Research Questions Prompted by the CitiSense Air Quality Monitoring System

William G. Griswold
Computer Science & Engineering
Air Quality? Didn’t we solve that in the 70’s?

89% ↓ ½ 29

50% ↑ 30% 4M
So how is it affecting you? No idea.
CitiSense: A Computing Approach
CitiSense Components

Nima Nikzad

Piero Zappi

Nakul Verma

Celal Ziftci
Pollution levels vary widely by locale
Individual exposures vary widely compared to reported EPA “AQI”
Same Road, Different Exposure

Those who are doing the most to clean the air are paying a price

Poor and youth more affected

Highlights potential of participatory systems like CitiSense
System-Level Research Problems

How can we affordably collect high-resolution, high-accuracy (air quality) data, at scale, to guide public policy?

• Preliminary answers from CitiSense:
  – Commodity sensors (connected to commodity mobile devices)
  – Personal mobile; also fixed, drone…
  – Machine learning

• Encountered two unanticipated research problems
  – Power management
  – Data quality
Mobile Power Management

- **Techniques**
  - duty cycling
  - batching
- **Tradeoffs**
  - availability
  - resolution
  - timeliness
  - code modularity

- Idea 1: Programming support for delaying expensive operations until cheap (piggybacking)
  - E.g., don’t use WiFi until another service turns it on
  - New operation, “WaitUntil” in front of key network calls, etc.
  - Runtime monitoring of system service states

- Idea 2: Govern by quality-of-service decls on objects
  - Data-flow analysis finds object paths through WaitUntil’s
  - At run-time, log the time a tracked object spends at WaitU
  - When object’s specified delay is used up, WaitUntil terminates immediately

- Incremental, but not really modular power management

Octav Chipara, Marjan Radi, and Farley Lai (U. Iowa)
Data Quality: The Dirty Secret

- Calibration
- Cross-sensitivity
- Air flow
- Rate of change
- Warm-up time
- Context
- …?

- Idea: *crowd-enabled field calibration*
  - Field data naturally varied by context
  - But less controlled

- More machine learning
  - use redundancy of readings among sensors and over time to identify readings not predicted by ML, hence “out of bounds”

- Context capture, fed into ML
- In early modelling phase
Promise, in progress

Mass commodity sensing of our environment shows potential

• We have the elements of an architecture

• Challenges remain, e.g.
  – Power management
  – Data quality

More at http://citisense.ucsd.edu