

## Anomaly detection in time series data

*Wednesday, October 11, 2017 5:00 PM (30 minutes)*

In the era of ubiquitous data, especially coming from sensors of all kind providing time-dependent data, applying analytics methods to characterize the value in the observed processes is a promising, but also challenging task. Not just the sheer amount of data, but also the integration and verification of data at hand needs to be handled prior to any analysis of time series data. As a general purpose set up, we present an infrastructure called DataHeap to handle sensor data, both for storage and further analysis. From the users perspective interfaces

to access those data for offline or online analysis (stream processing) needs to be efficient and straight forward to gain fast insights in available data. Especially in the analytics domain, frameworks, e.g. project from the Hadoop ecosystem, have emerged providing analytics methods applicable to large scale data. This demonstrator represents a workflow for collecting and evaluating sensor data. In principle, this workflow can be deployed on a general purpose cloud-based infrastructure. Using this infrastructure, we will show exemplary an analysis using energy monitoring data from an HPC-system. Here, the interest lies in anomaly detection that could indicate problems in the cooling system of the HPC-machine. Furthermore, we show the applicability of the system to other domains or areas of interest by actual examples from our research partners.

### Track

BDAHM

**Author:** Dr JÄCKEL, Rene (TU-Dresden)

**Presenter:** Mr FRENZEL, Jan (TU Dresden)

**Session Classification:** Analytics