

Abstract

The rapidly increasing growth and complexity of telecommunication networks, the exponential growth of data bring the need of more complex data analytic systems. These analytics help to optimize the network, increase customer experience, support operation and network planning and detecting problems in the network. Finding sudden and unexpected changes in the telco network is especially important for Network and Service Operations Centers, as they could lead to more significant failures and degradations in the network.

We can think of this enormous amount of data as a large-dimensional, constantly changing space. Whenever we use our cellphone, numerous measurements are made and a plenty of data is collected. The problem of this thesis can be thought as finding abnormal behavior in these millions of streamed time series as soon as possible, label the new data points as normal or anomaly. We do not know the distribution of the data, but we assume daily and weekly profiles in the time series. In this thesis we bring a solution that helps to highlight these unwanted anomalous events that otherwise might remain hidden in the massive volume of data. In our new solution we try to capture the normal behavior of the data (which probably change over time) and compare the new data points to a time dependent threshold. We capture seasonality and growth in the time series, and continuously learn and update the learned behavior and so the threshold. We decompose the time series and learn the normal behavior of each component.