Karlsruher Institut für Technologie

Institut für Programmstrukturen und Datenorganisation (IPD)
Lehrstuhl für Systeme der Informationsverwaltung, Prof. Böhm

Bachelor/Master Thesis

ENERGY STATUS DATA
COLLECT • ANALYSE • UTILISE

Efficient Reduction of Energy Time Series

In the near future, **wind** and **photovoltaic** systems will have the largest share in electricity production in Europe. Electricity production based on these technologies is weather dependent. In other words, it has a high spatial variability and temporal fluctuation, which is a challenge for energy systems and markets. It also is necessary to estimate power plant capacity requirements accurately and to analyze the impact of fluctuant electricity production via techno-economic models.

Long-term system optimization models are useful to analyze political, technical, and economic questions regarding the development of an energy system. For example, it is interesting to forecast the impact of future energy investments on energy market landscape. However, modelling the intermittent character of different sources increases the model complexity, as the temporal resolution increases. This leads to unreasonable execution times, ranging from several days to months.

Consequently, new challenges are arising in the reduction of the execution time. One solution to bring down execution times is to reduce energy time series without destroying their energy-related characteristics. The reduced time series are then used in system optimization models. The expectations is that this leads to acceptable runtimes without degrading the quality of the predictions significantly.

The focus of this thesis is on the development of efficient data-reduction methods for large multivariate energy time series (e.g., one full year of data) to smaller ones without destroying their energy-related characteristics.

Several data-reduction methods exist already (see references), however, they are prone to the so-called curse of dimensionality or are applicable to univariate time series only. In this thesis, the focus is on clustering methods for multivariate energy data time series. The following questions are of particular interest:

- How can we assess the quality of a time series reduction algorithm? (internal evaluation)
- How can we quantify the impact of reduction on energy-specific optimization models? (external evaluation)
- Which methods can cope with the curse of dimensionality?

This results in the following tasks:

- Review of existing approaches for time series reduction.
- Design and implementation of new time series reduction algorithms.
- Implementation of metrics to quantify the reduction quality.
- Experimental evaluation on real world energy production and demand data.

In this thesis, you are working on latest research questions and acquire an extensive knowledge in data clustering and time series. You train highly demanded skills in the development and evaluation of data-mining algorithms. If you are interested, you will be integrated in an interdisciplinary research team, with background in computer science, energy informatics and electrical engineering. Your results will contribute to the development of better energy systems and therefore to a better societal use of energy resources. Knowledge from a lecture such as “Big Data Analytics” is beneficial. Elementary statistical knowledge, programming skills and the ability to accomplish conceptual work are desired. The scope of this work can be adapted, such that it would be suitable for either a Bachelor or Master thesis.

Contacts: Edouard Fouché (IPD) edouard.fouche@kit.edu
Yilmaz Hasan Ümitcan (IIP) hasan.yilmaz@kit.edu

[1] ElNozahy et al. (2013). A probabilistic load modelling approach using clustering algorithms
[2] Nahmmacher et al. (2016). Carpe diem: A novel approach to select representative days for long-term power system modeling
[3] Poncelet et al. (2016). Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems.