TSrepr R package: Time Series Representations

Peter Laurinec¹

1 Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava

Summary

TSrepr (Laurinec 2018) is an R package for time series representations computing. Time series representations are, in other words, methods for dimensionality reduction, feature extraction or for the preprocessing of time series. They are used for (Esling and Agon 2012):

- Significant reduction of the time series dimensionality,
- Emphasis of fundamental (essential) shape characteristics,
- Implicit noise handling,
- Dimension reduction will reduce the memory requirements and computational complexity of consequent machine learning methods (e.g., classification or clustering).

The TSrepr package contains various methods and types of time series representations including the Piecewise Aggregate Approximation (PAA), the Discrete Fourier Transform (DFT), the Perceptually Important Points (PIP), the Symbolic Aggregate approXimation (SAX), the Piecewise Linear Approximation (PLA) and Clipping. Except for these well-known methods, additional methods suitable for seasonal time series are implemented. These methods are based on the model, for example multiple linear regression, robust regression, generalised additive model or triple exponential smoothing (Laurinec and Lucká 2016, Laurinec et al. (2016)). Own developed feature extraction methods from the Clipping representation are also implemented - FeaClip and FeaTrend. In Figure 1, the comparison of all eight available model-based representations in the TSrepr on electricity consumption time series from the randomly picked residential consumer is shown.

Figure 1. The comparison of model-based time series representations on electricity consumption time series. The length of representations is 48, the same as frequency of the daily season of the used time series.

Additional useful functions and methods related to time series representations were also implemented. The TSrepr package includes functions for normalisations and denormalisations of time series - z-score and min-max methods. It supports the simple computation...
of the windowing method, a matrix of representations and forecasting accuracy measures (MAE, RMSE, MAPE, sMAPE etc.). Methods (functions) were implemented in base R and also in C++ for fast computations. In R, C++ programmes can be written thanks to the package Rcpp (Eddelbuettel and Francois 2011).

So far, no general package for time series representations computations has been created. The CRAN’s time series task view proves the previous statement. Packages TSMining (Fan 2015) and jmotif (Senin 2016) both includes implementations of PAA and SAX time series representations methods. However, these packages are mainly focused on motif discovery in time series.

References


