Iterative Data Analysis for Sensing Applications

Ella Peltonen

Department of Computer Science
University of Helsinki

Smartphones and sensors can generate large amounts of data in a short time. Near real-time modeling of such data requires rich processing capabilities not provided by previous event-based systems. Complex machine learning algorithms are required to understand the rich context information, especially from mobile devices, and iterative learning is needed to cope with the rapidly changing context.

Sensing Solutions

- Sensors collect a complex set of data
- Machine learning gives a deep insight to the sensor environment
- Near real-time processing provides a tight feedback loop
- Live model can be used to give personalized, actionable recommendations

Distributed Big Data

Distributed environment to store, manage and process collaboratively collected data

- Distributed machine learning and context modeling
- Complex, but efficient algorithms
- Optimizations to support machine learning algorithms

Iterative Learning

Iterative, incremental, and fast cycle to

- Collect newest data from mobile devices
- Send data models and feedback from the analysis system

Context Factors

Mobile devices operate in a versatile context:

- System state, e.g. system settings (screen brightness, network status) and subsystem variables (CPU and memory load)
- Sensor data, e.g. accelometer, microphone, temperature, etc.
- Applications running and their context: services, functionality, permissions, price etc

Expected PhD Contributions

- Design principles for distributed, iterative data analytics algorithms...
- ...and their integration to previous real-time platforms for...
- ...real-time collaborative context modeling and machine learning

Example Case: Carat

Carat is a research project and a mobile application that gives energy recommendations nowadays over 750,000 Android and iOS users from circa 200 countries.

- The dataset has been used to analyze energy benefit of mobile applications [2] and system settings [1]
- Future work to combine all the context factors to one recommendation engine of iterative learning cycles
- Also published the Carat Context Factor Dataset: http://carat.cs.helsinki.fi/research

[1] Peltonen, Lagerspetz, Nurmi, and Tarkoma. Energy Modeling of System Settings: A Crowdsourced Approach. IEEE PerCom 2015.
[2] Oliner, Iyer, Stoica, Lagerspetz, and Tarkoma. Carat: Collaborative Energy Diagnosis for Mobile Devices. ACM SenSys 2013.

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI
MATEMAATTIS-LUONNONTIETEELLINEN TIEDEKUNTA
MATEMATISK-NATURVETENSKAPLIGA FAKULTETEN
FACULTY OF SCIENCE