Studienarbeit

Lehrstuhl: ITM Beigl, KIT

Titel: Big Data Predictive Analytics Tool on Smart Meters

Betreuer: Yong Ding

Einleitung:

Power load predictive analytics can have significant effects on power system planning and operation, in particular short-term load forecasting (STLF). STLF refers to load forecasts of power system loads with lead times ranging from a few minutes to seven days ahead. The aim of STLF is to predict future power consumption based on historical consumption data and other exogenous variables, in order to make the best use of electric energy and relax the conflict between supply and demand.

Within the scope of Smart Grid development, the introduction of smart metering systems has been rapidly conducted worldwide for providing more frequent information about consumption and prices, in order to improve the demand side management. For instance, the 3rd package of Single European Market commits all EU member states to equip at least 80% of the power consumers with Smart Meters. The ultimate goal of smart metering is to allow utility firms to forecast energy usage, to improve their performance on the settlement markets – where money can be lost through inaccurate predictions – and to match supply and demand more closely. However, the challenge of Smart Meter data is twofold. First is the enormous volume of the 15-minute reads and secondly, finding meaning in the content by combining the interval data with other influencing factors (e.g. climate and seasonal factors, social activities, etc.) to create actionable intelligence based on dynamic load patterns. Massive amounts of Smart Meter data and exogenous information that contain many terabytes and possibly many petabytes of data, require specialized tools and know-how to analyze.

Thema:

Topic of this Studienarbeit is to develop an analytics tool for Smart Meter data. The proposed analytics tool focuses on the STLF with different lead times at different aggregation levels. The processing chain of the tool consists of two functional modules, i.e. data preprocessing and load forecasting. In order to speed up the computation due to large Smart Meter datasets, GPU implementation of the entire process (incl. preprocessing and Machine Learning algorithms) in CUDA will be compared with the state-of-the-art load forecasting tools.

Aufgabe:

* Einarbeitung in die folgenden Themen
  + Parallel Computing
  + Datenvorverarbeitung
  + STLF mit statistischem Verfahren und Maschinellem Lernen (ML)
* Design: GUI und globale Plattform für die Entwicklung des Analytics-Tools
  + Konfigurierbare Parametrisierung für beide Datenvorverarbeitung und Lastprognose
* Funktionsmodul 1: ***Datenvorverarbeitung (Preprocessing)***
  + Daten-Cleanup: missing data, outliers, etc.
  + Wavelet-Transformation: wavelet selection
* Funktionsmodul 2: ***Lastprognose (Forecasting)***
  + Umsetzung der folgenden ML-Algorithmen in CUDA
    - Artificial Neural Networks (ANN)
    - SVR (Support Vector Regression)
    - RF (Random Forest)
  + Umsetzung des ARMAX-Algorithmus in CUDA
* Evaluierung: Performance-Test des Tools
  + Validierung mit freiverfügbaren Open-Source-Datensätzen aus anderen Lastprognose-Tools (e.g. Weka, LIBSVM, etc.)
  + Case Study: Smart Meter Daten aus dem EU NOBEL Projekt
    - Definition der Performance-Metrik
    - Vergleich mit den state-of-the-art Lastprognose Tools