Challenges of Drive-By IoT Sensing for Smart Cities: City Scanner Case Study

Amin Anjomshoaa
MIT Senseable City Laboratory
Cambridge, MA 02139, USA
amina@mit.edu

Carlo Ratti
MIT Senseable City Laboratory
Cambridge, MA 02139, USA
ratti@mit.edu

Simone Mora
MIT Senseable City Laboratory
Cambridge, MA 02139, USA
moras@mit.edu

Philip Schmitt
Parsons School of Design
The New School
New York, NY 10011, USA
mail@philippschmitt.com

UbiComp/ISWC'18 Adjunct, October 8-12, 2018, Singapore, Singapore.

ACM ISBN 978-1-4503-5966-5/18/10...$15.00.
https://doi.org/10.1145/3267305.3274167

Abstract
Fixed sensor stations are the primary means to collect environmental data in cities. Yet, their high cost of deployment and maintenance often result in an accurate but geographically-sparse monitoring. We advocate for a drive-by approach to urban sensing in which vehicles are used as sensors to scan the city with a high spatiotemporal resolution. We present City Scanner, a highly-customizable, self-sufficient platform that allows for cost-efficient drive-by sensing. City Scanner modules can be deployed on existing vehicles (e.g. busses and taxis) without interfering with their operations. We describe our first prototype that includes sensing modules for air quality, temperature, humidity, and thermal imaging. We discuss the challenges we encountered during an 8-months deployment of the platform on trash trucks in order to derive implications for the design of future drive-by sensing systems.

Author Keywords
Environmental monitoring; Internet of Things; Smart City

ACM Classification Keywords
C.2.4 [Distributed Systems]: Distributed applications; H.5.2 [User Interfaces]: Graphical user interfaces (GUI)

Introduction

A myriad of sensors are deployed in cities throughout the world, helping us to better understand urban environments in order to provide better services for citizens. Types of sensor range from air and water quality, to temperature and noise pollution, to traffic counters and infrastructure monitoring. The collected data, once processed by advanced analytics solutions, can be: (i) consumed directly by citizens in order to make informed decisions, (ii) made available to public agencies and urban managers for decision-making; and (iii) be fed into machines and systems (e.g. self-driving cars, smart buildings, traffic control systems, etc.) to create elaborated urban services.

In order to capture a dense spatiotemporal dataset, a large network of stationary sensors is needed, often leading to huge costs for sensor acquisition, deployment and maintenance [2]. Recently, low-cost sensors have become available; e.g. for air quality monitoring [5]. Fueled by these advances, a number of projects have developed portable and personal sensors, e.g. [10, 6], with the goal of democratizing environmental data and contribute to citizen science studies.

Likewise, a number of works, e.g. [4, 8], have developed *drive-by* approaches [1]: a family of opportunistic data collection methods that leverage traditional or dedicated vehicles as sensors. With a fleet of vehicles moving throughout the city in different areas and frequency, it is possible to create a networked and mobile approach to environmental sensing. Compared to traditional sensors stations, collected data can provide a higher spatial resolution, higher flexibility, and lower operational costs. For example, Anjomshoaa et al. [1] have estimated that by equipping only 20 random taxis with sensors, it is possible to cover 50% of street segments in Manhattan at least once a day.

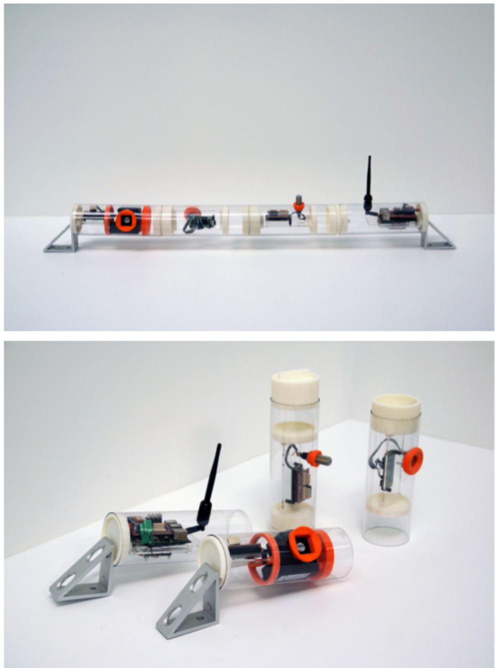

Figure 1: An early prototype of the City Scanner sensing platform

The City Scanner project aims at developing a cost-effective, drive-by modular sensing platform (Figure 1) to assess spatiotemporal environmental phenomena in cities; e.g., patterns in ambient temperature, air quality, road conditions etc. Differently from other drive-by platforms, our sensing modules can be deployed on existing urban vehicles requiring minor or no modification on the hosting vehicles and without interfering with their operations. The platform can be deployed on vehicles which follow either scheduled or unscheduled routes; e.g. municipal buses, garbage trucks, taxi, etc.

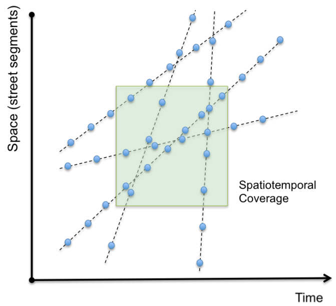

Figure 2: Spatiotemporal sensing quality. Each line represents one vehicle that visits a selected area during a specific time window. Dots are the sensed data and density of dots is considered as the Sensing Quality.

This paper presents the design and development of the first prototype of the City Scanner platform and highlights the challenges that we were facing to build a reliable mobile sensing platform. As a proof of concept, the platform has been deployed on municipal garbage trucks in Cambridge, MA for eight months. The outcomes of this application are discussed. We conclude the paper by deriving implication for the design of future drive-by systems.

Drive-by scanning approaches

Sensors attached to vehicles can achieve a dense spatial coverage using a minimal number of sensing devices. By navigating throughout the city in different areas and time of the day, it is possible to obtain a connected approach to comprehensive environmental sensing. For instance, a random deployment of sensors on urban vehicles such as taxis, can utilize the collective mobility of the network to effectively cover certain street segments. In this context, the density of sensing is defined by number of captured data points in a given time window and a specific urban area such as street segment (cf. Figure 2).

Drive-by approaches have been proven efficient for environmental-related use-cases, where spatial and temporal dimensions are of great importance for monitoring urban environments effectively. Most of the existing environmental datasets are built using sensors deployed in fixed locations brining to data which is limited in spatial, temporal, or both dimensions. Yet, ambient pollutants gradients sensed in precise locations are characterized by high variance even within locations at very close distance and during different time of the day [11].

On the other hand, satellite-based measurements can produce highly accurate spatial content, but with limited temporal coverage. For instance, satellite imagery used to measure surface temperature rely on robust, computationally-heavy mathematical models [9] to predict temperature changes over time.

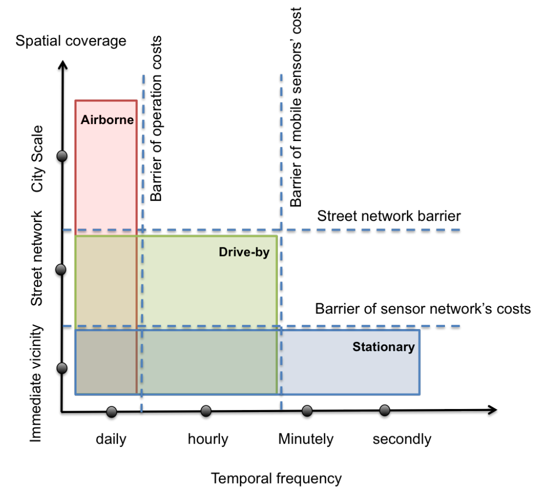

Figure 3: Comparing spatiotemporal sensing approaches.

Figure 3 depicts the limitation of stationary and remote sensing methods and how drive-by sensing can overcome these limitations.

Finally, most experiments with stationary and portable sensors have focused on specific aspects of the urban environment, in particular air quality [3, 7], potholes detection and gas leak detection and using vehicles that were specifically designed and deployed for a specific study.

Design challenges of drive-by sensing systems

Today, there is an increasing number of mobile and sensor-equipped devices in the urban environment. However, these devices are often not tailored for real-time environmental

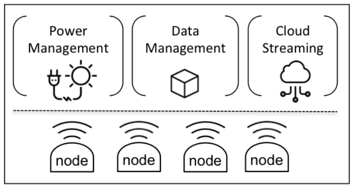

Figure 4: The sensing layer of City Scanner is composed of i) a core component that is responsible for power management, data management, and upstreaming captured data to the cloud servers and ii) a number of sensing nodes that contain a specific sensor and capture data and transmit it to the core component.

sensing. For instance, smart phones include half a dozen of sensors but they usually do not have any built-in environmental sensors. Modern vehicles are also equipped with thousands of sensors but they are merely used for maintenance and diagnostic purposes. In order to add environmental sensing capabilities to vehicles, several technical challenges need to be overcome. Some of these challenges are in common with other IoT systems, e.g. low-energy consumption or real-time data streaming; while other are specific for drive-by IoT systems.

For example, supplying power to a drive-by IoT system can be achieved with multiple competing strategies:

- The system can be connected to the power grid of the hosting vehicle. This usually requires either tapping the wiring system of target vehicle or connecting the device to an available power outlet the vehicle may provide (e.g. mobile charger ports). In both cases, there are concerns regarding the ease of deployment for large number of sensors as well as maintenance and safety issues.
- The system can be battery-powered. Yet, many environmental sensors (e.g. air quality) are not optimized for low-energy consumption. Further replacing batteries enforces a maintenance overhead.
- Renewable energy sources can address some of the described challenges. However, most traditional solutions (e.g. solar panels) are designed and tested for fixed deployment. In order to use them for mobile sensing various performance and feasibility concerns should be considered.

In addition, sensor restraints according to their fitness for drive-by sensing need to be evaluated. Although the rise

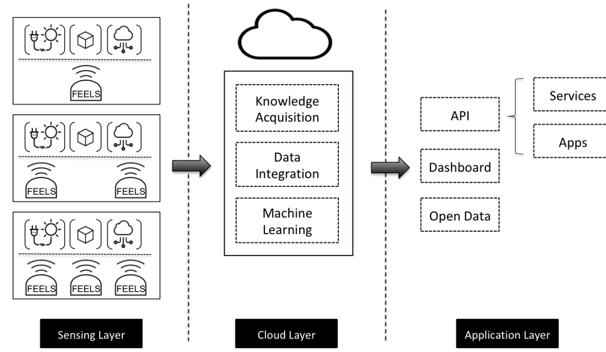

Figure 5: City Scanner platform architecture

of of affordable and portable sensors have lifted many of the hardware limitations that were difficult to cope with in early works; however, most of the sensors specifications do not specify the limits of their configuration for mobile sensing. For instance, low-cost particle counting sensors are vulnerable to airflow disturbance on moving vehicles and the temperature measurements are also impacted by the airflow.

City Scanner Platform

City Scanner aims at democratizing urban sensing with a platform (Figure 5) composed by: (i) low-cost networked and modular sensors for drive-by data capture, (ii) methods to process and aggregate data from multiple sources (iii) different visualization approaches including mobile, situated and ambient interfaces. The platform is currently composed by a prototype of a modular sensing device and a data analysis and visualization platform.

City Scanner's sensing layer (Figure 4) includes a core component and a number of node components which embed specific sensors. The core component provides basic services such as power management, GPS, and data communication for both upstreaming data to cloud servers as well as the node components.

During operation time, node components capture data and send it to the core module; which in turn annotates the captured data with location information, eventually streaming to cloud servers for storage and analysis. Data communication between node components and the core component is achieved over WiFi, with the core module acting as a hotspot. As such, multiple node components can be added to the City Scanner platform, with the number of sensors only limited by restriction dictated by power supply. Because City Scanner's modular architecture is geared towards resiliency, a failure in a node component does not impact or block other nodes or the overall functioning of the system.

Data visualization

As a proof of concept, we developed an interactive visualization, fed by data from the City Scanner platform, that offers discrete instruments to navigate and filter data both in time and space.

The web application, depicted in Figure 6 and available at http://senseable.mit.edu/cityscanner/app, allows to navigate the data in space via the interactive map and in time via a timeline scatter plot. The interface is arranged around the draggable map, which is overlaid with data points. Each dot visible in Figure 6 renders one data point with three characteristics: position on the map (location), color (sensed value) and opacity (age).

The data source to be displayed on the map can be selected by the user via a dedicated menu. In addition, clicking on an area in the map locks the visualization to a specific section. The user can further select points on the timeline, which in turn get highlighted on the map. The timeline also allows for selecting group of data points that can be further filtered for specific time intervals. In this way, only datapoints recorded within the selected timeframe are shown on the map. When the thermal imaging source is selected, it is possible to view instants from the thermal video feed by scrubbing the timeline through the days.

Finally, in the bottom right corner of the interface, a histogram of all sensor values (in a selected timeframe) gives a general overview of the distribution of sensor data. The graph also displays an overall average as well as the average of selected data points. When the user hovers over any datapoint in the visualization, a second line is displayed that allows to compare the selection against the average. As such, one can easily compare the selected neighborhood (e.g. work, home, etc.) with the overall status in the target city.

Deployment case-study

The City Scanner platform was deployed on trash trucks of City of Cambridge, Massachusetts for a period of 8 months. The goal of the case study was to assess viability of using City Scanner for drive-by sensing and reliability of data collected. Further, we wanted to understand feasibility and challenges of using trash trucks as hosting vehicles for our platform.

The trash trucks we employed during the study cover the entire city on a weekly basis, making it an ideal setting for our research goals. Yet during deployment we have encountered a number of technical challenges, which are summarized and discussed in the remaining of this section.

Data Transfer

The City Scanner prototype was equipped with thermal cameras to capture thermal flux variations in the built en-

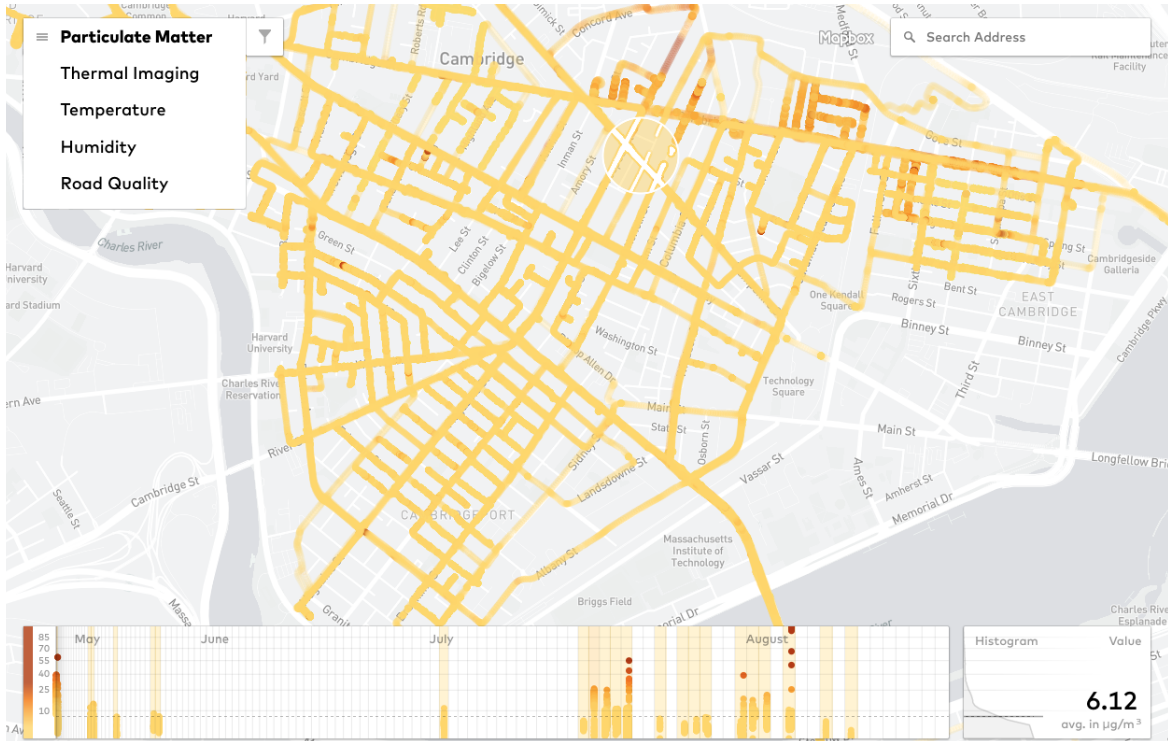

Figure 6: Interactive data visualization interface. Available at http://senseable.mit.edu/cityscanner/app

vironment. Because the captured data was rather heavy (about 1GB per day), we needed an efficient mechanism to transfer data between the sensing node and the core component; which in turn communicates with the cloud storage. For sake of modularity and to keep the platform independent from different proprietary interfaces and protocols, each node was therefore equipped with a dedicated micro-controller to handle the sensor operation and store the captured data locally. Eventually, the data is communicated via a local WiFi network to the core component which offers a WiFi hotspot service.

Although this design breaks the system down into smaller and more controllable components, yet the system consumes more power compared to having multiple sensors to share one micro-controller. Furthermore, mobile network availability and pricing for transferring big amount of data via cellular network to cloud server are still a challenging issue. For instance, fast 4G mobile networks are not

Figure 7: Cityscanner sensor modules prototype

yet available in all areas and the price for transferring big amount of data is not yet reasonable for urban-scale data capturing.

The data size can be reduced by:

- Compressing the data - according with our tests, we can achieve a data compression rate of 0.44. Yet, because data compression algorithms are CPU-intensive, this comes at a higher energy usage that impacts overall autonomy of the sensing modules.
- Discarding repeated frames - during data capture, because the vehicle might stop moving, this cases repeated shots of the same scene. We may drop the repeated frames by comparing each frame with the previous one.

Power

For the current City Scanner prototype we used off-the-shelf products which are typically designed for general purpose and not optimized for low-energy consumption; which impacted the prototype autonomy. For the development of future prototypes, we need to consider to either sourcing energy either from the vehicle power grid or from renewable sources. While the first option is the easiest to implement it creates dependencies that impacts employability. At the same time, a self-sufficient solution that benefits from renewable resources such as solar panels is preferable but more challenging.

Our calculations show that currently the efficiency of mobile solar panels (which reduces the effective solar gain) is not enough for the current requirements of the City Scanner framework. One solution to this problem, is to replace the general purpose micro-controllers with low energy alternatives. Furthermore, it is important to put rationing mechanism in place to control and adjust the sampling rate according to the battery charge.

For the solar irradiance of target area, we should use the Global Horizontal Irradiance (GHI) which is sum of both direct and diffused radiation received on a horizontal plane. For Boston, the average GHI is around 3910 (Wh per square meter per day). We need also to consider the power loss along the chain by 40% and consider only the number of hours the sun is pointed directly at our solar panels. Since the vehicles are moving, the trees and buildings will block the sun and only diffused radiation will be received by the solar panels which in turn increases the required time for charging the batteries. There are a lot of variables, but the best practice for choosing a solar panel is to have 2-3 times as much power as we need. In our case we would need 21 Wh per day so a 40-60 Watt solar panel should be enough; provided that it is exposed to enough solar radiation during the day. In some cases, it may be beneficial to reduce power consumption by programming dynamic sensing properties, e.g., a reduced sampling rate when the vehicle is idle or traveling below a certain speed.

Sensor limitations

Typically, environmental sensors are designed for stationary use. When deployed on a moving vehicle we should consider the impact of motion on the sensor and thus on the captured data. For instance the thermal camera we used in our experiment is very sensitive and fragile. Deploying sensors in a harsh environment such as on a trash truck might impact their reliability, e.g. because of vibrations and water leaks. Furthermore, a moving vehicle can add more bias to the captured dataset. For instance, the extra airflow caused by the vehicle's motion can impact the measured air quality

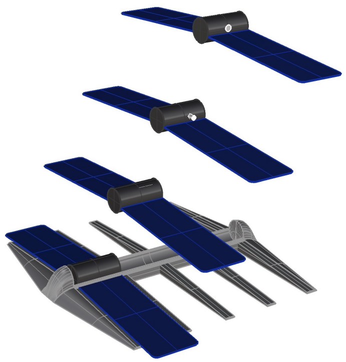

Figure 8: Powering drive-by sensing with solar panels: the power generated by a 50 Watt solar panel with 0.23 efficiency which is exposed to 5 hours daily sunlight is around 23 Wh [50 Watt x 4.5 hours x 0.23 x 0.4 (power loss)]. In case of mobile solar panels the generated power is also reduced by the inefficient tilt angle and shadow of building and trees.

data. To this end, we need to make sure that side effects are minimized and filtered out.

In the specific case of air quality, we selected a particulate matter sensor which considers the airflow rate and sampling time in order to calculate the normalized particle counts. As such, the bias introduced by the extra airflow caused by the moving vehicle is minimized. Another concern about the air quality sensor is the possibility of having biased measurements due the emissions from the trash truck on which the sensor is deployed. However, because emissions from the trash truck are particles of sizes 100 nm, and our selected sensor only records particles larger than 380 nm, we can assume that the truck emissions will not have a big impact on the sensed values.

Conclusions and Future Work

In this paper we reflected on our experience developing and deploying the City Scanner platform prototype. We have described how environmental phenomena in our cities can be efficiently monitored with a small number of urban vehicles. Our approach not only reduces deployment and maintenance costs by an order of magnitude or more, but also offers a way of democratizing data by bringing the data to the doorsteps of citizens. The project architecture follows a centralized IoT regime and is designed to to be self-sufficient, highly-customizable and cost-efficient. The captured data can be used in analytics and machine learning processes combined with multiple visualization approaches. We expect City Scanner to generate value for citizen e.g., by informing policy-making, optimizing city operation and developing innovative cross-domain services.

Building on the challenges highlighted in our case study, future works point in multiple directions. We aim at improving our platform by making it self-sufficient via the use of solar panel and optimizing the efficiency of our hardware and software architectures. At the same time, we plan to switch from WiFi to cellular connectivity in order to achieve near real-time data streaming; e.g. leveraging emerging low-energy IoT cellular protocols like LTE and NB-IoT.

Furthermore, we are planning to extend the platform with new types of environmental sensors and to develop calibration routines and data fusion mechanisms to improve reliability. By providing real-time APIs we intend to foster the development of an ecosystem of third-party application and services that will be served by City Scanner data.

Finally, we plan new deployments of our platform using different vehicles, including public transports, private car and autonomous vehicles; and in diverse cities. These studies will allow us to understand suitability of different vehicles for sensing urban phenomena and to derive sampling characteristics of specific cities.

Acknowledgements

The authors thank Allianz, Amsterdam Institute for Advanced Metropolitan Solutions, Brose, Cisco, Ericsson, Fraunhofer Institute, Liberty Mutual Institute, Kuwait-MIT Center for Natural Resources and the Environment, Shenzhen, Singapore- MIT Alliance for Research and Technology (SMART), UBER, Victoria State Government, Volkswagen Group America, and all the members of the MIT Senseable City Lab Consortium for supporting this research. Also, the financial support of the Austrian Science Foundation (FWF) through the Erwin-Schrödinger Grant No. J3693-N30 is gratefully acknowledged.

REFERENCES

1. A. Anjomshoaa, F. Duarte, D. Rennings, T. Matarazzo, P. de Souza, and C. Ratti. 2018. City Scanner: Building

and Scheduling a Mobile Sensing Platform for Smart City Services. *IEEE Internet of Things Journal* (2018), 1–1. DOI: http://dx.doi.org/10.1109/JIOT.2018.2839058

2. Nuria Castell, Franck R. Dauge, Philipp Schneider, Matthias Vogt, Uri Lerner, Barak Fishbain, David Broday, and Alena Bartonova. 2017. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environment International* 99 (2017), 293 – 302. DOI: http://dx.doi.org/https://doi.org/10.1016/j.envint.2016.12.007

3. Laure Deville Cavellin, Scott Weichenthal, Ryan Tack, Martina S Ragettli, Audrey Smargiassi, and Marianne Hatzopoulou. 2015. Investigating the use of portable air pollution sensors to capture the spatial variability of traffic-related air pollution. *Environmental science & technology* 50, 1 (2015), 313–320.

4. Bret Hull, Vladimir Bychkovsky, Yang Zhang, Kevin Chen, Michel Goraczko, Allen Miu, Eugene Shih, Hari Balakrishnan, and Samuel Madden. 2006. CarTel: A Distributed Mobile Sensor Computing System. In *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems (SenSys '06)*. ACM, New York, NY, USA, 125–138. DOI: http://dx.doi.org/10.1145/1182807.1182821

5. Prashant Kumar, Lidia Morawska, Claudio Martani, George Biskos, Marina Neophytou, Silvana Di Sabatino, Margaret Bell, Leslie Norford, and Rex Britter. 2015. The rise of low-cost sensing for managing air pollution in cities. *Environment International* 75 (Feb. 2015), 199–205.

6. Balz Maag, Zimu Zhou, and Lothar Thiele. 2018. W-Air: Enabling Personal Air Pollution Monitoring on Wearables. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 24 (March 2018), 25 pages. DOI: http://dx.doi.org/10.1145/3191756

7. Grant R McKercher, Jennifer A Salmond, and Jennifer K Vanos. 2017. Characteristics and applications of small, portable gaseous air pollution monitors. *Environmental Pollution* (2017).

8. Prashanth Mohan, Venkata N. Padmanabhan, and Ramachandran Ramjee. 2008. Nericell: Rich Monitoring of Road and Traffic Conditions Using Mobile Smartphones. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems (SenSys '08)*. ACM, New York, NY, USA, 323–336. DOI: http://dx.doi.org/10.1145/1460412.1460444

9. Adar Rosenfeld, Michael Dorman, Joel Schwartz, Victor Novack, Allan C Just, and Itai Kloog. 2017. Estimating daily minimum, maximum, and mean near surface air temperature using hybrid satellite models across Israel. *Environmental Research* 159 (2017), 297–312.

10. Rundong Tian, Christine Dierk, Christopher Myers, and Eric Paulos. 2016. MyPart: Personal, Portable, Accurate, Airborne Particle Counting. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1338–1348. DOI: http://dx.doi.org/10.1145/2858036.2858571

11. Sotiris Vardoulakis, Norbert Gonzalez-Flesca, Bernard EA Fisher, and Koulis Pericleous. 2005. Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris. *Atmospheric Environment* 39, 15 (2005), 2725–2736.