# Mobile Solutions to Air Quality Monitoring

You-Chiun Wang

**Abstract**—Air pollution is one of the most compelling global problems since it poses a serious threat on everyone's health. Governments and people thus put a premium on the reduction of air pollution in the living environment. Consequently, it draws considerable attention on how to efficiently collect air-quality data, especially in cities. In the past, the job of air quality monitoring was usually conducted by installing a few monitoring stations on fixed locations. However, this scheme provides just coarse-grained monitoring, where the resolution of air-quality samplings may be poor. Even worse, it is difficult to move monitoring stations after installation, but the monitoring mission could be often changed. To deal with the problems, many studies propose various mobile solutions to air quality monitoring by equipping gas sensors on mobile devices or vehicles, which allow people to actively and cooperatively detect air pollution in their surroundings. In the chapter, we provide a comprehensive survey of these mobile solutions, and our discussion has four parts. First, we introduce the techniques to evaluate air quality, including an index to report the quality of air and models to predict the dispersion of air pollution. Then, we present the mobile solutions to collect air quality, which can be realized by pedestrians, cyclists, and drivers. Afterwards, we discuss how to analyze raw data collected by smart phones, followed by the issue of reporting sensing data collected by cars. Some research directions and challenges for future mobile solutions to air quality monitoring will be also addressed in the chapter.

**Index Terms**—air quality monitoring, air pollution, mobile solution, participatory sensing, wireless sensor network.

## 1 Introduction

Since the industrial revolution, numerous factories and vehicles have been discharging a large amount of exhaust gases and airborne contaminants to the atmosphere. These air pollutants are harmful to humans, animals, and the ecosystem. The World Health Organization also warns that air pollution has become one of the most serious environmental health risks in the world [1]. Nowadays, people pay more and more attention to environmental protection and health, which prompts the governments and scientists to keep monitoring air quality in our living environment and strive to reduce air pollution.

A traditional solution to air quality monitoring is to install some large, expensive monitoring stations on the dedicated locations in a city [2]. These stations provide large-scale monitoring of air quality around their locations. However, the samples of air-quality data are pretty few, causing the resolution to be poor. Besides, it lacks flexibility to use static monitoring stations, since some sites chosen to install stations such as displaced plants or new parks may become redundant due to the development of a city. Unfortunately, these stations are not easy to move to support dynamic monitoring missions (e.g., to detect air pollution in a new industry district).

In recent decades, the rapid advances of micro-electromechanical systems and wireless communication technologies have made *wireless sensor networks (WSNs)* become popularized [3]. A WSN is composed of many tiny sensor nodes deployed in a region of interest, where each node contains sensing modules to detect events and a wireless transceiver to send its sensing data to a remote sink [4]. Therefore, WSNs provide a cheap and convenient manner to monitor the physical environment. Many WSN applications have been also developed to enrich our life, from health care [5], [6] to intelligent buildings [7], [8], light control [9], [10], security surveillance [11], [12], and smart shopping [13], [14].

*Y. C. Wang is with the Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, 80424, Taiwan. E-mail: ycwang@cse.nsysu.edu.tw*

Thanks to their context-aware sensing capabilities, many studies adopt WSNs in the applications of air quality monitoring. For example, Tsujitaa et al. [15] use a WSN along with monitoring stations to increase samplings of air quality. Each sensor compares its sensing data with the data collected by neighbors and nearby monitoring stations, so as to calibrate its sensing module. Wang et al. [16] deploy a WSN to monitor the concentration of CO (carbon monoxide) and PM (particulate matter) pollutants. Sensors are powered by solar batteries for energy harvesting. Besides, they can turn off transceivers during a suspended period to extend lifetime. Penza et al. [17] install gas sensors on some positions in a city to measure the variation of CO, PM, $H_2S$ (hydrogen sulfide), $NO_2$ (nitrogen dioxide), and $SO_2$ (sulfur dioxide) gases. The collected data are transferred to the format of *data quality objective* defined by European Directive [18]. Brienza et al. [19] develop a sensor suite for people to easily install gas sensors on their houses to monitor air quality in the community, and share their monitoring data through social networking.

Introducing mobility to a WSN further improves its flexibility and allows it to conduct different missions such as moving sensors to replace broken nodes or dispatching sensors to analyze events [20]. *Mobile sensors* can be implemented by putting sensing devices on mobile platforms like smart phones, robots, or vehicles [21]. With the concept of mobile sensors, a number of researchers develop their mobile solutions to air quality monitoring, which allows pedestrians, cyclists, or drivers to carry sensors to measure air quality when they move in a city. Two interesting issues are also arisen from these mobile solutions. In particular, people may prefer using simple (and cheap) gas sensors or even in-built (non-gas) sensors on their smart phones to collect air quality. In this case, how can we estimate the concentration of monitoring pollutants by analyzing the raw data collected by these sensors? In addition, some mobile solutions are implemented by equipping gas sensors on cars to collect air quality in a city. Because the mobility of cars is usually uncontrollable [22], how can we make cars collect air quality on desired positions and report their sensing data

TABLE 1: Six AQI classes defined by U.S. EPA.

| class | AQI range | dedicated color |
|---|---|---|
| good | 0 to 50 | green |
| moderate | 51 to 100 | yellow |
| unhealthy for sensitive groups | 101 to 150 | orange |
| unhealthy | 151 to 200 | red |
| very unhealthy | 201 to 300 | purple |
| hazardous | 301 to 500 | maroon |

accordingly?

This chapter gives a comprehensive survey of existing mobile solutions to the problem of monitoring air quality. It is organized as follows. In the next section, we give background knowledge of some techniques to evaluate air quality. Section 3 presents the mobile solutions to collecting air quality in cities. We discuss how to analyze raw data collected by smart phones in Section 4, and also how to report sensing data collected from cars in Section 5. Then, Section 6 addresses research directions and challenges. Finally, Section 7 concludes this chapter.

## 2 Techniques to Evaluate Air Quality

There are a number of common techniques to evaluate air quality and model air pollution. In this section, we first introduce the technique of *air quality index (AQI)* to measure the degree of air pollution. Afterwards, we present the mathematical models used to simulate the dispersion of air pollutants. Among these models, we detail one popular dispersion model, called *industrial source complex (ISC3)*, to evaluate the effect of air pollution.

### 2.1 The AQI Technique

AQI provides an intelligible index to measure and report to the public how clean or polluted the air is during one day. We take AQI defined by the U.S. Environmental Protection Agency (EPA) [23] as an example, whose range is within $[0, 500]$. It is divided into six classes, where each class is assigned with one color for ease of understanding, as presented in Table 1.

In addition, each AQI class is also associated with some health effects that people may experience when they are doing outdoor activities.

- **Good:** The air pollution poses little or no risk on health, so the outdoor air is basically safe to breathe.
- **Moderate:** Although the quality of air is acceptable, unusually sensitive people, for example, patients who have lung diseases, are suggested to reduce prolonged or heavy outdoor exertion.
- **Unhealthy for sensitive groups:** Most people are not likely to be affected by the air pollution, but members of *sensitive groups* such as the elderly, children, outdoor workers, and patients with asthma, need to reduce prolonged or heavy outdoor exertion.
- **Unhealthy:** Members of sensitive groups should avoid prolonged or heavy outdoor exertion. Everyone else has to reduce prolonged or heavy outdoor exertion for health concern.
- **Very Unhealthy:** Members of sensitive groups have to avoid all outdoor exertion. Everyone else should reduce outdoor exertion.
- **Hazardous:** The government will announce health warning of emergency conditions. Everybody should avoid possible outdoor activities.

To compute the value of AQI, EPA suggests sampling five common kinds of air pollutants, including ground-level $O_3$ (ozone, measured in parts per million, which is abbreviated to 'ppm'), CO (measured in ppm), $SO_2$ (measured in parts per billion, which is abbreviated to 'ppb'), $NO_2$ (measured in ppb), and PM (measured in microgram per cubic meter, which is denoted by $\mu\mathrm{g}/\mathrm{m}^3$). In particular, for each air pollutant $k$, its AQI value $A_k$ is calculated by

$$A_k = \frac{A_{high} - A_{low}}{B_{high} - B_{low}} \times (C_k - B_{low}) + A_{low}, \tag{1}$$

where $C_k$ is the rounded concentration of pollutant $k$, $B_{high}$ is a concentration breakpoint no smaller than $C_k$, $B_{low}$ is a concentration breakpoint no larger than $C_k$, $A_{high}$ is an AQI value corresponding to $B_{high}$, and $A_{low}$ is an AQI value corresponding to $B_{low}$. The suggested values for breakpoints can refer to EPA's technical assistance document in [24]. To estimate the value of $C_k$, EPA suggests taking the average concentration of different pollutants as follows:

- $O_3$ (short term), $SO_2$, and $NO_2$: one hour.
- $O_3$ (long term) and CO: eight hours.
- PM: 24 hours.

The overall AQI value is the maximum value of $A_k$ from all observing pollutants.

### 2.2 Air Pollution Dispersion Models

Air pollution is caused by the emission of pollutants such as particulates or harmful gases to the atmosphere from some sources. According to their dimensions, these emission sources can be divided into four categories. In particular, a *point source* has zero dimension (0D), where there is only one identifiable source which diffuses air pollutants. Examples of point sources include a single smokestack, a flue, or a furnace. On the other hand, a sequence of point sources together will form a 1D *line source*. One can image traffic congestion in a road, where each car is viewed as a point source but they together form a line source. Then, an *area source* is a 2D plane on which air pollutants are emitted, for instance, methane gases diffused from a landfill. Finally, a *volume source* can be treated as an area source but it has a third dimension (i.e., the height). One representative is the smoke emission caused by the forest fire in a mountain.

Given the emission sources in a monitoring region, there are five common mathematical models used to imitate the dispersion of air pollution in that region:

- **Box model:** This model considers a simplified environment, where the given volume of atmospheric air in the monitoring region is constrained to a box-shaped space. Based on the assumption that air pollutants are homogeneously distributed, the box model calculates the average concentration of pollutants in the space. Consequently, the box model cannot provide accurate estimation of air pollution dispersion due to its impractical assumptions.
- **Lagrangian model:** In the Lagrangian model, we assume that an observer follows along with the *emission plume* of air pollution, which is a flow of pollutants in the form of smoke or vapor discharged into the air. Suppose that the motion of each pollution plume parcel (i.e., a particle) complies with the random-walk mobility model [25]. The Lagrangian model associates

with a mobile reference system for particles to predict their trajectories when they move in the air. Then, the dispersion of air pollution can be estimated according to the statistics of moving trajectories caused by a great deal of particles.

- **Eulerian model:** The Eulerian model can be viewed as one variation of the Lagrangian model, where it also computes the moving paths of particles inside the emission plume of air pollution. However, this model considers that an observer is watching the emission plume of air pollution as plume goes by. Besides, instead of using a mobile reference system, the Eulerian model adopts a fixed 3D Cartesian coordinate system to track the trajectories of particles when they leave from their initial positions.
- **Dense-gas model:** As its name would suggest, the dense-gas model aims at simulating the diffusion of gas pollution plumes which are heavier than the general air (usually toxic gases such as the leakage of chlorine from a trunk). When a stream of dense gas is injected into the flowing air, it may produce a wide and flat plume at the ground level. There are many variations of this model, including ALOHA, HGSYSTEM, SLAB, SCIPUFF, PHAST, and TRACE. The work of [26] gives a comparison of these six variations in actual railcar accidents.
- **Gaussian model:** This model adopts a Gaussian distribution to evaluate the dispersion of air pollutants. In other words, the dissemination of pollutants has a normal probability distribution. The major application of the Gaussian model is to estimate the diffusion of continuous, floating air pollution plumes from ground-level or elevated emission sources. However, the Gaussian model can be also used to estimate the dispersion of non-continuous air pollution plumes, which is usually called the *puff model*.

These dispersion models are useful for scientists to evaluate the effect of industrial districts on air quality or forecast AQI after some disasters, for example, haze and smog caused by a large-scale wild fire. Among these models, the Gaussian model is the most popular one to evaluate the diffusion of air pollution due to its flexibility.

### 2.3 The ISC3 Model

ISC3 ('3' indicates the version) is developed from the Gaussian model to evaluate both diffusion and sedimentation of pollutants in the air, which is able to estimate different types of emission sources including point, line, area, and volume sources discussed in Section 2.2. It can also deal with the separation of point sources. In addition, ISC3 is usually adopted to analyze some characteristics of air pollutants, such as settling and dry deposition of particles, effect of down-wash, and limited terrain adjustment.

In ISC3, the concentration of air pollutants on a position $(x, y, z)$ in the 3D space is calculated as follows:

$$\hat{C}(x,y,z) = \frac{\varphi \xi R V}{2\pi \mu_s \delta_y \delta_z} \times \exp\left[-0.5 \times (y/\delta_y)^2\right], \qquad (2)$$

where $\varphi$ is a coefficient used to convert the output into concentration, $\xi$ is a coefficient of disintegration, which is used for certain pollutants with a half-life period (e.g., $SO_2$), $R$ is the discharge rate of the pollutant (measured in grams per second), $V$ denotes the reflection in the vertical direction, $\mu_s$ is the wind's speed (measured in meters per second), $\delta_y$ is the diffusion factor in the horizontal direction, $\delta_z$ is the diffusion factor in the vertical direction, and $\exp[\cdot]$ denotes the exponential function. In Eq. (2), we use parameter $V$ to measure the pollutant concentration in the earth surface or the inversion layer of the atmosphere. Given the height $H$ of an emission source, it can be derived by

$$V = \exp\left[-0.5 \times \left(\frac{z-H}{\delta_z}\right)^2\right] + \exp\left[-0.5 \times \left(\frac{z+H}{\delta_z}\right)^2\right]. \qquad (3)$$

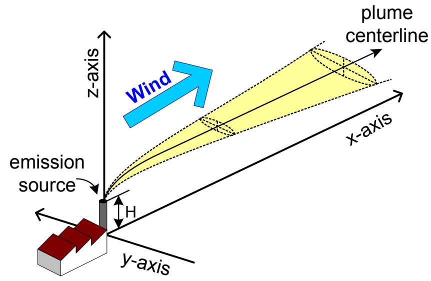


Fig. 1: An example of the ISC3 model.

Fig. 1 illustrates an example of ISC3, where the smokestack has a height of $H$ and the wind blows along the $x$ axis. In this case, the air pollutant spreads in both horizontal and vertical directions, which depends on the parameters of $\delta_y$ and $\delta_z$, respectively. The colored area points out the range of the emission plume of air pollution. It is worth of noting that the coefficients in Eqs. (2) and (3) will be determined by the weather and temperature. In addition, EPA suggests the minimum observing period of ISC3 to be one hour.

## 3 Mobile Solutions to Collecting Air Quality

In this section, we discuss the mobile solutions to collecting air quality. Based on their collecting methods, we classify these solutions into three categories: *pedestrian-based*, *bike-based*, and *car-based* solutions. The pedestrian-based solutions ask pedestrians to carry smart phones along with gas sensors and walk in the monitoring region to collect air-quality data. On the other hand, bike-based and car-based solutions equip gas sensors and communication devices on bikes and cars, respectively, which support the collection of air quality in a much larger region. These solutions also have different mobility models [25]. Generally speaking, pedestrians may follow either random waypoint or reference-point group mobility models. On the other hand, the Manhattan grid model is suitable to depict the moving behavior of bikes and cars, as they usually move along roads and streets.

### 3.1 Pedestrian-based Solutions

Many people choose smart phones as their major computing and communication devices. Smart phones are programmable, so most phone vendors provide their APP stores to allow developers delivering new applications to the public. Moreover, each smart phone can be viewed as a sensor suite, as it usually has accelerometer, camera, digital compass, GPS (global positioning system) receiver, and microphone. Therefore, a new sensing paradigm called *participatory sensing* [27] is proposed to allow people using smart phones to collect environmental information on their own. Based on the motivation, a number of studies develop their mobile solutions for pedestrians to collect air quality by their smart phones and exterior sensors.

Nikzad et al. [28] propose a participatory sensing system called *CitiSense*, which allows pedestrians to carry smart phones and wearable sensor boards to monitor air quality throughout one day, especially during times when exposure to air pollutants will be the highest, for instance, during a rush-hour commute. A lightweight sensor board that contains CO, $NO_2$, and $O_3$ detectors is developed for pedestrians to easily carry. It also includes weather-related sensing devices to monitor temperature, humidity, and barometric pressure. The sensor board can communicate with the user's smart phone wirelessly through Bluetooth, whose communication range is below 60 meters (in Bluetooth version 4.0). Besides, an Android-based APP is developed to display the current air quality on the smart phone and also help the user post the collected data to social networks such as facebook and twitter. The APP reports the last estimated AQI along with the health-effect class presented in Table 1. Each sensor reading collected by the sensor board is tagged with the user's position gotten from GPS. However, since the GPS receiver is an energy-consuming module, it is used only when the user is moving. Moreover, CitiSense has a web interface to let users browse through the readings of air quality that they collected on any given day, which is displayed on the Google Maps. Each sensor reading is placed on the map as a color-coded and numbered marker, which are ordered by its monitoring time. Therefore, users can identify the hotspots of air pollution in locations where they passed through during the commute. A prototype of CitiSense was deployed in San Diego, California, U.S., where 16 participators each had a commute journey for at least 20 minutes, and they were regular users of the same social network.

Yang and Li [29] develop a smart sensor system for air quality monitoring, where each pedestrian can carry a smart phone and a box of embedded sensors (called a *sensor unit*) to detect various air pollutants in the surroundings. The smart phone serves as the middleware between the sensor unit and the server. When a user wants to measure air quality, he/she can execute a specific APP installed on the smart phone, which triggers the sensor unit to detect air pollutants and report sensing data. The data of air quality will be immediately displayed on the smart phone, and the user can also feed back these data to the server. Afterwards, the server manages the collected data and present the monitoring result of air quality through a map-based interface. In [29], the following sensors are included in the sensor unit to monitor different types of air pollutants:

- *PM sensor*, which detects particles in the size of around one micrometer (i.e., $10^{-6}\,\mathrm{m}$) in diameter, with the detection range from 0 to $1.4\,\mathrm{mg/m^3}$;
- *CO sensor*, which can estimate the CO concentration from 20 to 2,000 ppm;
- *$CO_2$ sensor*, which measures the concentration of $CO_2$ (carbon dioxide) gas to at most 2,000 ppm, with the maximum inaccuracy of $\pm 50$ ppm;
- *temperature and humidity sensor*, where the humidity measurement range is between 0% and 100% with at most $\pm 2\%$ error, and the temperature measurement range is between -40°C and 80°C with no more than $\pm 0.5$°C inaccuracy;
- *hazard gas sensor*, which is used to detect noxious gases such as ammonia, benzene, nitrogen oxide, and smoke;
- *volatile organic compound (VOC) gas sensor*, which can detect acetone, alcohol, formaldehyde, methanol, nitrogen, styrene, sulfur, and toluene in the air.

With the hazard and VOC gas sensors, users can be warned in real time once they enter a region with high concentration of harmful gases. These sensors are coordinated by an ARM-based microcontroller and their readings are transmitted to the nearby smart phone through a Bluetooth chip. Since all modules in the sensor unit are powered by small batteries, they usually stay in the sleeping state to save energy unless the smart phone sends a command to wake them up. A prototype of the proposed system was demonstrated in Prairie View, Texas, U.S.

Dutta et al. [30] propose an *AirSense* system to let people participate in monitoring air quality in their neighborhood, which is composed of four tiers:

- **Crowd sensing tier:** Participators can carry both sensor suites and smart phones to collect air quality indoor and outdoor, which offers raw data to AirSense. They can also consume the service provided by AirSense (i.e., the analyzed result of air quality monitoring) through their smart phones.
- **Air quality sensing tier:** Each sensor suite sends the collected data to a nearby smart phone through Bluetooth. Instead of sending data periodically, the sensor suite reports its data only when there is a significant change in the measurement of sensor readings. Thus, the Bluetooth bandwidth can be saved accordingly.
- **Data forwarding tier:** Sensing data collected by smart phones will be transmitted to a cloud server via 4G or Wi-Fi connections. Since GPS receivers are the basic modules in most smart phones, these sensing data can be associated with the positions where they are monitored.
- **Data analysis tier:** The cloud server finally calculates the AQI value according to the collected data. It can also construct a pollution footprint for each participator, which is shown on the smart phone in the form of a mobile APP.

To implement the sensor suite, [30] adopts an Arduino board to integrate with multiple gas sensors to monitor PM, $O_3$, $NO_2$, $SO_2$ pollutants and also a Bluetooth device for data transmissions. Arduino is an open-source, single-board microcontroller kits for developers to easily build digital devices and embed sensing modules [31]. On the other hand, a free cloud service provider, called OPENSHIFT [32], is used to manage and analyze the collected data. OPENSHIFT adopts MySQL as its database to store sensing data, where the identification of each sensor suite is selected as a primary key to query data. A prototype of the AirSense system was deployed in Kolkata, India for demonstration.

### 3.2 Bike-based Solutions

Since bikes provide better mobility than human walking and it is easy to ask cyclists to ride along the pre-scheduled routes, some researchers suggest putting multiple sensors and communication devices on bikes to collect environmental information.

Eisenman et al. [33] propose a *BikeNet* framework, whose objective is to use multiple sensors installed on each bike to gather quantitative data related to the ride of a cyclist. In particular, BikeNet provides two types of information collected from these sensors. One is the context in terms of the cyclist's performance, for instance, the riding speed, distance traveled, and calories burned by the cyclist. The other is about the environmental conditions for the ride, for example, the degree of air pollution, allergen, noise, and terrain roughness of the given route. To do so, each bike carries the following sensors and devices to conduct the monitoring job:

- *microphone*, which detects the surrounding noise;
- *magnetometer*, which detects the moving direction;
- *pedal speed monitor*, which is used to estimate the amount of calories burned by the cyclist;
- *inclinometer*, which measures the angle of slope;
- *lateral tilt*, which has the similar purpose of inclinometer;
- *stress monitor*, which detects the galvanic skin response;
- *speedometer*, which measures the moving speed;
- *GPS receiver*, which acquires the cyclist's position;
- *$CO_2$ meter*, which detects the potential air pollution.

These sensors and devices are wirelessly connected via ZigBee, whose physical communication range is around 10 to 20 meters. In addition, cyclists carry smart phones with cameras to take snapshots from the surroundings and collect the readings from the sensors. There are multiple Wi-Fi and GSM (global system for mobile communications) base stations deployed along some pre-planned paths where cyclists will follow to ride. When a cyclist rides close to a base station, the smart phone transmits sensor readings and snapshots to the back-end servers for analysis. The analyzed data will be visually displayed on web portal. BikeNet aims at improving cyclist experience, but it also provides air quality monitoring. In particular, a $CO_2$ map of streets where a cyclist ever rode through is also shown on the web portal. The $CO_2$ map identifies the regions with high $CO_2$ concentration, which can be used to warn cyclists not to ride in these regions for health concern. A prototype of BikeNet was deployed in Handover, New Hampshire, U.S. for demonstration.

Vagnoli et al. [34] also use bikes to develop a *SensorWebBike* framework for air quality monitoring, which is composed of three major components:

- *Arduino-based sensor platforms*, which are installed on bikes to monitor urban air quality and weather parameters;
- *GeoDatabase*, which is a database to store and manage the collected data;
- *web application*, which helps users view, query, and analyze the information of air quality.

In particular, SensorWebBike adopts an Arduino board to integrate multiple gas sensors to monitor different types of air pollutants, including CO, $CO_2$, $O_3$, $NO_2$, and $CH_4$ (methane). In addition, the Arduino board also contains noise, humidity, and temperature sensors used to collect the weather information. Participators can ride bikes equipped with the Arduino-based sensor platforms to collect the data of air quality and weather condition in a city, and report their monitoring data through

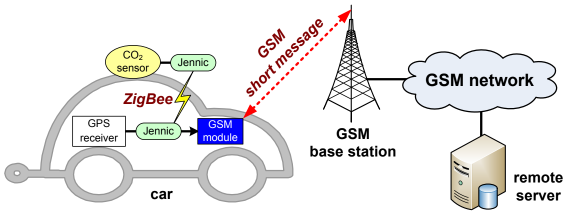


Fig. 2: The car-based mobile solution proposed in [36] for $CO_2$ monitoring.

GPRS (general packet radio service) communications. The collected data are maintained by GeoDatabase, which follows the data format defined by the open geospatial consortium (OGC) for inter-operability, where OGC is an international organization committed to making quality open standards for the global geospatial community [35]. Besides, the web application is developed by using Java2EE to provide the visualization of sensing data in GeoDatabase with multiple formats such as tabular, chart, and geographic map. SensorWebBike was adopted to monitor air quality in the city of Siracusa, Italy as a case study.

### 3.3 Car-based Solutions

In general, cars move much faster and in longer distances than pedestrians and bikes, and they can move in some regions where pedestrians or cyclists are prohibited to enter, for example, highways and freeways. So, it attracts attention to use cars as mobile platforms to carry sensors to conduct monitor jobs in urban areas [21]. Therefore, a number of studies also use cars for the application of air quality monitoring.

Hu et al. [36] develop a mobile solution by cars to monitor the concentration of $CO_2$ gas in urban areas, as shown in Fig. 2. Specifically, each car is equipped with four components: *$CO_2$ sensor*, *GPS receiver*, *GSM module*, and *Jennic board*. The $CO_2$ sensor is installed outside the car (e.g., the windscreen) to collect the surrounding $CO_2$ concentration. Both the GPS receiver and the GSM module are placed inside the car. The $CO_2$ sensor and the GPS receiver each connects with a Jennic board [37], which supports ZigBee communications. On the other hand, the GPS receiver adopts an RS232 interface to communicate with the GSM module. By combining the positioning information from the GPS receiver and the sensing data from the $CO_2$ sensor, the GSM module then periodically transmits the monitoring data to a nearby GSM base station through GSM short messages, which have the following data format [38]:

$$\langle \text{6-byte time, 6-byte } CO_2 \text{ reading, 11-byte latitude, 11-byte longitude} \rangle.$$

For example, when the base station acquires a GSM short message of "(151055, 000405, 2447.3630N, 12060.8732E)", it points out that a car detects the $CO_2$ concentration of 405 ppm on the geographic location of 2447.3630 degrees north latitude and 12060.8732 degrees east longitude at time 15:10:55 (in the format of hour:minute:second). A 16-node prototype was implemented to monitor $CO_2$ concentration in Hsinchu City, Taiwan. The monitoring result was presented on the Google Maps, on which each dot indicated the location where a car collected sensing data, and its color showed the range of detected $CO_2$ concentration.

Sivaraman et al. [39] propose a HazeWatch project that targets at fine-grained spatial measurement of air pollution by cars in Sydney, Australia. In HazeWatch, a driver can choose to mount either a cheap but simple metal oxide sensor or an expensive but sophisticated electrochemical sensor on the car to collect the concentration of CO, $NO_2$, and $O_3$ pollutants. Moreover, some commercial monitors which provide more accurate detection of the pollutants are installed on the roadside to calibrate the readings of sensors on cars. On the other hand, HazeWatch relies on the GPS and 3G capability of smart phones to report positions and sensing data to the server, where the collected data from sensors are transmitted to the smart phone via Bluetooth communications. The server software is composed of three layers:

- **Database layer:** It stores sensing data and provides a simple interface to extract and filter these data. The database layer is implemented by MySQL.
- **Model layer:** This layer gives an abstraction of the collected data, which can return the data of air quality from any location in the monitoring region.
- **Web-server layer:** The layer presents the data (from the model layer) to users, which is displayed in the form of web pages and maps.

However, it is infeasible to collect air quality on every point in the monitoring region, so two interpolation methods are used by the model layer to estimate the value of air quality on those points without actual sensing data. The *inverse-distance weighting method* estimates the pollutant's concentration on a point by assigning weights to all neighboring points, where a point farther away from the interpolation point has a smaller weight. On the other hand, the *kriging method* is based on the calculation of the empirical semi-variogram over the data, where variogram is usually used to describe the degree of spatial dependence of a spatial random field or stochastic process. This method can be implemented by clustering pairs of data points into bins that have similar distances, and plotting the semi-variance of each bin as a function of distance which corresponds to that bin. Then, the interpolation weights are estimated by solving a system of linear equations derived from these bins. The kriging method is more complicated, but it provides more accurate data estimation than the inverse-distance weighting method, especially when the points with actual sensing data are sparsely distributed in the monitoring region.

Devarakonda et al. [40] propose two mobile sensing models to use cars for air quality monitoring in a city:

- **Public transportation infrastructure:** In this model, buses are used as a mobile platform to collect air quality, where they will periodically move along the fixed routes (usually along high volume roads). To do so, each bus is installed with a mobile sensing box, which contains an Arduino board to integrate with multiple devices including a 3G communication module, a GPS receiver, a CO sensor, and a dust sensor (to monitor the PM pollutant).
- **Social community-based sensing:** Drivers can install a personal sensing device on their cars and register to participate in collecting air quality. The personal sensing device has a CO sensor and it can communicate with the driver's smart phone through a Bluetooth link.

All collected data are geo-tagged (by the GPS receiver in a mobile sensing box or a smart phone) and sent to a central server through the cellular network. Afterwards, the server translates these sensing data to AQI values and adopts two user interfaces, *maker map* and *heat map*, to display the degree of air pollution on a web page. The maker map is composed of data makers, each corresponding to a location where the sensing data are collected. When the user clicks on a data maker, it will show the related information such as the monitoring time, GPS coordinates, and pollutant concentration. On the other hand, the heat map illustrates all available measurements with gradient color display, where higher pollutant concentration is represented by higher ranked color in the color spectrum. The proposed system was demonstrated in two U.S. cities to monitor the CO pollutant, including Turnpike, New Jersey, and Staten island, New York.

### 3.4 Discussion

Table 2 compares the mobile solutions discussed in Section 3. Because most smart phones have Bluetooth modules, many solutions choose to use the Bluetooth protocol for sensors to transmit their sensing data to nearby smart phones. However, both [33], [36] adopt the ZigBee protocol for sensors and devices to communicate with each other, while [34] integrates all sensors into one single board. Unlike the bike-based or car-based solutions, pedestrians usually carry small batteries as the power supply for sensors and smart phones. Therefore, energy is a critical concern in the pedestrian-based solutions. To save energy of devices, [28] turns on the GPS receiver only when necessary. The solution in [29] makes sensors sleep until the smart phone wake them up. Besides, [30] allows sensors to report data only when there are significant changes in sensor readings.

For user interface, the solutions in [28]–[30], [39] develop APPs for users to submit their collected air-quality data and obtain the analyzed result through their smart phones. Most solutions adopt the Google Maps to display the monitoring result of air quality. However, [33] and [34] choose to use their own maps to demonstrate the result. In addition, the solutions in [28], [30], [40] adopt AQI discussed in Section 2.1 to display their measurement of air quality. Finally, both [33], [36] aim at collecting $CO_2$ concentration in a city. Other solutions allow users to monitor two or more types of air pollutants in their surroundings.

## 4 Analyzing Raw Data Collected by Smart Phones

Due to the budget consideration, some people may use simple gas sensors linked to their smart phones to measure air quality. Moreover, they may prefer directly using in-built sensors of smart phones (e.g., cameras) to monitor certain air pollutants. Therefore, a few research efforts propose different approaches to analyze raw data collected by smart phones to provide more accurate monitoring result of air quality.

Hasenfratz et al. [41] connect a smart phone with an $O_3$ sensor through its USB (universal serial bus) interface to support participatory air quality monitoring. To detect the ground-level $O_3$ concentration, they measure the resistance of the sensor's $SnO_2$ (tin dioxide) layer. In particular, the smart phone polls the $O_3$ sensor to acquire its raw sensor readings every 100 ms, which contain the $SnO_2$ layer's resistance $R$ and the on-board temperature $T$. Since the value of resistance highly depends

TABLE 2: Comparison on the mobile solutions to air quality monitoring.

| mobile solution | collecting method | device link | city for demo | energy saving | APP | map | AQI | air pollutants CO | $CO_2$ | $NO_2$ | $O_3$ | PM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [28] | pedestrians | Bluetooth | San Diego | ✓ | ✓ | Google | ✓ | ✓ | | ✓ | ✓ | |
| [29] | pedestrians | Bluetooth | Prairie View | ✓ | ✓ | Google | | ✓ | ✓ | | | ✓ |
| [30] | pedestrians | Bluetooth | Kolkata | ✓ | ✓ | Google | ✓ | | | ✓ | ✓ | ✓ |
| [33] | bikes | ZigBee | Handover | | | other | | | ✓ | | | |
| [34] | bikes | wired | Siracusa | | | other | | ✓ | ✓ | ✓ | ✓ | |
| [36] | cars | ZigBee | Hsinchu | | | Google | | | ✓ | | | |
| [39] | cars | Bluetooth | Sydney | | ✓ | Google | | ✓ | | ✓ | ✓ | |
| [40] | cars | Bluetooth | Turnpike | | | Google | ✓ | ✓ | | | | ✓ |

on the temperature, we can estimate the value of temperature-compensated resistance by

$$R_t = R \times \exp[K(T - T_0)], \quad (4)$$

where $T_0$ is the reference temperature and $K$ is a coefficient used to adjust the difference of temperature. In general, $T_0$ is set to 25°C and $K$ is set to 0.025. Because the response curve of the $O_3$ sensor is quasi-linear with respect to the concentration of $O_3$ pollutant, the concentration can be approximated by a first-order polynomial as follows:

$$f(R_t, \alpha, \beta) = \alpha + \beta R_t, \quad (5)$$

where $\alpha$ and $\beta$ are two parameters used to calibrate the sensor readings. According to the observation from [42], the spatial dispersion of $O_3$ pollutant in a street canyon is usually kept constant and its concentration would slowly change over time (specifically, in the order of minutes). Based on this observation, the smart phone can stream data sets of sensor readings through a data filter to construct tuples for data calibration. Specifically, given a data set $S$ with calibration tuples $(R_t, M)$, where $M$ is the reference measurement, for example, the sensing data acquired from the nearby monitoring station, we can adopt the least-squares method to determine both parameters $\alpha$ and $\beta$ in Eq. (5) such that the sum of squared differences between $f(R_t, \alpha, \beta)$ and $M$ is minimized:

$$\arg \min_{\alpha,\beta} \sum_{\forall (R_t, M) \in S} [f(R_t, \alpha, \beta) - M]^2. \quad (6)$$

Through the above scheme, we can calibrate the sensor's readings to improve the accurate of monitoring result. However, Eq. (6) is specific to $O_3$ pollutant. It may not be directly applied to other air pollutants.

Liu et al. [43] use in-built cameras on smart phones to estimate the concentration of PM 2.5 pollutant. Their proposed method works based on the relationship between the haze model and the photographic images [44]. Specifically, haze is one atmospheric phenomenon caused by dust, smoke, and PM 2.5 to obscure the clarification of sky, which makes the image look brownish and blurry. In the haze model, given a pixel $x$ on the image, we can estimate its observed image irradiance by

$$O(x) = I(x) \times t(x) + L(1 - t(x)), \quad (7)$$

where $I(x)$ denotes the scene irradiance and $L$ is the global atmospheric light. In Eq. (7), $t(x)$ is a meteorological parameter called *transmission*, which reveals the amount of light that can pass through the atmosphere. Its value is calculated by

$$t(x) = \exp[-\varepsilon d(x)], \quad (8)$$

where $\varepsilon$ denotes the light extinction and $d(x)$ is the scene depth that displays the distance between the object in the image and the participator who takes the photograph. It has been shown in [45] that PM 2.5 particulates have significant effect on the light extinction, so the value of $\varepsilon$ can be approximate to $pM_f$, where $p$ is a constant (usually set to 3.75 in the urban area) and $M_f$ is the concentration of PM 2.5 pollutant. Therefore, we can use Eq. (7) to measure PM 2.5 concentration from the image. In addition, three image features are considered to improve the accuracy of concentration estimation:

- **Spatial contrast:** We can use the decrease of spatial contrast to observe the degradation of image caused by haze, where distant objects in an image with haze will lose its acuity.
- **Dark channel:** The value of dark channel of a given pixel $x$ is the minimum intensity of the three color channels (i.e., red, green, and blue) of the image block around $x$. The dark channel of an image without haze should be zero in theory.
- **HSI color difference:** HSI is the acronym of three terms used in chromatology: hue, saturation, and intensity. The HSI color difference of the sky taken under different weather conditions will change with the visibility and hazy situation.

To construct the prediction model, we should not only take a sequence of photographs $P_{ts} = \{P_{ts}^1, P_{ts}^2, \cdots, P_{ts}^m\}$ at a location $L$ for $m$ days, but also acquire the data of PM 2.5 concentration $C_{ts} = \{C_{ts}^1, C_{ts}^2, \cdots, C_{ts}^m\}$ from PM 2.5 monitoring stations. Then, given a new photograph $P$ also taken at location $L$, we can use Eq. (7) and the above three image features to compare it with the photographs in $P_{ts}$, and estimate the concentration of PM 2.5 pollutant by consulting the data in $C_{ts}$ accordingly. The scheme in [43] provides a cheap way for people to monitor PM 2.5. However, it incurs a high cost to build the prediction model, as a user needs to take many photographs from the same locations in the monitoring region.

# 5 Reporting Sensing Data Collected by Cars

The mobile solutions discussed in Section 3.3 provide large-scale monitoring of air quality in a metropolitan area, since cars can move very long distances. However, since drivers have their own destinations, we may not control the moving directions and paths of cars. Besides, it is not a good idea to ask drivers to move to certain locations to collect air quality, as cars also discharge exhaust fume. Otherwise, the distortion of collected data and traffic jam will occur on these locations. Therefore, the mobility of cars is not intentional [46] but could follow some mobility models in VANET (vehicular ad hoc network) [22]. To provide better monitoring of air quality by cars under the above consideration of mobility, several studies propose adaptive algorithms to control the data reporting procedure of cars.

Mitra et al. [47] adopt *mobile agents* installed on some cars to conduct the mission of air quality monitoring. In particular, mobile agents are migratory programs capable of moving from one node to a neighboring node in the network and being executed at the destination node. Each mobile agent is composed of three components: 1) the program which implements the mobile agent, 2) the current status of the program, and 3) user data. Mobile agents can decide when and where to move on their own, so they are useful to collect sensing data in a mobile WSN. In [47], the remote server creates a few mobile agents in the beginning, each with the configuration information including the target monitoring region, the time period to collect sensing data, and the type of sensing data to be collected. These mobile agents are arbitrarily launched in some cars. During the movement of cars, mobile agents can migrate from one car to another to reach the monitoring region in time. Once the time period expires, mobile agents then transmit their collected sensing data to the remote server (e.g., through a cellular network). The advantage of using mobile agents is that they can move between different cars to collect only the relevant data. In this way, we can reduce network load because each car will discard unnecessary data after its mobile agent leaves [48]. However, since the movement of cars is uncontrollable, the route of a car that the mobile agent currently lodges may not be desirable. For example, the car may be driven away from the monitoring region, or it could be stuck in a traffic jam. In this case, it would be better for the mobile agent to immediately migrate to a neighboring car, or the number of air-quality samplings collected by the mobile agent may not be sufficient. To do so, each mobile agent periodically checks whether its lodging car is still moving towards the target region, or whether the car is stuck in the traffic jam by the following two strategies:

- **Distance strategy:** The mobile agent calculates the Euclidean distance between the lodging car and its destination. If the distance does not decrease as time goes by, there is a high possibility that the lodging car is stuck in a traffic jam. Therefore, the mobile agent will jump to another car within the communication range.
- **Angle strategy:** The mobile agent measures the angle between the moving vector of the lodging car and the straight line to the destination. When the angle increases as time goes by, it means that the lodging car may move away from the destination. Thus, the mobile agent should conduct the operation of migration.

The above two strategies are easy to implement, because the moving direction and geographic position of each car can be acquired by its GPS receiver. Therefore, mobile agents can determine whether to migrate to another car in a short time, and collect as many air-quality samplings as possible. However, when a mobile agent is lodged in a car that currently moves towards the destination but will become isolated soon (i.e., there are no neighboring cars), the mobile agent will have no chance to migrate to other cars.

Hu et al. [49] divide the monitoring region into a 2D array of homogeneous grids, as shown in Fig. 3, and then dynamically adjust the data reporting rates of cars in each grid based on its car density and the variation of pollutant concentration. Consider that it incurs a cost for car drivers to transmit sensing data to a remote server (e.g., via GSM short messages). The objective is to reduce the overall cost while ensuring the accuracy of monitoring result (in particular, obtaining sufficient air-quality samplings to calculate the distribution of pollutant in the monitoring region). Generally speaking, a higher reporting rate should be assigned to a grid where the variation of pollutant concentration increases, and vice versa. Fig. 3 presents some examples. Grids $(4, 2)$, $(5, 2)$, $(4, 3)$, and $(5, 3)$ have high variations of pollutant concentration, so higher data reporting rates should be imposed on these four grids to improve the monitoring accuracy. On the other hand, the pollutant concentration is almost flat in grid $(1, 1)$ but there are many cars in that grid. Thus, a lower reporting rate can be assigned to grid $(1, 1)$ to reduce the amount of data transmissions without significantly reducing the monitoring accuracy. With these observations, two schemes are proposed to dynamically adjust the reporting rate of cars in each grid.

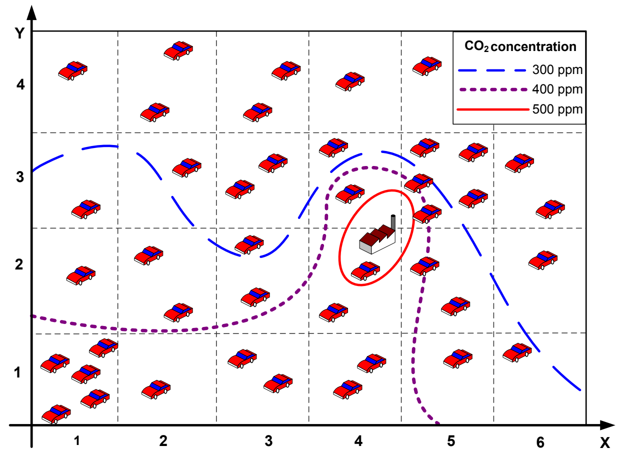


Fig. 3: Homogeneous grid partition by [49] for air quality monitoring.

- **Variation-based scheme:** Let us denote by $\sigma_i$ the standard deviation of pollutant-concentration values collected from grid $G_i$ in the previous time frame. Then, we can estimate the number of air-quality samplings that should be received from grid $G_i$ in the next time frame to keep its monitoring accuracy by

$$S_i^{\text{var}} = \alpha_i^{\text{var}} \times \sigma_i + \beta_i^{\text{var}}, \tag{9}$$

where $\alpha_i^{\text{var}}$ and $\beta_i^{\text{var}}$ are two constants based on the past experience, and larger values imply higher monitoring quality but larger message overhead. In Eq. (9), $\beta_i^{\text{var}}$ is the minimum number of air-quality samplings that we expect to receive from grid $G_i$. Then, the new reporting rate for grid $G_i$ will be set to $S_i^{\text{var}}/n_i$, where $n_i$ is the number of cars in grid $G_i$ which submitted their reports to the server in the previous time frame.
- **Gradient-based scheme:** Let $V_i^{\max}$ and $V_i^{\min}$ be the sets of the highest $\gamma$ ratio and the lowest $\gamma$ ratio of pollutant-concentration values collected from grid $G_i$ in the previous time frame, respectively. The gradient of two air-quality samplings $x \in V_i^{\max}$ and $y \in V_i^{\min}$ is defined by

$$\xi(x, y) = \frac{x - y}{D(x, y)}, \tag{10}$$

where $D(x, y)$ is the Euclidean distance between the two positions where $x$ and $y$ are sampled. We can also

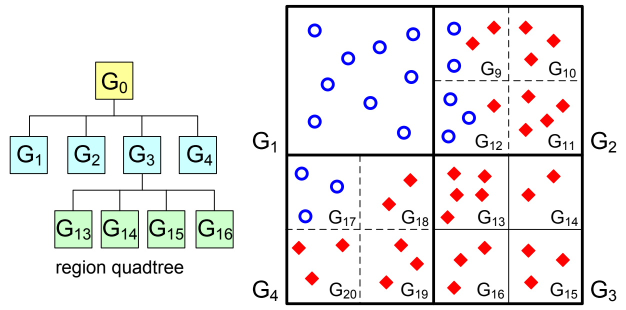


Fig. 4: Heterogeneous grid partition by [50] for air quality monitoring, where dotted-line grids do not appear in the current region quadtree.

measure the average gradient in grid $G_i$ as follows:

$$\xi_i^{\text{avg}} = \frac{\sum_{x \in V_i^{\max},\, y \in V_i^{\min}} \xi(x, y)}{|V_i^{\max}| \times |V_i^{\min}|}, \tag{11}$$

where $|\cdot|$ denotes the number of elements in a set. Then, the necessary number of air-quality samplings expected to be collected from grid $G_i$ in the next time frame is estimated by

$$S_i^{\text{gra}} = \alpha_i^{\text{gra}} \times \xi_i^{\text{avg}} + \beta_i^{\text{gra}}, \tag{12}$$

where both $\alpha_i^{\text{gra}}$ and $\beta_i^{\text{gra}}$ are constants based on the past experience, just like the variation-based scheme. Also, the new reporting rate for grid $G_i$ is set to $S_i^{\text{gra}}/n_i$.

We remark that the gradient-based scheme can provide higher monitoring accuracy than the variation-based scheme, since it takes the positions of air-quality samplings into consideration. In particular, consider two samplings with a fixed amount of pollutant-concentration difference. When these samplings are acquired from two very close positions, the drop of concentration is regarded as more significant than the fluctuation of concentration acquired from two farther away positions. Consequently, the gradient-based scheme collects high-value and low-value air-quality samplings and then measures the gradients of all pairs of samplings between the two sets $V_i^{\max}$ and $V_i^{\min}$, as given in Eq. (11), to increase the monitoring accuracy.

Nevertheless, the performance of variation-based and gradient-based schemes highly depend on the size of grids. In particular, the monitoring accuracy may decrease when the grid size increases, since a large grid results in lower resolution of air-quality samplings. On the other hand, reducing the grid size would increase the overall message overhead, because the cars in each small grid will report more data to the server. To conquer this problem, Wang and Chen [50] propose a heterogeneous grid partition, as illustrated in Fig. 4. Specifically, the monitoring region is recursively quartered and indexed by a *region quadtree*. It is a data structure popularly used to describe a partition of 2D space by iteratively decomposing the space into four equal quadrants. Each tree node in the region quadtree has either four children (i.e., an internal node) or zero child (i.e., a leaf node). To maintain the heterogenous grid partition, four operations are developed, where we call the set of air-quality samplings collected in a grid $G_i$ *$\lambda$-similar* if these samplings belong to the same AQI class discussed in Section 2.1 and the difference between the largest and the smallest samplings does not exceed $\lambda$.

- **No-change operation:** When all air-quality samplings in grid $G_i$ are $\lambda$-similar, it implies that the pollutant concentration keeps steady in grid $G_i$. Therefore, there is no need to change the grid. Grid $G_1$ in Fig. 4 gives an example.
- **Dividing operation:** If some child grids of grid $G_i$ have air-quality samplings that are not $\lambda$-similar, it means that the pollutant concentration may significantly change in grid $G_i$. Consequently, it is better to divide grid $G_i$ to acquire a more fine-grained observation. Grid $G_2$ in Fig. 4 shows this case. Because its child grids $G_9$ and $G_{12}$ have non-$\lambda$-similar air-quality samplings, grid $G_2$ should be further divided into four small grids.
- **Merging operation:** It is a special case of the no-change operation. When grid $G_i$ and its three sibling grids possess only $\lambda$-similar air-quality samplings, we can merge these four grids into the same one, as the current grid partition is too narrow. An example is given in Fig. 4, where grids $G_{13}$, $G_{14}$, $G_{15}$, and $G_{16}$ can be merged into a large grid $G_3$ on account of their similar air-quality samplings.
- **Marking operation:** The operation is a special case of the dividing operation. It is invoked when a grid has non-$\lambda$-similar air-quality samplings, but each of its child grids possesses only $\lambda$-similar air-quality samplings. Grid $G_4$ in Fig. 4 presents an example, where it has two types of $\lambda$-similar air-quality samplings, but its child grids $G_{17}$, $G_{18}$, $G_{19}$, and $G_{20}$ each has only one type of air-quality samplings. For this situation, we prefer not to divide the grid, because each child grid of grid $G_i$ in fact can share the same data reporting rate.

Once deciding the grid partition by the above four operations, we can calculate the data reporting rate of each grid by

$$R_i = \mu \times \frac{\omega(G_i)}{\phi(G_i) \times t_{\text{avg}}(G_i)}, \tag{13}$$

where $\mu$ controls the speed to sample air quality, $\omega(G_i)$ is a baseline for the number of air-quality samplings expected to be collected from grid $G_i$, $\phi(g_i)$ is the traffic density in grid $G_i$, and $t_{\text{avg}}(G_i)$ is the average time that cars stay in grid $G_i$. In Eq. (13), $\omega(G_i)$ is a constant which depends on the application requirement, and the coefficient $\mu$ can be set as follows:

$$\mu = \begin{cases} 0.5 & \text{if all air-quality samplings in grid } G_i \text{ are } \lambda\text{-similar} \\ 2 & \text{otherwise.} \end{cases} \tag{14}$$

Specifically, when the pollutant concentration keeps steady, it is unnecessary to collect a large number of similar air-quality samplings in that grid. Therefore, we halve the data reporting rate by taking $\mu = 0.5$. On the contrary, when there is significant variation in the pollutant concentration, we should double the data reporting rate by taking $\mu = 2$ to capture such high variation. Fig. 4 illustrates an example. We slow down the data reporting rates of grids $G_1$, $G_3$, $G_{10}$, and $G_{11}$, since they cover the subareas where the pollutant concentration remains stable. On the other hand, the data reporting rates of grids $G_4$, $G_9$, and $G_{12}$ should be speeded up to react to the significant change in the pollutant concentration. To verify the performance of this heterogenous grid partition, [50] uses the *simulation of urban mobility (SUMO)* [51] to imitate practical car traffic in a city and the ISC3 model discussed in Section 2.3 to simulate the dispersion of air pollution. Experimental results demonstrate that the proposed scheme significantly reduces message overhead while keeps the monitoring accuracy as compared with the variation-based and gradient-based schemes, which shows the superiority of heterogeneous grid partition.

## 6 Research Directions and Challenges

In this section, we discuss some research directions and challenges for the mobile solutions to air quality monitoring:

- The mobile solutions can combine with incentive mechanisms [52] to encourage more people to voluntarily participate in the monitoring missions. In this way, we can significantly increase sampling data of air quality and thus improve the accuracy of monitoring result. Moreover, it is useful to dynamically adjust the frequency to sample air-quality data in various situations, for example, increasing the sampling frequency when detecting abnormal air pollution, based on the designs of these incentive mechanisms [50].
- Collecting air quality on every position in the monitoring region is evidently infeasible. Therefore, it is a challenge to provide accurate estimation of air pollutant concentration for the positions with just little information or even without any sensing data. Some techniques like the dispersion models of air pollutants discussed in Sections 2.2 and 2.3 and the data mining approaches popularly used in big data analysis [53] would be helpful in the estimation of air quality.
- In practical applications, people may use various types of devices for communications (e.g., smart phones, laptops, or tablet computers). Moreover, they could carry different kinds of gas sensors to measure the concentration of different air pollutants. Consequently, it deserves further investigation on how to efficiently collect sensing data from heterogeneous devices and combine their data with different attributes. Interestingly, this issue has some similarities with the problem of dispatching multi-attribute mobile sensors discussed in [54].
- Since the technology of *unmanned aerial vehicles (UAVs)* [21] and autonomous cars [55] is evolving and getting mature, it is attractive to use these mobile platforms to carry gas sensors and communication devices to collect air quality. In particular, UAVs are able to provide 3D monitoring of air quality, while autonomous cars can move to certain places where people are difficult to enter, for example, the location of toxic gas leakage. In this way, more comprehensive measurement of air quality and detection of air pollution can be achieved.

## 7 Conclusion

Air pollution is a global problem, and it is beneficial for the residents living in cities to keep monitoring air quality and provide the detailed monitoring result to the public in real time. However, the traditional approach of using static monitoring stations to collect air quality may not meet the requirement. Thanks to the development of WSN technology, many researchers propose different mobile solutions to air quality monitoring by allowing people to carry gas sensors to detect air pollution on their own. This chapter discusses existing mobile solutions to collect air quality by pedestrians, cyclists, and drivers, which have different mobility models. Two issues arisen from these mobile solutions are also addressed, including how to analyze raw data collected by smart phones to estimate the concentration of the monitoring pollutants, and how to make cars efficiently collect air quality on desired locations and report their sensing data accordingly. We also point out some research directions and challenges for future mobile solutions to air quality monitoring in the chapter.

## References

[1] World Health Organization, "7 million premature deaths annually linked to air pollution," Mar. 2014. [Online]. Available: http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/

[2] B. Zou, J.G. Wilson, F.B. Zhan, and Y.N. Zeng, "Air pollution exposure assessment methods utilized in epidemiological studies," *J. Environmental Monitoring*, vol. 11, no. 3, pp. 475–490, 2009.

[3] Y.C. Wang, C.C. Hu, and Y.C. Tseng, "Efficient placement and dispatch of sensors in a wireless sensor network," *IEEE Trans. Mobile Computing*, vol. 7, no. 2, pp. 262–274, 2008.

[4] Y.C. Wang and Y.C. Tseng, "Distributed deployment schemes for mobile wireless sensor networks to ensure multilevel coverage," *IEEE Trans. Parallel and Distributed Systems*, vol. 19, no. 9, pp. 1280–1294, 2008.

[5] R. Paradiso, G. Loriga, and N. Taccini, "A wearable health care system based on knitted integrated sensors," *IEEE Trans. Information Technology in Biomedicine*, vol. 9, no. 3, pp. 337–344, 2005.

[6] S. Sarkar and S. Misra, "From micro to nano: the evolution of wireless sensor-based health care," *IEEE Pulse*, vol. 7, no. 1, pp. 21–25, 2016.

[7] L.W. Yeh, Y.C. Wang, and Y.C. Tseng, "iPower: an energy conservation system for intelligent buildings by wireless sensor networks," *Int'l J. Sensor Networks*, vol. 5, no. 1, pp. 1–10, 2009.

[8] N.K. Suryadevara, S.C. Mukhopadhyay, S.D.T. Kelly, and S.P.S. Gill, "WSN-based smart sensors and actuator for power management in intelligent buildings," *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 2, pp. 564–571, 2015.

[9] D. Caicedo and A. Pandharipande, "Sensor-driven lighting control with illumination and dimming constraints," *IEEE Sensors J.*, vol. 15, no. 9, pp. 5169–5176, 2015.

[10] Y.C. Wang and W.T. Chen, "An automatic and adaptive light control system by integrating wireless sensors and brain-computer interface," *Proc. IEEE Int'l Conf. Applied System Innovation*, 2017, pp. 1399–1402.

[11] Y.C. Tseng, Y.C. Wang, K.Y. Cheng, and Y.Y. Hsieh, "iMouse: an integrated mobile surveillance and wireless sensor system," *IEEE Computer*, vol. 40, no. 6, pp. 60–66, 2007.

[12] Y.C. Wang, Y.F. Chen, and Y.C. Tseng, "Using rotatable and directional (R&D) sensors to achieve temporal coverage of objects and its surveillance application," *IEEE Trans. Mobile Computing*, vol. 11, no. 8, pp. 1358–1371, 2012.

[13] P. Lopez-Iturri, L. Azpilicueta, J.J. Astrain, E. Aguirre, E. Salinero, J. Villadangos, and F. Falcone, "Implementation of wireless sensor network architecture for interactive shopping carts to enable context-aware commercial areas," *IEEE Sensors J.*, vol. 16, no. 13, pp. 5416–5425, 2016.

[14] Y.C. Wang and C.C. Yang, "3S-cart: a lightweight, interactive sensor-based cart for smart shopping in supermarkets," *IEEE Sensors J.*, vol. 16, no. 17, pp. 6774–6781, 2016.

[15] W. Tsujitaa, A. Yoshinoa, H. Ishidab, and T. Moriizumi, "Gas sensor network for air-pollution monitoring," *Sensors and Actuators B: Chemical*, vol. 110, no. 2, pp. 304–311, 2005.

[16] C.H. Wang, Y.K. Huang, X.Y. Zheng, T.S. Lin, C.L. Chuang, and J.A. Jiang, "A self sustainable air quality monitoring system using WSN," *Proc. IEEE Int'l Conf. Service-Oriented Computing and Applications*, 2012, pp. 1–6.

[17] M. Penza, D. Suriano, M.G. Villani, L. Spinelle, and M. Gerboles, "Towards air quality indices in smart cities by calibrated low-cost sensors applied to networks," *Proc. IEEE SENSORS*, 2014, pp. 2012–2017.

[18] European Commission, "Air quality – existing legislation," June 2016. [Online]. Available: http://ec.europa.eu/environment/air/quality/legislation/existing_leg.htm

[19] S. Brienza, A. Galli, G. Anastasi, and P. Bruschi, "A low-cost sensing system for cooperative air quality monitoring in urban areas," *Sensors*, vol. 15, no. 6, pp. 12 242–12 259, 2015.

[20] Y.C. Wang, F.J. Wu, and Y.C. Tseng, "Mobility management algorithms and applications for mobile sensor networks," *Wireless Comm. and Mobile Computing*, vol. 12, no. 1, pp. 7–21, 2012.

[21] Y.C. Wang, "Mobile sensor networks: system hardware and dispatch software," *ACM Computing Surveys*, vol. 47, no. 1, pp. 12:1–12:36, 2014.
[22] J. Harri, F. Filali, and C. Bonnet, "Mobility models for vehicular ad hoc networks: a survey and taxonomy," *IEEE Comm. Surveys & Tutorials*, vol. 11, no. 4, pp. 19–41, 2009.
[23] U.S. Environmental Protection Agency, "Air quality index (AQI) basics," Aug. 2016. [Online]. Available: https://www.airnow.gov/index.cfm?action=aqibasics.aqi
[24] U.S. Environmental Protection Agency, "Technical assistance document for the reporting of daily air quality–the air quality index (AQI)," May 2016. [Online]. Available: https://www3.epa.gov/airnow/aqi-technical-assistance-document-may2016.pdf
[25] W.H. Yang, Y.C. Wang, Y.C. Tseng, and B.S. Lin, "Energy-efficient network selection with mobility pattern awareness in an integrated WiMAX and WiFi network," *Int'l J. Comm. Systems*, vol. 23, no. 2, pp. 213–230, 2010.
[26] S. Hanna, S. Dharmavaram, J. Zhang, I. Sykes, H. Witlox, S. Khajehnajafi, and K. Koslan, "Comparison of six widely-used dense gas dispersion models for three actual railcar accidents," *Process Safety Progress*, vol. 27, no. 3, pp. 248–259, 2008.
[27] N.D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A.T. Campbell, "A survey of mobile phone sensing," *IEEE Comm. Magazine*, vol. 48, no. 9, pp. 140–150, 2010.
[28] N. Nikzad, N. Verma, C. Ziftci, E. Bales, N. Quick, P. Zappi, K. Patrick, S. Dasgupta, I. Krueger, T.S. Rosing, and W.G. Griswold, "CitiSense: improving geospatial environmental assessment of air quality using a wireless personal exposure monitoring system," *Proc. ACM Conf. Wireless Health*, 2012, pp. 1–8.
[29] Y. Yang and L. Li, "A smart sensor system for air quality monitoring and massive data collection," *Proc. Int'l Conf. Information and Comm. Technology Convergence*, 2015, pp. 147–152.
[30] J. Dutta, C. Chowdhury, S. Roy, A.I. Middya, and F. Gazi, "Towards smart city: sensing air quality in city based on opportunistic crowd-sensing," *Proc. ACM Int'l Conf. Distributed Computing and Networking*, 2017, pp. 42:1–42:6.
[31] Arduino. [Online]. Available: http://www.arduino.cc/
[32] OPENSHIFT. [Online]. Available: https://www.openshift.com/
[33] S.B. Eisenman, E. Miluzzo, N.D. Lane, R.A. Peterson, G.S. Ahn, and A.T. Campbell, "BikeNet: a mobile sensing system for cyclist experience mapping," *ACM Trans. Sensor Networks*, vol. 6, no. 1, pp. 6–39, 2009.
[34] C. Vagnoli, F. Martelli, T.D. Filippis, S.D. Lonardo, B. Gioli, G. Gualtieri, A. Matese, L. Rocchi, P. Toscano, and A. Zaldei, "The SensorWebBike for air quality monitoring in a smart city," *Proc. IET Conf. Future Intelligent Cities*, 2014, pp. 1–4.
[35] OGC. [Online]. Available: http://www.opengeospatial.org/
[36] S.C. Hu, Y.C. Wang, C.Y. Huang, and Y.C. Tseng, "A vehicular wireless sensor network for $CO_2$ monitoring," *Proc. IEEE Conf. Sensors*, 2009, pp. 1498–1501.
[37] Jennic board. [Online]. Available: https://www.nxp.com/
[38] S.C. Hu, Y.C. Wang, C.Y. Huang, Y.C. Tseng, L.C. Kuo, and C.Y. Chen, "Vehicular sensing system for $CO_2$ monitoring applications," *Proc. IEEE Asia Pacific Wireless Comm. Symp.*, 2009, pp. 168–171.
[39] V. Sivaraman, J. Carrapetta, K. Hu, and B.G. Luxan, "HazeWatch: a participatory sensor system for monitoring air pollution in Sydney," *Proc. IEEE Conf. Local Computer Networks*, 2013, pp. 56–64.
[40] S. Devarakonda, P. Sevusu, H. Liu, R. Liu, L. Iftode, and B. Nath, "Real-time air quality monitoring through mobile sensing in metropolitan areas," *Proc. ACM SIGKDD Int'l Workshop on Urban Computing*, 2013, pp. 1–8.
[41] D. Hasenfratz, O. Saukh, S. Sturzenegger, and L. Thiele, "Participatory air pollution monitoring using smartphones," *Proc. Int'l Workshop on Mobile Sensing*, 2012, pp. 1–5.
[42] S. Vardoulakis, B. Fisher, K. Pericleous, and N. Gonzalez-Flesca, "Modelling air quality in street canyons: a review," *Atmospheric Environment*, vol. 37, no. 2, pp. 155–182, 2003.
[43] X. Liu, Z. Song, E. Ngai, J. Ma, and W. Wang, "PM2.5 monitoring using images from smartphones in participatory sensing," *Proc. IEEE INFOCOM Workshop*, 2015, pp. 630–635.
[44] S.K. Nayar and S.G. Narasimhan, "Vision in bad weather," *Proc. IEEE Int'l Conf. Computer Vision*, 1999, pp. 820–827.
[45] H. Ozkaynak, A.D. Schatz, G.D. Thurston, R.G. Isaacs, and R.B. Husar, "Relationships between aerosol extinction coefficients derived from airport visual range observations and alternative measures of airborne particle mass," *J. the Air Pollution Control Association*, vol. 35, no. 11, pp. 1176–1185, 1985.
[46] Y.C. Wang and Y.C. Tseng, "Intentional mobility in wireless sensor networks," in *Wireless Networks: Research, Technology and Applications*. New York: Nova Science Publishers, 2009.
[47] G. Mitra, C. Chowdhury, and S. Neogy, "Application of mobile agent in VANET for measuring environmental data," *Proc. Int'l Conf. Applications and Innovations in Mobile Computing*, 2014, pp. 48–53.
[48] M. Chen, S. Gonzalez, and V.C.M. Leung, "Applications and design issues for mobile agents in wireless sensor networks," *IEEE Wireless Comm.*, vol. 14, no. 6, pp. 20–26, 2007.
[49] S.C. Hu, Y.C. Wang, C.Y. Huang, and Y.C. Tseng, "Measuring air quality in city areas by vehicular wireless sensor networks," *J. Systems and Software*, vol. 84, no. 11, pp. 2005–2012, 2011.
[50] Y.C. Wang and G.W. Chen, "Efficient data gathering and estimation for metropolitan air quality monitoring by using vehicular sensor networks," *IEEE Trans. Vehicular Technology*, vol. 66, no. 8, pp. 7234–7248, 2017.
[51] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO – simulation of urban mobility," *Int'l J. Advances in Systems and Measurements*, vol. 5, no. 3 & 4, pp. 128–138, 2012.
[52] H. Gao, C.H. Liu, W. Wang, J. Zhao, Z. Song, X. Su, J. Crowcroft, and K.K. Leung, "A survey of incentive mechanisms for participatory sensing," *IEEE Comm. Surveys & Tutorials*, vol. 17, no. 2, pp. 918–943, 2015.
[53] M. Marjani, F. Nasaruddin, A. Gani, A. Karim, I.A.T. Hashem, A. Siddiqa, and I. Yaqoob, "Big IoT data analytics: architecture, opportunities, and open research challenges," *IEEE Access*, vol. 5, pp. 5247–5261, 2017.
[54] Y.C. Wang, "A two-phase dispatch heuristic to schedule the movement of multi-attribute mobile sensors in a hybrid wireless sensor network," *IEEE Trans. Mobile Computing*, vol. 13, no. 4, pp. 709–722, 2014.
[55] G. Bresson, Z. Alsayed, L. Yu, and S. Glaser, "Simultaneous localization and mapping: a survey of current trends in autonomous driving," *IEEE Trans. Intelligent Vehicles*, vol. 2, no. 3, pp. 194–220, 2017.