ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

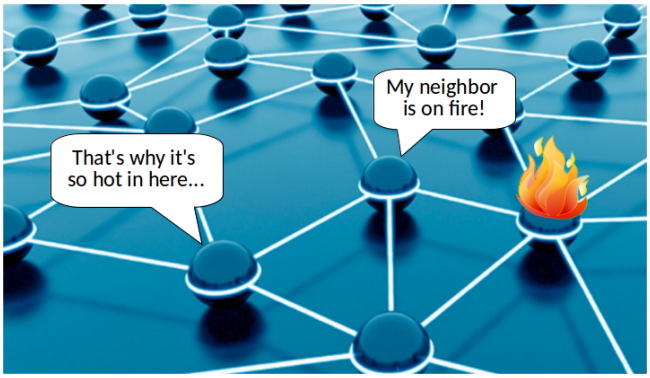
TIK
Institut für Technische Informatik
und Kommunikationsnetze
Computer Engineering and
Networks Laboratory

Prof. L. Thiele

Semester Thesis:

Anomaly Detection in Mobile Sensor Networks for Environmental Monitoring

Sensor networks for monitoring environmental conditions, such as air pollution, have received increasing interest in the last few years. Large scale sensor deployments form a valuable source of data for a large variety of applications, *e.g.* creating highly resoluted air pollution maps. Maintaining high data accuracy is a challenging task in such sensor networks. One aspect that can greatly decrease data quality is the presence of faulty or anomalous sensors, *i.e.* sensors that do not work ordinary and provide false measurements. Thus, it is important to detect anomalous sensors as fast and precise as possible.



The reasons why a sensor might not perform ordinary are many-fold: *(i)* The sensor's sensitivity has drifted over time due to ageing effects and, thus, needs recalibration. *(ii)* Depending on the context the sensor is reporting measurements that are of no use and do not coincide with measurements of other sensors. For instance in a crowd-sensing deployment a user places its sensor inside a backpack although it is intended to measure ambient temperature. *(iii)* The opposite of the previous reason, *i.e.* a sensor measures a situation that does not coincide with other measurements but is in fact of large interest. For instance, pollutant measurements next to a industrial plant are considerably higher compared to other areas in a city, but are in fact of great use for detecting highly polluted areas. *(iv)* The sensor has actually a technical defect and needs to be replaced.

Task: In this thesis we want to create a tool that detects extraordinary performing sensors and uncovers the underlying reason. The tool should base its decision on various features, for instance statistical assumptions or the context of a sensor node. The performance of the approach can be evaluated using the dataset from the OpenSense project (www.opensense.ethz.ch). Various sensor installed on top of ten trams in the city of Zürich have collected a wide range of measurements, such as, air pollution (O3, CO, NO2 and PM), GPS coordinates, and environmental conditions (humidity and temperature) over a period of over three years. Overall the dataset consists of over 100 Mio., possibly anomalous, samples.

Requirements: Fundamental mathematical knowledge. Experience with MATLAB or python might be beneficial, but is not required.

Interested? Please have a look at http://www.tec.ethz.ch/research.html and contact us for more details!

Contacts

- Balz Maag: balz.maag@tik.ee.ethz.ch, ETZ G75
- Zimu Zhou: zimu.zhou@tik.ee.ethz.ch, ETZ G85