Enabling Large-Scale Urban Air Quality Monitoring with Mobile Sensor Nodes

David Hasenfratz
Air Pollution Monitoring: Current Approach

- Static networks operated by national authorities
- Highly reliable and accurate measurements
- High acquisition and maintenance costs

Today’s pollution maps have limited spatial resolution
Air Pollution Monitoring: Our Approach

- Mobile sensor nodes with low-cost sensors
- Lower accuracy but much higher spatial coverage

Enables pollution maps with unprecedented spatial and temporal resolution facilitating novel applications
Challenges of Mobile Air Quality Monitoring

Mobile measurement system

Data quality assurance

High-resolution air quality models and its applicability
Thesis Contributions

Chapter 2
- MS’12
- AAAI’12

Mobile measurement system

Chapters 3 & 4
- EWSN’12
- DCOSS’13
- IPSN’15

Data quality assurance

Chapters 5 & 6
- PerCom’14 (best paper award)
- PMC’15

High-resolution air quality models and its applicability
Thesis Contributions

Chapter 2
- MS’12
- AAAI’12

Mobile measurement system

Chapter 3 & 4
- EWSN’12
- DCOSS’13
- IPSN’15

Data quality assurance

Chapter 5 & 6
- PerCom’14 (best paper award)
- PMC’15

High-resolution air quality models and its applicability
Mobile Measurement System

Air quality sensor node equipped with UFP, CO, O$_3$, and NO$_2$ sensors

- Based on existing hardware and software infrastructure
- Air quality sensor nodes monitor main pollutants
- Measurements sent in real-time to backend server
Streetcar Measurement Network

- Ten sensor nodes are installed on top of streetcars

Collected largest data set of spatially resolved air pollution measurements with over 100 million data points
Mobile Sensor Networks: Related Work

- Mobile air pollution measurements
  - Standard instruments, e.g.,
    - Hudda et al., Emissions from an international airport increase particle number concentrations 4-fold at 10 km downwind, Environmental Science & Technology, 2014
  - Low-cost gas sensor node prototypes, e.g.,
    - Dutta et al., Common sense: participatory urban sensing using a network of handheld air quality monitors, SenSys, 2009
    - Zwack et al., Characterizing local traffic contributions to particulate air pollution in street canyons using mobile monitoring techniques, Atmospheric Environment, 2011
    - Elen et al., The aeroflex: a bicycle for mobile air quality measurements, Sensors, 2013
- Sensor networks based on existing infrastructure, e.g.,
  - Chen et al., B-Planner: Night bus route planning using large-scale taxi GPS traces, PerCom, 2013
  - Shang et al., Inferring gas consumption and pollution emission of vehicles throughout a city, SIGKDD, 2014
Thesis Contributions

Chapter 2
- MS’12
- AAAI’12

Mobile measurement system

Chapter 3 & 4
- EWSN’12
- DCOSS’13
- IPSN’15

Data quality assurance

Chapters 5 & 6
- PerCom’14 (best paper award)
- PMC’15

High-resolution air quality models and its applicability
Calibration of Mobile Sensor Networks

- Use meeting points between low-cost sensors and reference stations to calibrate sensor network

Multi-hop calibration:

- Calibrated reference sensor (e.g., static station)
- Uncalibrated low-cost sensors (e.g., mobile sensors)
Calibration of Mobile Sensor Networks

- Use meeting points between low-cost sensors and reference stations to calibrate sensor network

Multi-hop calibration:

Calibrated reference sensor (e.g., static station) → Uncalibrated low-cost sensors (e.g., mobile sensors)
Calibration of Mobile Sensor Networks

- Use meeting points between low-cost sensors and reference stations to calibrate sensor network

Multi-hop calibration:

Calibrated reference sensor (e.g., static station)

Uncalibrated low-cost sensors (e.g., mobile sensors)
Calibration of Mobile Sensor Networks

- Use meeting points between low-cost sensors and reference stations to calibrate sensor network

Multi-hop calibration:

- Calibrated reference sensor (e.g., static station)
- Uncalibrated low-cost sensors (e.g., mobile sensors)

⇒ Main problem: Error accumulation in the network
Standard Calibration Approach

Use Ordinary Least Squares (OLS) to compute calibration coefficients (intercept $\alpha$ and slope $\beta$):

- Minimizes sum of squared vertical residuals $AB$
- Optimizes the root-mean-square error (RMSE)

Bias of Ordinary Least Squares (OLS)

Ordinary Least Squares regression has problem of regression dilution

- Calibration has bias towards zero if uncalibrated sensor measurements have noise

\[ \beta_{OLS} = \frac{\text{cov}(u, v)}{\text{var}(v) + \text{var}(e_v)} \]

- Single-hop calibration: regression dilution is not a problem
- Multi-hop calibration: regression dilution leads to high error accumulation in the network
Error Accumulation of OLS

Calibration of sensor z is strongly biased towards zero:

\[
\beta_{OLS} = \frac{\text{cov}(\text{ref}, a) \cdot \ldots \cdot \text{cov}(y, z)}{\text{var}(a) + \text{var}(e_a)) \cdot \ldots \cdot (\text{var}(z) + \text{var}(e_z))}
\]

Large error accumulation over number of calibration hops
Calibration with Low Error Propagation

Use Geometric Mean Regression (GMR) to compute calibration coefficients (intercept $\alpha$ and slope $\beta$):

- Minimizes the sum of the areas of triangles $ABC$
- No bias towards zero
No Error Accumulation with GMR

- Slope of sensor $z$ is independent of the noise and variance of sensors $a...z$:

$$\beta_{GMR} = \sqrt{\frac{\text{var}(\text{ref}) + \text{var}(e_{\text{ref}})}{\text{var}(z) + \text{var}(e_{z})}}$$

Calibration does not depend on the calibration path
What is the Assumption?

The calibration of sensor $z$ is independent of sensor $y$ if $\text{var}(m(x \leftrightarrow y)) = \text{var}(m(y \leftrightarrow z))$

$m(x \leftrightarrow y)$: Set of measurements of sensor $y$ at meeting points with sensor $x$

$m(y \leftrightarrow z)$: Set of measurements of sensor $y$ at meeting points with sensor $z$

In reality this assumption usually does not hold, but simulations and real sensor measurements show:

- GMR accumulates much less error even if the above assumption does not hold
Real Air Pollution Measurements

Calibration of ozone measurements from the mobile sensor network

Proposed algorithm is suitable for the calibration of large, mobile, and heterogeneous sensor networks
Real Air Pollution Measurements

Calibration of ozone measurements from the mobile sensor network

Proposed algorithm is suitable for the calibration of large, mobile, and heterogeneous sensor networks

- Assess quality of measurements and outlier detection by integrating **generic models** for the phenomena monitored and the **sensors** used

![Graph showing calibration error (RMSE) vs. number of nodes in calibration path (hop count).](image)
Thesis Contributions

Chapter 2
- MS’12
- AAAI’12

Mobile measurement system

Chapter 3 & 4
- EWSN’12
- DCOSS’13
- IPSN’15

Data quality assurance

Chapters 5 & 6
- PerCom’14 (best paper award)
- PMC’15

High-resolution air quality models and its applicability
From Single Measurements to High Resolution Pollution Maps

Processing steps:

- Single measurements
- Data filtering
- Data validation
- Land-use regression modeling
- Pollution maps
Developing Land-Use Regression Air Quality Models

- Land-use regression (LUR) models are widely used to assess spatial variation of air pollutants
- Pollution levels are modeled based on land-use and traffic characteristics (explanatory variables)
  1. Evaluate dependency between explanatory variables and monitored pollution levels
  2. Use found dependencies to model pollution levels at locations without measurements but land-use data
Increasing the Temporal Resolution

Challenging due to reduced number of measurements
Increasing the Temporal Resolution

Challenging due to reduced number of measurements

- Additional data from a database with historical measurements enhances the original data set

Our approach enables constructing pollution maps with good accuracy and high spatial and temporal resolution
Ultrafine Particle Pollution Maps for Zurich

- Spatial resolution: 100m x 100m
- Temporal resolution: seasonal (winter, spring, summer, fall)

Winter (Jan.—Mar.)
Spring (Apr.—June)
Summer (July—Sept.)
Fall (Oct.—Dec.)

Particle concentration \[ \text{particles/cm}^3 \]
The Health-Optimal Route Planner

High-resolution pollution maps make new useful applications possible and raise the citizens’ awareness.
Conclusions

- Built a mobile monitoring system and collected today’s largest spatially resolved air pollution data set

- Developed algorithms to calibrate mobile sensor networks and to assess data quality

- Derived pollution maps with an unprecedented spatio-temporal resolution and showed their utility

These findings can serve as stepping stone towards accurate, detailed, and real-time pollution assessment