



HAL
open science

Unsupervised Time Series Anomaly Detection: The Road to Effective Explainability

Michele Linardi, Vassilis Christophides

► **To cite this version:**

Michele Linardi, Vassilis Christophides. Unsupervised Time Series Anomaly Detection: The Road to Effective Explainability. Conférence francophone sur l'Extraction et la Gestion des Connaissances 24-28 janvier 2022 (Atelier EXPLAIN'AI), Jan 2022, Blois, France, France. hal-04316207

HAL Id: hal-04316207

<https://hal.science/hal-04316207>

Submitted on 30 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Unsupervised Time Series Anomaly Detection: The Road to Effective Explainability.

Michele Linardi*, Vassilis Christophides**

*ETIS UMR-8051 CY Cergy Paris Université

michele.linardi@cyu.fr

**ETIS UMR-8051, ENSEA,

Vassilis.Christophides@ensea.fr

1 Anomaly Detection in Multivariate Time Series

Introduction. Given a large collection of multivariate data sequences¹, or a continuously incoming time series in real-time systems, the anomaly detection task (AD) aims to find data occurrences that describe rare and anomalous events. Anomalies can have different forms; they can occur as single abnormal values (a.k.a. point anomaly) or as complex temporal events described by a contiguous sequence of individual data points (a.k.a sub-sequence anomaly).

State-of-the-art. Currently, we dispose of a plethora of heterogeneous and unsupervised AD techniques (Freeman et al. (2021)), which exhibit shaky performance according to the specificity of the analyzed sequences. In fact, accuracy performance may vary depending to data seasonality, trend, variables correlation drift over time, and the presence of missing values.

In this context, a user that wants to perform anomaly detection must have a deep understanding of the underlying data characteristics in order to know which algorithm represents the best fit in term of accuracy. This is clearly an unrealistic constraint, which severely limit the adoption of the current AD solutions. Moreover, the users require to have the necessary confidence to trust the decisions taken by the model, which cannot be applied as a black-box. In details, AD models must ideally provide explanations that indicate the specific data characteristics that triggered anomaly detection (Rojat et al. (2021)). In this sense, the ability to explain why a data point (or a sequence) was considered as abnormal (a.k.a Explainable Anomaly Detection) is a strict requirement in a variety of safety-critical domains.

Challenges. We consider two open challenges in AD over temporal data. Firstly, we note that typically multivariate time series, represent a difficult-to-interpret data type due to the presence of many different variables (i.e., monitored metrics) that are non-trivially correlated due to synchronization or sampling mismatches. Modern machine/deep learning (DL) models can efficiently learn temporal data evolution but do not provide a robust solution that can learn complex inter-metrics correlations. Furthermore, we need to account that human TS visual interpretation is a hard task also for an expert user. We should note that obtaining a satisfying level of insight remains a more natural task on images, video, or text.

1. Multivariate time series (TS) are sets of one-dimension variables (a.k.a metrics), namely sequences of real values recorded over time at a fix interval rate.

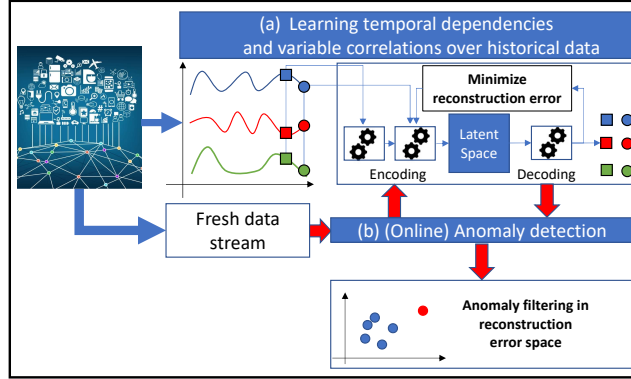


FIG. 1 – Architecture of unsupervised anomaly detection in multivariate time series. (a) Learning time series representation in a latent space (Variational autoencoder), (b) Detect anomaly based on reconstruction errors.

Several DL algorithms that obtain state-of-the-art performance in unsupervised TS-AD are black-box applications for the user. This is the case of variational autoencoders that adopt neural-network architectures like LSTM, GRU, and CNN (Su et al. (2019); Li et al. (2021); Pol et al. (2019); Nguyen et al. (2021)). Using these methods does not provide the possibility to shed any light on the critical factors that affect the decisions taken by the model. Specifically, to increase the confidence in the anomaly detection outcomes, the user may want to understand which feature sub-spaces actively contribute to anomaly classification and which are the root causes that trigger such abnormal event. Even though several recent research efforts (Rojat et al. (2021)) have tried to adapt explainability techniques used in DL models that were originally applied to images and videos, they currently fail to address the specific requirements of temporal data. To increase user confidence, a model should provide a holistic picture of system events. Hence, in the case of multi-variate TS, we want to obtain information about the temporal evolution of an anomaly, e.g., we want to detect and understand all the events in the past that trigger an anomaly, and along which metrics they occur. Exclusively detect the data occurrences that deviates from normality is a clear limitation of the currently available techniques.

In the next part, we briefly discuss the main elements of our vision of new generation anomaly detection in multivariate time series.

2 The Road To Explainable Anomaly Detection

Here, we introduce the architecture of the current AD models based on a DL architecture, discussing the research directions we want to follow.

Unsupervised state-of-the-art AD models adopts the architecture we sketch in Figure 1. The first part (Figure 1(a)) consists into an unsupervised learning step over historical multivariate time series using variational autoencoders and recurrent artificial neural networks. In the learning phase, the variational autoencoder maps real data variables to an (estimated) stochastic distribution in a latent space, while minimizing decoding loss, which corresponds to the

error of sampling data from the estimated distribution. In this respect, learning to accurately reconstruct the temporal evolution of normal data, while maintaining the existing correlations between the different variables in a latent space becomes crucial to detect anomalies.

In fact, online anomaly detection (Figure 1(b)) considers reconstruction errors of fresh data that are ideally high when decoding abnormal occurrences, and thus the model correctly reconstructs the structure of normal data. Hierarchical variational autoencoder initially proposed for image learning, are recently used to embed time series into a latent space that encodes both the temporal dependency and the inter-variable correlation of multivariate time series.

Learning Lagged Correlations. We note that an unsupervised autoencoder learns in essence correlations in a fixed time window. Hence, the user must know not only the window length parameter, but she must expect that inter-variable correlations occur in a well-known timeframe. We thus need to enlarge the capability of current learning algorithm that must be able to learn metrics relations irrespective of their time shift and scale. We consider this as a crucial requirement of models that implement explainability.

To that extent, Long-Run Variance Decomposition Networks (LVDN) are types of complex networks that are typically used to describe lagged relationships (Silva et al. (2021)). LVDN are graphs whose nodes represent TS modeled by vector autoregression, and edges get weighted by the information that one TS contributes to the other. Specifically, each weight connecting TS A and B measures how much the anomalies in A influence the variance of forecasting error of B. Statistical approaches based on LVDN have found high applicability in market analysis to capture long-run dependencies of shocks (abrupt events) (Lanne and Nyberg (2016)). We envision the adoption of complex network representation in AD models as it can serve the building of a discrete and finite structure that encodes scale-invariant correlations.

Effective User Interaction. Our second objective is to enable effective user interaction to the learning phase of anomaly detection. This is possible if the user can interpret the relationships learned by the model, e.g., learning complex networks data structures that encode the temporal evolution and relationship of time series metrics. In this sense, we should stress the growing research efforts towards interpretable (Yuan et al. (2020)) deep learning models applied to graphs. In contrast to applying AD models as black-boxes, graph modeling should permit us to design explanation strategies that highlight relationships between the detected anomalies and their precursors.

Anomaly Causal Analysis. One of the crucial requirements in safety-critical domains is causal analysis, namely the detection and study of the events chain that eventually brings the underlying system in an abnormal state. We observe that existing Causal Models (CMs) work over graph structures that encode temporal relationships (Moraffah et al. (2021)). In this sense, we should consider the theoretical connection between Graph Neural Network and CMs (Zecovic et al. (2021)) to propose new applications that can learn anomaly leveraging the powerful graph encoding.

References

- Freeman, C., J. Merriman, I. Beaver, and A. Mueen (2021). Experimental comparison and survey of twelve time series anomaly detection algorithms. *Journal of Artificial Intelligence Research (JAIR)*.
- Lanne, M. and H. Nyberg (2016). Generalized forecast error variance decomposition for linear and nonlinear multivariate models. *Oxford Bulletin of Economics and Statistics* 78.
- Li, Z., Y. Zhao, J. Han, Y. Su, R. Jiao, X. Wen, and D. Pei (2021). Multivariate time series anomaly detection and interpretation using hierarchical inter-metric and temporal embedding. In *KDD*, pp. 3220–3230. ACM.
- Moraffah, R., P. Sheth, M. Karami, A. Bhattacharya, Q. Wang, A. Tahir, A. Raglin, and H. Liu (2021). Causal inference for time series analysis: Problems, methods and evaluation. *CoRR abs/2102.05829*.
- Nguyen, H. D., K. P. Tran, S. Thomassey, and M. Hamad (2021). Forecasting and anomaly detection approaches using LSTM and LSTM autoencoder techniques with the applications in supply chain management. *Int. J. Inf. Manag.* 57, 102282.
- Pol, A. A., V. Berger, C. Germain, G. Cerminara, and M. Pierini (2019). Anomaly detection with conditional variational autoencoders. In *IEEE ICMLA 2019*, pp. 1651–1657. IEEE.
- Rojat, T., R. Puget, D. Filliat, J. D. Ser, R. Gelin, and N. D. Rodríguez (2021). Explainable artificial intelligence (XAI) on timeseries data: A survey. *CoRR abs/2104.00950*.
- Silva, V. F., M. E. Silva, P. Ribeiro, and F. Silva (2021). Time series analysis via network science: Concepts and algorithms. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 11(3).
- Su, Y., Y. Zhao, C. Niu, R. Liu, W. Sun, and D. Pei (2019). Robust anomaly detection for multivariate time series through stochastic recurrent neural network. In *SIGKDD KDD 2019*, pp. 2828–2837. ACM.
- Yuan, H., H. Yu, S. Gui, and S. Ji (2020). Explainability in graph neural networks: A taxonomic survey. *CoRR abs/2012.15445*.
- Zecevic, M., D. S. Dhimi, P. Velickovic, and K. Kersting (2021). Relating graph neural networks to structural causal models. *CoRR abs/2109.04173*.

Résumé

We study the problem of anomaly detection (AD) and explanation algorithms in multivariate time series. We introduce the main details of the state-of-the-art AD architecture, discussing several limitations of the existing techniques and the aspects we need to consider to address the open challenges. To date, we dispose of a plethora of AD techniques, which exhibit shaky performance depending on the data characteristics. In this context, we must unrealistically assume that the user would have a deep understanding of the data and she would trust the actions taken by the applied model. We thus envision algorithms that automatically learn complex data dependencies irrespective of time shift and scale of multivariate time series variables. Hence, this must enable users to interpret the evolution of the studied events to interact and refine the anomaly outcomes and perform an effective anomaly causal analysis.