

Time Series in HPC: An exploratory study with distributed time series data for energy use cases

Soner Steiner^a, Rosana de Oliveira Gomes^b, Gregor Beyerle^b, Dimitar Dimitrov^b,
Jan Zabloudil^{a,c}

^aVSC Research Center / EuroCC Austria, TU Wien, Austria

^bHAKOM Time Series, Vienna, Austria

^cBOKU, Vienna, Austria

This project resulted from a collaboration initiated from EuroCC1 between TU-Wien, University of Natural Resources and Life Sciences (BOKU) and HAKOM Time Series. These three parties submitted together to the call "HPC Exploratory Projects" at the Austrian research promotion agency (FFG). It has been successfully granted in 2022 and started in November of the same year.

The HAKOM Time Series Manager (TSM) technology is designed to meet the operational needs of all players in the energy industry. It offers users an automated data management tool, able to perform classical tasks such as creation, updating and aggregation of time series data.

The goal of this project is to deploy, benchmark, optimize the TSM framework on the Vienna Scientific Cluster and evaluate highly computationally demanding operations. The high computational processing requirements come both from a large data volume and also algorithmic calculations.

Dynamic time warping (DTW) in the field of time series analysis basically means a measure for time series with different time lags. See for example two time series graphically depicted in Fig. 1. An everyday example of two time series with different time lags would be two people walking next to each other with a different velocity. In such cases DTW is the better measure of choice, because it can handle vectors with different lengths. In Euclidian distance or cosine distance metric the vectors need to be of the same size.

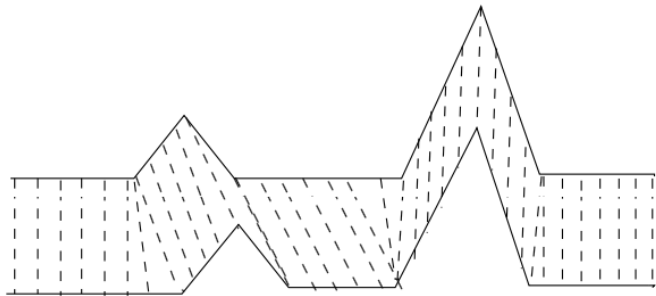


Fig. 1: Depicting DTW visually, taken from [1].

In this work, we apply DTW as a metric to classify time series into different groups. , e.g. load profiles of households, of small businesses like hairdresser, and many more.

The data set comes from a synthetic source, namely the *Last Profile* dataset [2], which simulates the profile of households measured from smart meters. The data set contains time series classified into 27 different profiles. The data volume is enhanced by adding white noise to such profiles. We finally test our algorithm by classifying the generated profiles and evaluating its performance against the original data.

References

[1] https://en.wikipedia.org/wiki/Dynamic_time_warping.

[2] <https://www.apcs.at/de/clearing/technisches-clearing/lastprofile>.