Feature-based visualization of big time series data

Souhaib Ben Taieb*
Monash University, Melbourne, Australia - souhaib.bentaieb@monash.edu

Earo Wang
Monash University, Melbourne, Australia - earo.wang@monash.edu
Rob J. Hyndman
Monash University, Melbourne, Australia - rob.hyndman@monash.edu
Di Cook
Monash University, Melbourne, Australia - dicook@monash.edu

We are in the era of big data where massive amounts of data are continuously produced and stored every day. Time series data represent an important fraction of the various types of data collected in many areas of scientific, industrial and economic activities. Data visualization is essential to uncover structures and relationships between series, as well as to identify unusual behaviors. However, traditional time series visualization methods are not appropriate when dealing with such a large collection of time series. For example, smart electricity meters are currently deployed in millions of households to collect detailed individual electricity consumption data. We wish to visualize all these time series, and identify households that are behaving unusually. I will present an approach to this problem based on feature extraction on each time series and dimensionality reduction of the feature matrix. The features measure characteristics of the series, and may include lag correlation, strength of seasonality or other features based on prior knowledge. The dimensionarlity reduction allows us to visualize and explore a lower dimensional space.

Keywords: time series; data visualization; dimension reduction; feature extraction.