Thesis Proposal

MACHINE LEARNING DEPARTMENT

Reconstructing and Mining Signals: Algorithms and Applications

Speaker

Hyun Ah Song

Thesis committee

Christos Faloutsos (Chair)
Roni Rosenfeld
Nicholas D. Sidiropoulos (University of Virginia)
Vladimir Zadorozhny (University of Pittsburgh)

Jan. 16, 2020
10:00 AM
GHC 6501

Abstract

Given two different brain-scan modalities, like fMRI and MEG, how can we combine them, to achieve better resolution in both space and time, than each? Given power grid measurements (voltage, current), how can we spot patterns, anomalies and do forecasts? We answer these types of questions in the two parts of the proposal.

In the first part of the thesis proposal, we discuss our work on signal reconstruction. We live in a world that is flooded by streaming signals that come from different sources in different resolutions which often account for the same information. One example is historical records on patient counts; one TV source may report the patient counts in weekly intervals, while a newspaper source may report in monthly intervals. Another example is the brain activity signals; while fMRI has high spatial resolution and low temporal resolution, it is vice versa for MEG/EEG. We are interested in constructing higher resolution signals that complement those signals from different sources or modes in different resolutions. We discuss our work on signal reconstruction in two applications: historical epidemiological data and brain data. 1) Historical epidemiological data: We introduce constraints such as smoothness and periodicity, utilizing the well-known epidemiological model 'susceptible-infected- susceptible (SIS)'. Experimental results on the Tycho dataset show that the best algorithm reduces the reconstruction error by 42% compared to the naive approach. 2) Brain data: We discuss our proposed approaches that employ various assumptions such as sparsity, low rank, or smoothness and show that reconstructed brain signal displays richer information of fMRI and MEG, interpolating in time and space in a principled way.

In the second part of the thesis proposal, we discuss our work on signal mining. Raw signals can contain unnoticeable hidden information that is not observable in their raw forms. One example is power grid signals (voltages, currents); signals in its raw form do not lead to straightforward interpretation. Another example is aircraft sensor signals; sensor signals are result of complex physical models which do not provide straight-forward interpretation. If we want to understand the data, we should explore it deeper (if you mind it, mine it!), preferably with domain knowledge. We are interested in mining signals that can provide us with better interpretation of the data, and aid us with various data mining tasks such as forecasting, anomaly detection, etc. We discuss our signal mining work on two different application domains: power grid data and aircraft data. 1) Power grid data: We introduce our works that incorporate a physics model, the BIG model, to better interpret the data, and utilize tensor factorization and Holt-Winters to model the data for anomaly detection and forecasting. Experimental results on CMU and Lawrence Berkeley National Laboratory (LBNL) power grid dataset demonstrate 32% and 27% error reduction in forecasting compared to the latest algorithm. Also we show that proposed algorithm successfully detects anomalous events. 2) Aircraft data: We discuss our proposed method on analyzing aircraft sensor signals for detecting anomalous events using coupled tensor factorization.