N-SMARTS: Networked Suite of Mobile Atmospheric Real-time Sensors

RJ Honicky (honicky@cs.berkeley.edu)
UC Berkeley Intel Research

1 Towards a societal scale sensor network

While industry analysts predict that cell phones will become the "next PC," we believe that the cell phone has the power to become something much more than a scaled down, connected IO and processing device. In addition to these standard PC traits, a cell phone is situated in an environment, mobile, and typically co-located with a user. These traits make the cell-phone ideally suited to track and understand the impact that the environment has on individuals, communities, cities and on a global scale, as well as understanding how humans effect their environment.

By attaching sensors to GPS-enabled cell phones, we can gather the raw data necessary to begin understand how, for example, urban air pollution impacts both individuals and communities. While integrating a sensor into a phone and transmitting the data that it gathers to a database is not very difficult, doing so at low cost, on a societal scale, with millions of phones of phones providing a data from hundreds of networks spread throughout the world makes the problem much more tricky.

On top of the systems challenges, understanding the raw data gathered from a network of cell phone-attached sensors presents significant challenges as well. Cell phone users are mobile, are unlikely to ever explicitly calibrate their sensors, typically put their phone in their pocket or handbag (thus obstructing the sensor from airflow), spend significant time indoors or in cars, and typically charge their phone at most once per day, often much less frequently. Even if users did calibrate their sensors, the very low-cost sensors we intend to use drift over time and environmental conditions anyway. Without knowing the location of a sensing event, automatically calibrating the sensors in the phone, detecting the environment of the phone, and intelligently managing power (by sampling at the right times) the data gathered by the phones will be next to useless.

Thus the N-SMARTS project focuses on:

- Developing a platform to understand the real-world challenges of sensing on a mobile phone, and to provide other researchers, both within and outside of computer science, with a platform for their own experiments. *(What do the sensor data look like? What are people's movement patterns? How do people's behaviors impact the data? How can the impact of those behaviors be minimized by platform design?)*

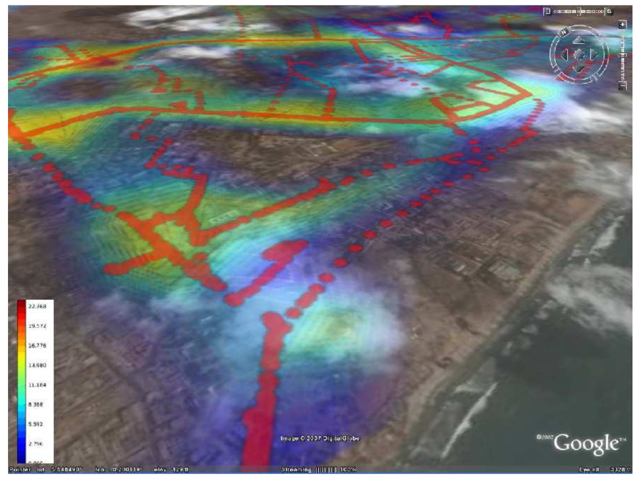

Figure 1: CO data collected from sensors in taxi cabs in Accra, Ghana on March 21st, 2007, overlaid on aerial photography using Google Earth.

- Building a system architecture that can scale to millions of phones *(What are the system bottlenecks? How can communication costs be minimized? How much computation should occur on the phone?)*
- Designing algorithms to scalably provide accurate estimates of pollution levels and other sensed data *(How can accuracy be increased by super-sampling? How can the phones be automatically calibrated to one another, or other sensors in the environment? How can those inferences be parallelized?)*
- Designing algorithms to detect and account for the user's behaviors *(Can we accurately detect when the phone is a user's pocket or purse, when the user is in a car, indoors, outdoors, etc.? Can we correct our readings? Can we accurately label data with the user's context, so that we can answer questions like "What is the median exposure to CO for bicycle commuters on Shattuck Avenue?")*
- Assembling and building a suite of useful sensors to integrate.

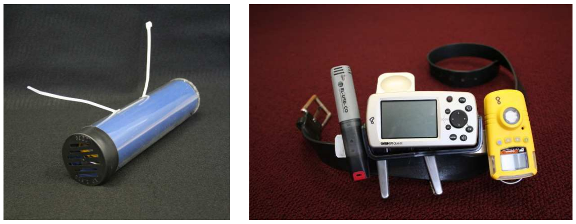

Figure 2: The automotive and personal version of the data-logging sensor platform

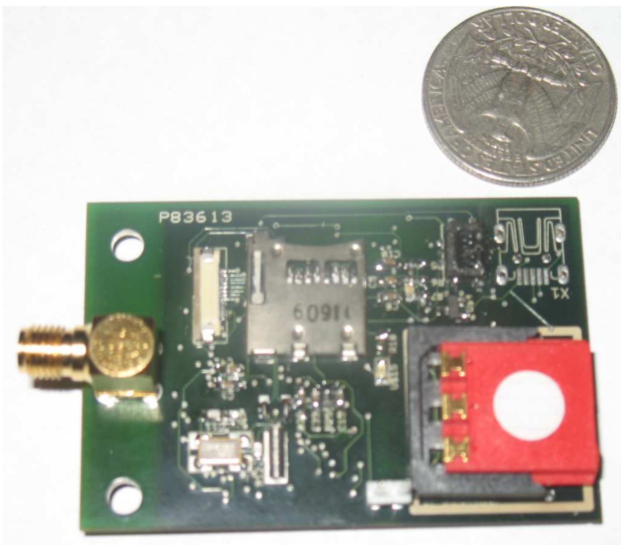

Figure 3: A prototype of the Bluetooth board with a CO sensor, a NOx/CO dual sensor, a temperature sensor and an accelerometer. This board will integrate directly with the phone.

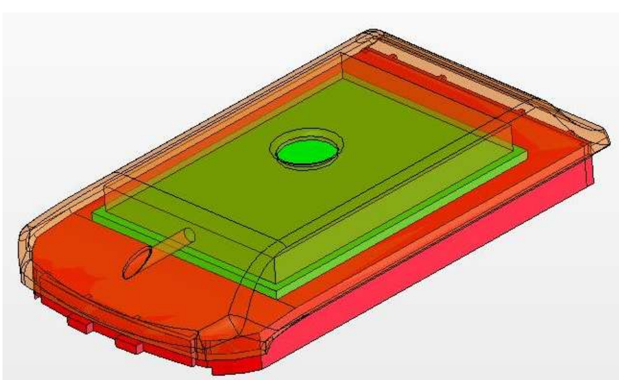

Figure 4: The battery of a LG VX9800 with a PCB mounted on top and covered by a new enclosure (outline shown for the new enclosure only).

2 NSMARTS platforms

Because we don't have a clear picture of what the pollution data and movement patters of users will be, we need to gather data up front, before we have an integrated sensor/phone platform available. For that reason, we have put together a portable sensor platform which can be carried around, allowing a person to gather data that is roughly similar to the data which will be gathered by the integrated platform. The data acquisition platform will allow us to develop and test the algorithms that make up the core of the N-SMART platform.

The data acquisition platform consists of off-the-shelf pollution sensors, and a GPS. Each unit contains

- A Lascar EL-USB-CO Carbon Monoxide data logger
- A Garmin Qwest GPS (with external antenna)
- A NO2, SO2 or O3 datalogger from BW Technologies

All of the devices log data, and their clocks are synchronized so that the data from each device can be correlated. See our web page for details on the sensors (see Section 7).

There are two version of this kit: a "automotive platform" which can be mounted near a car window or externally, and a "personal platform" which can be worn on a user's belt (see Figure 2).

We deployed six automotive platforms on taxis and four personal platforms on students in Accra, Ghana, West Africa, for two weeks in March, 2007. These data were uploaded into a database and can be viewed in a variety of formats, including an overlay on Google Earth (see Figure 1). The database will be publicly available soon.

We are also developing an integrated platform which will more closely approximate a phone manufactured with sensors integrated directly into the phone itself. This model will allow significant cost reduction with respect to less complete integrations. Rather than actually manufacturing a new phone and enclosure, however, we simply replace the battery pack of the phone with a module that clips in to the battery well of the phone, and contains both a battery and the sensor module (see Figure 4).

The current version of the integrated platform has:

- CO and NOx sensors
- A temperature sensor for calibration
- An accelerometer for activity inferencing
- A Bluetooth radio for communication with the phone

We chose to use Bluetooth to communicate with the phone to avoid mechanical problems with a direct serial link, and to make the design and software more generic.

3 Exploiting mobility

One of the main advantages N-SMARTS has over many other sensor networks is that the sensors are mobile, and are co-located with people. Not only does this mean that we will tend to have data in locations that are relevant to people, but it also means that sensors will tend to have spacio-temporal density similar to people. We can take advantage of this density in at least two interesting ways.

First, as the density of sensors at a given location increases, we can increase our precision by super-sampling, and averaging. For sensors with Gaussian noise (which our CO sensors exhibit) sampling in the same location, we expect the variance of the signal to be $\frac{C}{n}$ if we average the signals from n sensors with noise variance C. In Figure 5, we a experiment with six sensors in a chamber in which we can control the concentration of CO. In this case, we stepped the concentration of CO by 0.2ppm increments over an hour, and observed the response of the sensors. The light dots show the response of one sensor, and the dark dots show the averaged response of six sensors. Clearly the noise variance has decreased. Figure 6 show the variance of the signal versus the number of sensors averaged. The empirical results match the theoretical results closely!

In the real world, of course, sensors will not be located exactly in the same place at the same time. Theoretical results on the learning curves of Gaussian processes suggest, however, that if sensors are in proximity to one another signals from the sensors can be "averaged" using a Gaussian process, with the increase in precision related to both the number of sensors and their density in space-time.

Another interesting way in which we can exploit mobility is by calibrating the sensors when they are in close proximity to one another, or to an accurate reference sensor situated in the environment. A Gaussian process model can easily accommodate this type of inference. In Figure 7, we can see a signal sampled by two different sensors, one represented by circles, the other by triangles. Each sensor has a bias with respect to the signal, and noisy samples are taken from the signal using a Gaussian noise function. The sensors never sample in the same location, but the samples in the middle of the graph are in close proximity to one another. If we smooth the samples without considering the bias of each sensor, then we get significant error in locations that are not sampled by both sensors. If, however, model the bias, then the smoothed function closely tracks the underlying value.

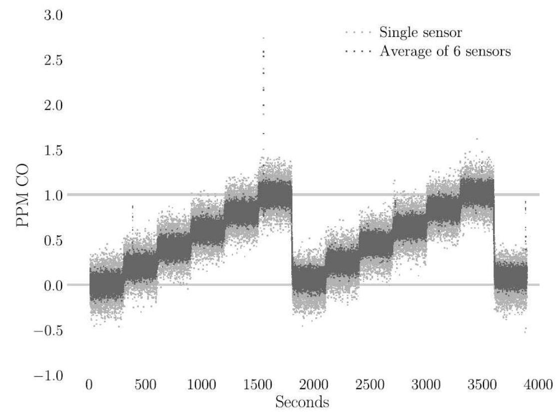

Figure 5: Sensor readings of concentration of CO(ppm) vs. time. Concentration is changed in 0.2ppm steps by a mass flow controller. The lighter dots represent readings from a single sensor. The darker dots represent the average of six sensors.

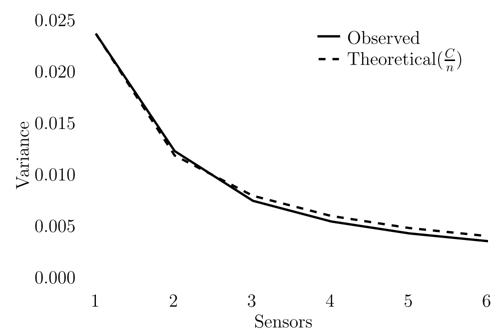

Figure 6: Variance of the average signal from a set of sensors (in ppm) vs. the number of sensors in the set. $\frac{C}{n}$ is show for reference.

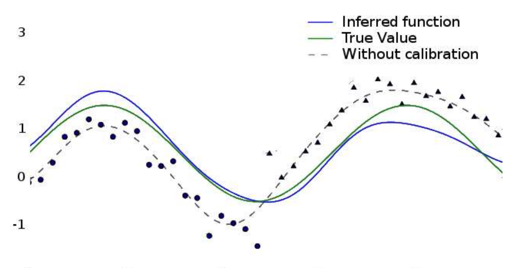

Figure 7: A simple example in which two sensors (represented by circles and triangles) each have some bias. Although the readings from each sensor are taken in different places, our model can infer the bias from each sensor and correct for it.

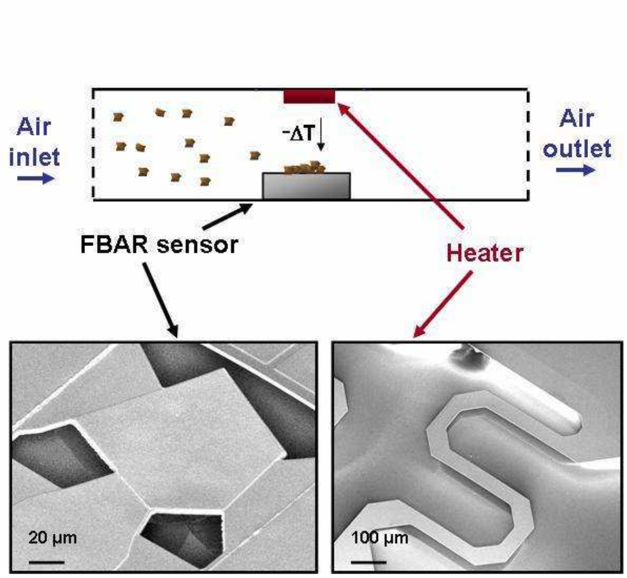

Figure 8: PM2.5 particles are deposited on a resonating FBAR via thermal phoresis, changing the resonance frequency of the FBAR.

4 Advanced Sensing

Since aerosol pollutant's constitute a major public health concern, causing an estimated 65,000 deaths annually in the United States, the N-SMARTS project is working to integrate a MEMS PM2.5 mass sensor developed at LBNL and Berkeley into our design. This design uses thermophoretic precipitation of particles onto the mass monitor, a thin-film bulk acoustic wave resonator (FBAR) / Pierce oscillator (See Figure 8. Our version uses an inertial impaction filter to select aerosol pollutants less than 2.5 microns in size.

This work was featured independently as part of the BSAC Advisory Board Meeting.

5 Participants and collaborators

N-SMARTS Platform

- Professor Eric Brewer - *Berkeley CS*
- Professor John Canny - *Berkeley CS*
- Professor Ronald C. Cohen - *Berkeley, Chemistry*
- R.J. Honicky - *Berkeley CS*
- Dr. John Huggins - *CEO BSAC*
- Dr. Eric Paulos - *Intel Research Berkeley*
- Professor Albert Pisano - *Berkeley ME*
- Dr. Ian Smith - *Intel Research Seattle*
- Paul Wooldridge - *Berkeley, Chemistry*

Enclosure Design

- Professor Paul Wright - *Berkeley ME*
- Kyle Yeates - *Berkeley ME*

Advanced Sensing

- Dr. Justin Black - *Berkeley EE*
- Alex Elium - *Berkeley EE*
- Professor Albert Pisano - *Berkeley ME*
- Professor Richard White - *Berkeley EE*

6 Future Directions

We've left security entirely out of this discussion, and that will be a serious concern, since users will be divulging private information about their location and context. We have some ideas about how to ensure users' privacy while still transmitting useful information to the database(s). We are seeking collaborators in this area!

Another area which we intend to persue is plume detection and warning. Users could be warned of a plume nearby and guided away from it. This has some very interesting public safety and emergency and disaster response applications.

7 Resources and further information

1. The N-SMARTS home page:
   http://www.cs.berkeley.edu/~honicky/nsmarts

2. Participatory Urbanism at Intel Research:
   http://www.urban-atmospheres.net/ParticipatoryUrbanism/index.html

3. A tech report on the algorithms for automatic calibration (UC Berkeley EECS Tech Report EECS-2007-34):
   http://www.eecs.berkeley.edu/Pubs/TechRpts/2007/EECS-2007-34.pdf

4. Justin Black's dissertation on FBAR-based particle measurement (UC Berkeley EECS Tech Report EECS-2006-193):
   http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-193.pdf