nano-tera.ch
swiss scientific initiative in health / security / environment systems

OpenSense2 RTD 2013

FNSNF

Infrastructure for Crowdsourcing Environmental Monitoring

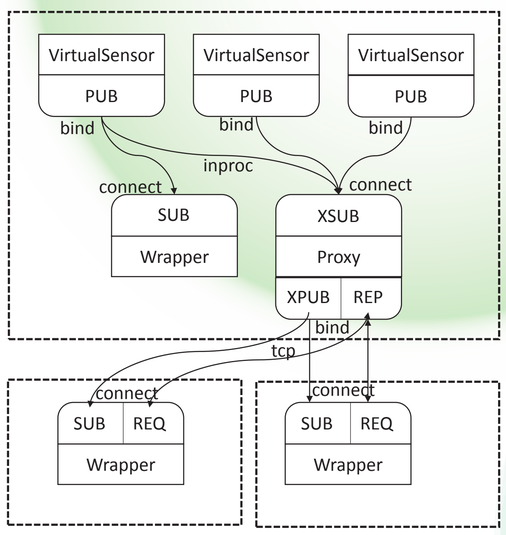
Julien Eberle, Jean-Paul Calbimonte, Karl Aberer
EPFL, Switzerland

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

GSN Global Sensor Networks

Server-side
Distributed and scalable infrastructure

GSN (Global Sensor Networks) [1]
- Extensible streaming middleware
- Support for semantic annotations
- Online modeling capabilities [2]
  - Queries over models or data streams
- Distributed processing through zeroMQ messaging
  - Fast and transparent communication between virtual sensors



queries

aggregated data

raw data + meta data

Client-side
Users' own devices for data collection

Leveraging user's own devices as part of the "infrastructure"
Rich sensors environment (internal + external + virtual)

Competing for system resources like CPU, memory, battery, ...
... but also user's attention
The goal is to minimize the load on the device and the user.

Two approaches :

Adaptively enable sensors
Driven by application need and sensor availability
Collaborate with other devices

Continuously sense everything
Richer dataset, training data
Performing data-mining offline

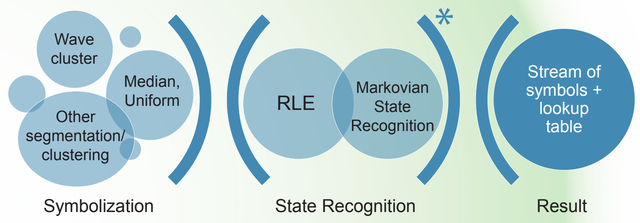
local queries

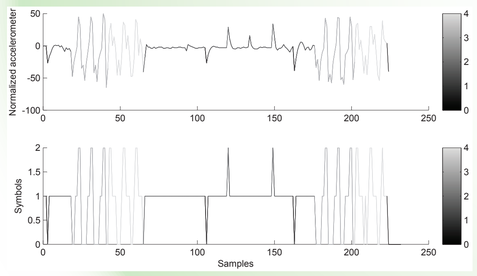
raw data

Local processing
Efficient on-device data processing

To reduce communication, processing and storage load, using the smartness of the sensors or sensing device
The idea: aggregating the data into meaningful symbols and perform the usual processing and machine learning tasks on them. [3]
Several levels of symbolic representation can be used (abstraction level)

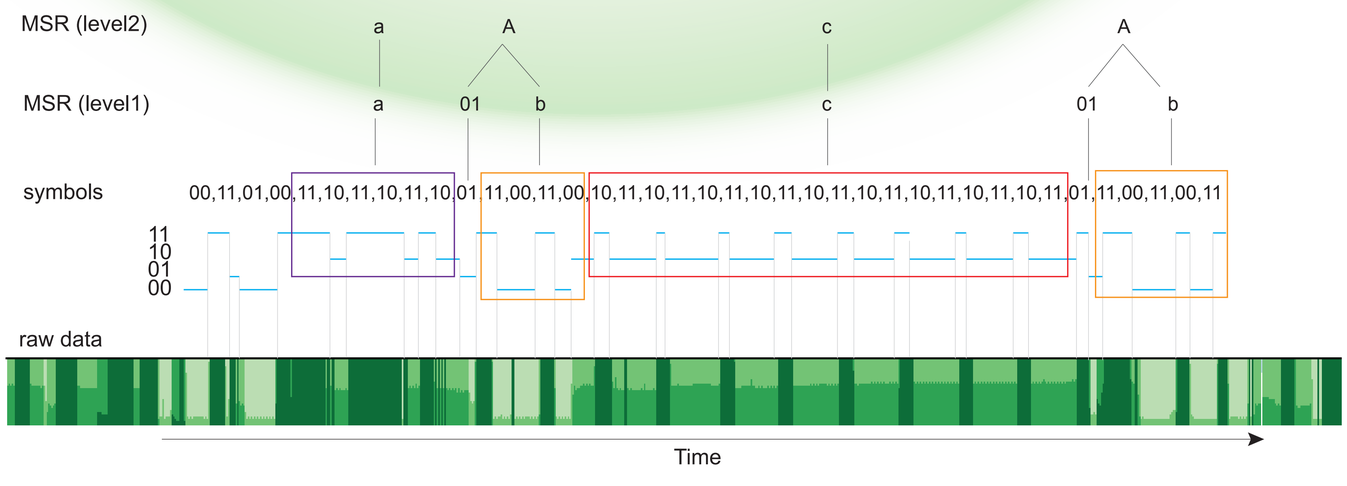
Constraints:
- Unsupervised learning
- Online processing
- Limited memory use and processing power





Symbols can be mapped to semantic meaning (not always)
Level-0: taken from global data distribution
- Using SAX, clustering, quantiles, expert knowledge,…
- Suitable for data-mining

Level-i: patterns composed of level-(i-1) symbols
- Using online state recognition
- Maintain the recent history in a transition matrix
- Predictability to infer state changes



OpenSense Deployments
And other sources of information

Mobile sensors on roof of trams and buses in Zürich, resp. Lausanne, collecting air quality information and providing meta information: location, sensor type, ...

Official federal references stations (NABEL)

Different interpolation models including land use information. [4]

Expert knowledge and regulatory limit values for setting thresholds and color scales.

References

[1] GSN, Global Sensor Networks, available on Github : http://gsn.epfl.ch, https://github.com/lsir/gsn
[2] Erol Can Un, Julien Eberle, Yongsung Kim, Karl Aberer. A model-based back-end for air quality data management. In Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication (pp. 1143-1150), ACM, 2013.
[3] Tri Kurniawan Wijaya, Julien Eberle, Karl Aberer. Symbolic representation of smart meter data. In Proceedings of the Joint EDBT/ICDT 2013 Workshops (pp. 242-248), ACM, 2012.
[4] Jason Jingshi Li, Arnaud Jutzeler, Boi Faltings. Estimating Urban Ultrafine Particle Distributions with Gaussian Process Models. In Research@Locate'14, Canberra, Australia, 07-09 April 2014, published at http://ceur-ws.org