ISSAQ: An Integrated Sensing Systems for Real-Time Indoor Air Quality Monitoring

Jung-Yoon Kim, Chao-Hsien Chu, and Sang-Moon Shin

***Abstract*—With growing transportation and population density, increasing global warming and sudden climate change, air quality is one of the critical measures that is needed to be monitored closely on a real-time basis in today's urban ecosystems. This paper examines the issues, infrastructure, information processing, and challenges of designing and implementing an integrated sensing system for real-time indoor air quality monitoring. The system aims to detect the level of seven gases, ozone (O_3), particulate matter, carbon monoxide (CO), nitrogen oxides (NO_2), sulfur dioxide (SO_2), volatile organic compound, and carbon dioxide (CO_2), on a real-time basis and provides overall air quality alert timely. Experiments are conducted to validate and support the development of the system for real-time monitoring and alerting.**

***Index Terms*—Air quality monitoring, calibration, energy consumption, indoor, wireless sensor network.**

I. Introduction

AIR quality is one of the important measures to be closely monitored in real-time for today's urban ecosystems, because air quality has major influence on the health, safety, productivity and comfort of people [1]–[3]. With growing transportation and population density, increasing global warming and sudden climate change, the air quality in many big cities worldwide have frequently below the environmental quality standards [4]–[7]. Many countries have developed their own infrastructure, methodology and standards to monitor and provide alert to their citizen. For example, the Environmental Protection Agency (US EPA) was charged to monitor and protect the human health and environment in the United States. They define the national outdoor air quality standards in the U.S. according to six common air pollutants, suggest using an Air Quality Index (AQI) to measure air quality and set its limits on human health [8]. However, the information provided is only in an aggregated index alerting the harmful condition. Also, the air quality information is limited to outdoor and within a wide range of region. Moreover, the information is static and is an average value; while, air quality is changeable and was impacted by different situations such as indoor, outdoor, inside of vehicle, fire, wind, population density, and so on.

In recent years, indoor air quality has received more attention and research than outdoor air quality, due to the fact [9] that, on average, the air quality level indoor is two to five times higher than outdoor and people in the U.S. spend about 65 to 90 percent of their time indoors. In fact, each day we eat, sleep, travel, work, shop and/or spend time in different enclosed environments, such as living room, bed room, kitchen, office building, bus, subway, tunnel, underground subway station and/or shopping centres etc. Also, people may work inside chemical, machinery, electronic assembly, oil refinery plant or underground mining field etc. Thus, for most people, the health risks due to exposure to indoors air pollution may be greater than outdoors, but they did not even aware the fact and its consequence [10].

To avoid the bad effect from the primary pollutants, accurate air quality monitoring system and follow-up actions are needed. Metal oxide semiconductor (MOS) sensors have been widely used for detecting types and levels of gases [11]. However, Oxide gas sensors have limitations, such as cross sensitivities, temperature and humidity dependence behaviour, and the lack of long-term stability [12]–[17]. It is very important to consider these factors during the systems design and implementation stages. Meanwhile, the existing portable air quality monitoring systems are not enough to detect the various gasses and particulate matters for determining the air quality level in various environments. Most of the previous systems focus on one or two pollutants. Even if the system measures more than three pollutants, information processing capability is weak or lacking. In addition, they didn't deal with power management issue for remote monitoring [17]–[24]. Therefore, there is an emerging need to develop novel methodology and system for real-time air quality monitoring.

In this paper, we examine the issues, infrastructure, information processing and challenges of designing and implementing an integrated sensing node or wireless sensor network system for monitoring indoor air quality. A prototype of sensor node, built upon Raspberry Pi architecture, for monitoring seven different gases as proposed by US EPA is developed. A smoothing algorithm is introduced for preventing temporary sensor errors. Temperature and humidity sensors are also

Manuscript received April 16, 2014; revised June 23, 2014 and September 9, 2014; accepted September 9, 2014. Date of publication September 23, 2014; date of current version October 21, 2014. This work was supported in the U.S. Department of Defense by a Capacity Building Fund and in part by the Smart Spaces Initiative Fund from Penn State's Smart Space Center and the College of Information Sciences and Technology, USA. The associate editor coordinating the review of this paper and approving it for publication was Dr. Ashish Pandharipande.

J.-Y. Kim was with Pennsylvania State University, University Park, PA 16801 USA. He is now with the iCity Laboratory, Singapore Management University, Singapore 178902 (e-mail: bassjykim@gmail.com).

C.-H. Chu is with Pennsylvania State University, University Park, PA 16802 USA, and also with the School of Information Systems, Singapore Management University, Singapore 178902 (e-mail: chu@ist.psu.edu).

S.-M. Shin is with the School of Electrical Engineering, University of Ulsan, Ulsan 110-810, Korea (e-mail: sangmoon.shin@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2014.2359832

TABLE I
AIR POLLUTANTS AND EFFECTS

Pollutant	Risk
CO	Fatigue, Headache, Impaired, vision, Coma, Brain damage, Death
CO_2	Headache, Dyspnea, Tremors, Increased heart rate and blood pressure, Convulsions, Coma, Death
O_3	Lung disease
NO_2	Adverse effects on the respiratory system
SO_2	Adverse effects on the respiratory system
Particulate Matters	Lung disease
VOC	Eye, nose, and throat irritation Headache Damage to liver, kidney and nervous system

TABLE II
INDOOR POLLUTANTS AND SENSOR TYPES

Room	Pollutants	Sensor Type
Bathroom	Mold	VOC Sensor
Bathroom	Dust	Dust Sensor
Kitchen	Volatile Organic Compounds Carbon Monoxide	VOC Sensor CO Sensor
Basement	Carbon Monoxide Volatile Organic Compounds Mold	CO Sensor VOC Sensor VOC Sensor

installed to manage cross sensitivity issues and automatically calibrate the sensors' data. To reduce the network traffic and power consumption, an aggregation algorithm is proposed. All the smoothed, calibrated, and aggregated data are then sent to a ubiquitous infrastructure for determining air quality level according to the modified EPA's Air Quality Index (AQI).

Experiments are conducted in three settings – a living room, class room and big church – with control of several factors such as population density, temperature & humidity, air flows etc. to demonstrate the feasibility and functionality of the proposed prototype, explore the design issues and challenges, and evaluate the potential impact and possible solutions. The remainder of this paper is organized as follows. Section II provides background information on US EPA standard and method of assessing indoor and outdoor air quality. Section III reviews related works. Section IV discusses scenarios, requirements and potential design issues. Section V proposes information processing solutions. Section VI describes experiments, analyses and results. Section VII provides the discussion and conclusions of this study.

II. BACKGROUND

US EPA defines the national air quality standards according to seven common air pollutants. See Table I. These pollutants can harm people's health and the environment, and cause property damage, depending on the level and location.

Meanwhile, US EPA categorizes the indoor air quality according to the pollutants of each room in a house. Table II shows the pollutants and sensor types for the pollutants based on the specific spaces. As shown, different types of pollutants require different types of sensor to detect. VOC, dust, and CO sensors are the three most popular sensors to monitor indoor air quality.

US EPA suggests to use the level of five pollutants – O_3, CO, NO_2, SO_2, and PM – to measure outdoor air quality. We can calculate the AQI using the pollutant concentration data, breakpoint table, and AQI equation (linear interpolation) developed by US EPA [25]. The pollutant concentration data can be gathered by using the sensing nodes. Table III shows the breakpoints for the AQI. The breakpoints assign values to the variables in AQI equation:

$$I_p = (C_p - BP_{Lo}) \times \frac{I_{Hi} - I_{Lo}}{BP_{Hi} - BP_{Lo}} + I_{Lo} \quad (1)$$

Where:

I_p = The Index for pollutant p
C_p = The rounded concentration of pollutant p
BP_{Hi} = The breakpoint that is greater than or equal to C_p
BP_{Lo} = The breakpoint that is less than or equal to C_p
I_{Hi} = The AQI value corresponding to BP_{Hi}
I_{Lo} = The AQI value corresponding to BP_{Lo}

The calculated AQI values from different gases are compared with each other, and the highest of AQI value becomes the final AQI value. For example, if CO is "Moderate", O_3 is "Good", and NO_2 is "Unhealthy", then the final AQI will be designated as "Unhealthy" status. Overall, the US EPA's AQI can provide a general sense of air quality but the AQI did not reflect the harm level of each gas on human health. Our system will not only provide the overall AQI value but also display and warn all the types of gas that have over the acceptable level that have harm to human body or environment.

Example 1: Suppose the value of CO from the CO sensor is 8.4 ppm (that is between 4.5 and 9.4). Based on the breakpoints table, C_p is 8.4, BP_{Hi} is 9.4, BP_{Lo} is 4.5, I_{Hi} is 100, and I_{Lo} is 51. The AQI value can be calculated using equation 1 as:

$$I_p = (8.4 - 4.5) \times \frac{(100 - 51)}{(9.4 - 4.5)} + 51 = 90$$

Since the I_p value is 90, we conclude that the level of CO on health is moderate according to the breakpoint table.

III. RELATED WORK

Several researches have devoted effort to address research issues related to indoor air quality monitoring. We review and compare these issues and summarize them in Table IV.

A. Sensor Network Infrastructure

Wireless networks provide a promising infrastructure for capturing information from monitoring areas and transmitting them back to the server for processing and actions. As shown, three types of networks – wireless local area networks (WLAN), mobile networks and wireless sensor networks (WSN) – have been adopted for indoor air quality monitoring.

Postolache et al. [18] suggest that two types of WLAN – Ad Hoc architecture and AP infrastructure network – can be used for air quality monitoring. They also argue that the Ad Hoc structure seems to be a better solution, particular

TABLE III
EPA's Breakpoint and AQI

O_3(ppm) 8-hour	O_3(ppm) 1-hour	PM_{10} (ug/m^3)	PM_{25} (ug/m^3)	CO (ppm)	SO_2 (ppm)	NO_2 (ppm)	AQI	Levels of Health
0.000 ~ 0.059	-	0 ~ 54	0.0 ~ 15.4	0.0 ~ 4.4	0.000 ~ 0.034	-	0 - 50	Good
0.060 ~ 0.075	-	55 ~ 154	15.5 ~ 40.4	4.5 ~ 9.4	0.035 ~ 0.144	-	51 - 100	Moderate
0.076 ~ 0.095	0.125 ~ 0.164	155 ~ 254	40.5 ~ 65.4	9.5 ~ 12.4	0.145 ~ 0.224	-	101 - 150	Unhealthy for Sensitive Groups
0.096 ~ 0.115	0.165 ~ 0.204	255 ~ 354	65.5 ~ 150.4	12.5 ~ 15.4	0.225 ~ 0.304	-	151 - 200	Unhealthy
0.116 ~ 0.374	0.205 ~ 0.404	355 ~ 424	150.5 ~ 250.4	15.5 ~ 30.4	0.305 ~ 0.604	0.65 -1.24	201 - 300	Very Unhealthy
-	0.405 ~ 0.504	425 ~ 504	250.5 ~ 350.4	30.5 ~ 40.4	0.605 ~ 0.804	1.25 - 1.64	301 - 400	Hazardous
-	0.505 ~ 0.604	505 ~ 604	350.5 ~ 500.4	40.5 ~ 50.4	0.805 ~1.004	1.65 - 2.04	401 - 500	Hazardous

* NO_2 can generate an AQI above a value of 200, and 8-hour O_3 AQI values of 301 or higher are calculated with 1-hour O_3 concentrations.

TABLE IV
Summary of Previous Research

Issues and Characteristics	[18]	[19]	[20]	[21]	[22]	[23]	[24]	This Study
Sensor Network Infrastructure	WLAN Infrastructure AD HOC	WSN Star (Extended)	Mobile Network	WSN Star	WSN Star	Hybrid	WSN Star, Mesh, Tree	WSN Star
Network Protocol Used	Wi Fi	ZigBee	Bluetooth, Wi Fi	ZigBee	ZigBee		ZigBee	ZigBee
VOC	TGS 822[a]	+				TGS 2602[a] MOCON[c]		TGS 2602[a]
CO	TGS 203[a]	+		TGS 2442[a]				TGS 0425[a]
CO_2		NDIR[c]	TGS2442[a]	TGS 4161[a]		S100[c]	MG-811[a]	T6613[c]
O_3		+	MiCS2611[a]					MiCS2610[a]
NO_2							MQ-135[a]	GSNT11[a]
SO_2							MQ-6[a]	SO2-AF[b]
Methane	TGS 842[a]							
GAC	TGS 800[a]		+		TGS 2600[a]			TGS 2600[a]
PM		+		+				GP2Y1010AU0F[c]
Light					NSL-19M51			
Sensor Types (number)	MOS (4)	Optical (1) Unknown (4)	MOS (2) Unknown (1)	MOS (2) Unknown (1)	MOS (1)	MOS (1) Optical (2)	MOS (3)	MOS (5) Electro (1) Optical (2)
Temperature	SMT160-30[d]	+			LM35DZ[d]	+		DHT11[d]
Humidity	HM1500[d]	+			HIH-4000[d]	+		DHT11[d]
Calibration	X					X		X
Data Processing (Model)	Neural Network		Bayesian	AQI		Optimal/ Heuristic	AQI	Heuristic, AQI
Energy Management	X		X					X
Energy Consumption*	High	Low[+]	Medium	Medium	Low	Medium	High	High
Overall Cost*	Medium	High[+]	Medium	Medium	Low	High	Medium	High
Indoor Use	X	X	X	X	X	X	X	X
Outdoor Use	X						X	
Implemented	X	X	X	X	X	X		X
Evaluation	X		X	X	X	X		X

(a) MOS, (b) Electrochemical, (c) Optical, and (d) thermo type sensors; + The reference didn't provide detail information on specific gas sensors used;
* The energy consumption and cost are dependent on the type and number of sensor used

for monitoring outdoor air quality, because it requires less components and, thus, consuming less power. Wi Fi (IEEE 802.11) is a common protocol used in WLAN, and as such, carries its weaknesses that the allowed communication distance is shorter and also it may get more interference because there are many WLANs been used in real world.

Kim et al. [19] adopt an extended star type WSN for indoor air quality monitoring. Each wireless transceiver transmits the data it sensed from the sensors to a gateway. The gateway then samples collected data and transmits them to a central base station. The number of gateways was determined based on the range and signal strength. The ZigBee protocol/transceiver was used due to its low-power usage and reasonable communication range (up to 30 m).

A personalized mobile sensing system, MAQS, was proposed in [20] to monitor indoor air quality because it is inexpensive, portable and energy efficient. The system integrates smartphones with portable sensing devices, M-pods, to deliver personalized air quality information. According to [20], it may lead to redundant information when users are located near each other, and may lead to coverage gaps when users are not carrying sensing devices.

A WSN with star architecture was deployed in [21] and [22] for indoor air quality monitoring. The base station or ZigBee coordinator receives data at a regular time interval from the ZigBee end nodes. The end nodes of [21] are based on the Lielium Waspmote with gas sensor board for sensing data and in [22], it is based on the XBee and XBee pro module.

In [23], instead of relying on an established sensor network architecture, a hybrid sensor network architecture, which contains both stationary sensors (for accurate readings and calibration) and mobile sensors (for coverage), is formulated. Similar to the Knapsack problem, the synthesis problem is decomposed into two sub-problems. The first sub-problem is designed to select and place the stationary sensors, using exhaustive search from all possible schemes. The second sub-problem involves the allocation and assignment of the mobile sensors using a heuristic approach. In [24], a ZigBee-based WSN is adopted for real time air quality monitoring. Three possible structures – star, mesh and hierarchic tree are proposed but no implementation and testing are provided.

Clearly, WSN based on the ZigBee standard is an appropriate or good option for indoor air quality monitoring, which is also the framework that we choose for our design.

B. Sensor Node

Sensor node or array with multiple sensors is used to collect air pollution readings, from which the overall air quality level can be derived. As shown, there are three major types of gas sensors – the metal oxide semiconductor (MOS), electrochemical and optical sensor – available for air quality monitoring. Each type of sensors has unique features. Table V summarizes the features of these sensor types from the technical specifications of Figaro Eng. Inc. [26]–[28] and other online sources.

The sensor types and number of sensors included in the sensor node varies significantly in each implementation. Most studies consider multiple pollutants in their systems except [22], which only consider GAC sensor. But none of them contains all five air pollutants as suggested by the US EPA. We consider all seven pollutants in our implementation, despite that we can also use one or two sensors to meet specific needs, and thus reduce cost and energy consumption. Also, MOS type sensor is the most popular type of sensors used because of its low cost and small size; however, it also has lower sensitivity and is sensitive to environmental changes.

TABLE V

SUMMARY OF SENSOR TYPES AND CORRESPONDING CHARACTERISTICS

Characteristics	MOS	Electrochemical	Optical
Life time (span)	Long	Relatively short	Long
Sensitivity/selectivity	Low	High	High
Energy consumption	High	Very Low	Low
Warm-up time	Long	Medium/Long	Short
Response time	Short	Long	Short
Cost	Low	Medium	High
Size	Small	Medium	Big
Sensitive to environment changes	Sensitive	Less sensitive	Less sensitive
Accessibility of Pollutants	GAC, VOCs, NO_2, O_3, NO_2	SO_2, CO	PM, CO_2

C. Sensor Noise Reduction

One major problem of using low-cost sensors such as MOS type of sensors is their unreliable readings caused by long-term drift. To reduce the sensor noise, [17] proposed using accurate mobile sensors or fixed stationary sensors to automatically calibrate others sensor nodes in the network. The readings of temperature and humidity are needed to overcome noises due to cross sensitivities and the temperature and humidity dependence behavior. Among the seven studies we examined, only two of them [18], [23] present detailed algorithm for auto-calibration. In our design, we propose two heuristics to reduce possible sensor noise.

D. Data Processing

In order to predict air quality value or index from multiple sensor readings, data processing is needed. [18] uses two versions of multilayer perception neural network to calculate air quality values. [20] uses Bayesian room localization model to estimate air quality value. [21] and [24] use some types of air quality index (AQI) to calculate air quality index, which is also the option that we will use. [19] and [22] did not indicate which type of computation they used to determine air quality level despite that they also used multiple sensors.

E. Energy Consumption and Management

As shown in Table V, the MOS type of sensors consumes much more power than the optical and electrochemical types of sensors. Thus, the sensor nodes that use more number of MOS type of sensors, such as [18], [21], [24], have higher energy consumption. On the other hand, the WSNs which use more number of optical or electrochemical type sensors, such as [19] and [22], will be more energy efficient. Sine we include more number of sensors in the prototype, it shows that the power consumption is high. Of course, if we use less number of sensors, the rating will be lower.

To proper manage energy, some studies (see [18], [20]) propose energy management scheme such as data aggregation algorithm (see [20]) to reduce energy consumption, which is also adopted in our design. Please note that the rating in table IV only reflects the energy consumption related to sensor type and numbers, the actual energy consumption needs to consider other factors (e.g. data communication, energy management scheme) and needs a comprehensive evaluation.

F. Overall Cost

The overall cost of the sensor network is also highly dependent on the sensor type and number of sensors used. As shown in Table V, the cost of MOS type sensor is much cheaper than the cost of the electrochemical and optical types of sensor. Therefore, the more number of electrochemical and optical sensors used, the higher cost of the sensor network. See for example, in [19] and [23] and this study (which use one optical and one electrochemical sensor) have higher overall cost. However, the network can also detect more different types of pollutants. It is the trade-off needs to be considered in design.

G. Others

Among the seven studies that we examined, only two sensor networks (such as [18] and [24]) can be used to monitor both indoor and outdoor air quality, as there are additional issues need to be addressed for outdoor air quality monitoring. Also, most studies deployed prototype and evaluate their performance except [19] and [24].

IV. Design Requirements and Considerations

Designing a sensor network for smart sensing applications is a complicated and challenging task because many factors or issues, such as usage scenarios, agency requirements, flexibility, energy management, sensor board design [29], network traffic, power consumption, data accuracy, and fault tolerance etc., may impact and need to be considered in system performance. Also, several performance measures, such as accuracy, cost, reliability, energy/power consumption, standardization, efficiency and security/privacy, etc. need to be considered and there is always a trade-off between different criteria.

A. Scenarios

Although wireless sensor networks can be adopted and use to monitor the air quality for a variety of indoor spaces, its design may need to tailor to the special needs of different scenarios. We show below some special situations and their requirements.

1) Community Health Care. With growing aging population, how to care a large number of elderly with a limited number of caring people is an important consideration in community healthcare. In addition to other types of monitoring such as daily activities, health condition, and fall detection, air quality monitoring has been a key player. The bad air quality in a room can be caused by gas leakage, bad living habit and behavior, malfunction of the facility, food/medicine spoils, ventilation etc. The sudden change in air quality may imply a status of emergency, either the elderly has health problem or the room has cleaning or leakage problems. The monitoring system in this case needs to be able to monitor multiple pollutants, easy to use, can trigger air cleaning mechanism and provides direct connection with caring people for immediate attention.
2) Construction or Maintenance Sites. There are many possible critical situations for workers who provide construction work or maintenance service in structures or workspace. Once workers are fully focusing on their given tasks, their general senses become blurred. In this situation, if there is a gas leakage, the workers who are exposed to bad air for a long time may lose their spirit. For example, two workers were working on maintenance for refrigerator in a basement room. There was a gas leakage from the refrigerator and they didn't sense it and were found to dead. The indoor air quality monitoring will prevent this kind of critical emergency from happening. The monitoring system needs to be able to detect the gas leakage in real time, trigger for quick ventilation and alarm for immediate rescue.
3) Hazardous Locations [19]. Extreme areas, which are difficult to reach or access including underground transportation tunnels facilities (e.g., subway), mechanical rooms, hazardous material storage, and cleaning supply storage etc., can be benefit from wireless, energy efficient and low maintenance air quality monitoring system.
4) Schools or Gathering Places. Students spend most of their time in the indoor classroom of school buildings. The indoor air quality is very important for students' health. Generally, each person also becomes a source of pollutant in a closed place for a long term stay. The air quality monitoring system needs to be able to detect indoor air quality change and trigger for automatic ventilation such that fresh air can be provided so as to improve the performance of studying.

B. Requirements

The first requirement involves determining the number and types of pollutants be evaluated in air quality monitoring. Selecting the type of pollutants is highly related to agency specific requirements, sources and materials of emitting bad air; while, the number is closely related to coverage, accuracy, cost, reliability, and efficiency. In general, the decision is highly dependent on the agency specific requirements, application location (e.g., indoor or outdoor and place of indoor) and sources of emitting bad air. For example, we can adopt the six pollutants, CO, O_3, NO_2, SO_2, PM, and CO_2 for outdoor setting and the three common pollutants, VOC, CO and dust for indoor setting as suggested by the US EPA. We should also consider formaldehyde, benzene and toluene as the major pollutants for new cars and living environment. Refer to subsection D for more details.

The second requirement involves the selection of sensor types and sources for deployment. This issue is closely related to lifetime, power consumption, price, accessibility, stability, drift, and accuracy (e.g., sensitivity and selectivity). There are three sensing types of gas sensors available: The metal oxide semiconductor (MOS), electrochemical and optical sensor. Each has unique features.

Overall, most of the current available gas sensors belong to MOS type, which has long lifetime, low cost and fast response time; however, they have low sensitivity and selectivity, high power consumption and sensitive to environmental changes. Few available gas sensors belong to electrochemical type, which, on the other hand, has high sensitivity and selectivity, low power consumption and less sensitive to environmental change; however, their life time are relatively short and cost more than MOS type. Please note that Table V provides a general guideline for sensor selection; while the performance or information may vary depending on the specific sensor used. Several companies manufactured various sensors for detecting air pollutants, there is no single company has all the sensors for air quality monitoring. The quality, life time and performance of sensors are also varied from company to company.

The third requirement involves determining the optimal number of sensor nodes and placement. Sensor quantity is related to reliability, cost and areas/coverage to be monitored. In general, if the number of sensor nodes is increased, data accuracy and fault tolerance are better. However, if the number of sensor nodes is decreased, network traffic and power consumption will be reduced. The decision will also affect the total cost of implementation. The placement of sensor module is closely related to accuracy, security, and safety. This involved in deciding where we should place the sensor node. The type of air pollutant emission sources and security can be the key criteria for determining sensor node placement. These two parameters need to be optimized simultaneously, called coverage optimization. Coverage optimization has been a major stream of research in wireless sensor network [30], [31], in which protocol standard has been developed.

The fourth requirement includes the needs of periodic calibration. Most sensors are susceptible to drift error or impact by environmental changes. Drift is the gradual deviation of a sensor's readings from the correct value for long term. This problem is likely caused by material degradation, exposure to sulfur compounds or acids, aging, or condensate on the sensor surface etc. [32], [33]. The problem can be corrected by periodic calibration, which can be done with regular maintenance or automatic software compensation [32], [33]. The deviation caused by environmental change is more or less a short term effect, which can be calibrated by dynamic sensor configuration [17] or software correction. In this paper, two new algorithms are proposed to perform auto-calibration due to environmental changes. Gas sensor has relative sensitivity to various gases. To obtain accurate data, each sensor's cross sensitivities characteristic table needs to be captured and stored in middleware in batch so that the calibration algorithm selects actual value based on the table.

The fifth major issue faced is energy consumption. Gas sensor's sensing material is typically tin dioxide (SnO_2). The metal Oxide needs high temperature heat (about 300°) to detect gases. This means most of the gas sensors are high power consumption. Using several gas sensors in the sensor node needs strong power management algorithms. Also, the gas sensors have initial action that *Rs* drops sharply for the first few seconds. In the case of TGS2601, about 5 minutes is needed to stabilize the *Rs*. Thus, on-off control cannot be a solution for the power consumption. One side effect from power consumption is that a large amount of heats generated from the sensors, which may result in measurement error or trigger frequent auto-calibration. See subsection F detailed power consumption data for different types of sensor.

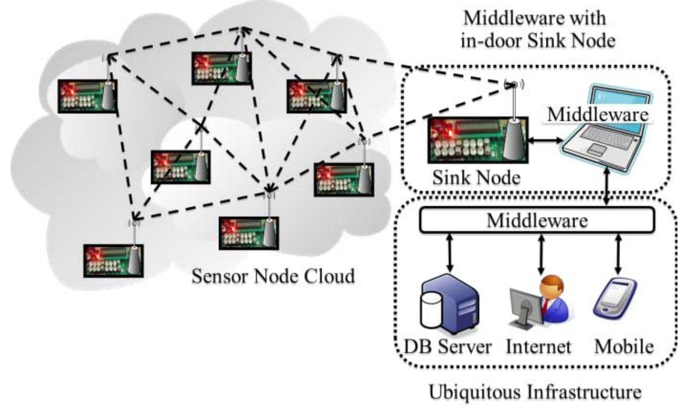


Fig. 1. Network configuration for remote indoor monitoring.

Finally, how to reliably transmit data with low power consumption from remote site can be an issue. We suggest using a data logger to store and transmitting data to the middleware. This is because the data logger will be installed without wall-power source. We design the embedded system for the data logger and power management unit.

C. Network Architecture

The sensing system we designed is capable of monitoring outdoor and indoor air quality. However, they are quite different in data communications. Figure 1 shows our proposed network configuration for remote indoor air quality monitoring. The sensor network cloud (SNC) senses the air quality data and sends the data to the sink node through wireless sensor network (WSN). The data logger keeps track of data from sink node and sends the data packets to the middleware using either wired or wireless (could be Bluetooth, Zing-Bee, or Wi-Fi) devices to transmit the captured data. The middleware calibrates the data and interprets the sensor data using AQI in high-level information processing. The middleware manages the processes of access, store, data/device management, alert and actions. Also, power consumption for sensors and communication is less a problem for indoors environment as regular electrical power sources can be used. A battery power is needed for system mobility or potential power breach.

D. Sensor Node

Figure 2 depicts a schematic diagram of the major components of a sensor node. For the measurement of primary pollutants, seven pollutant sensors are needed. Each sensor generates sensing signal based on the same environment, and the temperature and humidity sensors support the seven pollutant sensors to calibrate sensing data. The pre-calibrated data

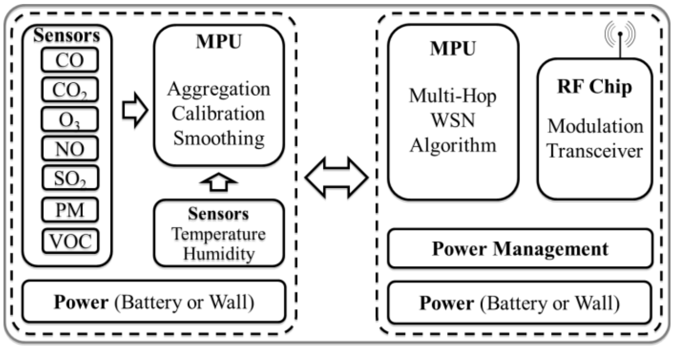


Fig. 2. Integrated air quality sensor node.

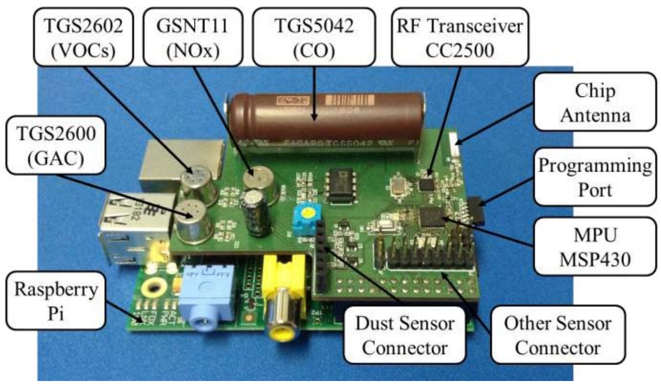


Fig. 3. Air quality monitoring sensor node.

is smoothed and aggregated by the algorithm in the embedded microprocessor. Through the Serial Peripheral Interface (SPI), the preprocessed data is transmitted to the communication unit. The communication unit transmits and receives the data packets using Multi-Hop Wireless Sensor Network (WSN) algorithm through the Radio Frequency (RF) chip (CC2500).

Each node with a microprocessor provides local processing, storage capability and low-power radio communication. The eZ430-RF2500 from Texas Instruments was selected for communication unit because of its good energy efficiency. The communication unit, which provides wireless communications, has separated power from the integrated sensing unit for reliable and efficient communication. The separated power makes the middleware aware of the problems whether low power or breakdown problem occurs in the integrated sensing board.

Figure 3 shows the prototype sensor module, consisting of four on-board gas sensors (TGS2600, TGS2602, GSNT11 and TGS5042), on-board WSN node (right side), temperature/ humidity sensors and a connector for expansion of T6613, MiCS-2610, SO2-AF, and GP2Y1010AUF. TGS2600 sensor was used to measure the GAC; TGS2602 sensor was used to monitor VOCs; GSNT11 was used to monitor NO_2; TGS5042 was used to monitor CO; T6613 sensor was used to measure CO_2; MiCS-2610 sensor was used to measure O_3; SO2-AF sensor was used to monitor SO_2; GP2Y1010AUF sensor was used to monitor PM, and temperature/humidity sensors (DHT11) are used to monitor room condition and for data calibration. The node was built upon Raspberry Pi

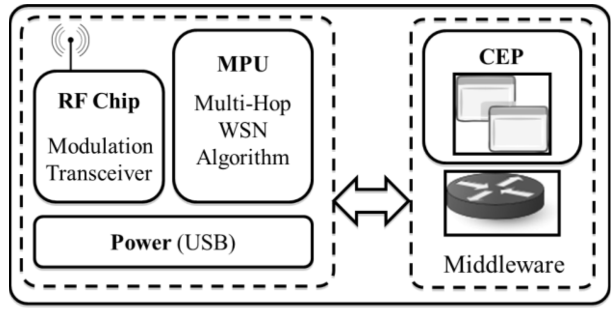


Fig. 4. Sink node and middleware for remote sensing.

making it easier to connect to cloud or Internet of Things framework.

Selecting a proper sensor is a relatively complicated issue as many factors need to be taken into consideration and always there is a trade-off among different factors such as life time, accuracy, power consumption, cost, and availability. Availability and accuracy are the driving determinants.

E. Sink Node

The sink node for sensor node cloud contains a communication unit and a data logger. Figure 4 depicts the middleware with the sink node. The communication unit is similar to the sensor node, which consists of a microprocessor (MSP430f2274) and low-power RF chip (CC2500). The data logger assigns the present time to each data using the real-time clock and stores the data into a non-volatile memory (Secure Digital (SD) card) in the data logger. The sink node sends data to middleware through Serial Communication Interface (SCI). The air quality data is stored in middleware and database (Internet of Things framework) and is processed using Complex Event Processing (CEP). This sink node is connected to the middleware directly. Thus, the sink node doesn’t need to consider the power consumption.

F. Power Consumption

There are three main sources that consume power/energy – sensors, data communications and data processing. Among which sensors consume much more power than the other two sources. Table VI summarizes the sensors used, sensor type and ranges of power consumption [26]–[28]. As shown, electrochemical type of sensor is more power efficient than MOS type. It doesn’t need power to operate (the power needed is for the peripheral circuit). MOS type of sensors needs significant amount of power to initiate and power to maintain their operations (thus, the max and min power levels are not much different). The optical type of sensors is more power efficient and once initiated, they don’t need much power to maintain. Therefore, proper selection of sensors is critical for power consumption as well as to reduce the amount of heats generated from the use of gas sensors.

The power needed for data communication and information processing involves the amount of data to be sent and frequency of sending the data. Therefore, reducing the data sampling frequency and using data aggregation algorithm to

TABLE VI
SUMMARY OF SENSOR USED, SENSOR TYPE*, DATA PROCESSING AND RANGE OF POWER CONSUMPTION+

Pollutant	Sensor Used	Sensor Type	Max [mA]	Min [mA]
GAC	TGS2600	MOS	50.04	41.54
VOCs	TGS2602	MOS	71.04	57.94
NOX	GSNT11	MOS	123.04	79.64
O_3	MiCS2610	MOS	45.44	35.24
CO	TGS5042	Electrochemical	0.96	0.96
SO_2	SO2-AF	Electrochemical	0.33	0.33
PM	GP2Y1010AU0F	Optical	20	11
CO_2	T6613	NDIR	141	9
Temperature/ Humidity	DHT11	Thermistor/ capacitive	1.21	1.21
Processing	MSP430 + CC2500		21.3	1.1
Processing	ADC + Computation		2.7	2.7

* Metal Oxide Semiconductor (MOS), Non Dispersive Infrared (NDIR);
\+ Power consumptions of the peripheral circuits are included, 5V power supply input.

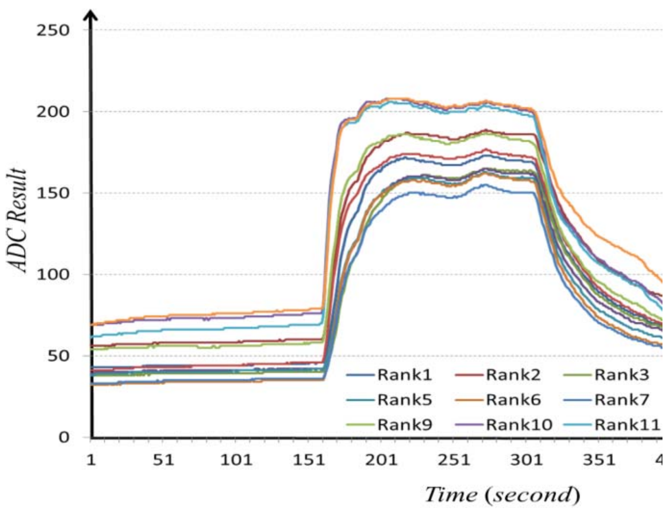


Fig. 5. Characteristic table of 12 TGS2601s for calibration.

reduce the amount of data to be sent can help to reduce power consumption. Please refer to Experiment and Results section for the impact of information processing.

Despite the fact that power source is less a problem for indoor than outdoor air quality monitoring, as AC power can be connected with the use of device or system for indoor use; battery power (regular or rechargeable) and backup battery are still needed to maintain system mobility and potential power breach.

G. Pre-Calibration

Characteristics of the MOS gas sensor vary from sensor to sensor and from production lot to production lot. Figure 5 shows the characteristic table for 12 TGS2601s [28]. As can be seen, twelve gas sensors have different characteristics. Thus, every gas sensor needs pre-calibration to accurately measure gas capacity. Gathered sensors' data are stored into characteristic tables in the integrated sensing units and the table and calibration algorithm [15] can resolve the pre-calibration problem.

TABLE VII
EXAMPLES OF ALERT USED

Context	AQI Level	Services / Actions
• Location • Time • Person • Pollutant Type	Good	Information
	Moderate	Attention / warning
	Unhealthy	Alert / warning
	Very Unhealthy	Alert, trigger for action
	Hazardous	Alert, trigger for action, contact for help

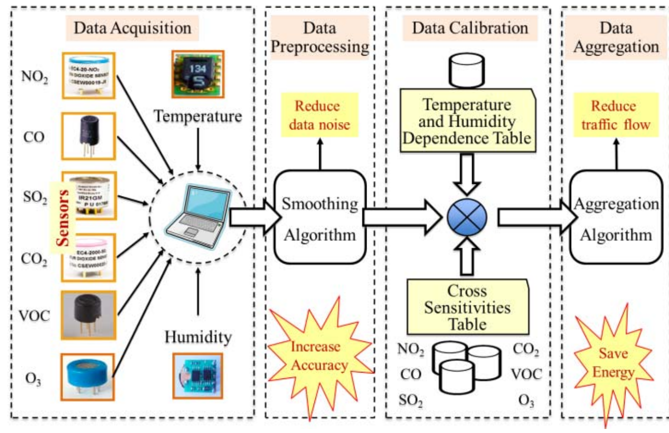


Fig. 6. Information flow for sensor information processing.

Another problem with sensory array is the lack of the long-term stability. This problem can be solved by auto-calibration method [17] and initialized-calibration. Every MOS gas sensor has distinctive characteristics in the beginning of sensing operation [15]. In our system, the initialized-calibration is used to resolve the long-term stability by checking the initialized action periodically.

H. Context Aware Services Design

A user-friendly alert scheme based on US EPA's AQI was developed to improve system usability and reduce the needs of technical expertise. Table VII shows the contexts and services that are provided for different levels of air quality based on the US EPA's AQI. Continuous visualized bars based on different color and size that indicates current air quality trend provides easier information.

V. INFORMATION PROCESSING

Information processing, which involves the process of extracting useful and timely information from the captured data to support decision making, is the cornerstone of any effective smart sensing applications. Figure 6 shows the common information processing tasks of sensor network: (1) data smoothing, in which a smoothing algorithm is introduced to prevent from sensor errors; (2) data calibration, in which temperature/humidity dependency and sensor cross sensitivity tables are used to calibrate the data captured from various gas sensors; and (3) data aggregation, in which an aggregation algorithm is used to reduce the network traffic and power consumption. The smoothed, calibrated and aggregated data is then sent to the middleware for inferencing and event processing.

Algorithm 1 An Algorithm for Smoothing Sensor Data

Rs (Sensor resistor)
Ro(Output load resistor)
Rs/Ro (Sensor resistance ratio)

```
if Rs/Ro ≥ ±3σ(Rs/Ro)
    then Filter out
else Keep Rs/Ro
end if
```

A. Smoothing Algorithm

It is common for sensors in an array generate suddenly temporary errors in electrical errors such as out of control. This kind of error comes up suddenly and disappeared. Ni et al. [34] defined the outlier as, "those measurements that significantly deviate from the normal pattern of sensed data". Outliers need to be detected and remove to get accurate measures. A smoothing algorithm is introduced for filtering out noise. To smooth gas sensor data, we capture important trends in repeated statistical survey and standard deviation with $\pm 3\sigma$ is used for limitation. Algorithm 1 illustrates the smoothing process.

B. Data Calibration

Gas sensors have the tendency of reacting to the presence of multiple gases. In other word, each sensor output does not exactly indicate the sensor's major target gas. Other gases may affect the sensor output. This phenomenon is known as cross sensitivity. Table VIII shows an example of cross sensitivity table for two gas sensors, MQ-6 and MQ-7. As can be seen, each gas sensor's output has multiple indications. For example, MQ-7 senses CO at 300 ppm and Hydrogen (H_2) at 190 ppm at 0.5 *Rs/Ro*. Therefore, using MQ-7 sensor alone cannot be used to accurately measure the type and level of gases due to cross sensitivity.

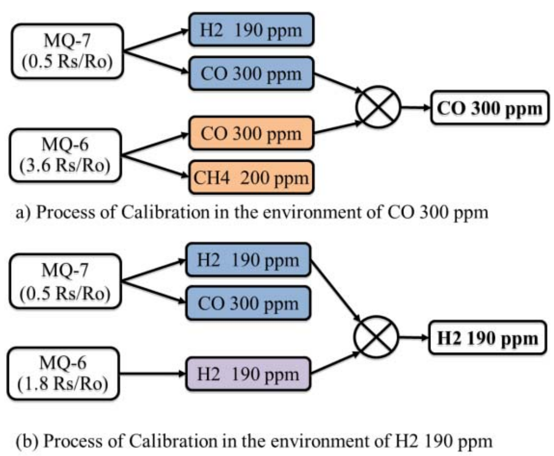
Example 2: Using the data from Table VIII as an example, if the MQ-7 gas sensor's output is 0.5 Rs/Ro, the output not only displays CO at 300 ppm but also shows H_2 at 190 ppm. Therefore, with only one sensor's output, it is not clear whether we have CO at 300 ppm or H_2 at 190 ppm. If at the same environment, MQ-6 provides a signal of 3.6 Rs/Ro. The sensor display CO at 300 ppm and Methane (CH_4) at 190. Since both of the sensor outputs display CO at 300 ppm, we can conclude that the environment has CO 300 ppm. Similarly, combining MQ-6 and MQ-7 data, we can conclude that the environment has H_2 gas 190 ppm. Figure 7 and Algorithm 2 depict the illustrative process of determining sensor result using cross sensitivity table.

C. Calibration With Temperature and Humidity Dependency

Thick film metal-oxide semiconductor sensors are often affected by temperature and humidity [16], [17]. For example, Table IX shows the resistance ratios (Rs/Ro) values under the same conditions for Figaro gas sensors [27]. Thus, pre-calibrated data needs to be adjusted periodically (or called running calibration) based on the temperature and humidity dependency table to ensure sensing accuracy.

TABLE VIII
SAMPLE CHARACTERISTIC FOR CROSS SENSITIVITIES

	Rs/Ro	CO	H2	CH4
MQ-7	0.1	4000	1800	-
	0.2	1200	600	-
	0.3	550	330	-
	0.4	400	250	-
	0.5	300	190	-
	0.6	200	150	-
	0.7	160	110	-
	1	100	70	-
MQ-6	0.68	-	10000	-
	0.72	-	4000	-
	0.9	-	1300	-
	1	-	1000	-
	1.1	-	800	-
	1.3	-	500	-
	1.8	-	190	-
	2.3	10000	-	2000
	2.6	2000	-	1000
	3	800	-	480
	3.6	300	-	200
	3.8	200	-	-



a) Process of Calibration in the environment of CO 300 ppm

(b) Process of Calibration in the environment of H2 190 ppm

Fig. 7. Process of calibration with cross sensitivities.

Algorithm 2 An Algorithm for Calibrating Sensor Data With Cross Sensitivities Table

S_{ID} (n=1,2,3,..6) , SV (# of Stored Value)

```
Parsing packet data
for ID from 0 to 5 by 1
    if S_out-ID has cross sensitivity
     then store the value and SV=SV+1
    end-if
 end-for
for SV from SV to 0 by 1
    Extract the multiple data
    Process of matching the table and values
end for
```

Algorithm 3 shows how we apply the auto-calibration process. The calibrated value of gas sensors (thick film metal oxide semiconductor type), Q_c, is calculated as:

$$Q_c = R_s/R_o \times T_H \tag{2}$$

TABLE IX
TEMPERATURE AND HUMIDITY DEPENDENCE FOR FIGARO GAS SENSORS (RS/RO)

R.H. (°C)	0% R.H.	20% R.H.	40% R.H.	65% R.H.	100% R.H.
-10	1.860	1.742	1.676	1.609	1.556
0	1.792	1.523	1.441	1.353	1.303
10	1.733	1.346	1.247	1.150	1.102
20	1.684	1.211	1.095	1.000	0.955
30	1.643	1.117	0.984	0.903	0.861
40	1.612	1.065	0.914	0.858	0.820

Algorithm 3 An Algorithm for Auto-Calibration of Temperature/Humidity Dependency

Q_c (Calibrated value of gas sensors), t (Temperature), λ_H (Humidity Value), D_c (Current ADC result value of gas sensor), R_L (The external resistance in an application circuit), D_M (The maximum ADC result value of a microcontroller), *Rs/Ro* (Sensor resistance ratio)

1: Capture sensors' values (D_c, Temperature, and Humidity)
2: Convert the captured values to dependency values (λ_H, γ, and ψ)
3: Calculate the value of temperature and humidity dependency, T_H, using (4)
4: Calculate the external sensor ratio, R_s / R_o,using (3)
5: Calculate the calibrated value of gas sensors, Q_c, using (2)

Where R_s is the sensor resistance in displayed gases at various concentrations, R_o is the sensor resistance in fresh air, and T_H is the temperature and humidity dependence value for calibration. The ratio of sensor resistances, R_s/R_o, is a converting value to quantity of gases using technical paper [27], [28], and is obtained by:

$$R_s/R_o = (R_L \times D_M/D_c) - R_L \tag{3}$$

Where R_L is the external resistance in an application circuit, D_M is the maximum ADC (Analog to digital converter) result value of a microcontroller, and D_c is the current ADC result value of a microcontroller. The temperature and humidity dependence value for calibration, T_H, can be obtained by:

$$T_H = \gamma t^2 - \psi t + \lambda \tag{4}$$

Where t is current temperature and λ is humidity dependency value. We set the factor values as follows: TGS2600 ($\gamma = 0.0047$, $\psi = 0.254$, $\lambda = 4.3633 + \lambda_H$) and TGS2602 ($\gamma = 0.001$, $\psi = 0.089$, $\lambda = 2.5252 + \lambda_H$) using table VIII and [27], [28].

D. Aggregation Algorithm

Energy consumption is a major concern with sensor network deployment. In this study, an aggregation algorithm is proposed to reduce power consumption. The key idea behind the aggregation algorithm is that each sensor node does not transmit packet data while the state (S_t) of sensing value is same as previous state (S_{t-1}). Algorithm 4 shows the operation for aggregating data. RF-transmit operation occurs only in the case of the transition from state to state.

Algorithm 4 An Algorithm for Aggregating Sensor Data

Rs/Ro (Sensor ratio)
S_t (States at time t = good, moderate, unhealthy for sensitive group, unhealthy, very unhealthy, hazardous)
t (1, 2, 3, 4, …T)

```
if S_{t-1} = S_t
  then No Transmit
else
  if S_{t-1} ≠ S_t
    then RF Transmit
  end if
end if
```

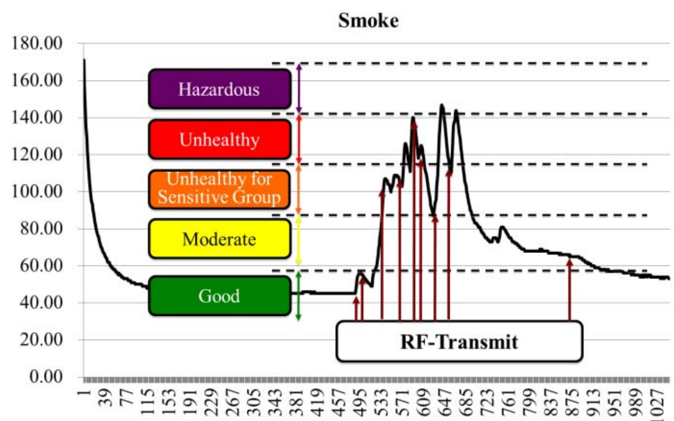


Fig. 8. Sample graph from TGS2601 with smoking.

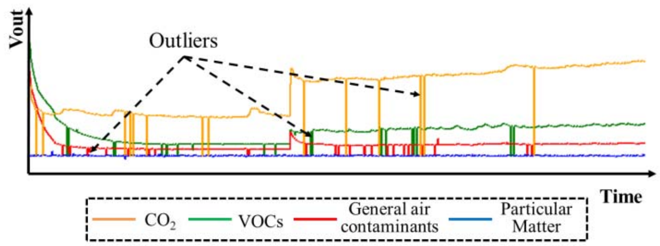
Figure 8 illustrates the sample graph when a person is smoking nearby TGS2601 gas sensor. At each state transition, data packets are transmitted. This algorithm can reduce the network traffic and power consumption.

VI. EXPERIMENTS AND RESULTS

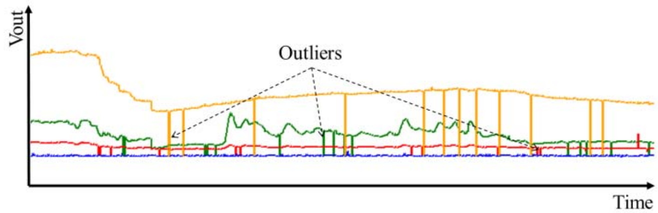
We conducted experiments in three different settings – a small size living room, a medium size classroom, and a big church – to illustrate the overall functionality of the prototype system, the factors impact air quality, the needs and effects of data smoothing, calibration and aggregation. The prototype monitor system that we constructed has four gas sensors built in and has an expansion socket to add additional sensors. The system was developed to closely monitor the indoor air quality level according to US EPA air quality standard (Note: The standard can be changed according to different regions/ countries requirements). To measure power consumption, we use the EZ GP4303D power supply device, configured to provide 5V to each sensor node. The EDM-4750 digital multimeter (DMM) is then used to measure the current flow between different sensors and peripheral circuits. This setup allows us to measure steady state current (via time integration) with transients.

A. Classroom Setting

In this experiment, we use a medium-size classroom (7m × 8.5m) as a target to test the feasibility of the system,

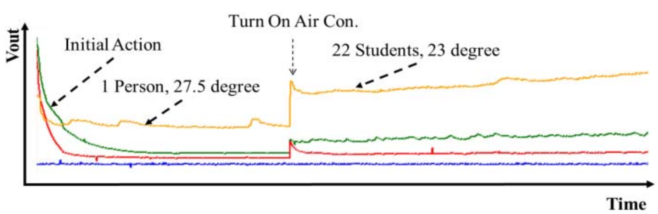


(a)

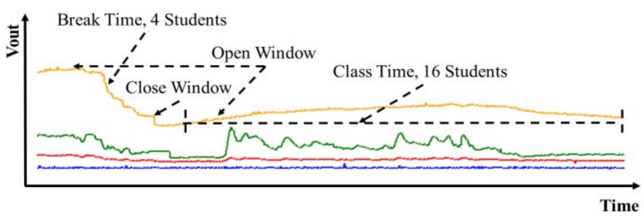


(b)

Fig. 9. Time series representation of air quality measurements in a classroom (before smoothing). (a) 1st phase of air quality (1 hour). (b) 2nd phase of air quality (1 hour).

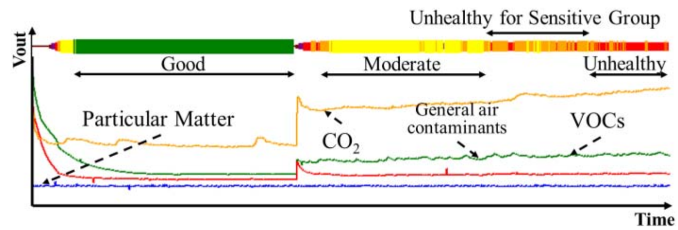


(a)

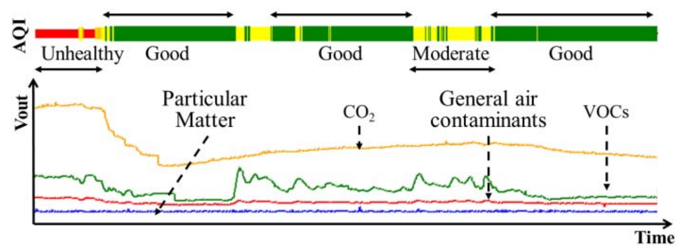


(b)

Fig. 10. Time series representation of air quality measurements and annotations in a classroom (after smoothing). (a) 1st phase of air quality (1 hour). (b) 2nd phase of air quality (1 hour).



(a)



(b)

Fig. 11. Time series representation of air quality measurements and AQI results in a classroom (after smoothing). (a) 1st phase of air quality (1 hour). (b) 2nd phase of air quality (1 hour).

the functionality of smoothing, calibration and aggregation algorithms, and the effect of impacting factors on air quality level. With this size of room, we estimate that one sensing node is sufficient, which was placed in the front side of the room.

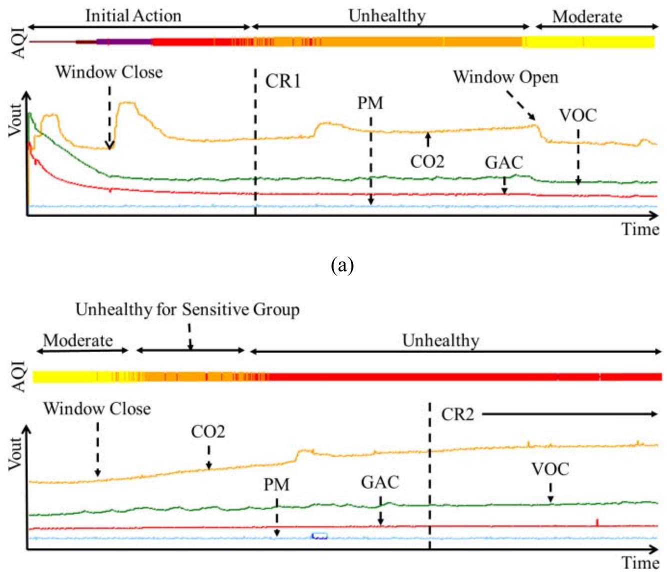
Figure 9 depicts the raw sensor signals of air quality monitoring without data smoothing. As expected, outliers, which cause detection errors, are commonly occurred in real time monitoring. Outliers need to be detected and remove to get accurate measures. Figure 10 shows the sensor signals after applying the proposed data smoothing algorithm. As can be seen, outliers are clearly removed so that the AQI index can be accurately computed.

Figure 10 also depicts the signal measures for air quality under four different scenarios: (1) initial sensor effect; (2) increase of people density; (3) temperature and humidity changes (by turning on air condition); and (4) close and open windows. First, when the sensor system is turned on, it needs about 5 to 10 minutes to reach stable time due to the initial action needed for any sensor. Second, CO_2 and VOCs level increase if people density increases. General air contaminants and Particular Matter are not much impacted by the increase of people density. Third, turning on air condition will have sudden change effects in CO_2, VOCs and general air contaminants, which will trigger the data calibration process to stabilize the air quality level. It doesn't have much impact on Particular Matter. Finally, opening windows will have major impact on the VOCs and will help to reduce CO_2 level, but it does not have much impact on the general air contaminants and Particular Matter (see Fig. 10(b)).

Figure 11 depicts the sensor signals along with the AQI results displayed according to the US EPA standard (Table III). As shown in Fig. 11(a), when the air quality monitoring system is turned on, it needs about 5 to 10 minutes to reach stable time as part of the sensor initiation. The AQI was initially marked as "good"; it was then gradually changed to "moderate" and then "unhealthy" when more people are in the room. The change in AQI will trigger to turn on air condition and in turn initiates the data calibration process to stabilize the measures and prevent air quality from deteriorating. During the break (as shown in Fig 11 (b)), the window was opened and the people density drop, the AQI improves to "good" level. But when the window was opened, it results in dynamic change in VOCs; thus, the AQI was marked as "moderate".

B. Living Room Setting

In this experiment, we use a small-size living room (3m $\times$ 2.5m) as a target to identify possible source of bad air quality. The room has a low quality sofa bed consisting of wood plastic composite (WPC) that continually emits VOC. Since the room is relatively small, we estimate that one sensing



(a)

(b)

Fig. 12. Time series representation of air quality measurements and AQI results in a living room (after smoothing). (a) 1st phase of air quality (1 hour). (b) 2nd phase of air quality (1 hour).

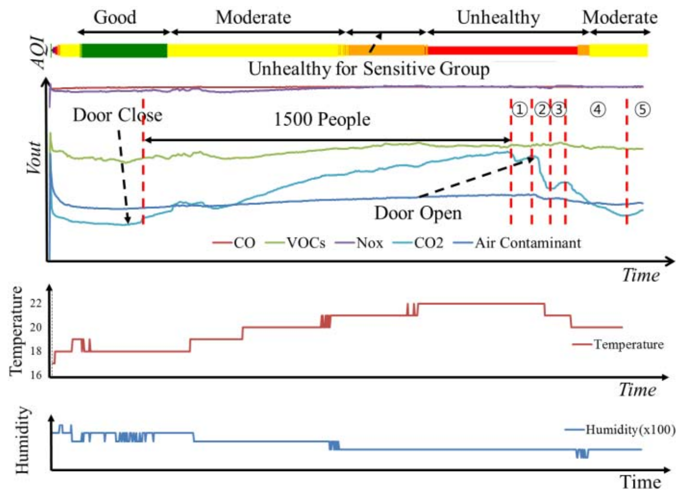


Fig. 13. Time series representation of air quality measurements, temperature, humidity and AQI results in a church space (after smoothing); (1) Worship is end, one door is open, and 300 people are out. (2) All doors are open and all people leave. (3) Cleaning and preparing for next event. (4) Less than 10 people (5) People are start to coming.

node is sufficient, which was placed in the front side of the room.

Figure 12 depicts the sensor signals along with the AQI results in this room. As shown in Fig. 12(a), when the air quality monitoring system is turned on, there is an initial period of unhealthy warning and the overall air quality level was worse than class room air quality. As shown, the AQI was initially marked as "Unhealthy for sensitive group"; it was gradually changed to "moderate" when widow is opened. Once the window is closed, it was gradually changed back to "Unhealthy for sensitive group" and "Unhealthy", and VOCs and CO_2 are increased.

C. Church Setting

We use a relatively large size of church (30m x 25m) to test the performance of the sensor network and its impact on people density. The church can host up to 1500 people in one time. With this size of room, we estimate that we may need to have two sensor nodes and a sink node to properly perform the duty. We found that there is not much different in the reading of the two sensor nodes. For simplicity, we use one sensor node placed at the left side (around middle) for experiment.

Figure 13 depicts the sensor signals along with the AQI results and temperature/humidity results in the church. As shown, when the air quality monitoring system is turned on, the AQI was initially marked as "good"; it was then gradually changed to "moderate" and then "unhealthy". Since the building uses gas heaters for warming the room, the room temperature was kept warm until the worship is end. Once the door is opened, it was gradually changed to "Unhealthy" and "Moderate", and NO_2 and CO_2 are decreased.

D. The Effects of Information Processing

Table X summarizes the average R_s and Q_c values of raw data (before smoothing), after smoothing, and after calibration (cali.) for sensor nodes under different settings. Classroom and living room tests were conducted during fall term when the temperature and humidity are more pleasant; therefore, the need for auto calibration is minor. While testing for church was conducted during winter time when heaters need to be turned on to maintain comfortable temperature. Therefore, periodic auto calibration is required to maintain accurate measurement.

In terms of air quality, as shown, the place where has more possible sources to emit bad airs constantly may result in bad air quality. For instance, church which has many bench-type chairs and big size gas heaters that emit VOC has overall high value of VOC and GAC. Similarly, the living room, which has a sofa bed that emits VOC constantly, has bad air quality than classroom. Also, because the area of living room is much smaller than the classroom and church, its CO_2 level is also the highest. The level of air quality often impacted by the existence of bad air sources/materials, room size, people density, and air flow etc.

The results also show the needs for calibration when there are significant changes in temperature and humidity. For example, the room temperature and humidity of classroom and living room did not change much during fall season so the change in Q_c values are not significant. As shown, for classroom, the Q_c values before calibration are 0.83 and 0.81 and the values after calibration are 0.80 and 0.92. However, the use of heaters in church during winter time triggered frequent auto-calibration; thus, there are major changes in Q_c values. The Q_c values before calibration are 0.13 and 0.13 but the values after calibration are 0.39 and 0.67, respectively.

Table XI summarizes the power consumption without and with data aggregation turns on for the three different settings used above. Table XI(a) shows the energy save with data sampling every 0.25 seconds (near real time) and XI(b) shows the energy save with data sampling every 4 second. As shown, reducing sampling frequency from 4Hz (0.25 seconds) to 0.25Hz (4 seconds) but without using data aggregation

TABLE X
AVERAGE SENSOR VALUES OF R_S AND Q_C FOR RAW, AFTER SMOOTHING, AND AFTER CALIBRATION UNDER DIFFERENT SETTINGS

Information Processing Results			VOC TGS2602	GAC TGS2600	CO_2 T6613	Temp. (°C)	Humidity (%)
Class Room	Before Smoothing	*Rs*	28.65	45.67	101.21	23.89	53.71
		Q_c	0.86	0.94	-		
	After Smoothing	*Rs*	28.83	46.09	102.39	Medium	High
		Q_c	0.83	0.81	-		
	After Cali.	Q_c	0.80	0.92	-		
Living Room	Before Smoothing	*Rs*	32.91	57.97	119.12	25.14	50.28
		Q_c	0.61	0.80	-		
	After Smoothing	*Rs*	33.31	58.83	**121.61**	High	Medium
		Q_c	0.61	0.77	-		
	After Cali.	Q_c	0.54	0.74	-		
Church	Before Smoothing	*Rs*	144.26	78.10	112.93	20.09	37.67
		Q_c	0.13	0.14	-		
	After Smoothing	*Rs*	**145.97**	**79.77**	114.18	Low	Low
		Q_c	0.13	0.13	-		
	After Cali.	Q_c	0.39	0.67	-		

TABLE XI
THE EFFECT OF DATA AGGREGATION. (a) 0.25 SECOND SAMPLING INTERVAL. (b) 4 SECOND SAMPLING INTERVAL

(a)

	Data Aggregation			
Location	Without (*mW*)	With (*mW*)	Saved (*mW*)	% Saved
Classroom	3.7160	3.3033	0.4127	11.107
Living Room	3.7160	3.3040	0.4121	11.089
Church	3.7160	3.3013	0.4147	11.160

(b)

	Data Aggregation			
Location	Without (*mW*)	With (*mW*)	Saved (*mW*)	% Saved
Classroom	3.3260	3.3033	0.0227	0.682
Living Room	3.3260	3.3040	0.0220	0.662
Church	3.3260	3.3013	0.0247	0.742

will lead to reduce power consumption for about 10.5% (=(3.716-3.326)/3.716); While, there is no power saved using data aggregation under this situation. Thus, reducing sampling frequency has more effect than using data aggregation algorithm. However, if we want to keep a real-time sampling, then we need to use data aggregation, which will save power consumption for about 11% and the value is not much impacted by the location setting.

E. Limitation and Challenge for Outdoor Sensing

There are several limitations and challenges to expand the indoor air quality sensing system to outdoor system.

First, communication distance or transmitting power needs to be considered for outdoor monitoring. The air quality of outdoor environment in local area is relatively similar comparing with the separated sections of indoor environment. In the outdoor air quality monitoring, sensing nodes are installed at a little bit distanced points that require high transmitting power and the sensitive antenna module. However, it will impact the power consumption and increases the node cost.

Second, there are various unexpected wireless traffics that may interrupt the wireless communications. Therefore, on site survey to identify possible interference sources is essential. Also, some sensor nodes may malfunction without notice, therefore, selecting communication protocols and structure that can perform automatic routing are important.

Third, sensing node has to rely on a limited capacity of battery. Most MOS type gas sensors consume a lot of energy and the low power electrochemical sensors have weaknesses such as they have relatively short lifespan and are short of target gases. Although solar panel recharging might be a solution, changing weather (rainy or cloudy) can be a variable for recharging scheme.

Finally, but not the least, the sensing node in outdoor environment is exposed to changing environment, and the sensors should be physically open to air flow. These characteristics require frequent maintenances in order to keep the quality of sensing. However, they also need to be weather prompt and need locked housing for protecting the electric circuit parts and board. Thus, appropriate level of packaging that meets both requirements is very important for outdoor air quality monitoring.

VII. CONCLUSION

In this paper, a smart integrated air quality monitoring system with multi-air pollutants sensor network for indoor environment is proposed. We describe the structure of an AQI-based monitoring with multiple gas sensors. Network architectures are proposed for remote indoor monitoring. Each sensing node includes an integrated sensor array that can measure Ozone, Particulate Matter, Carbon Monoxide, Nitrogen Oxides, Sulfur Dioxide, Carbon Dioxide, and Violate Organic Compounds. Temperature and humidity sensors support to calibrate the integrated sensor array's data. In the sensing node, a smoothing algorithm is introduced to prevent from temporary sensor errors and an aggregation algorithm is used to reduce the network traffic and power consumption. In the middleware, all data from sensor cloud are calibrated with the characteristic tables of the cross sensitivities and the temperature/humidity dependence.

Our experiments confirm that (1) many factors (e.g., the existence of sources or materials that may cause bad air, room size, people density, air flow, location, wind etc.) may impact air quality and air quality are changing constantly; thus, real time monitoring on air quality is essential; (2) sensor characteristics and environmental settings such as temperature and humidity may result in measuring errors; thus, pre-calibration and continual auto-calibration are needed; (3) using gas sensors for air quality monitoring consumes a lot of power; thus, how to properly select sensor type and improve energy efficiency during design and implementation stages are critical; and (4) identifying and removing the existence of sources or materials that may cause bad air and improving air flow are important to improve air quality.

The architecture and prototype of air quality monitoring system proposed in this study have the following advantages. First, it contains multi-air pollutants sensors so that it can be used for both indoor and outdoor environment under different settings. Secondly, its sensor node was built upon Raspberry Pi, making it easier for expansion and connecting with other Internet of Things applications. Thirdly, we develop a smoothing algorithm to reduce temporary sensor errors which may be caused by unexpected electrical errors and suggest two data calibration methods to reduce possible impact by environment such as temperature and humidity. These two developments help to ensure more accurate measurement. Finally, we propose an aggregation algorithm to reduce network traffic and power consumption; despite that more research is still needed to further explore the energy consumption issue.

Our work makes the following main technical contributions:

1) An energy efficient integrated sensing node or wireless sensor network system for monitoring indoor air quality in real-time was developed and tested. The sensing node was flexibly designed and integrated that it has multiple gas sensors built in yet has a socket for adding additional gas sensors.
2) Three simple yet effective information processing algorithms were proposed to increase monitoring accuracy and reduce energy consumption. First, a data smoothing algorithm was developed to reduce temporary sensor errors so as to increase measuring accuracy. Secondly, two types of data calibration methods were introduced to trigger auto-calibration so as to increase monitoring accuracy. And thirdly, a data aggregation algorithm was proposed to reduce network traffic and save energy consumption. These algorithms were integrated with the embedded microcontroller to achieve real-time monitoring alert and reduce power consumption.

REFERENCES

[1] P. F. Levelt *et al.*, "The ozone monitoring instrument," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1093–1101, May 2006.

[2] D. W. Dockery and C. A. Pope, "Acute respiratory effects of particulate air pollution," *Annu. Rev. Public Health*, vol. 15, pp. 107–132, May 1994.

[3] J. M. Samet, M. C. Marbury, and J. D. Spengler, "Health effects and sources of indoor air pollution. Part I," *Amer. Rev. Respirat. Disease*, vol. 136, no. 6, pp. 1486–508, Dec. 1987.

[4] *The Environment of Tokyo Metropolis, White Paper, Environmental Management Division*, Bureau of Environmental Protection, Tokyo, Japan, 2000.

[5] R. R. Ryder, "New York city automated air-quality data collection network," *IEEE Trans. Geosci. Electron.*, vol. 8, no. 2, pp. 81–88, Apr. 1970.

[6] W. Di Nicolantonio, A. Cacciari, and C. Tomasi, "Particulate matter at surface: Northern Italy monitoring based on satellite remote sensing, meteorological fields, and in-situ samplings," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 2, no. 4, pp. 284–292, Dec. 2009.

[7] W. Yanjia and H. Kebin, "The air pollution picture in China," *IEEE Spectrum*, vol. 36, no. 12, pp. 55–58, Dec. 1999.

[8] *U.S. Environmental Protection Agency*. [Online]. Available: http://www.epa.gov/air/airpollutants.html, accessed Aug. 15, 2013.

[9] EPA, *Buildings and Their Impact on the Environment: A Statistical Summary*, U.S. Environmental Protection Agency Green Building Workgroup, Washington, DC, USA. [Online]. Available: http://www.epa.gov/greenbuilding/pubs/gbstatspdf, accessed Oct. 1, 2014.

[10] S. Kim and E. Paulos, "InAir: Sharing indoor air quality measurements and visualizations," in *Proc. CHI*, Atlanta, GA, USA, Apr. 2010, pp. 1861–1869.

[11] Nemoto & Co., Ltd., Tokyo, Japan. *Nemoto Chemical Sensors Catalogue*. [Online]. Available: http://www.nemoto.co.jp/product

[12] F. Sarry and M. Lumbreras, "Gas discrimination in an air-conditioned system," *IEEE Trans. Instrum. Meas.*, vol. 49, no. 4, pp. 809–812, Aug. 2000.

[13] S. Marco, A. Ortega, A. Pardo, and J. Samitier, "Gas identification with tin oxide sensor array and self-organizing maps: Adaptive correction of sensor drifts," *IEEE Trans. Instrum. Meas.*, vol. 47, no. 1, pp. 316–320, Feb. 1998.

[14] A. P. Lee and B. J. Reedy, "Temperature modulation in semiconductor gas sensing," *Sens. Actuators B, Chem.*, vol. 60, no. 1, pp. 35–42, 1999.

[15] J.-Y. Kim, T.-Z. Shin, and M.-K. Yang, "Design of a smart gas sensor system for room air-cleaner of automobile (thick-film metal oxide semiconductor gas sensor)," *J. Elect. Eng. Technol.*, vol. 2, no. 3, pp. 408–412, 2007.

[16] J. H. Sohna, M. Atzeni, L. Zeller, and G. Pioggia, "Characterisation of humidity dependence of a metal oxide semiconductor sensor array using partial least squares," *Sens. Actuators B, Chem.*, vol. 131, no. 1, pp. 230–235, Apr. 2008.

[17] W. Tsujita, H. Ishida, and T. Moriizumi, "Dynamic gas sensor network for air pollution monitoring and its auto-calibration," in *Proc. IEEE Sensors*, vol. 1, Oct. 2004, pp. 56–59.

[18] O. A. Postolache, J. M. D. Pereira, and P. M. B. S. Girao, "Smart sensor network for air quality monitoring applications," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 9, pp. 3253–3262, Aug. 2009.

[19] J. J. Kim, S. K. Jung, and J. T. Kim, "Wireless monitoring of indoor air quality by a sensor network," *Indoor Built Environ.*, vol. 19, no. 1, pp. 145–150, 2010.

[20] Y. Jiang *et al.*, "MAQS: A personalized mobile sensing system for indoor air quality monitoring," in *Proc. 13th Int. Conf. UbiComp*, Beijing, China, 2011, pp. 271–280.

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

[22] J. Lozano, J. I. Suárez, P. Arroyo, J. M. Ordiales, and F. Álvarez, "Wireless sensor networks for indoor air quality monitoring," *Chem. Eng. Trans.*, vol. 30, pp. 319–324, 2012, doi: 10.3303/CET1230054.

[23] Y. Xiang, R. Piedrahita, R. P. Dick, M. Hannigan, Q. Lv, and L. Shang, "A hybrid sensor system for indoor air quality monitoring," in *Proc. IEEE Int. Conf. Distrib. Comput. Sensor Syst.*, May 2013, pp. 96–104.

[24] A. R. Kasar, D. S. Khemnar, and N. P. Tembhurnikar, "WSN based air pollution monitoring system," *Int. J. Sci. Eng. Appl.*, vol. 2, no. 4, pp. 55–59, 2013.

[25] EPA, "Air quality index—A guide to air quality and your health," Environmental Protection Agency, Research Triangle Park, NC, USA, Tech. Rep. EPA-454/K-03-002, 15.06.2008. [Online]. Available: http://www.njaqinow.net/App_AQI/AQI.en-US.pdf, accessed Oct. 1, 2014.

[26] *Technical Information for TGS2602*, FIGARO Engineering Inc., Osaka, Japan, Jan. 2005.

[27] *Technical Information for TGS2600*, FIGARO Engineering Inc., Osaka, Japan, Jan. 2005.

[28] *Technical Information for TGS5042*, FIGARO Engineering Inc., Osaka, Japan, Apr. 2013.

[29] S. Choi, N. Kim, H. Cha, and R. Ha, "Micro sensor node for air pollutant monitoring: Hardware and software issues," *Sensors*, vol. 9, no. 10, pp. 7970–7987, 2009.

[30] C. Liu and G. Cao, "Spatial-temporal coverage optimization in wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 5, pp. 465–478, Apr. 2011.

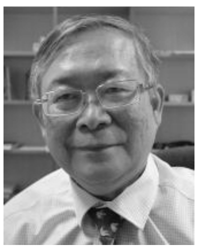
[31] Z. Liu, Q. Zheng, L. Xue, and X. Guan, "A distributed energy-efficient clustering algorithm with improved coverage in wireless sensor networks," *Future Generat. Comput. Syst.*, vol. 28, no. 5, pp. 780–790, 2012.

[32] T. Artursson, T. Eklöv, I. Lundström, P. Mårtensson, M. Sjöstrom, and M. Holmberg, "Drift correction for gas sensors using multivariate methods," *J. Chemometrics*, vol. 14, nos. 5–6, pp. 711–723, 2000.

[33] M. J. Wenzel, A. Mensah-Brown, F. Josse, and E. E. Yaz, "Online drift compensation for chemical sensors using estimation theory," *IEEE Sensors J.*, vol. 11, no. 1, pp. 225–232, Jan. 2011.

[34] K. Ni *et al.*, "Sensor network data fault types," *ACM Trans. Sensor Netw.*, vol. 5, no. 3, May 2009, Art. ID 25.

Jung-Yoon Kim received the B.S. degree in electronics and the M.S. degree in electrical and computer engineering from the University of Ulsan, Ulsan, Korea, in 2004 and 2006, respectively, and the Ph.D. degree in information sciences and technology from Pennsylvania State University, University Park, PA, USA, in 2014. He is currently a Research Fellow with the iCity Laboratory, Singapore Management University, Singapore. His areas of experience and expertise are smart health and wellbeing, especially in real-time cardiovascular disease and stress monitoring, physiological sensor design, and infrastructure and intelligent decision supports, environmental monitoring and assessment, especially in air quality monitoring, cut slope movement monitoring, and intelligent information processing based on machine learning and data mining, and ubiquitous computing, especially in embedded system design, energy-efficient processing, and programming model for networking performance.



Chao-Hsien Chu is currently a Professor of Information Sciences and Technology with Pennsylvania State University, University Park, PA, USA, where he received the Ph.D. degree in business in 1984. He is the Founding Director of the Smart Sensing Laboratory at Pennsylvania State University. His research interests include technologies and applications of Internet of things and cyber physical systems, intelligent technologies and their applications, and information security and privacy assurance. He has authored over 180 papers, many of them are in top-ranking journals or proceeding, such as the IEEE SENSORS JOURNAL, the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATIONS, the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, *INFORMS Journal on Computing*, the IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, *Decision Sciences*, *IIE Transactions*, ACM Conference on Computer and Communications Security, and the IEEE Conference on Computer Communications. Four of his papers have received the best paper award from different professional societies.

Sang-Moon Shin received the B.S., M.S., and Ph.D. degrees from the School of Electrical Engineering, University of Ulsan, Ulsan, Korea, in 2000, 2002, and 2014, respectively. He is currently a Research and Development Engineer with ARTECH Company, Ltd., Gumi, Korea. His research interests include various topics in indoor air quality, sensor networks, microwave/millimeter-wave communication, optical communication, high-voltage discharge, dielectric barrier discharge, ion generation, and lightning detection.