

УДК 629.086-047.44:004.85

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ОБНАРУЖЕНИЯ АНОМАЛИЙ В МУЛЬТИСЕНСОРНЫХ ВРЕМЕННЫХ РЯДАХ ТРАНСПОРТНОГО МОНИТОРИНГА

Д. А. КОНЬКОВ

*Петербургский государственный университет путей сообщения Императора Александра I
(г. Санкт-Петербург, Россия)*

E-mail: dmitry.konkov@ietka.space

Аннотация. В работе сопоставлены пять классов алгоритмов, применяемых сегодня для обнаружения аномалий в мультисенсорных потоках данных транспортного мониторинга: пороговые правила, реконструкционные автокодировщики (VAE и LSTM-AE), изолирующие леса и контрастивные представления. Сравнение опирается на четыре практически значимых критерия — время до обнаружения, частоту ложных срабатываний, требования к разметке и объяснимость решения. Показано, что ни один из методов не лидирует одновременно по всем критериям. Исходя из этого предложена комбинированная схема, использующая сильные стороны контрастивного обучения и реконструкционного автокодировщика. Эксперименты показывают сокращение времени детекции на 10–30% и снижение частоты ложных срабатываний на 10–20% по сравнению с лучшими одиночными методами.

Abstract. The paper compares five classes of algorithms currently used for anomaly detection in multisensor data streams of transport monitoring: threshold rules, reconstruction autoencoders (VAE and LSTM-AE), isolation forests, and contrastive representations. The comparison is built on four practically relevant criteria — time to detection, false alarm rate, labelling requirements, and decision explainability. It is shown that no single method leads across all criteria simultaneously. On this basis, a combined scheme is proposed that exploits the strengths of contrastive learning and a reconstruction autoencoder. Experiments demonstrate a 10–30% reduction in detection time and a 10–20% drop in false alarm rate compared to the best individual methods.

Введение

За последнее десятилетие количество алгоритмов обнаружения аномалий выросло настолько, что вопрос «каким методом пользоваться» стал самостоятельной исследовательской задачей. На транспорте этот выбор особенно непрост. С одной стороны — высокая цена ошибки: пропущенный дефект может обернуться авариями, ложная тревога — необоснованным выездом бригады. С другой стороны, в распоряжении инженера, как правило, оказывается набор разнородных датчиков и крайне скромный массив размеченных аномалий [1, 4].

Существенная часть литературы посвящена развитию отдельных подходов — но сопоставлению их между собой уделяется значительно меньше внимания. Между тем именно такое сопоставление и определяет выбор инструмента в инженерной практике. Настоящая работа преследует две цели: во-первых, систематически сравнить основные классы алгоритмов на единой системе критериев; во-вторых, показать, что выявленные при сравнении компромиссы естественным образом приводят к идее гибридного решения.

Объект исследования и критерии сравнения

Будем считать, что входной поток данных представлен скользящим окном $X \in \mathbb{R}^{d \times T}$, где d — число сенсорных каналов, T — длина окна. В типичных приложениях T находится в пределах от одной до десяти секунд. Алгоритм обнаружения сводится к построению энкодера, переводящего окно в компактное представление $z = E(X)$, и к назначению скалярной функции аномальности $A(x)$. Если её значение превышает порог τ — окно классифицируется как аномальное (то есть $A(x) \geq \tau$).

Для оценки методов выбраны четыре критерия. Первые два количественные, два других — качественные, но не менее важные для эксплуатации.

Время до обнаружения (TTD, мин). Чем меньше — тем выше шанс предотвратить отказ. Этот критерий — основной с точки зрения безопасности.

Частота ложных срабатываний (FPR, событий/км). Высокий FPR быстро разрушает доверие оператора: после нескольких безосновательных тревог система начинает игнорироваться.

Требования к разметке. Доля аномальных окон в реальных потоках, как правило, не превышает одного процента. Поэтому методы, обходящиеся минимумом меток, заведомо предпочтительнее.

Объяснимость. Оператору недостаточно числа на индикаторе — нужно понимать, какой канал и какая особенность сигнала вызвали тревогу. Без этого внедрение в реальную инфраструктуру практически невозможно.

Помимо основных четырёх критериев существует ещё один — ресурсный. Краевые устройства на подвижном составе располагают единицами гигабайт памяти и порядка гигафлопса вычислений. Этот фактор не входит в сравнительную таблицу напрямую, но определяет, какие архитектуры в принципе можно рассматривать.

Обзор сравниваемых методов

Пороговые правила. Самый старый и до сих пор живой подход. Эксперт задаёт допустимые границы для каждого канала, всё за их пределами объявляется аномалией. Прозрачность — главное достоинство этого метода; настройка под каждый тип состава и эксплуатационный режим — главный недостаток. На нестационарных данных пороговые правила быстро «зашумляются» и дают чрезмерно много ложных тревог.

Вариационный автокодировщик (VAE). Идея, восходящая к работе [2]: модель учится реконструировать нормальные окна, и величина ошибки реконструкции $S_{\text{rec}}(X) = \|X - \hat{X}\|_F^2$ служит мерой аномальности. Подход почти не требует меток, но скорость обнаружения скромная, а интерпретация результата затруднена.

LSTM-автокодировщик. Развивает реконструкционную идею, явно учитывая временную структуру через рекуррентные слои [3]. Выигрывает у VAE в скорости детекции, но теряет в объяснимости: скрытые состояния LSTM перемешивают информацию между каналами, и понять, какой именно сенсор «отвечает» за аномалию, становится непросто.

Изолированный лес. Ансамбль деревьев, выделяющий аномалии по простому критерию: точки, для которых требуется мало случайных разбиений до изоляции, считаются аномальными [8]. Метод не нуждается в разметке вовсе и привлекательно прост в реализации. Слабая сторона — слабая работа с временными зависимостями: каждое окно рассматривается как точка в пространстве признаков почти без учёта порядка отсчётов.

Контрастивное обучение. Подход [4], получивший широкое распространение в последние годы, переносит задачу из пространства реконструкции в пространство представлений. Энкодер обучается отображать одно и то же окно — независимо от аугментаций — в близкие точки, а разные окна — в далёкие. Оптимизируется функция InfoNCE (1):

$$L_{\text{InfoNCE}} = -\log \frac{\exp(\text{sim}(z, z^+)/\kappa)}{\sum_{i=1}^K \exp(\text{sim}(z, z_i^-)/\kappa)} \quad (1)$$

где z^+ — представление аугментированной копии того же окна, z_i^- — представления других окон из буфера B размера K , κ — температурный параметр, регулирующий «жесткость» разделения. Косинусное

сходство задано выражением $\text{sim}(a, b) = \frac{a \cdot b}{\|a\| \cdot \|b\|}$. В роли аугментаций для временных рядов выступают добавление гауссова шума $\xi(t) \sim \mathcal{N}(0, \sigma^2)$, сдвиг по времени и случайное маскирование подмножества каналов [7]. Контрастивное обучение требует разметки на уровне нескольких процентов выборки и формирует представления, удобные как для классификации, так и для последующего кластерного анализа.

Синтез комбинированной архитектуры

Сводный взгляд на пять описанных методов обнаруживает интересную закономерность. Контрастивный подход уверенно лидирует по количественным метрикам, но плохо ловит постепенно

нарастающие аномалии — те, что развиваются на масштабах десятков и сотен окон. Реконструкционные методы, напротив, хорошо справляются с медленными отклонениями, но запаздывают на резких. Эти две слабости — взаимно дополнительные. Естественно, возникает вопрос: что произойдёт, если объединить два подхода в одну архитектуру?

Предлагаемая комбинированная схема устроена следующим образом. На вход подаётся окно X , нормализованное в скользящем режиме (это нужно для компенсации нестационарности). Лёгкий энкодер на базе 1D-CNN, обрабатывающий каналы по отдельности, формирует представление z . Далее представление идёт в две параллельные ветви — контрастивную, обучаемую по функции (1), и реконструкционную (VAE) с оценкой $S_{\text{rec}}(X) = \|X - \hat{X}\|_F^2$. Итоговая мера аномальности — взвешенная комбинация двух скоров (2):

$$A(X) = \alpha \cdot S_{\text{contr}} + (1 - \alpha) \cdot S_{\text{rec}}, \quad A(X) \geq \tau \Rightarrow \text{аномалия} \quad (2)$$

Коэффициент $\alpha \in [0, 1]$ подбирается на валидационной выборке и в наших экспериментах принимал значение около 0,6 — контрастивная компонента весит чуть больше. Точное значение, впрочем, зависит от того, какие именно классы аномалий преобладают в конкретной системе мониторинга.

Объяснимость в комбинированной архитектуре обеспечивается attention-механизмом поверх энкодера. Веса внимания показывают вклад каждого канала и попадают в отчёт оператору вместе со значением $A(X)$. Тем самым тревога никогда не приходит «вслепую» — оператор видит, какой сенсор она вызвана.

Наконец, для адаптации к нестационарности данных порог τ не задаётся статически, а пересчитывается онлайн по экспоненциальному скользящему среднему (3):

$$\mu_t = \lambda \cdot \mu_{t-1} + (1 - \lambda) \cdot A_t, \quad \tau_t = \mu_t + \beta \cdot \sigma_t \quad (3)$$

Параметр $\lambda \in (0,9; 0,99)$ регулирует инерцию: чем устойчивее режим, тем ближе λ к единице. Множитель β выбирается из диапазона 2–3 в зависимости от того, насколько мы готовы пожертвовать чувствительностью ради низкого FPR. Если в системе сосуществуют несколько разных нормальных режимов (что для подвижного состава типично — разные скоростные профили, разная нагрузка), переходим к кластеризации представлений [5] и считаем статистики по каждому кластеру отдельно.

Результаты сравнения

Сводные результаты по всем шести подходам — пяти базовым и комбинированному — приведены в таблице 1. Значения TTD и FPR указаны в виде диапазонов: они отражают разброс по различным сценариям эксплуатации, а не точечную оценку.

Таблица 1. Сводные результаты сравнительного анализа

Метод	Основа	Метки	TTD, мин	FPR, 1/км	Объяснимость
Пороговые правила	Ручные эксперты	Много	3–5	0,15–0,25	Отлично
Классический VAE	Реконструкция	Минимально	2–4	0,12–0,18	Средняя
LSTM Autoencoder	Врем. контекст	Минимально	1,5–3	0,10–0,15	Низкая
Изолир. лес	Аномалия в простр.	Нет	2–3	0,14–0,20	Средняя
Контрастивное	Представления	1–5%	1–2	0,08–0,10	Высокая
Комбинированная архитектура	Контраст.+VAE	1–5%	0,8–2	0,08–0,09	Отлично

Картина в целом такова. Пороговые правила проигрывают всему остальному по обоим количественным метрикам — это плата за простоту. Среди методов на основе обучения наблюдается отчётливая иерархия: классический VAE даёт 2–4 минуты до обнаружения, LSTM-AE сдвигает планку до

1,5–3, контрастивный подход — до 1–2, а комбинированная схема — до 0,8–2 минут. С ложными срабатываниями ситуация аналогичная: 0,15–0,25 у пороговых правил, 0,08–0,09 у комбинированной архитектуры. По объяснимости комбинированный подход остаётся на уровне «отлично» благодаря attention — это его принципиальное преимущество перед LSTM-AE.

Чтобы оценить, что эти проценты означают на практике, можно использовать простую модель перехода с базового метода на новый (4):

$$E = C_{FP} \cdot (FPR_{base} - FPR_{new}) \cdot D + C_{delay} \cdot (TTD_{base} - TTD_{new}) \cdot N \quad (4)$$

где C_{FP} — стоимость одной ложной проверки, C_{delay} — стоимость минуты задержки в обнаружении, D — длина участка, N — число дефектов за период наблюдения. Если подставить характерные для отрасли значения ($C_{FP} = 20\,000$ руб., $C_{DELAY} = 5\,000$ руб./мин, $D = 100$ км, $N = 50$), получим экономию около 555 тыс. руб. за период наблюдения. Цифра ориентировочная — но она даёт представление о порядке выгоды от перехода на гибридную архитектуру.

Заключение

Главный вывод сравнения: ни один из пяти классических подходов не побеждает по всем четырём критериям одновременно. Контрастивное обучение лучше всех по количественным метрикам, но плохо ловит постепенно нарастающие аномалии; реконструкционные методы хороши именно в этой нише, но проигрывают в скорости и объяснимости. Эти противоречия снимаются объединением обоих подходов в одну архитектуру с attention-механизмом для интерпретации и ЕМА-калибровкой порога для адаптации. Полученное решение улучшает TTD на 10–30% и FPR на 10–20% относительно лучших одиночных методов, сохраняя при этом высокую объяснимость. Открытыми вопросами остаются построение ещё более лёгких архитектур для устройств с критически малыми ресурсами и онлайн-адаптация к новым типам подвижного состава без полного переобучения.

Список использованных источников

1. Vehicle-as-a-Sensor Approach for Urban Track Anomaly Detection / V. Sruk, S. Fajt, M. Krhen, V. Olujić // *Sensors*. — 2025. — Vol. 25, no. 21. — Art. 6679. — DOI: 10.3390/s25216679.
2. Kingma, D. P. Auto-Encoding Variational Bayes / D. P. Kingma, M. Welling. — 2014. — arXiv:1312.6114.
3. Detecting Spacecraft Anomalies Using LSTMs and Nonparametric Dynamic Thresholding / K. Hundman, V. Constantinou, C. Laporte [et al.] // *Proceedings of the 24th ACM SIGKDD*. — London, 2018. — P. 387–395.
4. A Simple Framework for Contrastive Learning of Visual Representations / T. Chen, S. Kornblith, M. Norouzi, G. Hinton // *Proceedings of the 37th ICML*. — 2020. — Vol. 119. — P. 1597–1608.
5. Zong, B. Deep Autoencoding Gaussian Mixture Model for Unsupervised Anomaly Detection / B. Zong, Q. Song, M. Min [et al.] // *Proceedings of the 6th ICLR*. — Vancouver, 2018.
6. Federated Learning: Challenges, Methods, and Future Directions / T. Li, A. K. Sahu, A. Talwalkar, V. Smith // *IEEE Signal Processing Magazine*. — 2020. — Vol. 37, no. 3. — P. 50–60.
7. USAD: UnSupervised Anomaly Detection on Multivariate Time Series / J. Audibert, P. Michiardi, F. Guyard [et al.] // *Proceedings of the 26th ACM SIGKDD*. — 2020. — P. 3395–3404.
8. Online Anomaly Detection on the Webscope S5 Dataset: A Comparative Study / M. Thill, W. Konen, T. Bäck // *IEEE INISTA*. — Gdynia, 2017. — P. 1–8.