Design and Evaluation of a Metropolitan Air Pollution Sensing System

Ke Hu, Vijay Sivaraman, *Member, IEEE*, Blanca Gallego Luxan, and Ashfaqur Rahman, *Senior Member, IEEE*

***Abstract*— Urban air pollution is believed to be a major contributor to premature deaths and chronic illnesses worldwide. Current systems for urban air pollution monitoring rely on static sites with low spatial resolution, and moreover, lack the means to estimate exposures for (potentially mobile) individuals in order to make medical inferences. This paper describes the design and evaluation of a low-cost participatory sensing system called HazeWatch that uses a combination of portable mobile sensor units, smart-phones, cloud computing, and mobile apps to measure, model, and personalize air pollution information for individuals. Our contributions are three-fold: 1) we architect, prototype, and compare multiple hardware devices and software applications for collecting urban air pollution data with high spatial density in real-time; 2) we develop web-based tools and mobile apps for the visualization and estimation of air pollution exposure customized to individuals; and 3) we conduct field trials to validate our system and demonstrate that it yields much more accurate exposure estimates than current systems. We believe our system can increase user engagement in exposure management, and better inform medical studies linking air pollution with human health.**

***Index Terms*— Wireless sensor networks, mobile applications, air pollution, participatory sensing.**

I. Introduction

ONE OF the basic requirements of human health and well-being is clean air. However, the World Health Organization (WHO) estimates that around 1.4 billion urban residents worldwide are living in areas with air pollution above recommended air quality guidelines [1], and reports that air pollution kills about 7 million people a year [2]. Chronic exposure to air pollution increases the risk of cardiovascular and respiratory mortality and morbidity [3], while acute short-term inhalation of pollutants can induce changes in lung function and the cardiovascular system exacerbating existing conditions such as asthma, and ischemic heart disease [4], [5]. Monitoring and controlling air pollution is high on the public consciousness in both developing and developed countries.

Manuscript received August 14, 2015; revised October 29, 2015; accepted November 4, 2015. Date of publication November 10, 2015; date of current version February 4, 2016. This is an expanded paper from the IEEE SENSORS 2013 Conference. The associate editor coordinating the review of this paper and approving it for publication was Dr. M. R. Yuce.

K. Hu and V. Sivaraman are with the University of New South Wales, NSW 2052, Australia (e-mail: z3265076@student.unsw.edu.au; vijay@unsw.edu.au).

B. G. Luxan is with the Centre for Health Informatics, Australian Institute for Health Innovation, University of New South Wales, NSW 2052, Australia (e-mail: blanca.gallegoluxan@mq.edu.au).

A. Rahman is with the Commonwealth Scientific and Industrial Research Organisation, Computational Informatics, Hobart Tasmania 7005, Australia (e-mail: ashfaqur.rahman@csiro.au).

Digital Object Identifier 10.1109/JSEN.2015.2499308

Several governments operate air quality monitoring stations and publish the data [6]. These stations are generally outfitted with several high-quality monitoring devices that can measure a wide range of air pollutants (such as CO, NO_x, SO_2, ozone, particulate matter, etc.). However, the high costs of installing and maintaining these sites limits their number – for example, the greater Sydney area in Australia has approximately 15 active monitoring sites, separated from each other by tens of kilometres. The low spatial sampling resolution necessitates the use of mathematical models to estimate pollutant concentrations over vast sections of the metropolis, which can be both complex (requiring inputs such as land topography, meteorological variables and chemical compositions) and inaccurate (e.g. due to highly variable meteorological conditions [7]), leading to incorrect inferences [8], [9].

Current epidemiological studies rely on air pollution exposure data obtained from the home suburbs of their subject, implicitly assuming that the user is at home at all times. This can be inaccurate, as it does not account for mobility whereby the user spends time at home, at work, commuting, etc., at locations with very heterogeneous pollutant concentrations. Estimating personal inhalation intake is essential not only to inform risk assessment for epidemiological studies but also for the individuals to manage risk, both by retrospectively understanding the pollutant levels that affect their health, and in prospectively choosing commuting routes and timings that reduce their risk [10].

To address the above two concerns, we leverage new developments in portable sensor and communication technologies to develop a participatory sensing system – “HazeWatch”, which aims to crowd-source fine-grained spatial measurements of air pollution, and to engage users in managing their pollution exposure via personalized tools. Our specific contributions are: (1) We architect and prototype a low-cost system for users to contribute air pollution data. This includes design, prototyping, and comparison of multiple portable sensing units, coupled with mobile phone applications for data tagging/uploading and a cloud-based repository for hosting the data. (2) We show how the data can be analyzed and consumed by users. This includes appropriate models for interpolating the spatio-temporal data points, visualization of pollution over a geographical map of the area, and mobile apps that show personal exposure. (3) We validate and evaluate our system with a small number of users to show that it yields much more accurate estimates of personal exposure than existing systems based on coarse-grained data from static sensors, demonstrating the potential benefits that larger scale deployments can bring to our understanding of the relationship between pollution exposure and health.

The rest of this paper is organized as follows: §II describes prior efforts to build systems for urban air pollution monitoring. In §III we describe our system architecture and how the data is visualized and personalized. §IV describes our deployment experiences, and the paper concludes in §V.

II. Related Work

The idea of crowd-sourcing pollution data from users has been investigated by several projects around the world in the past few years. Among the first projects with this vision is the MESSAGE (Mobile Environmental Sensing System Across Grid Environments) project [11] from Cambridge University and partners in the UK, which aims to develop fixed and portable devices for high-density measurement of concentrations of carbon monoxide and nitrogen oxides in urban areas. They have very recently reported their development and deployment experience [12] in the Cambridge area, and demonstrated that the use of low-cost fixed and portable devices deployed in high densities can give a much more accurate picture of the spatial and temporal structure of air quality in the urban environment. The scale and scope of this project is commendable, and the contributions in building the devices, deploying them city-wide, and modelling the collected data are noteworthy; however, we believe that the portable devices still remain relatively expensive and bulky (at around 445 grams) for regular use by pedestrians/bicyclists, and personalized tools (e.g. mobile apps) for estimating and managing exposure remain under-explored.

Vanderbilt University, supported by Microsoft, embarked upon a similar project, called MAQUMON [13], that developed portable wireless sensor units for measuring ozone, nitrogen dioxide and carbon monoxide. Their units are autonomous, having on-board flash (for storage), GPS (for location) and GSM (for communication) capabilities, making them much more bulky and expensive compared to our design (as described later in this paper). They also developed innovative web-based visualization (e.g. contour-maps) and personalization (e.g. route planning) tools [13], making it more accessible for lay users. To the best of our knowledge, this project has not undertaken any long-term deployments. Intel has also been developing as part of the CommonSense project [14] a prototype that is a portable handheld device capable of measuring various air pollutants. This data can be uploaded in real time and viewed on Google Maps.

While several other projects, such as ExposureSense [15], have similar goals to ours, we would like to make particular note of the ongoing OpenSense project [16] at EPFL Switzerland that seems to have successfully deployed several air monitoring units on top of public buses. In spite of the replication of effort across these several projects, we believe they are all worthwhile efforts since they collectively explore different deployment scenarios (e.g. buses versus private cars) in different regions of the world. Another similar project EveryAware seventh framework programme (FP7) [17] at Torino Italy has just finished. They also developed an air pollution participatory sensing system including sensorbox, mobile apps and server. To the best of our knowledge, they focused on sensorbox design and calibration aspects, rather than system performance, and no field test has been implemented in their paper.

In addition to the above large-scale projects, several smaller efforts have looked at various individual aspects of the system. A participatory sensing system for air pollution monitoring and control called P-sense was developed in [18]. This paper discusses several challenges in large-scale deployment of sensors and applications, emphasizing aspects such as privacy and security. The authors in [19] design an indoor air quality monitoring system that uses Zigbee-based devices and base stations – this system has battery life limited to a few hours since the sensors use resistive heating. Two mobile platforms for real-time pollution monitoring were introduced in [20], with the aim of fusing data from portable devices with data from larger sensors to create a social air pollution network. In another project which is described in [21], the authors designed a wireless sensing system that gathers air pollution data from their sensor equipment and uploads data to the back-end server via mobile network. With their system, users only recorded and visualised their personal exposure, and no interpolation model was introduced into the system, which means users cannot share (and benefit from) data from others. Another shortcoming is that they don't show any field test results.

A mobile measurement unit was developed and tested in Bologna, Italy, and models such as Voronoi diagrams and ordinary kriging variograms were used to that estimate air pollution distribution. In [22] a tool is developed and trialled in Barcelona, Spain, for estimating personal exposure for mobile individuals with varying levels of activity. This is very much aligned with our objectives; however, they derive their pollution estimates from a model, the Atmospheric Dispersion Modelling System (ADSM), developed from a previous year, and their estimates are hence neither real-time nor accurate.

III. System Architecture

The data collection architecture in the HazeWatch project is based upon the idea of "crowd-sourcing" or "participatory sensing". Users collect and contribute air pollution data obtained from personal sensing units, and the greater spatial density of data thus obtained from many users in turn gives each user more accurate estimates of their pollution exposure. Our overall system architecture is shown in Fig. 1, and consists of (1) portable sensor units that monitor air pollution, (2) applications on the driver's mobile phone that harvests the data from the sensor unit, tags it with location and time information, and uploads it in real-time to our server, (3) the cloud-based server that stores the data, and applies interpolation models to generate spatio-temporal estimates, and (4) visualization tools that map pollution levels and personalize the information for the individual user. The first two steps constitute data collection, while the latter two steps comprise data consumption.

A. Pollution Measurement Node

We designed and built our own hardware platform for air pollution measurement, and compared it against sensor

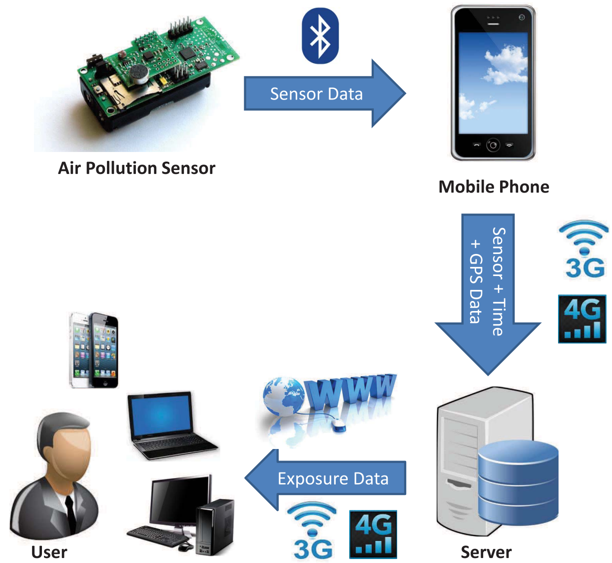


Fig. 1. HazeWatch System Architecture.

nodes that are starting to emerge in the market. We begin by describing our experiences with building the hardware (we call it the *HazeWatch node*), and then describe the features of comparable devices such as *Node* and *SensorDrone* sensors.

HazeWatch Node: We faced several challenges in the design and manufacture of our air pollution monitoring sensor node, and had to make several design decisions with a view towards maximizing chances of mass adoption. The challenges we had to overcome are briefly summarized below:

Portability: If the device is bulky, as the one used by the government monitoring stations – this condemns it to be fixed at a location, reducing spatial coverage. Therefore we decide the device must can be made portable enough for a user to carry on their person, as intended in [11] and [14].

Complexity: The next major decision we confronted was regarding target cost and complexity of the device. In order to operate autonomously, the device needs to have pollution sensors, a GPS module to time- and location-stamp the measurements, and a 3G/4G module to upload data in real-time. Indeed such a design was used for projects such as [23] and [24], and is suitable for mounting on public vehicles. However, in order to keep costs low, we chose a minimalist design that does not have GPS or 3G/4G capability. Instead, in our design the unit communicates via BlueTooth with the user's smart phone, which is assumed to be equipped with GPS for time and location tagging the pollution measurements, and with 3G/4G capability for uploading in real-time to our server. This offloading of capability to the mobile phone allows us to keep the unit cost low for the consumer market.

Sensor Type: The sensor unit therefore consists broadly of the (a) gas sensors, (b) micro-controller with built-in ADC to digitize the sensor readings and package them into messages, (c) BlueTooth module to transmit the readings to the user's mobile phone, and (d) battery power supply. The choice of gas sensors presented different trade-offs. For typical pollutant gases such as carbon monoxide (CO) and nitrogen dioxide (NO_2), our first version of the unit, shown in Fig. 2(a), used Metal Oxide Sensors (MOS) (Sensor model: CO-e2V MiCS5521, NO_2-e2V MiCS2710, O_3-e2V MiCS2610). These operate on the principle that when a semiconductor material is heated and when a gaseous pollutant is introduced into the chamber, electrons are freed from the semiconductor, which decreases its effective resistance proportional to the level of pollution. For instance, Fig. 3 shows a simple circuit that we use to convert heating voltage V_H to output resistance voltage V_S. R_L is a load resistor that converts the resistance R_S to output voltage V_S. Then V_S is converted to pollution concentration using the following equation:

$$Concentration = C_0 \times V_S^2 + C_1 \times V_S + C_2, \quad (1)$$

where C_0, C_1 and C_2 are calibration coefficients. MOS are compact and cheap (as low as $5 each), but have low accuracy and are non-linear. The use of MOS allowed us to built our unit housing three sensors (CO, NO_2 and O_3) at a cost price close to $50 (refer to [26] for a detailed description of the hardware design), but posed many performance problems related to non-linearity and influence of temperature and pressure. We therefore designed a second version of our unit (detailed in [27]), shown in Fig. 2(b), using electrochemical (EC) sensors (Sensor model: e2V EC4-500-CO). These operate by passing the pollutant gas through the inner membrane of a gas chamber where it is oxidized, producing an electric current proportional to the level of concentration. EC sensors are sensitive, accurate, and linear, but expensive ($50-100 each) and require more complex circuitry. We therefore designed out unit to house only one sensor at a time (the figure shows the CO unit), at an overall cost of about $150.

Node sensor: Concurrent to our development effort, we noted that commercial devices (funded by KickStarter) were starting to emerge that promised similar capabilities. One such device is the Node sensor as shown in Fig. 2(c). The Node sensor platform is designed with plug-in modules mode. It comprises body platform part and interchangeable OXA gas sensor header part. With changing the OXA headers, Carbon Monoxide (CO), Nitric Oxide (NO), Nitrogen Dioxide (NO_2) and other three pollutants can be monitored. Smart phones can connect to the body platform with Bluetooth 4.0 up to 250 feet away. It has to be calibrated in six months by mobile app. The cost of Node device is about $150 for body platform and $150 for one OAX header each.

SensorDrone sensor: Another sensor device we used is the SensorDrone which is shown in Fig. 2(d). There are more than 11 sensors in one SensorDrone device and it can measure various factors, e.g. CO, CO_2, pressure, and ambient temperature. We can connect mobile phone with SensorDrone via Bluetooth 2.1 or 4.0. The price is $200 for each SensorDrone platform.

B. Sensor Calibration

1) Initial Calibration Method: Once built, we needed to calibrate each unit, which entails converting the current measurements into pollutant concentrations. We designed our initial calibration method with reference to one on field calibration approach [28] and one common sensor calibration method [29], and this method requires us to determine, at each known concentration of the gas, the reading output by our unit.

(a) (b) (c) (d)

Fig. 2. Air pollution sensor: (a) Metal Oxide sensor; (b) Electrochemical Sensor; (c) Node Sensor; and (d) SensorDrone Sensor.

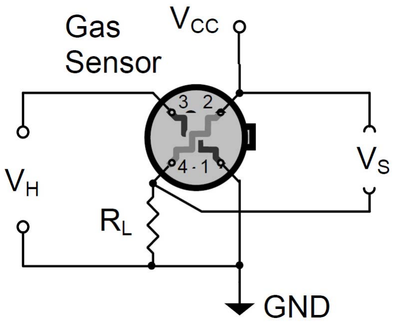


Fig. 3. Metal Oxide Sensor circuit [25].

This posed a major challenge for us since we did not have facilities for controlled experimentation with known concentrations of the gas. To address this problem, we procured a commercial monitor called the GasAlert Micro 5 (Sensor model: 4COSH-CiTiceL) built by BW Technologies, that could tell us the true pollutant concentrations. We then custom-built an air-tight container, as shown in Fig. 4(a), into which we put our unit (to be calibrated) along with the commercial monitor. Since we did not have a license to operate toxic gas cylinders, we had to resort to crude measures to obtain the pollutant gases. For CO we simply captured car exhaust fumes and dumped them inside the chamber, while for NO_2 we added copper shavings to a beaker of nitric acid inside the chamber that caused a chemical reaction in which the gas was released. Repeated experiments yielded varying concentrations of the pollutant gas, as indicated by the commercial monitor, and noting down the corresponding current from our unit allowed us to plot the current versus concentration curve for each unit, yielding the calibration coefficients.

To calibrate CO sensors, we first connect the calibration circuit and wait for circuit to stabilize (multimeter voltage reading will drop to approximately around 10mV that equates to 1-2ppm discrepancy, takes anywhere between 5-15mins). After that, we allow the circuit to read values at zero gas concentration. As the circuit has a reference voltage of 2.5V, this needs to be converted into the corresponding ppm and subtracted from within Matlab. While CO gas is transferred into the chamber through a plastic gasbag, the gas concentration will stabilize within 5mins, and gas will need to be let out periodically to be able to record a collection of value for calibration. When the concentration moderately drops by approximately 10ppm we note down the corresponding voltage, and collection of more than 15-20 data values is sufficient for analysis. Finally, when the concentration of gas reaches close to zero, we need to confirm that voltage readings from the multimeter have fallen below 20mV, and Repeat steps above to achieve multiple sets of data for calibration (all tests must be repeatable so that any discrepancies are noted). The concert equation between voltage and ppm values are shown below:

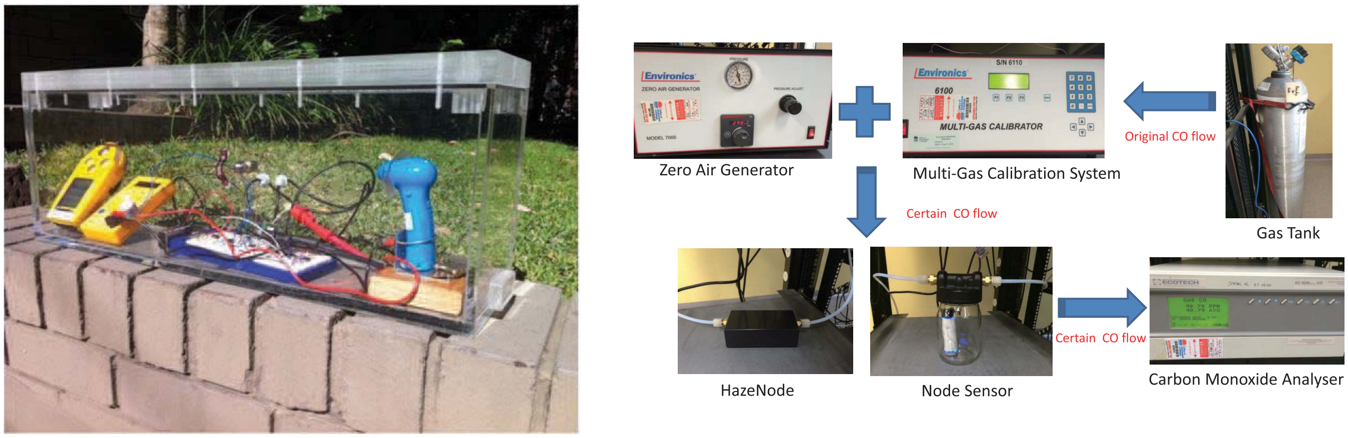
$$Concentration = \frac{1}{Sensitivity}\left(\frac{V_{out} - V_{ref}}{R_{gain}} - A\right), \quad (2)$$

where the unit of concentration is ppm; A is offset; and V_{ref} and R_{gain} is 2.5V and 100Kohms respectively. In spite of the relatively crude nature of our calibration, the curves we got for the electrochemical sensors were remarkably linear, giving us confidence in the calibration. These were further validated via field tests as described later. We refer the reader to our report [30] that outlines our calibration procedure and outcomes in great detail.

2) Improved Calibration Method: Although we successfully calibrated all the HazeNode sensor using the initial calibration method, we still had some concerns about the calibration accuracy. The use of an extra haze detection analyzer to calibrate the sensor is questionable, and uncontrolled and unknown production of pollutants by car exhausts sampling raises the calibration uncertainty. To address this, we partnered with the New South Wales Government Office of Environment and Heritage in Australia [31], and designed an improved calibration procedure. The whole calibration system has three parts as shown in Fig. 4(b). The first part is the gas generation part, which contains a gas tank, a multi-gas calibration system (Environics Series 6100) and a zero air generator (Environics Series 7000). The gas tank is loaded with a concentration of 500 ppm CO, and the multi-gas calibration system can suck the original CO flow from the tank with certain gas flow rate controlled by management software. The zero air generator can continuously deliver dry, contaminant-free air with fixed flow rate. We can adjust the gas flow rate from the multi-gas calibration system to get a certain CO concentration and feed it to the second part – the calibration part. For different sensors, we designed different containers to expose the sensor to the gas steadily as shown in the figure. Using Node sensor as an example, for different concentration rate, we use the following formula to calibrate the sensor:

$$PPM = \frac{Raw\_Reading - Base\_Line}{0.37736 \times (Gain \times Ratio)} \times 10^9 \times Rate, \quad (3)$$

where Base_ Line denotes the calibration baseline, and Rate represents the calibration factor, while Gain and Ratio are constants and the values are 35000 and 39 respectively.



(a) (b)

Fig. 4. (a) Initial calibration chamber setup and (b) improved calibration system.

(a) (b) (c)

Fig. 5. (a) HazeWatch node on top of a car. (b) Node sensor inside a car. (c) SensorDrone on a bike.

Because we can acquire raw readings from the sensor directly, we firstly use 0 ppm and let Rate equal 1 to get Base_ Line, and then use 2 ppm, 5 ppm, 10 ppm, 20 ppm, 30 ppm and 40 ppm to calculate the Rate values respectively. Finally we choose the optimal value among these values as the Rate value. All these concentrations are chosen based on certain ratios of the sensor measurement range. The third part of the calibration system is the confirmation part, in which we use a CO analyzer (Ecotech EC9830) to validate the real certain gas concentration we get in the system.

To quantity the accuracy of our calibration, we show the experiment we conducted at the New South Wales Government Office of Environment and Heritage site that indicates the relationship between the standard calibration system and one of our portable sensor in Fig. 6. From the plot we can see that the response of Node sensor to the change in CO concentration is linear with the R^2 (coefficient of determination) values above 0.995. Mean absolute error can reach 0.7662 while mean absolute percentage error is close to 0.12. This calibration result indicates that the portable Node sensor is able to sense the CO concentrations in a stable and reasonable accuracy.

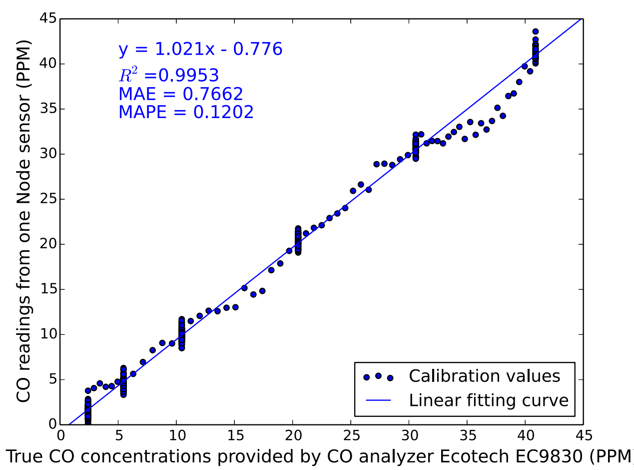


Fig. 6. Correlations between measurements from CO analyzer and one Node sensor.

C. Mounting the Sensor Device

Mounting the sensor devices (on a vehicle or a person) posed another significant challenge. The primary objective in our project was to mount these sensor nodes on vehicles, and accordingly Fig. 5 shows various mounting positions we tried: on top of a car, inside the car, and on a bicycle. The Node sensor and SensorDrone device have smaller form-factor than our HazeWatch node, which enables them to be carried on the person, such as clipped to the belt or backpack. Because turbulence flow effects are significant if there is forced flow [32], we place the sensors so the orientation is across the wind rather than into it when mounted on the top of the car, so wind does not directly blow in via the vent holes (on the side of the unit). This avoids large changes in

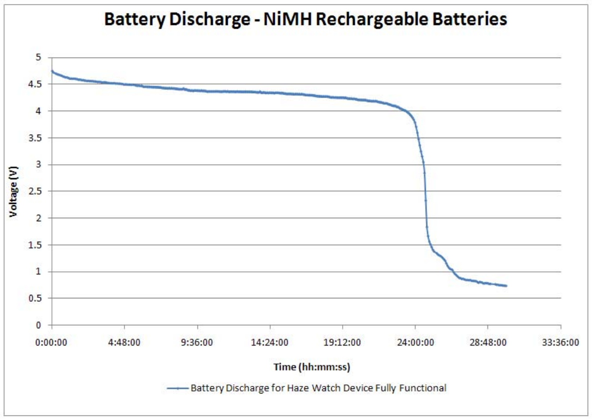


Fig. 7. Discharge trend for NiMH rechargeable batteries.

air pressure. When we designed the trials below, we also put the inlets of all the sensors together (maximum 10cm away from each other) to avoid forced flow thus to improve confidence in the results. Further, the casing on our HazeWatch unit prevents the sensor from being directly exposed to the sun or rain. However, we found that the Node sensor and SensorDrone device are not weatherproof and are hence better mounted inside the car when it is raining or snowing.

Removing and recharging the sensor also have to be considered. The electrical performance of the batteries is vital as this will determine how long the batteries will last and hence how long the sensors can be used for before get removed and recharged. The sensors should be able to last for a whole week (minimum 14 hours of operation) without the batteries being recharged or replaced. This minimum requirement is realized by considering that a typical data contributor would experience travel times of up to an hour per trip, twice per day. This ensures that certain number of air pollution data is captured and recorded in our system. Furthermore it is very important to make sure that the batteries are able to supply enough voltage to power all the various components on the sensor board. We used four Energizer AA nickelmetal hydride (NiMH) rechargeable batteries to check the battery performance of our HazeWatch node; Fig. 7 shows that the voltage level of the batteries slowly drops off during the first 22 hours of operation and drops off suddenly at approximately 23 hours. During the first 22 hour-period the wireless sensor was completely functional, and at about 23 hours of operation the Bluetooth module stops functioning and the wireless sensor board becomes inoperable. We also incorporate a low battery detector into our sensor to alert the data contributor through a red solid LED when the batteries are low on charge and require recharging. Node sensor and SensorDrone device can all keep functional over 30 hours operating after fully recharged.

D. Mobile App for Data Upload

The sensor unit tethers with the user's mobile phone via BlueTooth. As explained earlier, we rely on GPS and 3G capability in the phone, rather than replicating these functionalities on the sensor unit. As of 2011, 46% of all Australians are estimated to own a smart-phone, and this number is rising rapidly, so we do not expect this requirement to be onerous on the user. We developed several apps for Android-based and iOS-based phones to interact with the sensor unit over BlueTooth. Screenshots of these app interfaces are shown in Fig. 8. For example, application that connect mobile phone and HazeWatch node is shown in Fig. 8(a), from which we can see that upon startup the apps scans for Bluetooth devices and shows a list of sensor units that are within communication range. Upon connecting to the appropriate unit (unit 104 in this case), the app downloads the calibration constants for that unit from our server (the calibration constants are not hard-coded into the app so that drifts can be easily adjusted at the server end without requiring any change to the code in the app). Thereafter, the app then constantly displays information to the user, such as current location, pollutant type, current pollutant values reported by the unit, current time, up to five samples recorded in the past, etc. Note that our design requires minimal input from the user, who is required only to start the app and connect to the sensor unit at the commencement of each record; all actions thereafter are automatic.

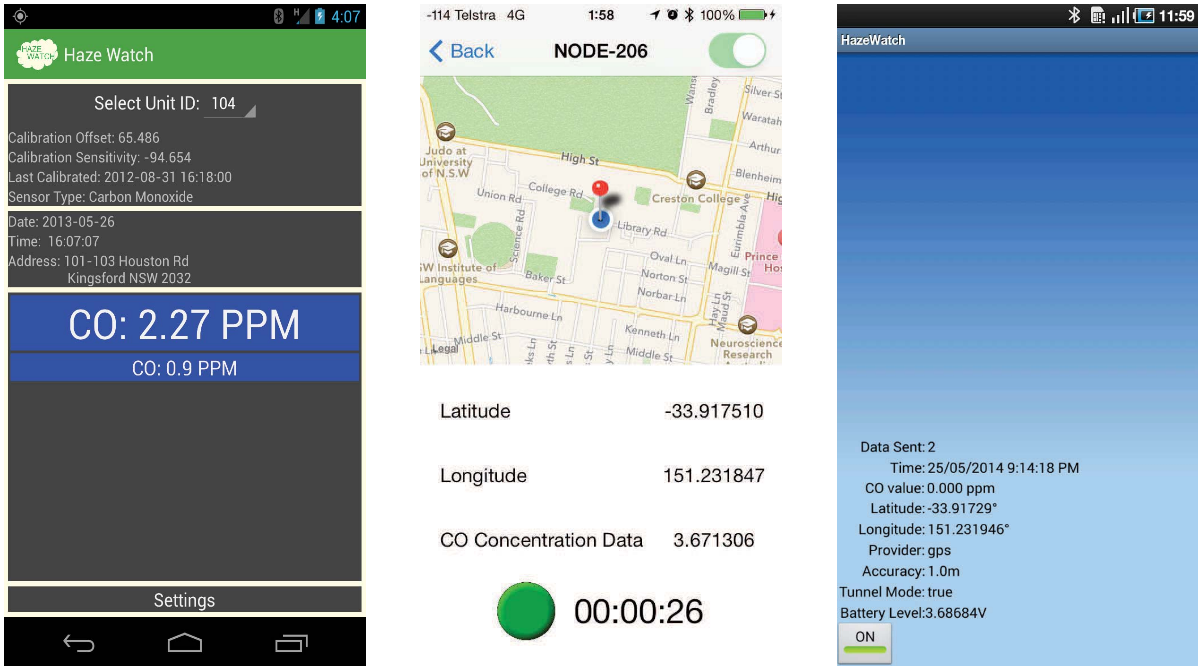
A second app we developed that works with the Node sensor device is shown in Fig. 8(b). On the top is a visual map which indicates the recording location and route. It also shows the GPS information and pollutant values along with the recording time. Based on energy efficiency consideration, we set that the app will not call for the GPS data unless the user click the record button on the bottom. This app can run in the background and upload data continuously. Air pollution data is collected per five seconds, and uploaded to the server per 25 seconds. A similar app for the Sensorcon device is shown in Fig. 8(c).

One of the challenges we faced was that location-stamping of the pollution data is done by the mobile phone using GPS information, which was lost inside the numerous tunnels in Sydney. To overcome this problem, we developed a simple interpolation algorithm in the mobile phone app so that pollution data is stored locally while GPS is lost (when one enters the tunnel), and whenever GPS gets re-acquired (when one exits the tunnel), the stored samples are equally spaced between the entry and exit points via simple linear interpolation.

E. Server Database

The last component of our data collection architecture is the database server itself. This is the central repository, hosted in our data center, to which all our data contributor users (who carry sensor unit devices along with the mobile apps) automatically upload data. We also wrote automated scripts on our server so it harvests data published hourly by the state Department of Environment on pollution levels at their fixed stations (around 12 in number) in and around Sydney.

The architecture of our server software comprises three layers: the web-server layer, the model layer, and the database layer. The *database layer* forms the core of the system, by storing all readings and providing a simple interface for extracting and filtering readings. We use MySQL, chosen for



(a) (b) (c)

Fig. 8. Mobile applications for uploading air pollution data with: (a) HazeWatch node and Android system, (b) Node device and iOS system, and (c) SensorDrone sensor and Android system.

its efficiency, reliability, and ease of use when searching and filtering over large sets of data. The *model layer* provides an abstraction of the data, whereby it can return the air pollution level for any arbitrary point in location and time, by employing an underlying interpolation model (discussed in the next section) over the collected data. This conceptually separates the production of data from its consumption, allowing an application to be written without constraints on the underlying data density or continuity. Note that all interaction with the data occurs via the model, so that a consistent interface is presented to any application seeking to use the data. The *web-server layer* presents the data (via the model) to the outside world, in the form of web-pages, maps, and applications that access it via an API. A detailed description of the server design and implementation can be found in our report [33].

F. Data Modelling, Visualization, and Personalization

1) Interpolation Models: By using mass produced mobile units, we expect to measure air pollution at much finer spatial granularity than available from the government's fixed monitoring stations today. Nevertheless, since no system can measure pollution over all points in space and time, we need to employ models that can estimate concentrations covering the full urban space under consideration. The available methodological approaches to estimate the spatial distribution of air pollution range from simple empirical techniques such as interpolation [34], to various statistical regression methods or data-driven models such as land use regression [35], [36] and neural networks [37], to more complex models including atmospheric chemistry and dispersion [38], [39]. In most instances, progression from a simpler empirical model to a more complex forecasting model entails increased data requirements (other than direct measurements), more specialized software, and a corresponding higher number of sources of uncertainty. Our initial effort in this project has been to use simple interpolation models, and we hope to refine these in our subsequent work.

Even so, there are many different forms of interpolation which process data differently. The speed and accuracy of various techniques, as well as the variability and density of the original dataset, must all be taken into account. Interpolated data generally has greater reliability when sampled data locations are densely and uniformly distributed; conversely if data locations are clustered with large gaps between sites, inaccurate estimates will be obtained. This holds true regardless of the method we choose. We must also be aware of the fact that interpolation inherently underestimates the peaks and overestimates the dips due to the nature of averaging. We implemented two interpolation methods: inverse-distance weighting, and ordinary kriging, as briefly described next.

Using inverse distance weighting (IDW) to estimate concentration at a point in space involves allocation of weights to all neighbouring points, based on the distance between the points. A point that is further away from the interpolation point therefore has less significance than one closer. IDW can be implemented easily and quickly, and is the default option for our model. However, it can have high error rates when points are sparsely distributed, and the contour maps thus generated are not very smooth (known as bull's eye effect). We therefore also implemented ordinary kriging, which is more complex but yields more robust results. Kriging involves computing the empirical semivariogram over the data, which is done by clustering pairs of data points into bins that have similar distance, and plotting the semi-variance of each

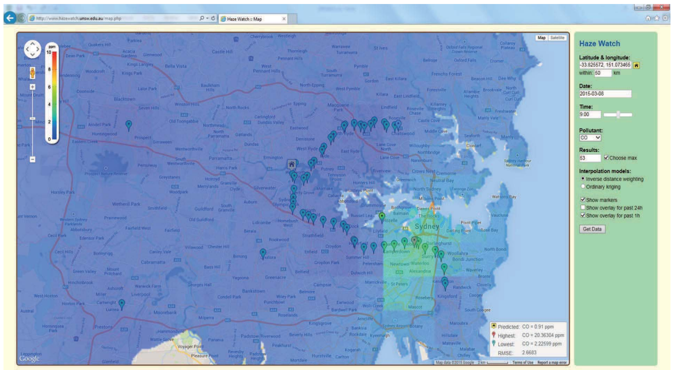
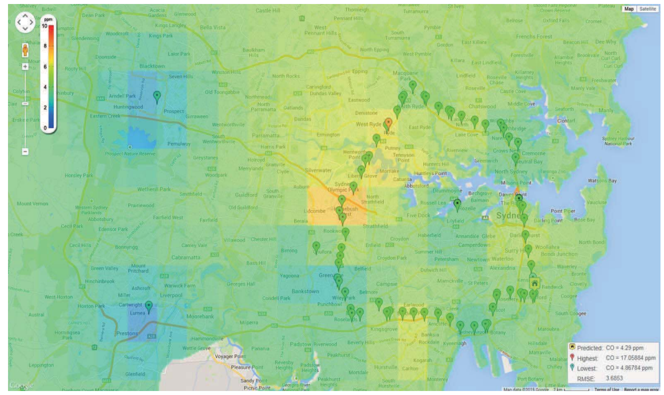


Fig. 9. Contour map of CO concentration overlaid on Google maps.

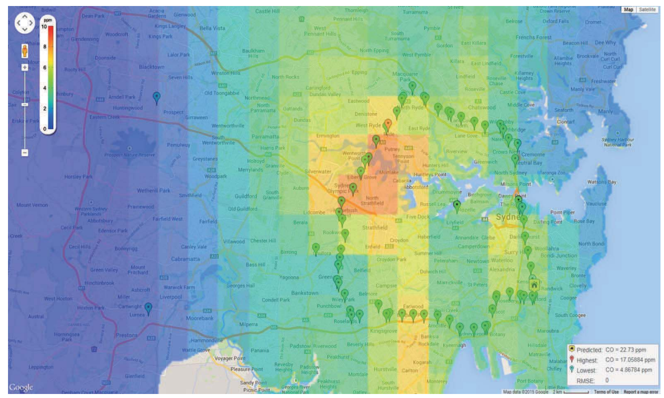
bin as a function of the distance corresponding to the bin. The interpolation weights are derived by solving a system of linear equations relating the weights to the semi-variance determined from the model variogram. An important benefit of this technique is that it provides the ability to assess error or uncertainty of the estimated point, and is a widely accepted method in air quality studies. We also found that it presents a much smoother and natural-looking contour plot in our maps. However, the maps takes several seconds to render on our web-page when this interpolation method is used. A detailed discussion of the interpolation methods and its implementation in this project can be found in our report [40].

2) Web Based Visualization: The web application consists of a client-side component and a server-side component, separated by a network. As described earlier, the server stores the geo-referenced data in the MySQL database, runs the interpolation models, generates a contour map for selected datasets, and exposes query processing APIs to outside applications. The client side runs a web-based form input that allows users to enter position, time, and other parameters, and pass those to the server. The pollution contour map generated is overlaid on Google maps, chosen for its ease-of-use, popularity, and well-documented API. Our client implementation uses standard web technologies of HTML, CSS and Javascript, and also leverages the power of AJAX (asynchronous JavaScript and XML) with PHP server-side scripting to deliver the maximum data modelling and visualization capabilities.

A sample screen-shot depicting a contour map of CO concentration on the web-page is shown in Fig. 9. The panel on the right allows the user to input data such as location (latitude/longitude) and radius for the map, the date and time of interest, the pollutant that needs to be mapped, the number of measurement points, the interpolation model, and the time at which the map is created. The panel on the left shows the contour map, along with labels with the data points. In a particular enquiry time, we use 40 minutes (20 minutes before enquiry time and 20 minutes after enquiry time) as a time window and use all the values within these 40 minutes, along with the interpolation model to compute this contour map. Hovering over a label opens a pop-up showing the details of the data point such as date/time and value. The bottom right on this panel also shows the minimum and maximum values, along with the estimated value at any point where the yellow marker is dropped.



(a)



(b)

Fig. 10. Contour map over same data points obtained from: (a) Inverse Distance Weighting interpolation and (b) Ordinary Kriging interpolation.

To contrast the results we obtain from the two interpolation models, the corresponding maps, obtained from identical data on CO measurements, are shown in Fig. 10. The inverse distance weighting (IDW) contour map in Fig. 10(a) shows high pollution is tightly concentrated in the tunnels, whereas the ordinary kriging contour map in Fig. 10(b) shows the CO pollution spreading around the tunnels and city CBD, with the air getting cleaner as one moves west. While the relative performance of these models depends on data density and distribution, we found that kriging usually present a smoother gradient and better aesthetics than inverse distance weighting.

3) Mobile App for Health Impact: We believe that though a relatively small of users in an area may carry our sensor units and contribute pollution data, everyone (including people who do not have a unit) should benefit from the data, and be empowered with personal tools to estimate and manage their pollution exposure. To this end we developed an iPhone application that tracks the user, and computes their exposure *a posteriori* based on their location trace. Our app allows the user to start and stop tracking their route, which get recorded as a trip. The user can see a list of their trips, and for each trip, compute the average exposure to each pollution. The trip can also be seen overlaid on a map, and the pollution exposure can be seen as a graph. For example, in Fig. 11 we show a screen-shot of the pollution graph, showing how the exposure to CO varied over time as the user was driving in a large loop around Sydney from approximately 1:40pm to 3:20pm on a work-day. The graph also shows the user, via a red line,

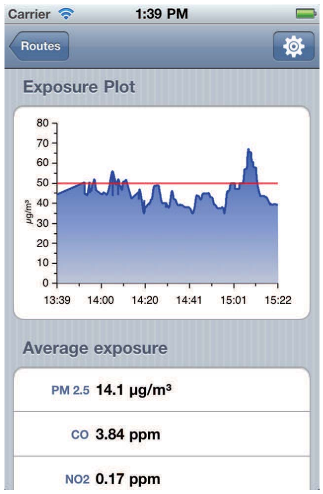


Fig. 11. Personal health impact estimation mobile application interface.

how their exposure compares with the long-term value deemed safe by the WHO. We are currently working on enhancing the app to provide prospective route mapping, namely to guide users on alternative driving paths that have lower pollution exposure. One can see that a mobile app like the one we are developing can not only help users who do not carry a sensing unit, but also personalize the data and make it more relevant to them.

IV. Validation and Field Trials

We briefly outline our experiments to validate the sensor platforms, and field trials that illustrate the value of our system in getting better estimates of personal air pollution exposure.

A. Validation of Sensor Platforms

We conducted several experiments to validate the correctness of the various sensor units, including the HazeWatch node, the Node sensor and the SensorDrone sensor. In our first experiment both the commercial monitor (GasAlert Micro 5 unit used in our calibration) as well as our HazeWatch sensor unit were mounted on the car, and the CO measurements from both were recorded these are shown in Fig. 12(a) as the blue and orange curves respectively. Two immediate observations can be made - first, that the pollution on Sydney roads shows significant spatial variation, with pollution peaking in tunnels, often reaching dangerously high levels, as shown by annotations in the figure. The second observation is that the measurements from our unit closely follow the commercial meter, validating that our construction, calibration, and software are working correctly, giving us reasonable confidence that our measurements are correct. Another observation that emerges from this plot is that the green curve, which corresponds to the values obtained from 12 government monitoring station readings and an interpolation model, indicate a very low level of pollution (often below 1 ppm). This large discrepancy illustrates the need for finer grained monitoring, as envisaged by our system. In the second experiment, two sensor devices (Node and SensorDrone) are attached to a car along with the commercial (GasAlert Micro 5) monitor, and CO measurements were taken from 7am to 8:30am along a typical commute route in Sydney. The results shown in Fig. 12(b) again confirm that values from Node and SensorDrone sensors correspond reasonably well with data from the GasAlert Micro 5 commercial monitor. Nevertheless, we observe that the Node device more accurately follows the readings from the commercial monitor, while the SensorDrone can depict higher values of pollution, particularly at high concentration values.

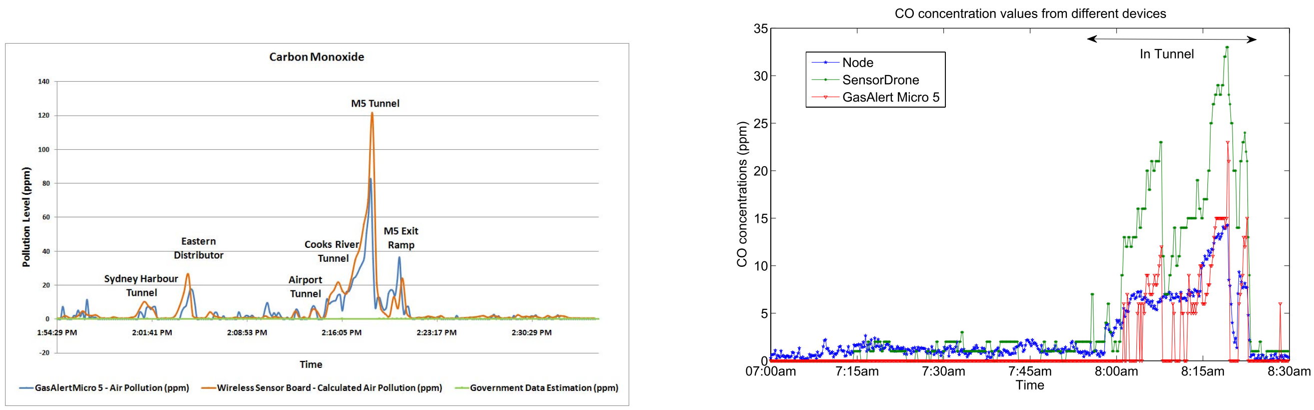
B. Trials With Exposure Estimation

Once validity of the data from the various sensor devices is established, we conducted several field trials to determine how our system facilitates estimation of personal air pollution exposure. Our methodology is as follows: we have multiple users carry our pollution sensors and contribute data in real-time using our apps (these are known as "data contributors"). Meanwhile, one subject user, for whom we wish to estimate air pollution exposure, is made to carry the commercial monitor (as an indicator of ground truth), and our mobile app that computes his personal exposure based on his movement pattern during the course of the trial. Our objective is to see how accurately we can estimate this subject user's exposure (ground truth being derived from the commercial monitor) based purely on software on this user's phone (in other words the subject user is not required to carry an air pollution sensor) - if our estimates are accurate enough, it will show that exposure can be estimated for large populations based on data uploaded by relatively few number of contributors who carry the sensors.

The first field trial, shown in Fig. 13(a), corresponds to multiple data contributors who carry the HazeWatch sensor unit and contribute data, while the subject user only has the air pollution estimation iPhone app (to obtain ground truth data, the subject user also carries the GasAlert Micro 5 commercial monitor). The second trial, shown in Fig. 13(b), is similar but uses the Node sensor device. The routes used for these trials contain tunnels and motorways. They key observation from these trials is that even though the subject user does not carry a sensor device, their estimated exposure (orange curve) derived from data from other contributors in our system, is able to capture the periods of high pollution (typically tunnels and congested motorways), shown by peaks in the blue curve from the commercial monitor; by contrast, the data from the government monitoring sites (green curve) is always low and flat, unable to capture the spatial variation as the subject user drives around. This establishes value in having our participatory sensing system both in terms of spatial density and in terms of user movement patterns. In the first and second trial of exposure estimation, we observe that peak points are shifted between estimation values from our mobile application and true values. There are multiple data contributors who carry the sensors and contribute data, and our server uses average values to represent the concentration on each point, while true values are from commercial sensor instantly. This is the reason why the time shift happens.

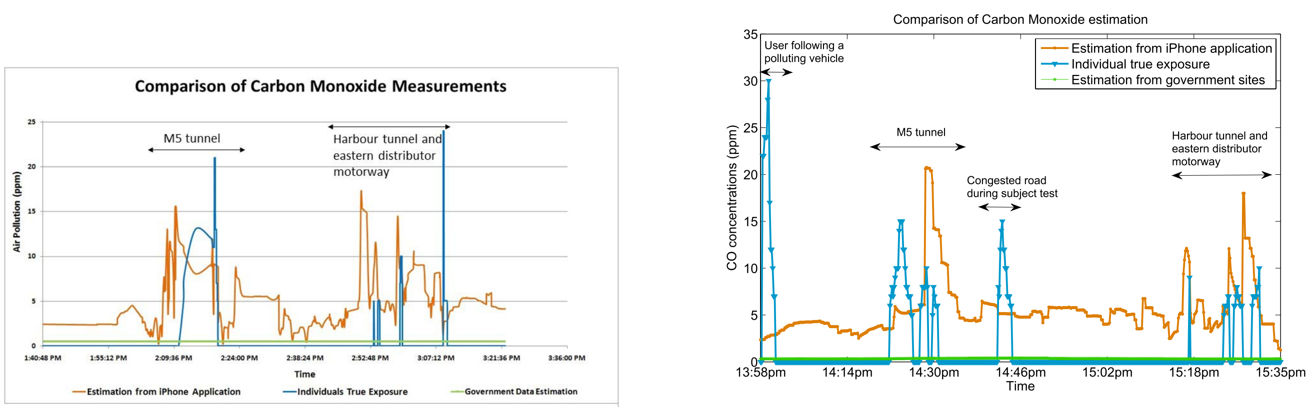
C. Impact of Mobility on Personal Dosage

We acknowledge several limitations in our system: currently we have low density in deployment (each of our trials only



(a) (b)

Fig. 12. Validating the performance of: (a) HazeWatch node and (b) Node sensor and SensorDrone sensor.



(a) (b)

Fig. 13. Exposure estimation in our personalized app with (a) our HazeWatch sensor, and (b) Node sensor.

had two or three data contributors), which explains why our estimates, though much better than the government monitoring system, are still not that close to the ground truth. We believe these results can be improved significantly with deployment density; indeed we are in conversation with the government department to procure, deploy, and integrate our system with their current monitoring systems. The other limitation of our system is that it does not capture sporadic pollution events, such as the subject in trial 2 who happened to be stuck behind a truck on one of the motorways near the very beginning of the trial – this highly localized phenomenon cannot really be estimated from data from other contributors (one can see that the orange curve does not capture this spike in the blue curve). These limitations not withstanding, we believe it is still worthwhile that our system gives much better estimates of exposure than available from current government systems.

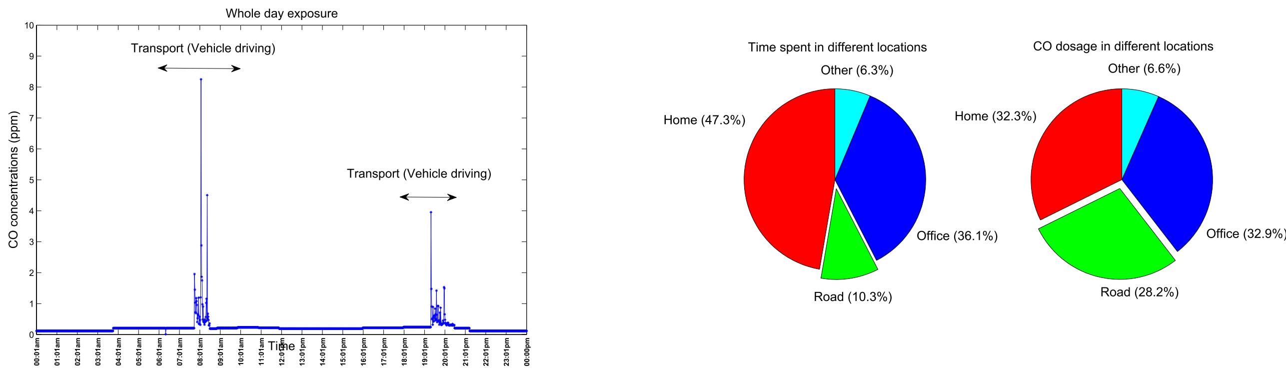
We did longer-term trials in which the individual's air pollution dosage was studied over the course of a day to determine how it is impacted by their daily movement patterns. Fig. 14(a) shows the individual's exposure over the 24-hour period, from which we can see that concentrations peak during morning and evening commute to/from work by car. During the daytime, CO concentrations stays low, though higher than at home. In order to determine how much pollution is inhaled by this user over the 24-hour period (aka their pollution "dosage"), we show in Fig. 14(b) the time spent and the dosage in the various locations. The inhaled dosage is calculated using the following simple formula:

$$\begin{aligned} Inhaled\_dose = {} & Respiratory\_minute\_volume \\ & \times CO\_concentration \times time \\ & \times conversion\_factor, \end{aligned} \tag{4}$$

where respiratory minute volume (RMV) refers to the volume of air inhaled by a person per minute, and is chosen as 12 L/min for a typical adult male [41], and the conversion factor (to change parts-per-million to $\mu g/L$) for carbon monoxide is 1.145. What is interesting from this figure is that even though the individual spends less than 10% of his time on the road, nearly 30% of his daily pollution dosage comes from his commutes. Of course this would vary by individual and where they live/work, but indicates that road travel in a metropolitan area can be a significant contributor to air pollution inhalation.

D. Deployment Challenges

Our experience with building and trialling the system over the past 2-3 years has taught us that the highly interdisciplinary nature of this project makes it full of challenges



(a) (b)

Fig. 14. (a) Whole day exposure. (b) Time spent and CO dosage in different locations in one day.

on many fronts. The most significant challenges have arisen with the hardware: (a) the metal oxide sensors are cheap but non-linear and unreliable, while the electrochemical are expensive and extremely sensitive, operating at nano-amps, requiring very careful circuitry, particularly for stabilization of the sensor, (b) the calibration of the sensor units has been challenging, given that we do not have proper facilities and certification to store and handle toxic gases, (c) the packaging of the unit, and mounting it on the car has also presented difficulties, and (d) mass production of these units also requires careful consideration, bearing in mind cost, aesthetics, battery life, etc., that we have not currently optimized for. Some of the other challenges we have faced in this project include finding the right user base to target it to, ranging from bicyclist groups and asthma sufferers to transportation operators and members of the general public. Getting a dedicated user-base of data contributors is non-trivial but necessary if the system is to become useful to the general public at large.

V. Conclusion

In this paper we have described the architecture, prototype and evaluation of a low-cost participatory metropolitan air pollution sensing system. Our system compared multiple portable sensing units, including ones designed by us, and used mobile phone applications and cloud-based service to obtain higher resolution pollution surface for the metropolitan area in real time. We developed mobile apps for personalising the pollution estimates for individuals based on their mobility patterns, allowing them to better understand how they are impacted by pollution. We validated our system with a handful of users to demonstrate that our system yields more accurate estimates of personal exposure than current systems based on government monitoring data. We believe that our system can be applied world-over, particularly in pollution-heavy countries, to better understand the relationship between urban air pollution and its health impacts.

Acknowledgment

The authors would like thank Yvonne Scorgie and John Kirkwood who worked in the New South Wales Government Office of Environment and Heritage for the kind and selfless help in improving our calibration procedure. They would also like to thank the reviewers for their helpful and inspirational comments.

References

[1] World-Health-Organization. *Air Pollution*. [Online]. Available: http://www.who.int/topics/air_pollution/en/

[2] World-Health-Organization. *7 Million Premature Deaths Annually Linked to Air Pollution*. [Online]. Available: http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/

[3] S. Weichenthal *et al.*, "Long-term exposure to fine particulate matter: Association with nonaccidental and cardiovascular mortality in the agricultural health study cohort," *Environ. Health Perspect.*, vol. 122, no. 6, pp. 609–615, 2014.

[4] J. K. Wendt, E. Symanski, T. H. Stock, W. Chan, and X. L. Du, "Association of short-term increases in ambient air pollution and timing of initial asthma diagnosis among medicaid-enrolled children in a metropolitan area," *Environ. Res.*, vol. 131, pp. 50–58, May 2014.

[5] B. A. Franklin, R. Brook, and C. A. Pope, III, "Air pollution and cardiovascular disease," *Current Problems Cardiol.*, vol. 40, no. 5, pp. 207–238, 2015.

[6] The-Weather-Channel. *Air Quality Forecast*. [Online]. Available: http://www.weather.com/activities/health/airquality/

[7] G. Hoek *et al.*, "Satellite NO_2 data improve national land use regression models for ambient NO_2 in a small densely populated country," *Atmos. Environ.*, vol. 105, pp. 173–180, Mar. 2015.

[8] T. K. Beatty and J. P. Shimshack, "Air pollution and children's respiratory health: A cohort analysis," *J. Environ. Econ. Manage.*, vol. 67, no. 1, pp. 39–57, 2014.

[9] R. Beelen *et al.*, "Effects of long-term exposure to air pollution on natural-cause mortality: An analysis of 22 European cohorts within the multicentre ESCAPE project," *Lancet*, vol. 383, no. 9919, pp. 785–795, 2014.

[10] N. M. Gatto *et al.*, "Components of air pollution and cognitive function in middle-aged and older adults in Los Angeles," *NeuroToxicology*, vol. 40, pp. 1–7, Jan. 2014.

[11] Imperial and Cambridge. *MESSAGE*. [Online]. Available: http://research.cs.ncl.ac.uk/message/

[12] M. I. Mead *et al.*, "The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks," *Atmos. Environ.*, vol. 70, pp. 186–203, May 2013.

[13] Vanderbilt University. *MAQUMON*. [Online]. Available: http://www.isis.vanderbilt.edu/projects/maqumon/

[14] P. Dutta *et al.*, "Common sense: Participatory urban sensing using a network of handheld air quality monitors," in *Proc. 7th SenSys*, Berkeley, CA, USA, Nov. 2009, pp. 349–350. [Online]. Available: http://www.communitysensing.org/

[15] B. Predic, Z. Yan, J. Eberle, D. Stojanovic, and K. Aberer, "ExposureSense: Integrating daily activities with air quality using mobile participatory sensing," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PERCOM Workshops)*, Mar. 2013, pp. 303–305.

[16] EPFL. *Open Sense*. [Online]. Available: http://www.opensense.ethz.ch/

[17] Fondazione ISI. *Enhance Environmental Awareness Through Social Information Technologies (EveryAware)*. [Online]. Available: http://www.everyaware.eu/resources/deliverables/D7.3.pdf

[18] D. Mendez, A. J. Perez, M. A. Labrador, and J. J. Marron, "P-sense: A participatory sensing system for air pollution monitoring and control," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PERCOM Workshops)*, Mar. 2011, pp. 344–347.

[19] S. Bhattacharya, S. Sridevi, and R. Pitchiah, "Indoor air quality monitoring using wireless sensor network," in *Proc. 6th Int. Conf. Sens. Technol. (ICST)*, Dec. 2012, pp. 422–427.

[20] S. Devarakonda, P. Sevusu, H. Liu, R. Liu, L. Iftode, and B. Nath, "Real-time air quality monitoring through mobile sensing in metropolitan areas," in *Proc. 2nd ACM SIGKDD Int. Workshop Urban Comput. (UrbComp)*, 2013, Art. ID 15.

[21] L. Capezzuto *et al.*, "A maker friendly mobile and social sensing approach to urban air quality monitoring," in *Proc. IEEE SENSORS*, Nov. 2014, pp. 12–16.

[22] A. de Nazelle *et al.*, "Improving estimates of air pollution exposure through ubiquitous sensing technologies," *Environ. Pollution*, vol. 176, pp. 92–99, May 2013.

[23] P. Völgyesi, A. Nádas, X. Koutsoukos, and. Á. Lédeczi, "Air quality monitoring with SensorMap," in *Proc. 7th Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, St. Louis, MO, USA, Apr. 2008, pp. 529–530. [Online]. Available: http://www.isis.vanderbilt.edu/projects/maqumon

[24] K. Aberer *et al.*, "OpenSense: Open community driven sensing of environment," in *Proc. ACM SIGSPATIAL Intl. Workshop GeoStreaming (IWGS)*, San Jose, CA, USA, Nov. 2010, pp. 39–42. [Online]. Available: http://opensense.epfl.ch

[25] e2v. *MiCS5521 Data Sheet*. [Online]. Available: http://goo.gl/ezAFHo

[26] J. Carrapetta, "Haze watch: Design of a wireless sensor board for measuring air pollution," Elect. Eng. Telecommun. School, Univ. New South Wales, Sydney, NSW, Australia, Tech. Rep., 2010.

[27] M. Kelly, "Design of a wireless sensor board for an air pollution monitoring system," Elect. Eng. Telecommun. School, Univ. New South Wales, Sydney, NSW, Australia, Tech. Rep., 2012.

[28] S. De Vito, E. Massera, M. Piga, L. Martinotto, and G. D. Francia, "On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario," *Sens. Actuators B, Chem.*, vol. 129, no. 2, pp. 750–757, 2008.

[29] European Commission Joint Research Centre. *Report of Laboratory and In-Situ Validation of Micro-Sensor for Monitoring Ambient Air—Ozone Micro-Sensors, aSense, Model B4 O3 Sensors*. [Online]. Available: https://ec.europa.eu/jrc/sites/default/files/lb-na-26681-en-n.pdf

[30] J. Liu, "Calibration of a wireless sensor board for measuring air pollution," Elect. Eng. Telecommun. School, Univ. New South Wales, Sydney, NSW, Australia, Tech. Rep., 2012.

[31] OEH. *Office of Environment and Heritage—NSW*. [Online]. Available: http://www.environment.nsw.gov.au/

[32] J. Fonollosa, I. Rodríguez-Luján, M. Trincavelli, A. Vergara, and R. Huerta, "Chemical discrimination in turbulent gas mixtures with mox sensors validated by gas chromatography-mass spectrometry," *Sensors*, vol. 14, no. 10, pp. 19336–19353, 2014.

[33] N. Youdale, "Haze watch: Database server and mobile applications for measuring and evaluating air pollution exposure," Elect. Eng. Telecommun. School, Univ. New South Wales, Sydney, NSW, Australia, Tech. Rep., 2010.

[34] A. A. Szpiro, L. Sheppard, P. D. Sampson, and S.-Y. Kim, "Validating national kriging exposure estimation," *J. Environ. Health Perspect.*, vol. 115, no. 7, pp. 338–339, 2007.

[35] M. Jerrett *et al.*, "A review and evaluation of intraurban air pollution exposure models," *J. Exposure Anal. Environ. Epidemiol.*, vol. 15, no. 2, pp. 185–204, 2005.

[36] P. H. Ryan and G. K. LeMasters, "A review of land-use regression models for characterizing intraurban air pollution exposure," *J. Inhalation Toxicol.*, vol. 19, no. 1, pp. 127–133, 2007.

[37] M. Z. Boznar and P. Mlakar, "Use of neural networks in the field of air pollution modelling," in *Air Pollution Modeling and Its Application XV*. Amsterdam, The Netherlands: Springer-Verlag, 2004, pp. 375–383.

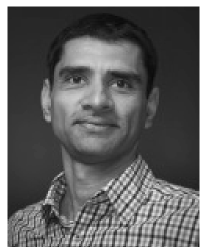
[38] K. Schere, "The emergence of numerical air quality forecasting models and their application," in *Proc. AMS Annu. Meeting*, Atlanta, GA, USA, 2006, pp. 1–6.

[39] W. L. Physick, M. E. Cope, S. Lee, and P. J. Hurley, "An approach for estimating exposure to ambient concentrations," *J. Exposure Sci. Environ. Epidemiol.*, vol. 17, no. 1, pp. 76–83, 2007.

[40] A. Chow, "Haze watch—Data modelling and visualisation using Google maps," Elect. Eng. Telecommun. School, Univ. New South Wales, Sydney, NSW, Australia, Tech. Rep., 2010.

[41] P. Young *et al.*, "Delivery of titrated oxygen via a self-inflating resuscitation bag," *Resuscitation*, vol. 84, no. 3, pp. 391–394, 2013.

Ke Hu received the B.Eng. degree from Beijing Jiaotong University, and the M.Eng. degree from Xinjiang University, in 2007 and 2012, respectively. He is currently pursuing the Ph.D. degree with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia. His research interests include wireless sensor network, data mining, and air pollution.



Vijay Sivaraman received the B.Tech. degree from the Indian Institute of Technology, Delhi, India, in 1994, the M.S. degree from North Carolina State University in 1996, and the Ph.D. degree from the University of California at Los Angeles, in 2000. He was with Bell Laboratories as a Student Fellow, in a silicon-valley start-up manufacturing optical switch-routers, and as a Senior Research Engineer at CSIRO, Australia. He is currently an Associate Professor with the University of New South Wales, Sydney, Australia. His research interests include optical networking, packet switching and routing, network architectures, and sensor networks for the environment, health-care, and sports monitoring.

Blanca Gallego Luxan received the B.S. degree from the Universidad Autónoma Metropolitana, and the Ph.D. degree from the University of California, Los Angeles. She is currently a Senior Research Fellow with the Australian Institute of Health Innovation and leads the modeling and simulation in health team. She has extensive international research experience in data analysis and computational modeling and has made significant and innovative contributions to the design, analysis, and development of models derived from complex empirical data for a wide range of applications, such as patient safety, biosurveillance, corporate sustainability reporting, ecological footprint analysis, and climate variability.

Ashfaqur Rahman (M'04–SM'12) received the Ph.D. degree in information technology from Monash University, Australia. He was involved in specific machine learning problems, including ensemble learning and fusion, feature selection/weighting methods, genetic algorithm based optimization, and Image segmentation and classification. He is currently a Project Leader and the Research Scientist with Commonwealth Scientific and Industrial Research Organisation Computational Informatics, Dickson. He is also a Researcher in ICT for more than ten years. He has authored around 50 peer-reviewed journal articles, book chapters, and conference papers. His research focus is on computational intelligence and machine learning. He serves as a Reviewer of prestigious conferences and journals. He was a Program Committee Chair of DICTA 2013.