HazeWatch: A Participatory Sensor System for Monitoring Air Pollution in Sydney

Vijay Sivaraman [#1], James Carrapetta [#2], Ke Hu [#3], Blanca Gallego Luxan [*4]

 # Electrical Engineering and Telecommunications, UNSW
* Centre for Health Informatics, UNSW
[1] vijay@unsw.edu.au [2] j.carrapetta@unsw.edu.au
[3] ke.hu@student.unsw.edu.au [4] b.gallego@unsw.edu.au

Abstract—This paper describes our experiences with building and deploying a low-cost participatory system for urban air pollution monitoring in Sydney. Though air pollution imposes significant health costs on the urban community globally, today's published data on pollution concentrations is spatially too sparse, and does not allow for sufficiently accurate estimation of exposures for (potentially mobile) individuals in order to make medical inferences. The HazeWatch project described in this paper uses several low-cost mobile sensor units attached to vehicles to measure air pollution concentrations, and users' mobile phones to tag and upload the data in real time. The greater spatial granularity of data thus collected enables creation of pollution maps of metropolitan Sydney viewable in real-time over the web, as well as personalized apps that show the individual's exposure history and allow for route planning to reduce future exposure. We share the insights we obtained from building and trialling such a system in Sydney, and highlight challenges that can be addressed collaboratively by groups developing similar systems world-wide.

I. Introduction

One of the basic requirements of human health and well-being is clean air. However, the World Health Organisation (WHO) estimates that around 1.4 billion urban residents worldwide are living in areas with air pollution above recommended air quality guidelines [28]. Chronic exposure to air pollution increases the risk of cardiovascular and respiratory mortality and morbidity, while acute short-term inhalation of pollutants can induce changes in lung function and the cardiovascular system exacerbating existing conditions such as asthma, chronic obstructive pulmonary disease (COPD), and ischemic heart disease [5], [21]. In fact, in Australia, air pollution is estimated to kill more people each year than road accidents [9]. Monitoring and controlling air pollution is high on the public consciousness in both developing and developed countries, including Australia that has among the highest prevalence of asthma in the west [2].

While some government agencies do monitor and publish metropolitan air quality (e.g. the New South Wales state government air quality index [22]), there are two fundamental limitations to the existing approach: **First**, the spatial resolution of the pollution sampling is very poor. For example, the greater Sydney area has approximately 15 active monitoring sites, separated from each other by tens of kilometres. This necessitates the use of mathematical models to estimate pollu-

tant 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 [12]), leading to incorrect inferences [30], [31]. **Second**, current observations of concentrations do not reflect actual exposures experienced by individuals, due to spatial heterogeneity of pollutant concentrations and individuals mobility patterns, such as time spent and activity levels at home, at work, and commuting. 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 [23].

In this paper we describe project "HazeWatch", which aims to crowd-source fine-grained spatial measurements of air pollution in Sydney, 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 development of a portable sensing unit mounted on vehicles, a mobile phone application for data tagging and uploading, and a centralized repository for hosting the data. (2) We show how the data can be analysed and consumed by users. This includes appropriate models for interpolating the spatio-temporal data points, visualisation of pollution over a geographical map of the area, and mobile apps that show personal exposure and low pollution travel routes. (3) We test and deploy 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 for data collection, and in §IV we describe how the data is visualized and personalized. §V describes our deployment experiences, and the paper concludes in §VI.
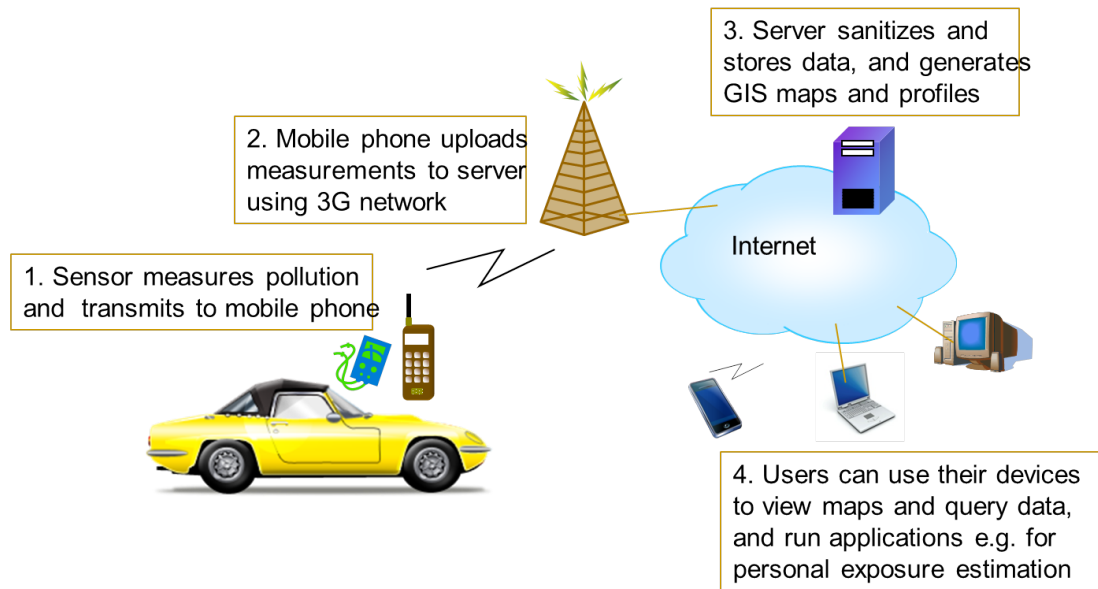
Fig. 1. System Architecture

II. PRIOR 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 [15] 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 [20] 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 modeling 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 [27], that developed portable wireless sensor units for measuring ozone, nitrogen dioxide and carbon monoxide. Their units are autonomous, having onboard 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 did

experiments has not undertaken any long-term deployments.

Intel has also been developing as part of the CommonSense project [11] 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. We believe the Common Sense project is currently running trials with these devices attached to the rooftops of street cleaners in the city of San Francisco. Several other projects, such as Sensaris, iSniff, etc. have similar goals to ours, but probably the most noteworthy is the well-funded OpenSense project [1] that is ongoing at EPFL Switzerland. To the best of our knowledge they 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.

In addition to the above large-scale projects, several smaller efforts have looked at various individual aspects of the system. Low-cost sensors and hardware devices for measuring air pollution have been explored in [7], and the design of the APOLLO (Air POLLutants mOnitoring system) sensor node using MEMS/infrared gas sensors is undertaken. Their paper is noteworthy for the detailed discussion of the issues encountered in the hardware design, including sensor selection, impact of humidity and temperature, different supply voltage required by the different sensors, power management, wake-up latencies for sensors, etc. The prototype developed is however somewhat crude, and does not address issues around packaging and mounting. Further, the sensor device developed requires communication with a WSN node that runs heavy-duty software running a multi-threaded operating system called RE-TOS. Experimental testing was done with tobacco smoke, and no road-tests were conducted. Another relevant work is [14]

that proposes methods to reduce communication overheads from sensor nodes when a large number are deployed in a city. While they do develop some ZigBee-based hardware, it is rather clunky (with CO2 sensor outside car and GPS/GSM modules inside the car) and likely to be expensive since they do not leverage mobile phone capabilities. Further, much of their evaluation is predominantly simulation based rather than via experimentation.

A mobile measurement system was also developed in [3]. 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 [10] 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 Modeling System (ADSM), developed from a previous year, and their estimates are hence neither real-time nor accurate.

III. Data Collection 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 collected 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 attached to cars that monitor air pollution as users drive, (2) an application 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) visualisation tools that map pollution levels and personalize the information for the individual user. The first two steps constitute data collection (described in this section), while the latter two steps comprise data consumption (discussed in the next section).

A. Pollution Measurement Node

Arguably the most critical and challenging component of the architecture is the device that enables users to measure air pollution. There are several options here, and a careful selection is needed in order to maximize chances of mass adoption. The challenges are briefly summarised below:

Portability: On the one extreme, the device can be bulky, as the one used by the government monitoring stations – however this condemns it to be fixed at a location, reducing spatial coverage. On the other extreme, the device can be made portable enough for a user to carry on their person, as intended in [15], [11] – however we believe that getting users to consistently carry a device, even if it is as small as a mobile phone, is very challenging. An alternative solution, which we feel is more viable, is to fit a portable device on the user vehicle. This frees the user from having to carry it

on their person, enables large spatial coverage as the user drives, and captures pollution near a significant source, namely motor vehicles on the road. Note that obtaining pollution measurements solely on and near roads will bias the sampling towards high-concentration areas; in order to temper this effect we also harvest the readings reported hourly in the Office of Environment web-page [22], which include sites that are situated away from roadways.

Complexity: Having decided to make our devices suitable for vehicle-mounting, 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 module to upload data in real-time. Indeed such a design was used for projects such as [27] and [1], and is suitable for mounting on public vehicles such as buses. However, our target users are consumers driving private cars, and in order to keep costs low, we chose a minimalist design that does not have GPS or 3G 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 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 (MO) Sensors. 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. MO sensors are compact and cheap (as low as $5 each), but have low accuracy and are non-linear. The use of MO sensors allowed us to built our unit housing three sensors (CO, NO_2 and O_3) at a cost price close to $50 (refer to [6] 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 [17]), shown in Fig. 2(b), using electrochemical (EC) sensors. These operate by passing the pollutant gas through the inner membrane of a gas chamber where it is oxidised, 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 our unit to house only one sensor at a time (the figure shows the CO unit), at an overall cost of about $150.

Calibration: Once built, we needed to calibrate each unit, which entails converting the current measurements into
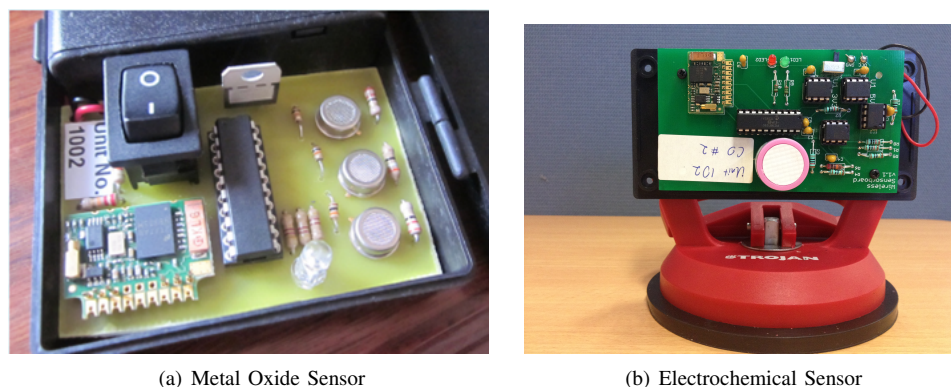
(a) Metal Oxide Sensor

(b) Electrochemical Sensor

Fig. 2.   Sensor unit using (a) Metal Oxide Sensor and (b) Electrochemical Sensor

(a) Commercial Monitor
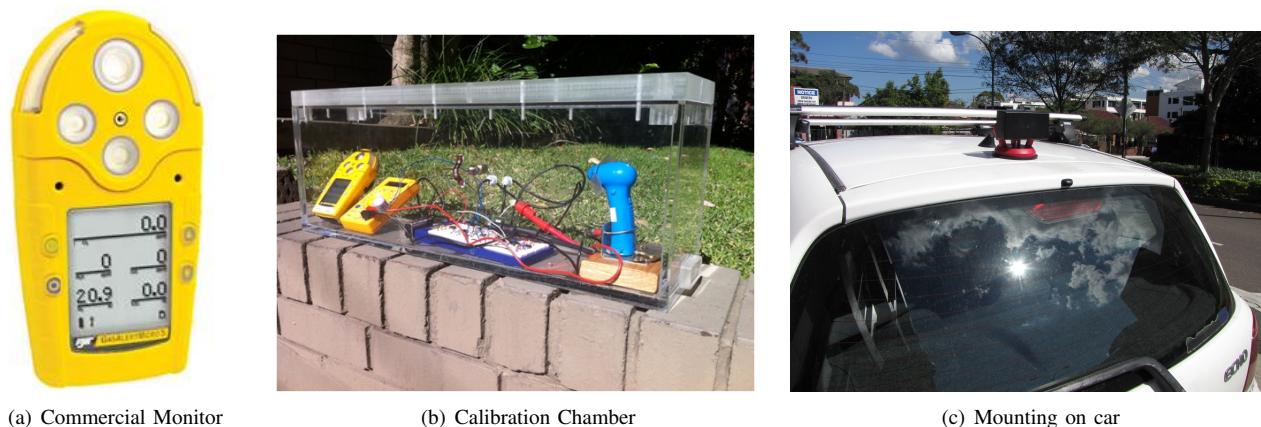
(b) Calibration Chamber

(c) Mounting on car

Fig. 3.   (a) Commerical monitor used for calibration, (b) the calibration chamber setup, and (c) Mounting our unit on the car

pollutant concentrations. This requires us to determine, at each known concentration of the gas, the reading output by our unit. This posed a major challenge for us since we did not have facilities for controlled experimentation with known concentrations of the gas. To this end we procured a commercial monitor called the GasAlert Micro 5 built by BW Technologies, as shown in Fig. 3(a), that could tell us the true pollutant concentrations. We then had custom-built an air-tight container, as shown in Fig. 3(b), 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 used added copper shavings to a beaker nitric acid inside the chamber that causes a chemical reaction in which the gas is 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. 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 [19] that outlines our calibration procedure and outcomes in great detail.

Mounting: Lastly, the mounting of the device on the car posed another challenge. Aspects of mounting position such as near or far to the exhaust, into or across the wind, high or low from the road, are significant. We chose to mount the unit on top of the car, as shown in Fig. 3(c), so it is not directly in line with the exhaust fumes. Our mounting is across the wind, rather than into it, so that wind does not directly blow in via the vent holes (on the side of the unit) that can cause large changes in air pressure. Further, the casing on our unit prevents the sensor from being directly exposed to the sun or rain.

B. 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 an app for Android-based phones to interact with the sensor unit over BlueTooth (the reason we could not do so for iPhones is that the APIs for
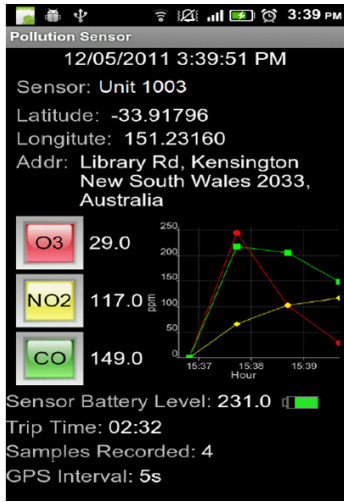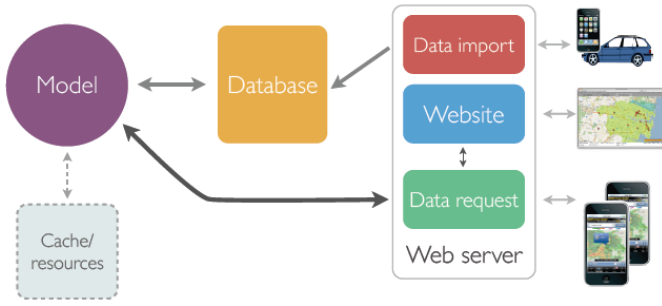
Fig. 4. App to Upload Pollution Data

Fig. 5. Server software architecture

C. 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 our sensor unit device along with the Android app) automatically upload data as they drive around. 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 – these provide measurements at locations away from roads, and reduce the sampling bias in our system arising from measurements being taken in and around roadways.

The architecture of our server software is depicted in Fig. 5, and 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 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 [29].

IV. Data Modeling, Visualization, and Personalization

In this section we describe how the data, once collected, is consumed. We briefly describe the models we use to interpolate the data, the web-based mapping to visualise the pollution in real-time, and the personalized tools we have developed to make the data meaningful to the user.

A. 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 [18], to various statistical regression methods or data-driven models such as land use regression [16], [25] and neural networks [4], to more complex models including atmospheric chemistry and dispersion [26],

BlueTooth access have not been made accessible by Apple to Universities). A screenshot of the app interface is shown in Fig. 4. Upon startup the app scans for bluetooth devices and shows a list of sensor units that are within communication range. Upon connecting to the appropriate unit (unit 103 in this case), the app downloads the calibration constants for that unit from our server (the calibration constants are not hardcoded 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, current pollutant values reported by the unit, sensor battery level, trip time, number of samples recorded in this trip, GPS sampling interval (the user can adjust this to control battery drain on the phone), 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 drive; all actions thereafter are automatic. We did try to eliminate even this action on the part of the user, by having the app run in the background and periodically scan the BlueTooth interface for presence of the sensor unit; however, this resulted in excessive battery drain, causing us to abandon this approach.
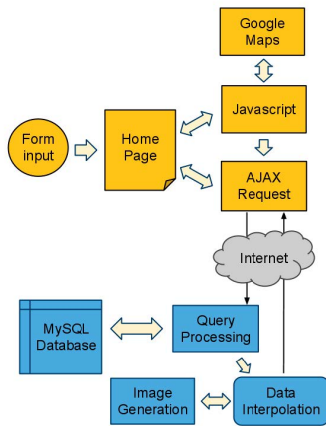
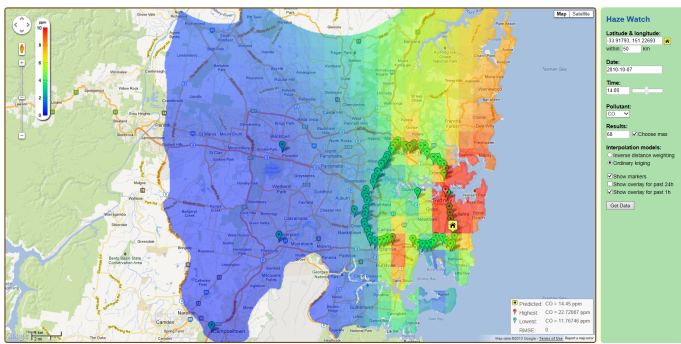Fig. 6.   Architecture of the web application

Fig. 7.   Contour map of CO concentration overlaid on Google maps

[24]. In most instances, progression from a simpler empirical model to a more complex forecasting model entails increased data requirements (other than direct measurements), more specialised 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 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 [8].

B. Web Based Visualization

The web application consists of a client-side component and a server-side component, separated by a network, as shown in Fig. 6. As described earlier, the server (blue boxes) 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 (shown in yellow boxes) 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 visualisation capabilities.

A sample screen-shot depicting a contour map of CO concentration on the web-page is shown in Fig. 7. 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 window for which the map is created. The panel on the left shows the contour map, along with labels with the data points. 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. We note that the map shows the pollution to be very low towards the west of Sydney, generated from government data, whereas the CO concentration is much higher in and around the city, based on readings obtained from our mobile sensor.

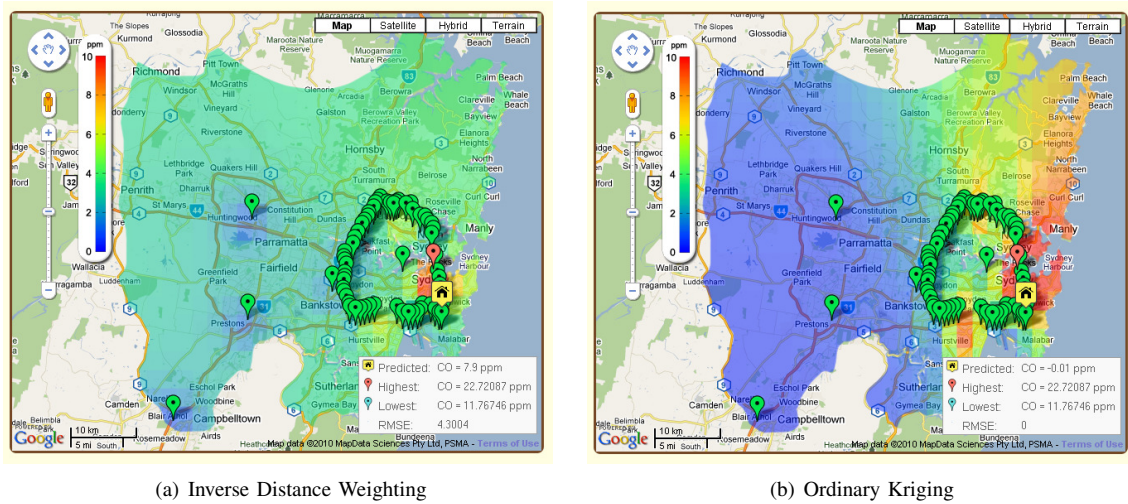(a) Inverse Distance Weighting
(b) Ordinary Kriging

Fig. 8. Contour map over same data points obtained from (a) Inverse Distance Weighting interpolation, and (b) Ordinary Kriging interpolation
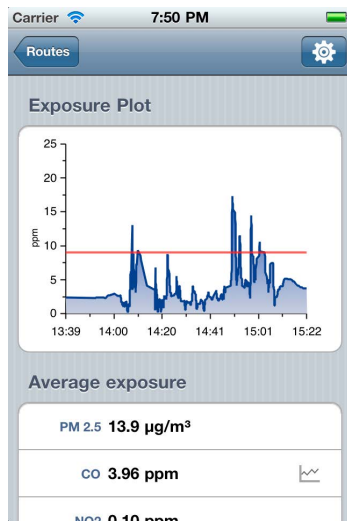

Fig. 9. iPhone app showing personal exposure to CO over a trip

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. 8. The inverse distance weighting (IDW) contour map in Fig. 8(a) shows the bulls eye effect whereby high pollution is tightly concentrated in the city CBD, whereas the ordinary kriging contour map in Fig. 8(b) shows the CO pollution spreading around the harbour, 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.

C. 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. 9 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 1:30pm to 3:30pm on a workday. The graph also shows the user, via a red line, 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.

V. Field Trials and Deployment Challenges

We briefly outline our field trial experiences and outline challenges for successful large-scale deployment.

A. Field Trials

We did tens of field trials with our system, in which we drove around Sydney with our sensor units collecting pollution data much the same way people would on their regular commutes. We highlight two specific trials, the first validating the correctness of our unit, and the second validating the goodness of our app that estimates personal exposure. In our first trial both the commercial monitor as well as our sensor unit were mounted on the car, and the CO measurements from both were recorded – these are shown in Fig. 10(a) as the blue and orange curves respectively. Two immediate observations can be made
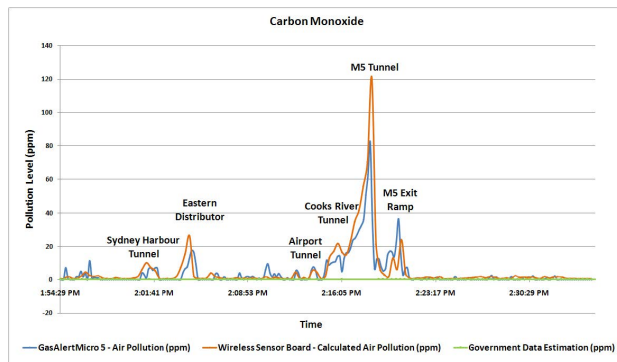
(a) Field Trial 1
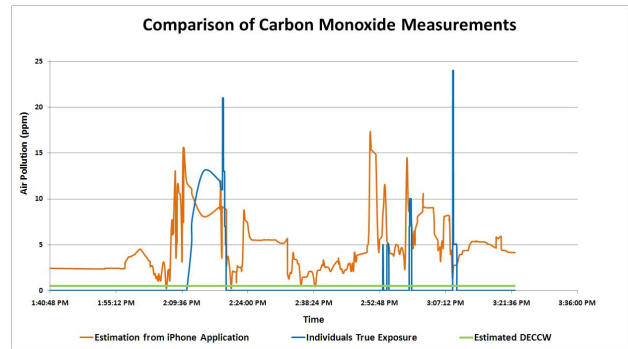
(b) Field Trial 2

Fig. 10.   (a) Field trial 1: validating our sensor unit, and (a) Field trial 2: validating our personalized estimation app

– 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 readings obtained from the nearest government monitoring stations, 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.

The second field trial we illustrate in Fig. 10(b) corresponds to multiple drivers who carry our sensor units and contribute data, while the subject user does not have a unit, and merely uses their iPhone app. The objective is to see if the pollution exposure for this subject user can be estimated from the data contributed by other users. The figure shows that the estimated values (orange curve) corroborate to some extent (though not very accurately) with the true exposure recorded by the commercial meter (blue curve). We believe this fit can improve significantly with deployment density, and we hope to achieve this as we construct more units that can be given out to people who are keen to contribute pollution data. The inaccuracy of our estimates not withstanding, it is worthwhile to note that the exposure estimated from the government data (green curve obtained from data from the department of Environment Climate Change and Water or DECCW) is close to zero and hardly representative of pollution individuals are exposed to on Sydney's roads.

B. 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 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 stabilisation 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 we have in fact lost two units that dropped off and got smashed on the road, 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.

VI. Conclusions

In this paper we have described our efforts over the past 2-3 years in building and deploying a low-cost participatory system for urban air pollution monitoring in Sydney. The aims of the project are noble, to get a better understanding of urban air pollution and to empower citizens with information about their personal exposure, and what steps they can take (e.g. route planning) to better manage their air pollution exposure. In spite of no funding for this project, the collective efforts of over a dozen students over this period have enabled us to build a working system, including hardware units, server database software, and mobile apps. We have demonstrated that the system, though fraught with challenges, is realizable, and can be applied world-over, particularly in pollution-heavy countries, to better understand urban air pollution and its health impacts. The economic model underlying large-scale deployment is yet to be validated (crowd-funding may provide a partial answer), but the potential of a low-cost crowd-sourced pollution monitoring system has been demonstrated, and might provide a more viable alternative to waiting for governments of the world to act on this important but ignored problem.

REFERENCES

[1] K. Aberer, S. Sathe, D. Chakraborty, A. Martinoli, G. Barrenetxea, B. Faltings, and L. Thiele, "OpenSense: Open Community Driven Sensing of Environment," in *Proc. ACM SIGSPATIAL Intl. Workshop on GeoStreaming (IWGS)*, San Jose, CA, USA, Nov 2010, p. Nov. [Online]. Available: http://opensense.epfl.ch

[2] ACAM, "Australian Centre for Asthma Monitoring." [Online]. Available: http://www.asthmamonitoring.org/

[3] A. Bernieri, D. Capriglione, L. Ferrigno, and M. Laracca, "Design of an Efficient Mobile Measurement System for Urban Pollution Monitoring," pp. 77–84, Jul 2012.

[4] M. Boznar and P. Mlakar, *Air Pollution Modeling and Its Application XV*. Springer Netherlands, 2004, ch. Use of Neural Networks in the Field of Air Pollution Modelling, pp. 375–383.

[5] B. Brunekreef and S. Holgate, "Air Pollution and Health," *Lancet*, vol. 360, no. 9341, pp. 1233–42, Oct 2002.

[6] J. Carrapetta, "Haze Watch: Design of a Wireless Sensor Board for Measuring Air Pollution," Unversity of New South Wales, Sydney, Australia, Tech. Rep., 2010.

[7] S. Choi, N. Kim, H. Cha, and R. Ha, "Micro Sensor Node for Air Pollutant Monitoring: Hardware and Software Issues," *Sensors*, vol. 9, pp. 7970–87, 2009.

[8] A. Chow, "Haze Watch - Data Modelling and Visualisation using Google Maps," Unversity of New South Wales, Sydney, Australia, Tech. Rep., 2010.

[9] CSIRO, "Air Pollution Bigger Killer than Road Accidents," 2004. [Online]. Available: http://www.smh.com.au/articles/2004/03/02/1078117396880.html

[10] A. de Nazelle, E. Seto, D. Donaire-Gonzalez, M. Mendez, J. Matamala, M. Nieuwenhuijsen, and M. Jerrett, "Improving Estimates of Air Pollution Exposure through Ubiquitous Sensing Technologies," *Environmental Pollution*, vol. 176, pp. 92–99, May 2013.

[11] P. Dutta, P. Aoki, N. Kumar, A. Mainwaring, C. Myers, W. Willett, and A. Woodruff, "Common Sense: Participatory Urban Sensing Using a Network of Handheld Air Quality Monitors," in *Proc. SenSys Demonstration*, Berkeley, CA, USA, Nov 2009. [Online]. Available: http://www.communitysensing.org/

[12] S. Greaves, T. Issarayangyun, and Q. Liu, "Exploring Variability in Pedestrian Exposure to Fine Particulates (PM2.5) along a Busy Road," *Atmospheric Environment*, vol. 42, no. 8, pp. 1665–76, Mar 2008.

[13] W. Hedgecock, P. Vlgyesi, . Ldeczi, and X. Koutsoukos, "Dissemination and Presentation of High Resolution Air Pollution Data from Mobile Sensor Nodes," in *Proc. ACM SE*, New York, NY, USA, 2010.

[14] S.-C. Hu, Y.-C. Wang, C.-Y. Huang, and Y.-C. Tseng, "Measuring Air Quality in City Areas by Vehicular Wireless Sensor Networks," *The Journal of Systems and Software*, vol. 84, no. 11, pp. 2005–12, Nov 2011.

[15] Imperial and Cambridge, "MESSAGE." [Online]. Available: http://research.cs.ncl.ac.uk/message/

[16] M. Jerrett, A. Arain, P. Kanaroglou, B. Beckerman, D. Potoglou *et al.*, "A Review and Evaluation of Intraurban Air Pollution Exposure Models," *Journal of Exposure Analysis and Environmental Epidemiology*, vol. 15, pp. 185–204, 2005.

[17] M. Kelly, "Design of a Wireless Sensor Board for an Air Pollution Monitoring System," Unversity of New South Wales, Sydney, Australia, Tech. Rep., 2012.

[18] D. Liao, D. Peuquet, H.-M. Lin, Y. Duan, E. Whitsel, R. Smith, and G. Heiss, "National Kriging Exposure Estimation," *Journal of Environmental Health Perspectives*, vol. 115, pp. 338–339, 2007.

[19] J. Liu, "Calibration of a Wireless Sensor Board for Measuring Air Pollution," Unversity of New South Wales, Sydney, Australia, Tech. Rep., 2012.

[20] M. Mead, O. Popoola, G. Stewart, P. Landshoff, M. Calleja, M. Hayesb, J. Baldovi, M. McLeod, T. Hodgson, J. Dicks, A. Lewis, J. Cohen, R. Baron, J. Saffell, and R. Jones, "The Use of Electrochemical Sensors for Monitoring Urban Air Quality in Low-Cost, High-Density Networks," *Atmospheric Environment*, vol. 70, pp. 186–203, May 2013.

[21] N. Mills, K. Donaldson, P. Hadoke, N. Boon, W. MacNee *et al.*, "Adverse Cardiovascular Effects of Air Pollution," *Nature Clinical Practice Cardiovascular Medicine*, vol. 6, no. 1, pp. 36–44, Jan 2009.

[22] NSW-Environment-Heritage, "Air Quality Index." [Online]. Available: http://www.environment.nsw.gov.au/aqms/aqi.htm

[23] D. Olaru and J. Powell, *Air Pollution Modeling and Its Application XIX*. Springer Netherlands, 2008, ch. What Activity-Based Analysis and Personal Sampling Can Do for Assessments of Exposure to Air Pollutants, pp. 717–718.

[24] W. Physick, M. Cope, S. Lee, and P. Hurley, "An Approach for Estimating Exposure to Ambient Concentrations," *Journal of Exposure Science and Environmental Epidemiology*, vol. 17, pp. 76–83, 2007.

[25] P. Ryan and G. LeMasters, "A Review of Land-use Regression Models for Characterizing Intraurban Air Pollution Exposure," *Journal of Inhalation Toxicology*, vol. 19, no. 1, pp. 127–133, 2007.

[26] K. Schere, "Emergence of Numerical Air Quality Forecast Models and their Applications," in *AMS Annual Meeting*, Atlanta, GA, USA, 2006.

[27] P. Vlgyesi, A. Ndas, X. Koutsoukos, and . Ldeczi, "Air Quality Monitoring with SensorMap," in *Proc. Intl. Conf. Information Processing in Sensor Networks (IPSN)*, St. Louis, MO, USA, Apr 2008. [Online]. Available: http://www.isis.vanderbilt.edu/projects/maqumon

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

[29] N. Youdale, "Haze Watch: Database Server and Mobile Applications for Measuring and Evaluating Air Pollution Exposure," Unversity of New South Wales, Sydney, Australia, Tech. Rep., 2010.

[30] L. Young, C. Gotway, J. Yang, G. Kearney, and C. DuClos, "Linking Health and Environmental Data in Geographical Analysis: Its So Much More than Centroids," *Spatial and Spatio-Temporal Epidemiology*, vol. 1, no. 1, pp. 73–84, Oct-Dec 2009.

[31] B. Zou, J. Wilson, F. Zhan, and Y. Zeng, "Air Pollution Exposure Assessment Methods Utilized in Epidemiological Studies," *Journal of Environmental Monitoring*, vol. 11, no. 3, pp. 475–490, Mar 2009.