# Unsupervised anomaly detection in multivariate time series

**Damian Horna**

## What is an anomaly?

> “An anomaly is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism.”
> D. M. Hawkins

## Multivariate time series

Many today’s systems are monitored with multivariate time series. Some primary use cases include IoT, DevOps and Real-Time Analytics.

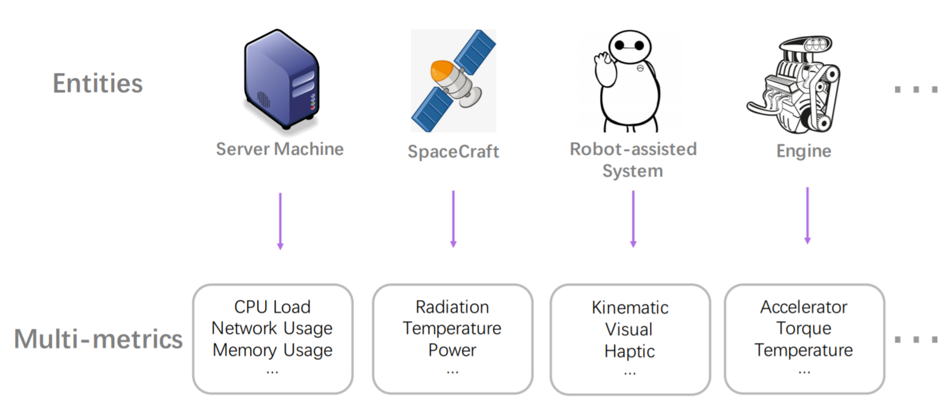


Figure 1: Entities with monitored multivariate time series

## Problematic characteristics of anomalous data

Anomaly detection especially in an unsupervised case and on multivariate time series data where time aspect is of key significance may be very challenging. Apart from severly skewed data distribution, we may face other problems, e.g. constantly changing definition of an anomaly. In our dataset we may have only a few anomalous examples and in most cases they won’t cover the broad spectrum of anomalies that can actually appear in practice. Exemplary multivariate time series fragment with relatively-easy-to-spot anomaly is presented below:

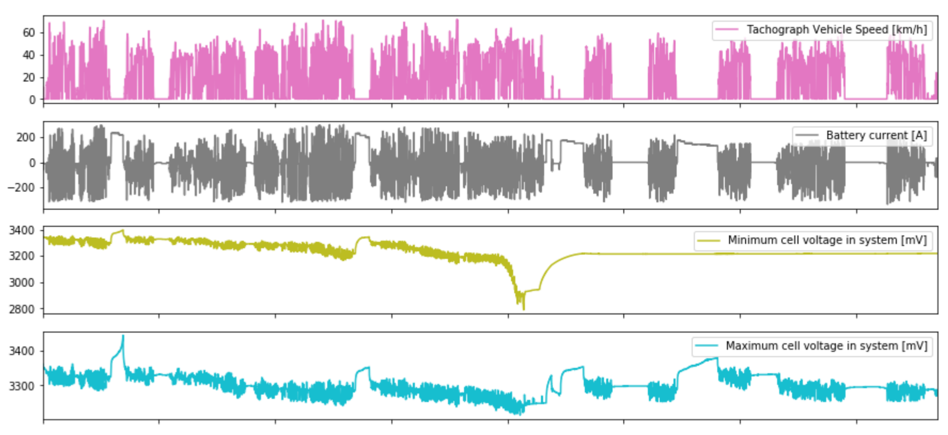

Figure 2: Exemplary multivariate time series with an anomaly

## Our approach

1. Exploratory data analysis and feature engineering - to get the intuition for what is normal and what is anomalous.
2. Generation of synthetic anomalies - using knowledge we gained during EDA.
3. Test-bed construction - with the use of both observed and generated synthetic anomalies.
4. Evaluation of unsupervised anomaly detection algorithms - on the test-bed we prepared before.

## Prediction-based anomaly detection: LASSO regression

LASSO model takes advantage of autoregressive features (similar to well-known AR model) in multivariate scenario:

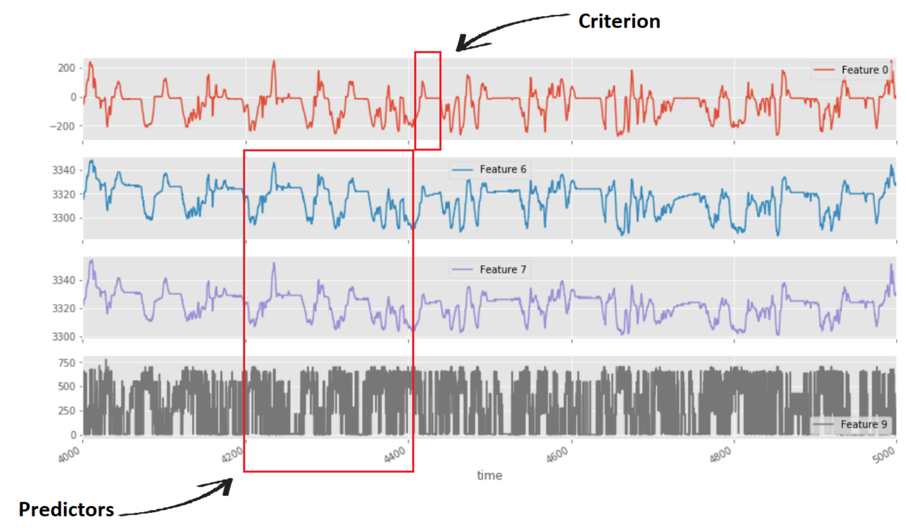


Figure 3: LASSO model

## Reconstruction-based anomaly detection: Autoencoder

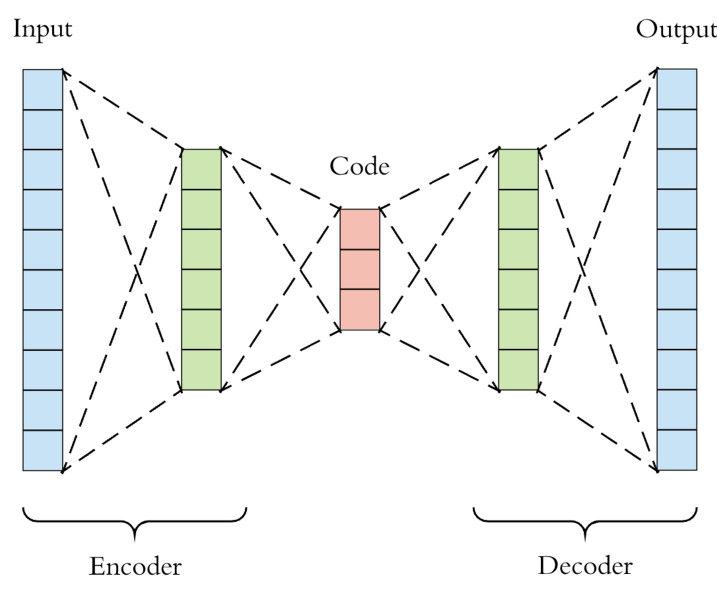


Figure 4: Autoencoder architecture

## Multi-Scale Convolutional Recurrent Encoder-Decoder

Recently proposed model uses correlation matrices tensor as input to the autoencoder [1]. The mentioned matrices are computed over subsequences of varying length $\delta$ on multivariate time series. En element $m_t$ of such a matrix $M^t$ is calculated with:

$$m_{ij}^{t} = \frac{\sum_{\delta=0}^{w} x_i^{t-\delta} x_j^{t-\delta}}{\kappa}$$

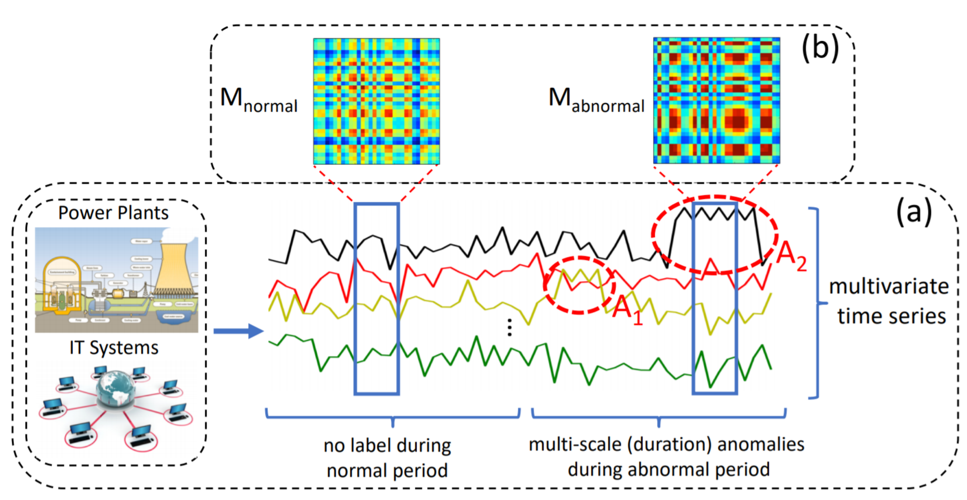


Figure 5: MSCRED

Proposed architecture employs convolutional encoder to encode the spatial information and model the temporal information via an attention based ConvLSTM. Square loss was utilized to perform end-to-end learning.

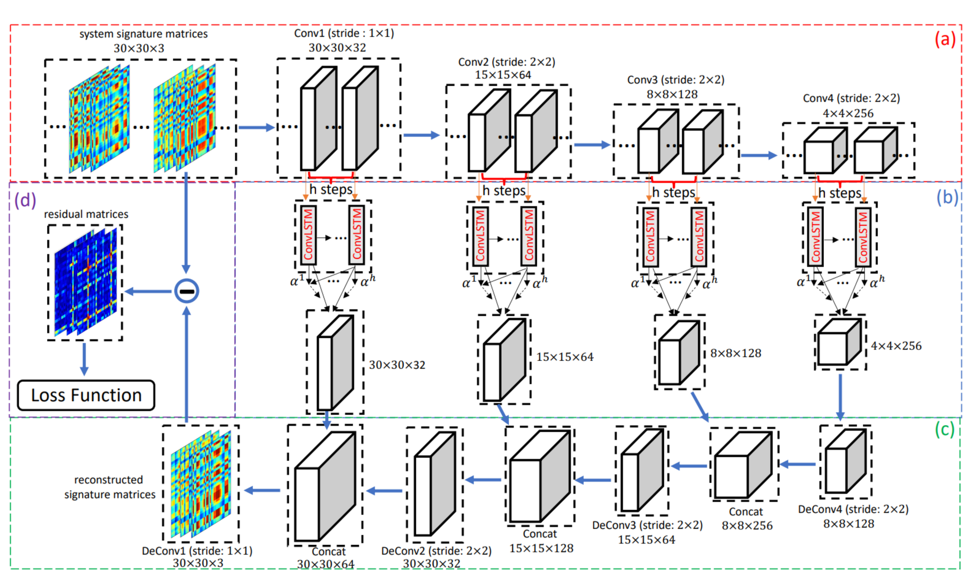

Figure 6: MSCRED architecture

## SR-CNN method

Researchers from Microsoft proposed recently a novel algorithm based on Spectral Residual (SR) and Convolutional Neural Network (CNN) [2].

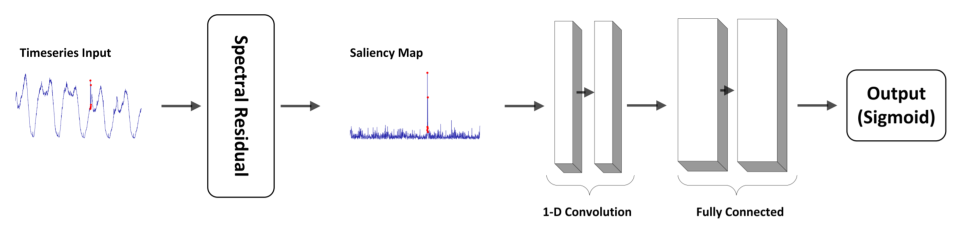


Figure 7: SR-CNN architecture

Spectral Residual method (based on FFT) used frequently in computer vision tasks for saliency detection was adopted:

$$\begin{aligned}
A(f) &= Amplitude(\mathfrak{F}(\mathbf{x}))\\
P(f) &= Phrase(\mathfrak{F}(\mathbf{x}))\\
L(f) &= log(A(f))\\
AL(f) &= h_q(f) \cdot L(f)\\
R(f) &= L(f) - AL(f)\\
S(\mathbf{x}) &= \left\|\mathfrak{F}^{-1}(exp(R(f) + iP(f)))\right\|
\end{aligned}$$

Figure 8: Saliency map calculation

## Spectral Residual step by step

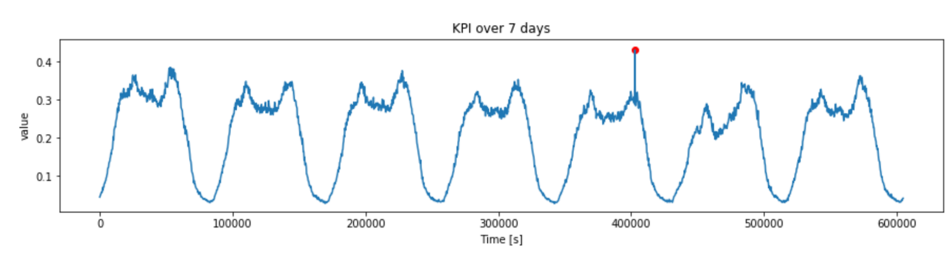

Figure 9: Time series input

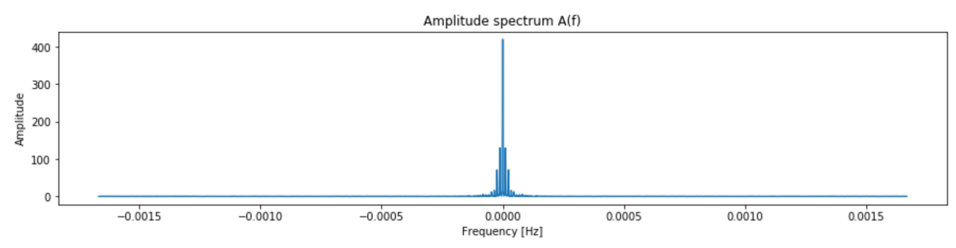

Figure 10: Amplitude spectrum

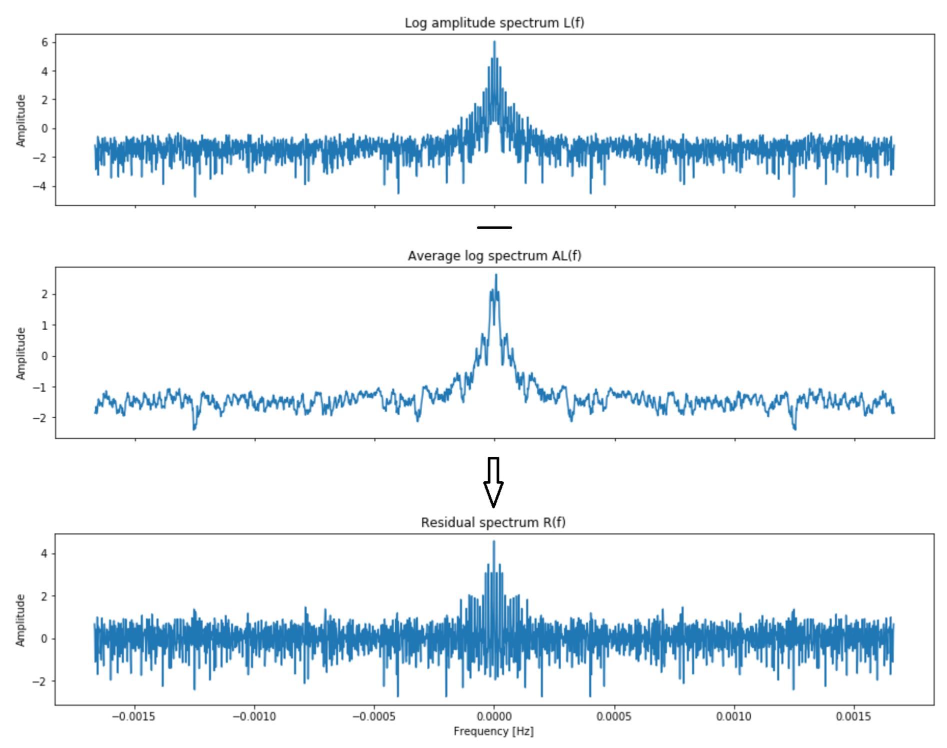

Figure 11: Calculation of residuals

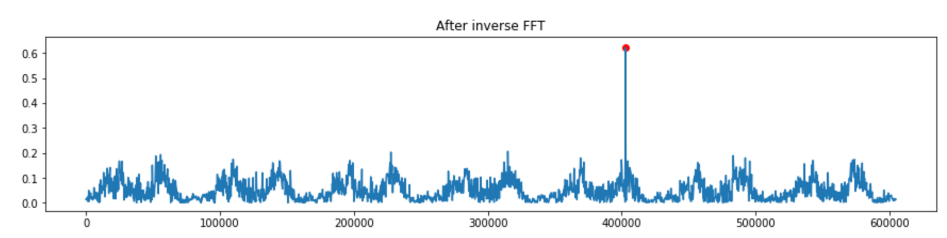

Figure 12: Inverse FFT

## References


[1] Zhang C., Song D., Chen Y., Feng X., Lumezanu C., Cheng W., Ni J., Zong B., Chen H., Chawla N. A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data 2018, Nov 20

[2] Ren H., Xu B., Wang Y., Yi C., Huang C., Kou X. Xing T., Yang M., Tong J., Zhang Q. Time-Series Anomaly Detection Service at Microsof 2019 Jun 10

[3] Aggarwal C. Outlier Analysis Second Edition, Springer 2017



PUT
Poznan University of Technology

GHOST
Group of Horribly Optimistic STatisticians

PSNC
Poznan Supercomputing and Networking Center