Realtime IoT Stream Processing and Large-scale Data Analytics for Smart City Applications

Ralf Tönjes, Muhammad Intizar Ali, Payam Barnaghi, Sorin Ganea, Frieder Ganz, Manfred Hauswirth, Brigitte Kjærgaard, Daniel Kümpel, Alessandra Mileo, Septimiu Nechifor, Amit Seth, Vlasios Tsiatsis and Lasse Vestergaard

**Project Objectives**

- Semantic modelling and knowledge representation of IoT and social data streams
- IoT and relevant social data stream processing
- Query processing and stream discovery
- Aggregation, summarisation and abstraction of streamed data
- Knowledge extraction and events processing on distributed data streams
- Resource virtualisation for IoT data/service framework
- Quality, trust and privacy-aware data stream provisioning with monitoring and test support
- Data mash-up and self-configuration

CityPulse will develop, build and test a distributed framework for the semantic discovery and processing of large-scale real-time IoT and relevant social data streams for knowledge extraction in a city environment.

**Concept**

- Large-Scale Data Analysis
- Applications
- Virtualisation
- Knowledge Based
- Reliability Testing

**Virtualisation**

- Heterogeneous data sources
- Overcome silo architectures and provide common abstract interface
- Assigning semantic annotations to data streams

**Federation**

- Sensorfusion
- Combines heterogeneous data streams to one unified view

**Aggregation**

- Datafusion
- Reduce amount of data
  - Clustering
  - Summarisation
  - Filtering
  - Pattern recognition

**Smart Adaptation**

- Higher level information processing
  - Interpretation of semantic data
  - Transforming lower level dynamic information to higher level abstractions
  - Enables adaptation of the data processing pipeline

**User centric decision support**

- Goal: provide optimal configuration of smart city applications
- Social and context analysis
  - Matchmaking and discovery mechanisms
  - Match data according to users preferences and context

**Reliable Information Processing**

- Challenge: Dynamic environments, changes and prone to errors
- Reliable data processing requires accuracy and trust
- Cope with malfunctions, disappearing sensors, conflicting data, ... monitoring of streams (runtime) testing of applications (design-time)

**Smart City Applications**

- Cities challenge:
  - Rapidly growing digital economy requires new applications and information systems
- Provide an API for faster prototyping and access to CityPulse framework and information

**Processing steps during Lifecycle**

- Three consecutive layers: federation of heterogeneous data streams, large-scale IoT stream processing and real-time information processing
- Knowledge-based methods, reliability monitoring and testing to achieve reliability
- Provide solutions for different life-cycle stages (design-time, run-time, testing)

**Contact and Partners**

Duration: September 2013 - August 2016
Website: http://ict-citypulse.eu
Contact: Ralf Tönjes (r.toenjes@hs-osnabrueck.de)