный эксперимент и сравнительный анализ предложенного метода с несколькими существующими;
- 3) продемонстрирована высокая точность (90%) предложенного подхода на нескольких видах аномалий.

Полученный результат дает четкое представление об области применимости предложенного подхода. В целом, точность такого метода в сравнении с другим методом, основанным на LSTM, выше. Но в то же время ему не удастся обойти по точности метод обучения с учителем (основанный на тех же LSTM). С другой стороны, метод обучения с учителем зачастую не может быть применен из-за высокого дисбаланса классов в реальных технологических данных. Предложенный подход может быть применен на реальном производстве с тем условием, что модель должна обучиться на данных оборудования, на сигналах которого будет происходить работа. Также может возникнуть необходимость изменять пороговое значение подхода в зависимости от данных. Одномерный случай при данном подходе может быть расширен на многомерный.

Список литературы

1. *Замятин А.В., Тренькаев В.Н., Острасть П.М., Телицын Е.А. и др.* Высокопроизводительный сервер истории системы диспетчерского управления и сбора данных//Промышленные АСУ и контроллеры. — 2017. — № 9. — С. 20–28.
2. *Gavrin S.* Detection of Change Point in Process Signals by Cascade Classification / Gavrin S., Murzagulov D., Zamyatin A. // 2018 International Russian Automation Conference (RusAutoCon 2018). — 2018. — Vol. 1–2. — P. 515–518.
3. *Мурзагулов Д.А., Замятин А.В.* Адаптивные алгоритмы машинного обучения в управлении технологическими процессами // Автоматизация. Современные технологии. 2018. № 8. С. 354–361.
4. *Ивченко Р.А., Купин А.И.* Исследование и разработка актуальных приемов, технологий и методик, применяемых при решении прикладных задач машинного обучения в целях предиктивного анализа // Системні технології (“System technologies”). — 2020. — №2 (127). — С. 55–79.
5. *Series Pankaj Malhotra, Lovekesh Vig, Gautam Shroff, Puneet Agarwal.* Long Short Term Memory Networks for Anomaly Detection in Time// ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), April 2015.
6. *Baihong Jin, Yingshui Tan, Alexander Nettekoven, Yuxin Chen.* Encoder-Decoder Based Approach for Anomaly Detection with Application in Additive Manufacturing// <https://arxiv.org/pdf/1907.11778.pdf>.
7. *Yong Shean Chong, Yong Haur Tay.* Abnormal Event Detection in Videos Using Spatiotemporal Autoencoder// International Symposium on Neural Networks. Advances in Neural Networks - ISSN 2017 pp 189–196.
8. *Нечахин В.А., Пищик Б.Н.* Применение методов глубокого обучения для обнаружения вторжений // Вестник Новосибирского государственного университета. Серия: Информационные технологии, 2019.
9. *Hochreiter S., Schmidhuber J.* Long short-term memory// Neural Computation. — 1997. — Vol. 9, no. 8. — P. 1735–1780.
10. *Simon Fong, Anke Meyer-Baese.* Multi-stage optimization of a deep model: A case study on ground motion modeling// PLoS One, 2018, doi:10.1371/journal.pone.0203829.
11. *Гаврин С.В.* Алгоритмическое обеспечение для обнаружения аномалий в многомерных технологических сигналах методом градиентного бустинга: маг. дис. [Электронный ресурс] / Гаврин С.В. — Электрон. дан. — Томск: [б.и.], 2019. — URL: <http://vital.lib.tsu.ru/vital/access/manager/Repository/vital:9875> (дата обращения: 27.02.2020). — Доступ из локальной сети Науч. б-ки Том. гос. ун-та.
12. *Мальцев В.А., Мурзагулов Д.А., Замятин А.В.* Применение неконтролируемой классификации в задаче обнаружения аномалий в технологических сигналах// Прикладная математика и информатика: современные исследования в области естественных и технических наук : Тр. V международной науч.-практич. конф. (школы-семинара) молодых ученых. — Тольятти. 2019. — С. 408–412.

Мальцев Вячеслав Андреевич — программист Improvado.io, техник лаборатории программно-аппаратных средств автоматизации ТГУ,

Мурзагулов Дамир Альбертович — главный аналитик Improvado.io, ассистент,

Замятин Александр Владимирович — д-р техн. наук, доцент, директор института прикладной математики и компьютерных наук, заведующий кафедрой теоретических основ информатики ТГУ.

E-mail: murzagulov.damir@gmail.com zamyatin@mail.tsu.ru

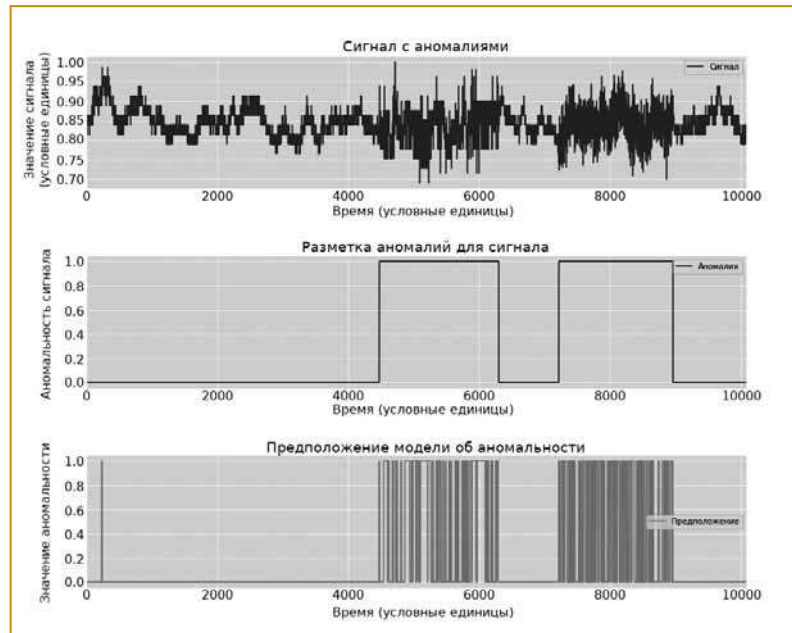


Рис. 5. Пример работы подхода на основе автоэнкодера