# MAQS: A Personalized Mobile Sensing System for Indoor Air Quality Monitoring

**Yifei Jiang**[†], **Kun Li**[‡], **Lei Tian**[†], **Ricardo Piedrahita**[*], **Xiang Yun**[§], **Omkar Mansata**[§], **Qin Lv**[†], **Robert P. Dick**[§], **Michael Hannigan**[*], **Li Shang**[‡]
[†] Dept. of CS, [‡]Dept. of ECEE, [*]Dept. of ME, University of Colorado at Boulder, CO 80309, U.S.A.
[§] EECS Department, University of Michigan, Ann Arbor, MI 48109, U.S.A.
[† * ‡] {yifei.jiang, kun.li, lei.tian, ricardo.piedrahita, qin.lv, hannigan, li.shang}@colorado.edu,
[§]{xiangyun,omkar}@umich.edu, [§]dickrp@eecs.umich.edu

## ABSTRACT
Most people spend more than 90% of their time indoors; indoor air quality (IAQ) influences human health, safety, productivity, and comfort. This paper describes *MAQS*, a personalized mobile sensing system for IAQ monitoring. In contrast with existing stationary or outdoor air quality sensing systems, MAQS users carry portable, indoor location tracking sensors that provide personalized IAQ information. To improve accuracy and energy efficiency, MAQS incorporates three novel techniques: (1) an accurate temporal n-gram augmented Bayesian room localization method that requires few Wi-Fi fingerprints; (2) an air exchange rate based IAQ sensing method, which measures general IAQ using only $CO_2$ sensors; and (3) a zone-based proximity detection method for collaborative sensing, which saves energy and enables data sharing among users. MAQS has been deployed and evaluated via user study. Detailed evaluation results demonstrate that MAQS supports accurate personalized IAQ monitoring and quantitative analysis with high energy efficiency.

## Author Keywords
Air quality sensing, location based service, smartphone

## ACM Classification Keywords
C.3.3 Special-Purpose and Application-Based Systems: Real-Time and Embedded Systems

## General Terms
Algorithms, Design, Experimentation, Performance

## INTRODUCTION
In recent years, indoor air quality (IAQ) has drawn considerable attention in both the public and scientific domains, due to the fact that most buildings appear to fall far short of reasonable air quality goals [15]. Statistics [28] from the U.S. Environmental Protection Agency (EPA) indicate that, on average, the indoor levels of pollutants are two to five times higher than outdoor levels and people in the U.S. spend about 90% of their time indoors. Bad indoor air quality influences human health, safety, productivity, and comfort [31, 9]. IAQ is important and different people have different exposure to pollutants. Providing personalized IAQ information has the potential to increase public awareness of the relationship between their behavior and air quality; help people to improve their living environments; and also provide valuable information to building managers, policy makers, health professionals, and scientific researchers.

IAQ monitoring is challenging because indoor air pollutants concentration and human motion patterns each vary spatially and temporally within and across rooms. Existing solutions that require stationary sensors or target mobile outdoor sensing scenarios are inappropriate for personalized IAQ monitoring. Stationary sensing [13] has several limitations: (a) it can only measure the IAQ experienced by those who happen to be near the sensors, and there can be substantial variation in IAQ even within one room and (b) when locations or rooms outnumber people, achieving full coverage with stationary sensors is more expensive than doing so with personalized mobile sensors. Outdoor mobile sensing solutions use GPS localization, which fails indoors and is therefore inappropriate for IAQ applications. Furthermore, existing air quality sensing solutions require multiple types of sensors, each of which covers a subset of pollutants. This can be prohibitively expensive for personalized mobile IAQ sensing.

A mobile sensing system designed for personalized IAQ monitoring must address the following three challenges. First, some existing approaches use proprietary radio frequency and ultrasound technologies for room localization, which require investment in infrastructure and special hardware worn by all users. Others use Wi-Fi based fingerprinting, which requires time consuming pre-characterization and is hampered by device or environment heterogeneity. Second, the mobile sensing devices must be inexpensive, portable, and energy efficient. This limits the number and types of sensors that

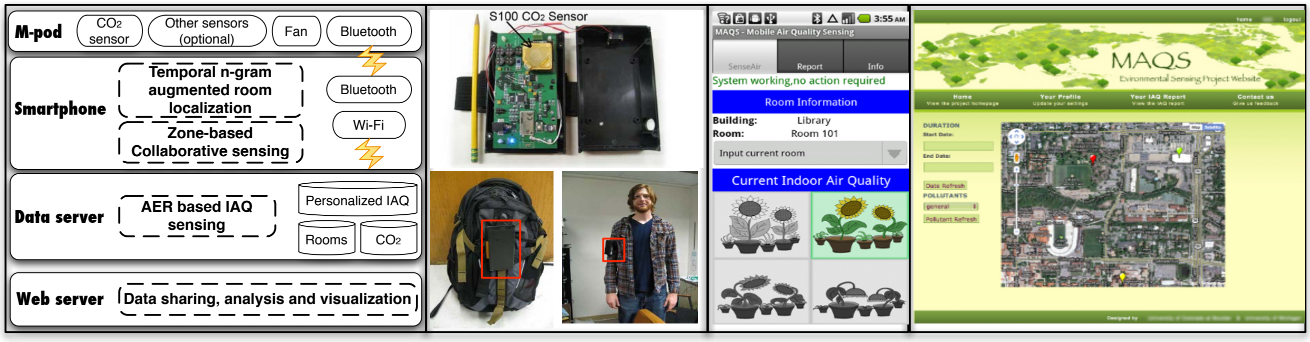



**Figure 1. MAQS: A mobile sensing system for personalized indoor air quality (IAQ) monitoring.**

can be integrated within each mobile device. Achieving high-quality IAQ monitoring with few sensors is challenging. Third, IAQ sensing depends largely on the motion patterns of individual user. This leads to redundant IAQ information when users are near each other, and may lead to gaps in coverage for users who are not presently carrying sensing devices.

This paper describes *MAQS*[1], a personalized mobile sensing system for IAQ monitoring. MAQS estimates human-dependent air quality factors (e.g., $CO_2$ and contagious viruses) using $CO_2$ concentration, and estimates other air quality factors (e.g., volatile organic compounds (VOCs)) using air exchange rates. MAQS integrates smartphones and portable sensing devices to deliver personalized, energy-efficient, IAQ information. MAQS is the first mobile air quality sensing system that achieves high coverage of people in indoor environments. Our work makes the following main technical contributions.

1. A temporal n-gram augmented Bayesian room localization method that is accurate and requires few Wi-Fi fingerprints;
2. An air exchange rate based IAQ sensing method, which measures general IAQ without requiring sensors for various types of air pollutants; and
3. A zone-based proximity detection method for collaborative sensing, which saves energy and enables data sharing among multiple users.

MAQS was evaluated via real-world system deployment and a user study. Our results demonstrate high accuracy (over 96% for room localization and 89% for zone detection) and $2\times$–$8\times$ better energy efficiency. Our quantitative IAQ analysis also reveals that most users are subject to poor IAQ (i.e., high $CO_2$ concentration and low air exchange rate) in a number of rooms.

## SYSTEM OVERVIEW

This section gives a high-level overview of the MAQS system architecture and describes the key components.

As illustrated in Figure 1, MAQS consists of four types of components: (1) *M-pods*, the portable IAQ sensing devices; (2) *smartphones*; (3) a *data server*; and (4) a *web server*. MAQS users carry smartphones and optionally M-pods. The data server communicates with clients and maintains room air quality, $CO_2$, and personalized IAQ data. The web sever allows users to view, analyze, and share IAQ data. There are three main functional units in MAQS: (a) temporal n-gram augmented room localization, (b) air exchange rate based IAQ sensing, and (c) zone-based collaborative sensing.

A MAQS client runs on each smartphone. It monitors the phone's accelerometer readings to detect room entrance and departure events. For the purpose of IAQ monitoring, rooms are defined as enclosed building units with walls, doors, and windows where people spend substantial time (e.g., office, classroom, bedroom) and we ignore transitional spaces indoor (e.g., hallway). Once the client detects that the user has entered a room, the room localization function collects Wi-Fi signals from nearby access points and uses the subsequences of Wi-Fi signals (spatial information) and the user's mobility pattern (temporal information) to determine the current room. The collaborative sensing unit then uses zone-based proximity detection to select specific sensing devices for (collaborative) IAQ monitoring of the room. This is useful since not all smartphone users carry IAQ sensing devices, and sensing devices close to each other (i.e., in the same zone) are largely redundant. As concentration readings of $CO_2$, VOCs, and other air pollutants are collected and transmitted to the server, they are stored in databases and combined with room information (e.g., room ID, volume) for air exchange rate calculation and personalized IAQ analysis. MAQS stops IAQ sensing after detecting a room departure and restarts when another room is entered.

## M-POD: THE PORTABLE IAQ SENSING DEVICE

In this section, we describe the design of our portable IAQ sensing device, the *M-pod*. The M-pod is a wireless embedded sensing, computation, and communication device based on the Arduino BT [2]. It is capable of sensing the concentrations of a number of air pollutants and either storing these data or transmitting them to nearby smartphones via its Bluetooth interface. The main requirements for the M-pod were to accurately sense pollutant concentrations relevant

[1]MAQS stands for Mobile Air Quality Sensing.

**Table 1. M-pod Processor, Wireless Interface, and Sensors**

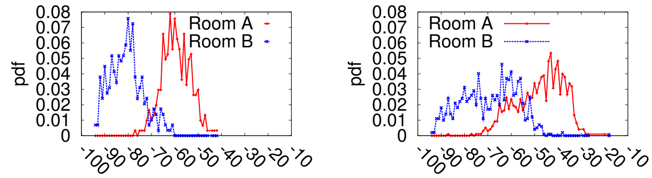
| Hardware specs | MCU | Bluetooth | Battery | Size (inch) |
|---|---|---|---|---|
| | ATMEGA 168 | WT11 | CS HDE160XL | 4.8x2.6 |
| On-board sensors | Temperature | $CO_2$ | Humidity | Light |
| | TMP101 | S100 | HYT271 | GL5528 |

to IAQ and transmit these data, within a compact and long battery life package.

**M-pod Hardware Design.** The M-pod's major components are an 8-bit, 32-pin microcontroller, a Bluetooth module, and up to ten on-board sensors, which are mounted on a custom-fabricated 4-layer printed circuit board. Table 1 lists the processor, wireless interface, and sensors. The M-pod has a humidity sensor, light sensor, two temperature sensors – one upstream to measure ambient air temperature and the other downstream to measure the temperature near the sensors, a $CO_2$ sensor, and low-cost metal oxide gas sensors such as Figaro TGS2600 and E2V MICS2611. The S100 is an accurate, low-power non-dispersive infrared based $CO_2$ sensor. The humidity sensor, temperature sensors, and $CO_2$ sensor are connected to the microcontroller via the I2C interface. The other sensors are attached to the microcontroller's analog to digital converter interface. The metal oxide gas sensors are power gated using PMOSFETs. The M-pod supports in-system and in-field wireless programming via its Bluetooth interface. When developing and evaluating MAQS, we primarily used data from the M-pod's $CO_2$ sensor. The M-pod case is a low-cost off-the-shelf enclosure that has been machined for this application. It can be carried using an armband, or attached to a backpack or briefcase. A 5 V DC fan is mounted to the case. The sensors are positioned to enable uniform airflow. When sensing, the fan moves 2 liters of air per minute, thereby minimizing sensing latency when IAQ changes.

**Energy Consumption.** The M-pod is powered by a 2,200 mA-H Lithium-ion battery, which can be recharged using a standard wall-mounted AC–DC converter. The Lithium-ion battery is protected by an interlock that halts the system when the battery voltage drops below 2.9 V. The M-pod requires 240 mW in low-power mode, in which only the processor and $CO_2$, humidity, and temperature sensors are enabled. It requires 1,080 mW when, in addition, four metal oxide gas sensors are activated. The fan requires up to 105 mW, but this can be reduced via pulse-width modulation. The Bluetooth interface needs to transmit so infrequently in this application that its power consumption has little impact on battery lifespan. The battery lifespan is approximately 5.5 hours if an M-pod is continuously on and greater than 24 hours when in low-power mode.

**Command Processing.** The M-pod processes *low power mode*, *full sensing mode*, *power state inquiry*, and *transmit data* commands, which are generally received from smartphones. IAQ sensor readings are collected every six seconds and stored in the microcontroller's SRAM.




(a) RSS without noise (b) RSS with noise

**Figure 2. Wireless RSS distributions in two adjacent rooms: (a) without noise and (b) with environment and device noise.**

In the low-power mode, the metal oxide sensors are power gated; these sensors contain resistive heating elements and therefore have high power consumptions. The power state inquiry command causes the current mode (low power or full sensing) to be transmitted to the requester. The data transmit command causes the M-pod to upload stored sensor data to the requester (smartphone), which generally takes three seconds.

## ROOM LOCALIZATION

IAQ data is most valuable when it can be associated with the appropriate source room. Room characteristics correlate closely with IAQ and rooms are the basic control units in building management. MAQS users review their IAQ histories and sharing data with others by reference to rooms. Unlike stationary IAQ sensing, M-pods are carried by users, i.e., their locations change as users move. Hence accurate room localization is required for personalized mobile IAQ monitoring.

Researchers have proposed room localization techniques based on Wi-Fi access point received signal strength (RSS) [12, 14]. These methods share two common stages. In the first (training) stage, a database that associates ambient Wi-Fi RSS fingerprints with physical rooms is constructed. In the second (operating) stage, the system identifies the stored Wi-Fi fingerprint that most similar to the one currently being measured, and returns the associated room.

The first stage of our room localization technique is similar to that of Park et al. [12]. All users contribute their Wi-Fi RSS and room information to create a shared database of room fingerprints. This is beneficial as it (1) eliminates the deployment cost for fingerprint presampling and (2) reduces individual users' effort to build the database.

In the second stage, Bayesian room localization models are commonly used [14, 12]. Given a database of fingerprinted rooms $R$ and a Wi-Fi RSS fingerprint represented by a set of access point (AP)-specific signal strengths ($w_i$ for the $i$-th AP), the mobile device (and user) is most likely in room $\hat{r}$:

$$\hat{r} = \arg\max_{r \in R} \left[ \prod_i P(w_i|r)P(r) \right] \quad (1)$$

This model is based on the assumption that the signal strengths observed by the mobile device from different APs are conditionally independent. However, the model fails to address the following two challenges: (1) *device heterogeneity* – different devices may be used for gathering RSS fingerprints and devices might be held differently (e.g., in hand, pocket, or bag) and (2) *environment heterogeneity* – the wireless environment of a room may change over time, due to motion of people and other room contents, influencing the RSS fingerprints gathered by mobile devices. As demonstrated in Figure 2, noise induced by device and environment heterogeneity significantly increase the RSS overlap between adjacent rooms, leading to much lower room localization accuracy. To address these problems, we propose a novel temporal n-gram augmented Bayesian room localization method, which is robust to both environment and device noises.

### N-gram Augmented Bayesian Room Localization

Our key observation is that, although the exact RSS values of each AP may change substantially for different devices and environments, *the ordered sequence of APs based on their RSS values tends to be similar for the same room and inconsistent among adjacent rooms*. For example, the ordering may be $[ap_1, ap_2, ap_3, ap_4, ap_5, ap_6]$ at one time, and $[ap_2, ap_1, ap_3, ap_4, ap_5, ap_6]$ at another time, for the same room. The ordered AP sequences of adjacent rooms are less similar, especially when different APs are observed in these rooms. Intuitively, the ordered AP sequence is useful for room localization because: (1) it captures the inherent correlations among APs, which are stable for the same room yet different for adjacent rooms and (2) it uses the order of RSS instead of their exact values, allowing many sources of device variation and wireless environment variation to be tolerated.

Based on the observations above, we propose an n-gram augmented Bayesian room localization model, which works as follows. Let $s$ be a sequence of $N$ APs ordered in descending RSS values:

$$s = (ap_1, ..., ap_N) \quad w_i \geq w_j (1 \leq i < j \leq N), \tag{2}$$

where $w_i$ is the RSS value of $ap_i$ observed by the mobile device. An n-gram is then defined as a subsequence of length $n$ extracted from the sequence $s$ at position $i$:

$$ngram_i(s) = (ap_i, \cdots, ap_{i+n-1}). \tag{3}$$

The most likely room $\hat{r}$ is determined as follows:

$$\hat{r} = \arg\max_{r \in R} \left[ \prod_i P(ngram_i(s)|r)P(r) \right]. \tag{4}$$

In other words, $\hat{r}$ is the room with the highest probability of having the same ordering of APs in subsequence $ngram_i(s)$.

### Temporal User Mobility for Room Localization

As shown in the experimental results, our n-gram augmented Bayesian room localization model achieves high accuracy when the room has enough fingerprints (more

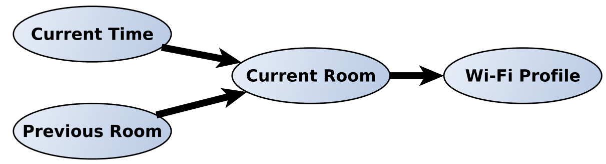



**Figure 3. Bayesian network for room localization.**

than 50 or 100). However, when the number of fingerprints is low, the model accuracy is poorer and users tend to be misclassified into nearby rooms. To remove such spatial errors, we propose to incorporate temporal user mobility information. This is motivated by the following observations.

- A user's current room is closely related to time and weekday, e.g., the user has weekly meetings in the conference room on Tuesday mornings.
- Users can only move among adjacent rooms, and their paths tend to contain patterns. For example, a user usually goes to the conference room from her office instead of from a classroom.

Based on these observations, a user's current room can be predicted based on current time and previous room. As shown in Figure 3, the Bayesian network has three layers: current time and user's previous room (first layer) indicate the user's current room (second layer), and user's current room determines the observed Wi-Fi RSS fingerprint. We also define a set of values to represent some semantic concepts of time, including "day of week", "morning", "afternoon", and "evening". Given a Wi-Fi scan observation $s$, the user's previous room $r'$, and current time $t$ , the user is most likely in room

$$\hat{r} = \arg\max_{r \in R} \left[ P(s, r, t, r') \right] \tag{5}$$

$$= \arg\max_{r \in R} \left[ P(s|r)P(r|t, r')P(t)P(r') \right]. \tag{6}$$

$P(s|r)$ can be computed using our n-gram augmented Bayesian room localization model. $P(r')$ and $P(r|t, r')$ are calculated from the user's mobility history. $P(t)$ can be ignored since it is the same for any room $r$.

### Room Entrance and Departure Detection

Our room localization method requires Wi-Fi scanning, which can be power intensive. For example, Wi-Fi scanning at 1/6 Hz requires 80 mW on average. Using a lower Wi-Fi scanning frequency improves energy efficiency but increases room localization latency.

To address this problem, our MAQS system leverages the smartphone accelerometer to monitor room entrance and departure, and triggers room localization only when a room entrance event is detected. Previous works have used accelerometer to detect arriving at or departing from a place (e.g., a building or outdoor place) [16] , while our work requires the finer-granularity of indoor room-level entrance/departure detection. Specifically, room entrance/departure detection helps to (1) reduce energy use for room localization; (2) locate rooms in a timely fashion; and (3) start/stop IAQ sensing quickly. We sample acceleration at a low frequency mode, 3–10 samples per second, to reduce the energy consumption of the accelerometer. The magnitude of acceleration is calculated over all three axes and the variance of the magnitude within a 5-second time window is used to detect motion. Variance larger than a threshold $\theta$ is taken to indicate motion. The threshold $\theta$ was empirically determined based on measured data. If a node is stationary for 60 seconds, a room entrance is detected. if a node is moving for 10 seconds, a room departure is detected. Our experimental results show that this approach achieves high detection accuracy. The false negative rate is 0.01%, which happens when a user moves between two adjacent rooms quickly. The false positive rate is 7% on average, and false positives can be corrected by our room localization algorithm.

## AIR EXCHANGE RATE BASED IAQ SENSING

Indoor air quality (IAQ) is influenced by multiple air pollutants and sources, including (1) air pollutants generated indoor, such as volatile organic compounds (VOCs), from combustion and off-gassing of paint and building materials; (2) air pollutants introduced from outside via ventilation, e.g., ozone; and (3) air pollutants generated by people, e.g., $CO_2$. It is impractical to install sensors on our portable IAQ sensing devices to monitor all pollutants of interest, as the sensing device would become unreasonably large and require too much power. Additionally, not all pollutant sensors are portable yet, and portable sensors are typically less accurate than stationary sensors.

As shown in previous studies [25, 24, 9, 11], $CO_2$ concentration and ventilation rate are strongly correlated with general IAQ. Therefore, the M-pod monitors $CO_2$ concentration, which is then used to calculate the air exchange rate, i.e., how quickly air is cycled through a room. This rate is used to estimate general IAQ in a room. Specifically, personalized air exchange rates are modeled using changes in $CO_2$ concentration and $CO_2$ generation rate. The rate of change in $CO_2$ concentration depends on the concentration of in-flowing air, the concentration of the out-flowing air, and the internal generation of $CO_2$ in a room. The time derivative of the monitored concentration is given by [24]:

$$V\frac{dC_t}{dt} = G + QC_{ex} - QC_t, \tag{7}$$

where $C_t$ is the internal concentration of $CO_2$ at time $t$, measured in units of ppm (parts per million, i.e., the volume of $CO_2$ over total volume of air). $C_{ex}$ is the external concentration of $CO_2$ (ppm). $G$ is the generation rate of $CO_2$ in the room ($cm^3/s$). V is the room volume ($m^3$). Q is the air change rate ($m^3/s$).

Solving the equation above gives us the formula to calculate air exchange rate $Q$:

$$Q = \frac{V\frac{dC_t}{dt} - G}{C_{ex} - C_t}, \tag{8}$$

where $C_t$ are the continuous $CO_2$ readings from the M-pod. External concentration of $CO_2$, $C_{ex}$, is set to 390 ppm, the globally averaged $CO_2$ concentration at the surface [4], unless local outdoor $CO_2$ concentration is available. To calculate the $CO_2$ generation rate $G$, we assume each person's generation rate is equal to 0.0052 L/s[24], which corresponds to an average-sized adult engaged in office work. At this time, we do not incorporate other possible sources of $CO_2$ such as cooking or smoking. Room volume $V$ can be provided by the user through our system, or it can be calculated from $CO_2$ data based on the Steady-State Concentration Balance Equation, $\frac{dC_t}{dt} = 0$:

$$Q = \frac{G}{C_{ex} - C_t}. \tag{9}$$

After determining $Q$, we apply it to the user's data, in which $\frac{dC_t}{dt} \neq 0$ and $Q$ do not change. Then we can calculate the value of $V$.

The air exchange rate required for good IAQ depends on the size and occupancy of each room. In our system, three metrics are considered for IAQ:

- *Indoor $CO_2$ concentration* is a surrogate for indoor pollutants emitted by humans and correlates with human metabolic activity. The ASHRAE Standard is at most 700 ppm above outdoor $CO_2$ concentration [3].
- *Air changes per hour* is a measure of how many times the air within a defined space (normally a room or house) is replaced per hour. Its value equals the air exchange rate of room divided by room volume. The ASHRAE Standard is at least 0.35 1/h [26].
- *Air flow per person* is the room air exchange rate divided by the number of people in the room. The ASHRAE Standard is at least 7.5 l/s/person [26].

## ZONE-BASED COLLABORATIVE SENSING

In real-world usage scenarios, multiple users are likely to stay in the same room, e.g., in meeting rooms or the library. Such user groups tend to be concentrated in small areas, leading to similar $CO_2$ concentration and IAQ within each group. Through collaborative sensing, we aim to reduce the number of sensing devices that have to run concurrently (thus saving energy), and also enable IAQ data sharing with people who do not carry sensing devices (thus increasing system coverage and utility).

Specifically, we propose a zone-based proximity detection and information sharing mechanism. Concentration gradients are driven by transport via molecular diffusion and convection. Both transport processes have random and non-random components. Whatever the process, the spatial gradients dictate that two points close in proximity will likely have more similar concentrations than two points that are more distant. The smaller the distance, the more similar the $CO_2$ concentration readings. In MAQS, we define an area with

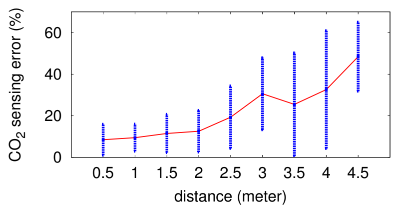



Figure 4. Sensing error of $CO_2$ concentration at different distances.

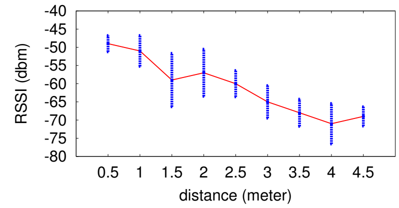



Figure 5. RSSI measurements (*avg* and *stdev*) at different distances.

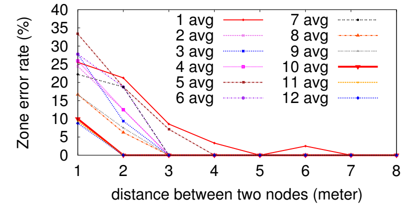



Figure 6. Zone detection error rate decreases by averaging multiple RSSI readings.

high density of people as a *zone*. All people within the same zone can share one M-pod for IAQ monitoring. When a new user without an M-pod joins this zone, the user's smartphone initiates a scan to determine if there is already an M-pod in the zone. If so, a communication link between the phone and the M-pod is established and the IAQ values reported by the M-pod is used to estimate the IAQ of this user.

Zone-based sharing would incur some error, which is defined as the difference between the $CO_2$ concentration reported by the shared M-pod nearby and the true $CO_2$ concentration at the user's location. Again, according to the diffusion equation, this error is highly correlated with the distance. Therefore, this error determines the effective range of each zone.

Over 50 experiments have been conducted to determine the relationship between distance and $CO_2$ error in public rooms, including classrooms and a library. In each experiment, two M-pods are placed 1–10 meters apart for over 20 minutes. The $CO_2$ readings from both M-pods are monitored and the corresponding air exchange rates are calculated. Figure 4 shows the sensing error rate of $CO_2$ concentration at different distances. Ranges less than two meters enable better and more consistent results. Two meters is therefore uses as the range threshold of zones in the MAQS system.

Given the zone range threshold (2 meters), each smartphone still needs to determine how far away a specific M-pod is. In MAQS, we use the Received Signal Strength Indication (RSSI) from the Bluetooth radio as the distance metric. The signal power of Bluetooth communication decays proportionally to $d^{-2}$ [18], where $d$ is the distance between transmitter (M-pod) and receiver (smartphone). MAQS is meant to be used in a broad range of unknown environments over long periods of time. It it therefore subject to environment-dependent and time-varying noise. Figure 5 shows the average and standard deviation of RSSI measurements obtained at different distances in real-world experiments. The illustrated noise can result in large errors even for two meter zones. Since this noise can be reasonably well modeled as additive white Gaussian noise, in MAQS, multiple readings are used to detect outliers and average values are used to improve distance estimates. Figure 6 shows that the accuracy of proximity detection can be significantly improved when 10 readings are averaged.

The MAQS zone-based information sharing procedure works as follows. When a user enters a room, the carried mobile phone first identifies the room. Then, the phone scans nearby M-pods 10 times and filters out all M-pods with RSSI readings exceeding a threshold corresponding to 2 meters (-64 dB in our experiments). The remaining M-pod with the smallest RSSI average is selected. The user joins the zone formed by this M-pod and shares its $CO_2$ concentration and air exchange rate information. Figure 7 illustrates the concept of zone-based information sharing. Numbers along the lines indicate the RSSI between phone and M-pod. In this scenario, phones A and B belong to the zone occupied by M-pod S1, because they are within the required range of S1. Since the RSSI between phone C and M-pod S1 is lower than the threshold, they belong to different zones.

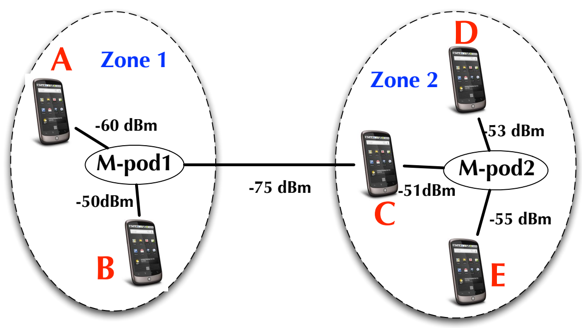



Figure 7. Zone-based collaborative sensing.

In addition to IAQ data sharing by people in the same room, users may also be interested in rooms which they have not yet visited; building managers would like to know the IAQ in their buildings; and medical personnel may need the information to diagnose their patients. MAQS provides a permission control mechanism allowing users to specify the type, location, and time periods of data to be shared publicly. They can also define groups and share their IAQ data with other members in each group, e.g., their family members.

## EVALUATIONS

The MAQS system has been implemented and deployed for user study. In this section, we describe the system deployment and user study. We experimentally evaluated the effectiveness and efficiency of the proposed system. We also make some observations based on the data gathered during the user study.

### System Deployment and User Studies

The deployed MAQS system includes M-pod sensing devices, Android-based smartphones, as well as the data and web servers. We have conducted two phases of user study with 17 participants, including faculty and graduate students, who share some workplaces and classrooms. The first phase was designed to evaluate our room localization method. In this phase, users were asked to carry their Android phones for 12 weeks. Our MAQS phone application continuously collected Wi-Fi signals, and requested manual labeling when users entered rooms. Weekly meetings with the users were held to verify the accuracy of the motion traces. In addition, users were asked to collect Wi-Fi signals in rooms that are adjacent to those they have visited, which further increased the noise and complexity of our room location task but better represents real-world scenarios.

In the second phase, users carried both their smartphones and the M-pod sensing devices for 3 weeks. The MAQS system collects both room information and IAQ data for all users. This phase of user study allows us to evaluate the entire MAQS system. We collected localization data for 171 rooms, and IAQ data for 56 rooms.

### Evaluation of Room Localization Technique

To evaluate our temporal n-gram augmented Bayesian room localization method, we determined the room localization accuracy as a function of the number of Wi-Fi fingerprints. A good room localization method should achieve high accuracy with few fingerprints. Specifically, we conduct experiments to answer the following questions: (1) How does the n-gram length affect overall room localization performance? (2) How does the n-gram augmented Bayesian model perform compared with existing approaches? (3) To what extent does temporal user mobility information improve room localization accuracy?

The subsequence length $n$ of n-gram has an important impact on room localization performance. If $n$ is too small, adjacent rooms may have many similar subsequences and be difficult to distinguish. For example, when $n = 1$, our n-gram augmented model is similar to the original Bayesian model, and the only difference is the binary discretization of signal strength: AP equals 1 (or 0) if its signal can (or cannot) be observed by the device. If $n$ is too big, then the room signature may change for each scan, since longer subsequences have a low probability of being repeated in each scan. For example, when $n$ is set to the total number of APs that can be scanned in a room, it is very unlikely to receive the same sequence for each scan. Therefore, a good $n$ value should make the n-gram signature for each room stable for different scans, and make the n-gram signatures of adjacent rooms distinguishable. Figure 8 shows the accuracy of our n-gram Bayesian model with different $n$ values and different number of fingerprints per room. As shown in the figure, when $n = 2$, our model achieves much better accuracy with fewer fingerprints.

**Table 2. Energy Consumption for Room Localization**

| Method | Wi-Fi (#scans/day) | Acc (#samples/sec) | Energy (mw) |
|---|---|---|---|
| Wi-Fi only | 1,440 | off | 8 |
| Wi-Fi + Acc | 5–30 | 3–10 | 1–3 |

We now compare our n-gram augmented Bayesian room localization model with four state-of-the-art algorithms: (1) Bayesian room localization [14, 12], (2) Delta signal Bayesian room localization, which is similar to Bayesian room localization but uses the difference of signal strength (instead of RSS values directly) between each pair of APs to calculate probability, (3) vector-based room localization [5], which uses AP RSS vector as the room signature and Euclidean distance to locate the nearest room, and (4) Delta signal vector-based room localization, which is similar to vector-based room localization but uses the difference in signal strength between each pair of APs to build the vectors. As shown in Figure 9(a), our n-gram augmented Bayesian model achieves the best accuracy, especially when the number of fingerprints per room is above 50.

By incorporating the temporal user mobility information, our temporal n-gram augmented Bayesian model can achieve better accuracy even when the number of fingerprints per room is limited. As shown in Figure 9 (b), our temporal n-gram model further improves the room localization accuracy over our n-gram model, especially when the number of fingerprints per room is less than 50. The Delta Bayesian room localization model, which performed second best in Figure 9(a), can also benefit from the use of temporal user mobility information, but it is still not comparable to our temporal n-gram model.

Finally, we evaluate the energy consumption of our room localization method. By using an accelerometer to detect room entrance/departure and performing a Wi-Fi scan only when entering a new room, we can significantly reduce the energy consumption. As shown in Table 2, our Wi-Fi + accelerometer (Acc) approach consumes only 1–3 mW on the mobile phone, while the Wi-Fi only approach consumes 8 mW.

### Evaluation of Air Exchange Rate Based IAQ Sensing

To estimate the general IAQ without using sensors for each specific air pollutant, we propose to calculate the air exchange rate from temporal $CO_2$ concentration readings, and use this air exchange rate to estimate IAQ. Here, we evaluate the accuracy of the air exchange rate model. We used the *Alnor EBT721 EBT 721 Air Balancing Balometer Flow Capture Hood* [1] to measure the air flow rate directly from vents as the ground truth. During each experiment, we change the forced ventilation rate and the number of people in the room to evaluate

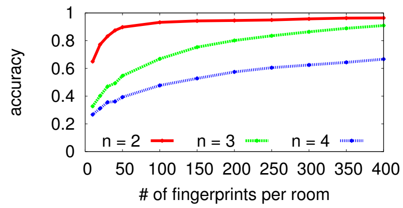



Figure 8. Comparison of different $n$ for n-gram augmented Bayesian model.

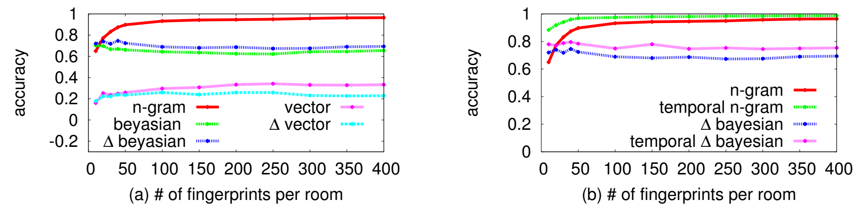



Figure 9. Performance comparison of room localization methods: (a) n-gram model and (b) temporal n-gram model.

Table 3. Zone Detection Accuracy

| Seat Map | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Accuracy (%) | 97 | 100 | 100 | 78 | 100 |
| Seat Map | 6 | 7 | 8 | 9 | 10 |
| Accuracy (%) | 100 | 88 | 100 | 42 | 85 |

the accuracy and responsiveness of our $CO_2$ based air exchange rate model.

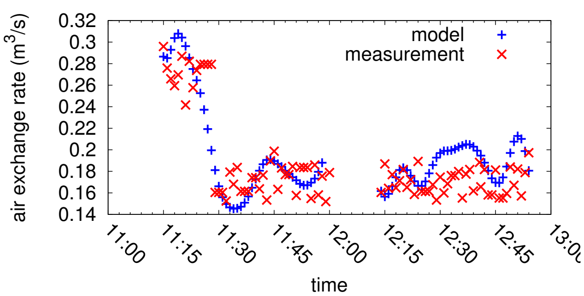



Figure 10. Air exchange rate model evaluation.

Figure 10 shows the results of one experiment. We started at 11:15 with a relatively high forced ventilation rate and lowered the ventilation rate at 11:30. We can see that the air exchange rate calculated by our model followed the actual rate drop quickly and stayed within the same range. Starting from 12:15, we kept the ventilation rate low and changed the number of people in the room every 5 minutes. Again, we can see that the AER calculated by the model well approximates the measured value. It is slightly higher than the measured value, since the door was opened and closed when we changed the number of people in the room, leading to additional air exchange, which is not captured by the vent hood, as it measures only the air exchange at the vent. We conducted multiple experiments in rooms with different sizes and vents, and obtained similar results.

### Evaluation of Zone-Based Collaborative Sensing

Zone-based collaborative sensing helps to reduce the number of M-pods needed, saves energy, and allow users to share IAQ data. Since the energy consumption of scanning Bluetooth RSSI for a short period of time (less than 30 seconds) is negligible compared with that of continuous IAQ sensing, a zone with $k$ devices can generally achieve $k\times$ better energy efficiency since only one device needs to be running. Here, we evaluate the accuracy of our zone-based proximity detection, i.e., whether we can identify the correct zone based on the Bluetooth RSSI readings. Given three M-pods and a mobile phone, we selected 10 seat maps that represent real-world user-sitting scenarios. For each seat map, we conducted 10 experiments at different times to capture potential temporal variations. For each experiment, the mobile phone used 10 RSSI readings from each M-pod to determine which zone (i.e., M-pod) it belonged to. The average zone detection accuracy for each seat map is shown in Table 3. We can see that our RSSI-based zone detection method achieved high accuracy for most seat maps with an average of 89%.

### IAQ Data Analysis

We analyzed the IAQ distributions for the data gathered during our user study. Specifically, we would like to answer quantitatively what IAQ distributions the study participants experienced. Figure 11 shows the IAQ distributions for all users. The dashed lines indicate the standard limits, blue solid lines represent good IAQ, and red solid lines represent bad IAQ. According to the figure, (1) 67% of the time, indoor $CO_2$ concentration is higher than the reasonable limit of 1,000 ppm; (2) 30% of the time, air changes per hour do not meet the minimum requirement of 0.35 1/h; and (3) 58% of the time, flow rate per person does not meet the minimum standard of 7 l/s/person. We conclude that our study participants frequently spent time in environments with poor air quality.

Figure 12 shows the distributions of $CO_2$ concentration, air changes per hour, and air flow per person for different users. The users are ordered by their average $CO_2$ concentration. We can observe that users are subject to different IAQ at different times, and almost all users are subject to a fairly high percentage of times at which they experienced poor IAQ (high $CO_2$, low air changes per hour, or low air flow per person).

Figure 13 shows the distributions of $CO_2$ concentration, air changes per hour, and air flow per person in different rooms. The rooms are ordered by their average $CO_2$ concentrations. We can see that rooms have dynamic and diverse IAQ profiles. Although the exact fraction is different, many rooms have poor IAQ some of the time. Table 4 compares the IAQ in different types of rooms: office, public place (e.g., library, lobby), classroom, and home (apartment or house). The office has good IAQ in general, since most office rooms are small, sparsely occupied, and have good ventilation systems. IAQ in the other three types of rooms is not good and can be very bad at times (indicated

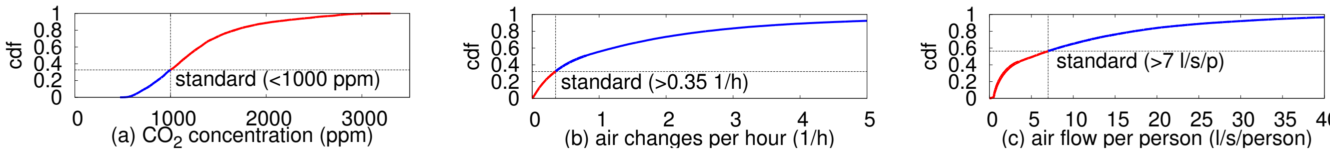



**Figure 11. IAQ data distribution for all users: (a) $CO_2$ concentration; (b) air changes per hour; and (3) flow rate per person.**

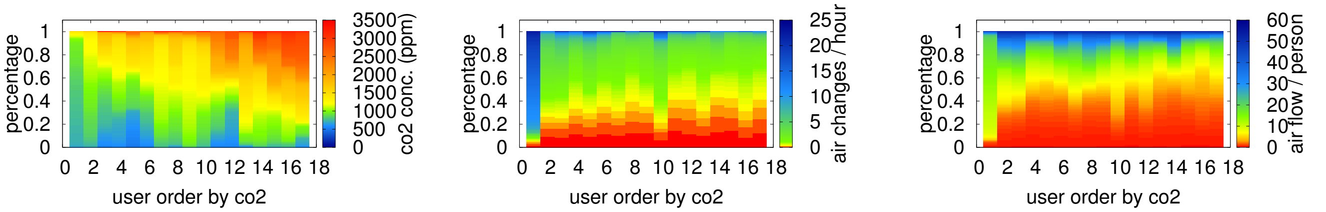



**Figure 12. User-specific distributions of $CO_2$ concentration, air changes per hour, and air flow per person.**

**Table 4. IAQ Comparison by Room Type**

| IAQ | | Office | Public | Classroom | Home |
|---|---|---|---|---|---|
| $CO_2$ | avg | 887 | 1238 | 1163 | 1491 |
| | stdev | 274 | 207 | 303 | 479 |
| Air changes per hour | avg | 2.6 | 2.9 | 1.6 | 1.3 |
| | stdev | 3.0 | 3.8 | 2.0 | 1.8 |
| Air flow per person | avg | 13.6 | 11.7 | 10.5 | 9.1 |
| | stdev | 13.9 | 14.9 | 11.5 | 10.6 |

by large standard deviation). For example, large classrooms can have insufficient ventilation for large classes. Also, apartments tend to be small and have no or limited forced ventilation, leading to poor IAQ.

## RELATED WORK

Our interdisciplinary work of building a personalized mobile system for IAQ monitoring draws upon research in a number of related fields. In this section, we survey research most relevant to ours.

**IAQ monitoring**. Previous studies [13] have focused on monitoring or identifying air pollutants in different types of rooms, such as classrooms, offices, and residential rooms. However, without knowing when and how long a user stayed in a room, it is difficult to estimate the user's exposure to indoor air pollutants. Moreover, it is difficult to measure all rooms a user stayed in. Other studies have focused on indoor air pollutants in certain areas (e.g., city or countryside) using statistical methods. This approach ignores the fact that users' indoor air pollution exposures differ from each other. Recently, Kim et al. proposed a mobile system for sharing IAQ measurements and visualizations within one's social network [17]. The shared information is room based and not user-specific.

**Indoor localization**. This has been a topic of active research, some focusing on indoor intra-room positioning and others (such as ours) focusing on inter-room positioning. Proprietary systems based on radio frequency [29], FM radio signal [19, 22] and ultrasound [30] have been implemented. Newer systems include DOLPHIN [23] which is based on ultrasound devices, and the Zigbee-based system proposed by Sugano [27]. They have good accuracy but require substantial investment in infrastructure and special hardware worn by all users. Our system can be deployed in any off-the-shelf smartphones with Bluetooth communication capability. Other techniques are fingerprint-based and leverage existing wireless infrastructure. Haeberlen et al. proposed a localization method over large-scale 802.11 wireless networks, which can be accurate within a few meters in regions with high infrastructure coverage [14] . However, this method has high deployment cost. Other methods leverage user collaboration [7, 8], i.e., users train the system while using it. Issues such as conveying uncertainty, determining when user input is actually required, and discounting erroneous and stale data are addressed by the work of Park et al. [12]. However, none of these methods addresses the challenges associated with environment heterogeneity, device-induced noise, and the requirement for large user inputs.

**Proximity detection**. Previous peer-based indoor positioning systems attempt to infer either the proximity of a pair of devices, or the actual distances between multiple pairs of devices in order to place them in a virtual map. Most of them techniques use anchors with available position information as references. Other nodes refer to the anchors to determine their own positions. PeopleTones [21] uses a GSM-based approach to detect proximity for mobile phone users. NearMe, proposed by Krumm et al. [20], provides a complete framework for clients equipped with Wi-Fi devices to obtain information about people and things that are physically close (30 to 100 meters). Banerjee et al. proposed Virtual Compass [6], which measures the distances between multiple nearby nodes and generates a map on a 2D plane. In our system, precise absolute position is not required and relative proximity information is sufficient. It is conceptually similar to the reality mining system proposed by Nathan et al. [10]. However, mobile devices within a room are distributed more densely and the amount of noise is greater. We propose zone-based proximity detection to tackle the unique problem of personalized IAQ monitoring.

## CONCLUSIONS

This paper has described MAQS, a mobile system for personalized IAQ monitoring. To achieve high accu-

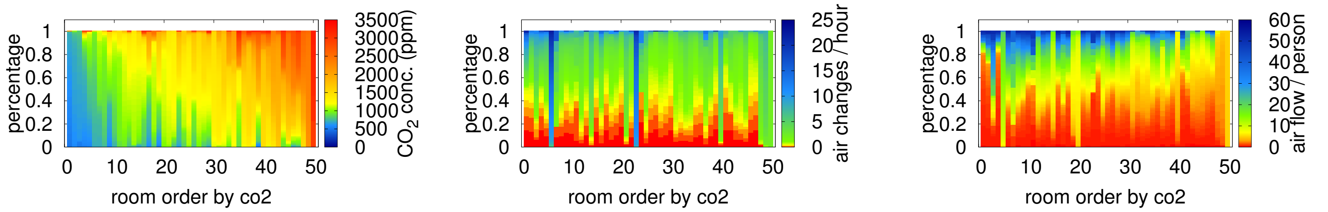



**Figure 13. Room-specific distributions of $CO_2$ concentration, air changes per hour, and air flow per person.**

racy and energy efficiency under diverse sensing scenarios, we developed a number of novel techniques: (1) a temporal n-gram augmented Bayesian room localization method that achieves high accuracy with a small number of Wi-Fi fingerprints; (2) an air exchange rate based IAQ sensing method that measures general IAQ using only $CO_2$ sensors; and (3) a zone-based proximity detection method for collaborative sensing, which saves energy and enables data sharing among multiple users. MAQS has been deployed and evaluated via user study. Detailed evaluation results demonstrate the feasibility, effectiveness, and efficiency of MAQS for personalized IAQ monitoring. We also found that study participants frequently experienced poor IAQ.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under awards CNS-0910995, CNS-0910816, and CNS-0347941. We thank Simon Olivieri and Lan Bai for help with the experiments. We also thank the anonymous reviewers for constructive comments.

## REFERENCES

1. Alnor ebt721 ebt 721 air balancing balometer flow capture hood. http://www.aikencolon.com/Alnor-EBT721-EBT-721-Air-Balancing-Balometer-Flow-Capture-Hood_p_2549.html.
2. Arduino BT. http://www.arduino.cc/en/Main/ArduinoBoardBluetooth.
3. Indoor environmental quality. http://www.cdc.gov/niosh/topics/indoorenv/BuildingVentilation.html.
4. Trends in atmospheric carbon dioxide. http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html.
5. P. Bahl and V. N. Padmanabhan. RADAR: An in-building rf-based user location and tracking system. In *INFOCOM*, pages 775–784, 2000.
6. N. Banerjee, S. Agarwal, P. Bahl, R. Chandra, A. Wolman, and M. Corner. Virtual Compass: Relative positioning to sense mobile social interactions. In *Pervasive Computing*, 2010.
7. E. S. Bhasker, S. W. Brown, and W. G. Griswold. Employing user feedback for fast, accurate, low-maintenance geolocationing. In *PERCOM*, page 111, 2004.
8. P. Bolliger. Redpin - adaptive, zero-configuration indoor localization through user collaboration. In *MELT*, 2008.
9. J. M. Daisey, W. J. Angell, and M. G. Apte. Indoor air quality, ventilation and health symptoms in schools: an analysis of existing information. *Indoor Air*, 13, 2003.
10. N. Eagle and A. (Sandy) Pentland. Reality mining: sensing complex social systems. *Personal Ubiquitous Comput.*, 10, 2006.
11. W. J. Fisk, A. G. Mirer, and M. J. Mendell. Quantitative relationship of sick building syndrome symptoms with ventilation rates. *Indoor Air*, 19, 2009.
12. J. geun Park, B. Charrow, D. Curtis, J. Battat, E. Minkov, J. Hicks, S. Teller, and J. Ledlie. Growing an organic indoor location system. In *MobiSys*, pages 271–284, 2010.
13. C. Godwin and S. Batterman. Indoor air quality in michigan schools. *Indoor Air*, 17(4), 2007.
14. A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki. Practical robust localization over large-scale 802.11 wireless networks. In *MobiCom*, pages 70–84, 2004.
15. C. Huizenga, S. Abbaszadeh, L. Zagreus, and E. Arens. Air quality and thermal comfort in office buildings: Results of a large indoor environmental quality survey. In *Healthy Buildings 2006*, 2006.
16. D. H. Kim, Y. Kim, D. Estrin, and M. B. Srivastava. SensLoc: Sensing everyday places and paths using less energy. In *SenSys*, pages 315–330, 2010.
17. S. Kim and E. Paulos. InAir: sharing indoor air quality measurements and visualizations. In *CHI*, pages 1861–1870, 2010.
18. J. D. Kraus. *Antennas*. McGraw-Hill, 1988.
19. J. Krumm, G. Cermak, and E. Horvitz. Rightspot: A novel sense of location for a smart personal object. In *UBICOMP*, 2003.
20. J. Krumm and K. Hinckley. The NearMe wireless proximity server. In *Ubicomp*, pages 283–300. Springer, 2004.
21. K. A. Li, T. Y. Sohn, S. Huang, and W. G. Griswold. Peopletones: a system for the detection and notification of buddy proximity on mobile phones. In *MobiSys*, pages 160–173, 2008.
22. A. Matic, A. Papliatseyeu, V. Osmani, and O. Mayora-Ibarra. Tuning to your position: Fm radio based indoor localization with spontaneous recalibration. In *IEEE PerCom*, 2010.
23. M. Minami, Y. Fukuju, K. Hirasawa, S. Yokoyama, M. Mizumachi, H. Morikawa, and T. Aoyama. DOLPHIN: A practical approach for implementing a fully distributed indoor ultrasonic positioning system. In *UbiComp*, volume 3205, pages 347–365, 2004.
24. A. Persily. Evaluating building IAQ and ventilation with indoor carbon dioxide. *ASHRAE Transactions*, 1997.
25. O. A. Seppänen, W. J. Fisk, and M. J. Mendell. Association of ventilation rates and co2 concentrations with health and other responses in commercial and institutional buildings. *Indoor Air*, 9, 1999.
26. M. H. Sherman. ASHRAEs first residential ventilation standard. *LBNL-54331*, 2004.
27. M. Sugano. Indoor localization system using rssi measurement of wireless sensor network based on zigbee standard. In *Wireless and Optical Communications*, pages 1–6, 2006.
28. U.S. Environmental Protection Agency Green Building Workgroup. Buildings and their impact on the environment: A statistical summary, 2009.
29. R. Want, A. Hopper, V. Falcão, and J. Gibbons. The active badge location system. *ACM Trans. Inf. Syst.*, 10:91–102, January 1992.
30. A. Ward, A. Jones, and A. Hopper. A new location technique for the active office. *IEEE Personal Communications*, 4:42–47, 1997.
31. D. P. Wyon. The effects of indoor air quality on performance and productivity. *Indoor Air*, 2004.