Occupancy Detection via Environmental Sensing

Ming Jin, Nikolaos Bekiaris-Liberis, Kevin Weekly, Costas J. Spanos, and Alexandre M. Bayen

***Abstract*— Sensing by proxy (SbP) is proposed in this paper as a sensing paradigm for occupancy detection, where the inference is based on "proxy" measurements such as temperature and CO_2 concentrations. The effects of occupants on indoor environments are captured by constitutive models comprising a coupled partial differential equation–ordinary differential equation system that exploits the spatial and physical features. Sensor fusion of multiple environmental parameters is enabled in the proposed framework. We report on experiments conducted under simulated conditions and real-life circumstances, when the variation of occupancy follows a schedule as the ground truth. The inference of the number of occupants in the room based on CO_2 concentration at the air return and air supply vents by our approach achieves an overall mean squared error of 0.6044 (fractional person), while the best alternative by Bayes net is 1.2061 (fractional person). Results from the projected ventilation analysis show that SbP can potentially save 55% of total ventilation compared with the traditional fixed schedule ventilation strategy, while at the same time maintain a reasonably comfort profile for the occupants.**

***Note to Practitioners*—Building indoor occupancy is essential to facilitate heating, ventilation, and air conditioning (HVAC) control, lighting adjustment, and geofencing to achieve occupancy comfort and energy efficiency. The significance of this paper is the proposal of a paradigm of sensing that results in a parsimonious and accurate occupancy inference model, which holds considerable potential for energy saving and improvement of HVAC operations. Parameters of the model are data-driven, which exhibit long-term stability and robustness across all the occupants' experiments. The proposed framework can also be applied to other tasks, such as indoor pollutants source identification, while requiring minimal infrastructure expenses. The data set and algorithm code are available to assist the comparison study.**

***Index Terms*— Building energy efficiency, differential equations, occupancy detection, sensing by proxy (SbP).**

Manuscript received April 14, 2016; revised July 17, 2016; accepted October 9, 2016. Date of publication November 14, 2016; date of current version April 5, 2018. This paper was recommended for publication by Associate Editor Q.-S. Jia and Editor M. P. Fanti upon evaluation of the reviewers' comments. This work was supported by the National Research Foundation, Singapore, through a Grant to the Berkeley Education Alliance for Research in Singapore for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics Program.

M. Jin, K. Weekly, C. J. Spanos, and A. M. Bayen are with the Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720 USA (e-mail: jinming@berkeley.edu; kweekly@berkeley.edu; spanos@berkeley.edu; bayen@berkeley.edu).

N. Bekiaris-Liberis is with the Department of Production Engineering and Management, Technical University of Crete, 73100 Chania, Greece (e-mail: nikos.bekiaris@dssl.tuc.gr).

This paper has supplementary downloadable multimedia material available at http://ieeexplore.ieee.org provided by the authors. The Supplementary Material contains a brief description of the data archive. This material is 16.1 MB in size.

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/TASE.2016.2619720

I. Introduction

THE thorough understanding of the interaction of occupants and the indoor environment has been the key component toward occupancy comforts and energy efficiency of buildings, which account for 40% of the total energy usage in the U.S. [1]. Intelligent buildings are conscious of both their occupancy and environment, in order to take control over their physical systems, such as heating, ventilation, and air conditioning (HVAC) and lighting, to optimize user comforts and energy consumption. The knowledge of zone-based occupancy coupled with adaptive building services offers considerable potential for energy reduction [2]–[7].

Existing approaches to indoor occupancy estimation rely on diverse sources of information, which can be broadly categorized into *direct* and *indirect* methods. For the direct methods, which prioritize detection accuracy over occupant privacy [8], vision- [2], [5], [9]–[13], tag- and smart-phone [3], [14]-based systems are typically employed. Erickson *et al.* [2] employed a wireless camera sensor network to count the number of occupants for real-time climate control, which is shown to achieve 42% annual energy savings while still maintaining American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) comfort standards. Meyn *et al.* [10] formulated a convex optimization-based estimator with the fusion of digital video cameras, passive infrared (PIR) detection, and CO_2 sensors. Information from thermal cameras and PIR sensors were used by k-nearest neighbor, linear regression, and artificial neural networks in the system of [11]. A framework depending on only depth image was demonstrated to mitigate privacy issues and scale well with the number of people [5]. The task has also been conducted by patrol robots with vision capability in a separate [12] or collaborative manner [13]. Radio frequency identification tags [3] and iBeacon on smart phones [14] have also been established as a reliable approach.

Indirect methods, in comparison, make use of less intrusive sensors, such as PIR [4], [15], pressure sensors [16], [17], electricity meters [18], [19], and environmental measurements such as acoustics, carbon monoxide (CO), total volatile organic compounds, small particulates (PM2.5), CO_2, illumination, temperature, and humidity [10], [20]–[26], as a privacy-performance tradeoff. PIR is able to give binary occupancy estimation [15], and can operate in conjunction with other sensors, such as magnetic reed switches [4], telephone hook sensors [27], and acoustic sensors [23] in a multisensor fusion model. Labeodan *et al.* [16] evaluated the use of chair sensors in an office building, and reported 0% error in the 8-h detection period. Nag and Mukhopadhyay [17]

developed a real-time dynamic thresholding scheme for flexiforce sensor in a residential house. Electricity usage often reflects indoor activities, and can be exploited for occupancy detection [18], [19], where [19] proposed a "zero training" algorithm without the need of collecting ground truth. Occupancy estimation using real (motion, door closure) and virtual (personal computer activity detector) sensors was presented in a small office based on the decision tree and artificial neural network models [28]. Several works have investigated the use of occupancy profiles for prediction [7], [15], [29]–[31]; in particular, Adamopoulou *et al.* [15] included contextual information to rapidly adjust to current conditions and capture unexpected events, and Jia *et al.* [31] introduced a decentralized stay-time-based occupant distribution estimation method facilitated by infrared beam systems.

Occupancy estimation through environmental monitoring is a promising approach, as parameters such as indoor CO_2 concentration are indicative of the presence of humans, which are the main source of variations; in addition, it can leverage existing sensing infrastructures without introducing significant privacy risks. A complex sensor network was established [20], [21] comprising ambient-sensing (lighting, temperature, relative humidity, motion detection, and acoustics), CO_2 sensing, and air quality sensing systems, which were incorporated into a hidden Markov model. Calì *et al.* [22] presented a CO_2-based binary output detection system, evaluated in office and residential buildings. Algorithms like conditional random field (CRF) [23], linear discriminant analysis (LDA), classification and regression trees, and random forest (RF) models [32] have been trained with humidity, temperature, light, and CO_2 measurements. Ebadat *et al.* [24] posed the estimation task as a deconvolution problem solved by the *fused-lasso* estimator using observations of CO_2 concentration, temperature, fresh air inflow, and door opening/closing events. Existing approaches, nevertheless, often require an extended training phase when data of ground truth occupancy are collected through surveys or camera recordings, which tends to limit its deployment. Another major drawback, especially for CO_2-based systems, is *the delay of detection* as a result of the relatively long time (10–15 min) it takes for the effect of human presence to build up to the detection threshold [10].

It is, therefore, the object of this paper to develop sensing by proxy (SbP), a sensing paradigm based on constitutive models that can promptly respond to occupancy changes, which makes it suitable for real-time indoor environment control [33]. The key contributions are as follows.

1) We develop a model based on partial differential equation (PDE) coupled with ordinary differential equation (ODE) that captures the spatial and temporal features of the system to uncover latent occupancy from environmental observations (Section II).
2) Our most significant contribution is the design, implementation, and evaluation of occupancy detection algorithm based on the SbP methodology in controlled and field experiments (Sections III and IV).
3) We examine the projected ventilation saving based on the ASHRAE standards [34], showing that our method can potentially save 55% of total ventilation compared with the traditional fixed schedule strategy, while maintaining similar comfort profile (Section V).

II. Sensing by Proxy: Methodology

A. Proxy Design and Modeling

We model the dynamics of the environmental parameters in the room using convection PDE with a source term that models the effect of human presence.

The source term, $\mathbf{X}(t) \in \mathbb{R}^m$, comprises m environmental measurements (e.g., temperature, CO_2 concentration) in their respective units (e.g., Celsius, part per million, or ppm). The state $\mathbf{X}(t)$ is the output of a linear, time-invariant, stable ODE system whose input $\mathbf{V}(t) \in \mathbb{R}^m$ represents the unknown humans' effect inside the room (within the vicinity of human bodies). For instance, $\mathbf{V}(t)$ can incorporate the occupants' rates of heat generation and CO_2 emission.

The relation is characterized by the following ODE:

$$\dot{\mathbf{X}}(t) = -A\mathbf{X}(t) + \mathbf{V}(t) \tag{1}$$

where we assume that the unmeasured environmental change rates due to the occupancy $\mathbf{V}(t)$ have the form of a piecewise constant signal

$$\dot{\mathbf{V}}(t) = \mathbf{0} \tag{2}$$

which is based on our experimental observation that the response of the environmental parameters in the room due to changes in the human's presence has some similarities with the step response of a low-pass filter. The matrix $A \in \mathbb{R}^{m\times m}$ characterizes the inertia of environments.

The ODE is coupled with a PDE that models the evolution of the environmental variables in the room, which is given as

$$\mathbf{u}_t(x,t) = -B\mathbf{u}_x(x,t) + B_X\mathbf{X}(t) \tag{3}$$

$$\mathbf{u}(0,t) = \mathbf{U}(t) \tag{4}$$

where $\mathbf{u}(x,t) \in \mathbb{R}^m$ denotes the environment at a time $t \geq 0$ and for $0 \leq x \leq 1$, $\mathbf{u}_t(x,t), \mathbf{u}_x(x,t)$ are standard notations for partial derivatives with respect to t and x, the steady-state environmental condition of the fresh incoming air is $\mathbf{U}_e \in \mathbb{R}^m$, and that of the fresh incoming air at the air supply vent is the input $\mathbf{U} \in \mathbb{R}^m$, which can be measured by sensors placed on the air supply vent. Parameter $B = \text{diag}(b_1, \ldots, b_m)$ is *a diagonal matrix* representing the speed of air convection in the room. The rate of dispersion from the local vicinity of the human to the room is reflected in B_X. We scale and center the dimension along the supply-return path so that the air supply is located at $x = 0$ and the air return is at $x = 1$; therefore, the spatial variable x is unitless and represents a normalized distance along the path. The environmental condition inside the room at the location of the air supply is represented by $\mathbf{u}(0,t)$, and that inside the room at the location of the air return is given as $\mathbf{u}(1,t)$.

The evolution of the environmental parameters in the room is thus modeled as a linear PDE–ODE system, one of whose inputs is the environmental condition in the fresh incoming

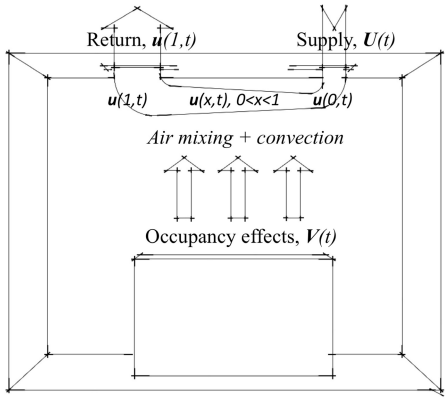

Fig. 1. Physical representation of the model. Fresh air with environmental condition $\mathbf{U}(t)$ enters the room from the supply vent, and exits the room after convection and mixing with human breath $\mathbf{V}(t)$, which rises to the ceilings, and the condition of the air at the return vent is $\mathbf{u}(1,t)$.

air at the location of the air supply, and the other input is the occupancy effects, if any. The output of the system can be viewed as the environmental condition of the air at the return vent, which is a mixture of air that convects from the air supply toward the air return and the local air influenced by humans. The environmental condition at the ceiling in a (nonratiometric) normalized distance along an axis from the supply to the return vent is indicated by the value of the PDE on the corresponding interior point of its spatial domain.

The physical representation of our model is illustrated in Fig. 1. The convection of air from air supply to the air return vent near the ceiling is represented by the PDE part. The diffusive term is intentionally omitted as it plays a relatively minor role in dispersing indoor pollutants as suggested in [35]. The model is applicable to environmental parameters that are transported mainly through the convection of air, such as temperature and CO_2. Those parameters are treated as the "proxy" for indoor occupancy.

Another design consideration involved is the modeling of the environmental condition near the ceiling, as this is where we see most effect from occupancy. One explanation is that the warm breath from a human occupant acts as a "bubble" of gas that rises to the ceiling, as it is more buoyant than the ambient, cooler air [36]. Thus, the air coming from lower in the room is modeled as a source term on the PDE across its entire path. The fact that this bubble of air does not immediately rise to the ceiling but is only gradually captured by the ODE part of the model behaves as a filter between the unknown human effects and the environmental conditions in the room.

B. Proxy Inference

The latent factors that are not directly observable are sensed by proxy based on the constitutive model that describes the evolution of proxy under the effects of latent factors. The temporal and spatial dynamics captured by the PDE–ODE model effectively regularizes the inference output. The approach is clearly different from discriminative models, e.g., LDA, support vector regression, and RF [32], [37], which assume that samples are independent and identically distributed. It shares some similarities with dynamic Bayesian models such as particle filters (PFs) and CRF [23], which account for time evolution of the underlying phenomenon. Nevertheless, proxy inference is directly derived from physical modeling and is thus more accurate and reliable with provable behavior as we show next.

The central task in this chapter is to derive an estimation strategy for latent factors, namely, the human effect on the environment $\mathbf{V}(t)$, based on proxy measurements at the supply vent $\mathbf{U}(t)$, and return vent $\mathbf{u}(1,t)$. For compactness, we define $\mathbf{Z}(t) = \begin{pmatrix}\mathbf{X}(t)\\ \mathbf{V}(t)\end{pmatrix} \in \mathbb{R}^{2m\times 1}$, so (1) and (2) can be written as

$$\dot{\mathbf{Z}}(t) = \bar{A}\mathbf{Z}(t) \tag{5}$$

where $\bar{A} = \begin{pmatrix} -A & I_{m\times m} \\ 0_{m\times m} & 0_{m\times m}\end{pmatrix}$ is defined using A from (1). Similarly, (3) can be recast as

$$\mathbf{u}_t(x,t) = -B\mathbf{u}_x(x,t) + B_Z\mathbf{Z}(t) \tag{6}$$

where $B_Z = (B_X \ 0_{m\times m})$ is the augmented matrix of B_X. We use $[\hat{B}_X]_{i,:}$ to denote the ith row of $\hat{B}_X$, and hat to indicate the estimated quantities.

We consider the following observer:

$$\hat{\mathbf{u}}_t(x,t) = -B\hat{\mathbf{u}}_x(x,t) + B_Z\hat{\mathbf{Z}}(t) + \mathbf{r}(x)\cdot\mathbf{L}(\mathbf{u}(1,t) - \hat{\mathbf{u}}(1,t)\big) \tag{7}$$

$$\hat{\mathbf{u}}(0,t) = \mathbf{U}(t) \tag{8}$$

$$\dot{\hat{\mathbf{Z}}}(t) = A\hat{\mathbf{Z}}(t) + \mathbf{L}(\mathbf{u}(1,t) - \hat{\mathbf{u}}(1,t)) \tag{9}$$

where $\mathbf{r}(x) \in \mathbb{R}^{m\times 2m}$ and $\mathbf{L} \in \mathbf{R}^{2m\times m}$ are yet to be determined, whereas $\mathbf{U}(t) \in \mathbb{R}^{m\times 1}$ and $\mathbf{u}(1,t) \in \mathbb{R}^{m\times 1}$ are the measurements of environmental parameters at the supply and return vents, respectively. The observer design for our PDE–ODE model is based on the design in [38], specifically Theorem 2. We refer the interested reader to [38] for the proof of the following corollary.

Corollary 1: Consider the system (7)–(9), where

$$\mathbf{r}(x) = (r_1 \ \cdots \ r_m)^\top \tag{10}$$

$$r_i(x) = \left(C_i - \int_0^{(1-x)/b_i} [B_Z]_{i,:}e^{-\bar{A}y}dy\right)e^{\bar{A}(1-x)/b_i} \tag{11}$$

$$C_i = \int_0^{1/b_i} [B_Z]_{i,:}e^{-\bar{A}\sigma}d\sigma, \quad i = 1,\ldots,m. \tag{12}$$

Let the pair $(\bar{A}, \bar{C})$, where

$$\bar{C} = \begin{pmatrix} C_1 \\ \vdots \\ C_m\end{pmatrix} \in \mathbb{R}^{m\times 2m}$$

be observable, and choose $\mathbf{L}$ such that the matrix $\bar{A} - \mathbf{L}\bar{C}$ is Hurwitz. Then, for any $\mathbf{Z}(0) \in \mathbb{R}^{2m}$, $u_i(x,t)$, $\hat{u}_i(x,t) \in L^2(0,1)$, $i = 1,\ldots,m$, where u_i is the ith component of $\mathbf{u}$, there exist positive constants λ and κ such that the following holds for all $t \geq 0$:

$$\Omega(t) \leq \kappa\Omega(0)e^{-\lambda t} \tag{13}$$

where

$$\Omega(t) = \int_0^1 \|\mathbf{u}(x,t) - \hat{\mathbf{u}}(x,t)\|^2 dx + \|\mathbf{Z}(t) - \hat{\mathbf{Z}}(t)\|^2. \tag{14}$$

C. Study Case: CO_2 Concentrations

We consider the case when there is only one key parameter, e.g., CO_2 concentration that is measured, in which case, (1)–(4) are reduced to the PDE–ODE systems

$$\dot{X}(t) = -aX(t) + V(t) \tag{15}$$

$$\dot{V}(t) = 0 \tag{16}$$

$$u_t(x,t) = -bu_x(x,t) + b_X X(t) \tag{17}$$

$$u(0,t) = U(t) \tag{18}$$

where both $X(t)$ and $V(t)$ are scalars, and the measure of how fast changes to the CO_2 emission rate by the humans affect the CO_2 concentration in the room is specified by the time constant $(1/a)$ in units of 100s. Positive parameter, b, in $(1/(100\text{ s}))$, represents the speed of air convection in the room. The rate of dispersion of CO_2 from the local vicinity of the human to the room is measured by b_X in $(1/(100\text{ s}))$, which is a positive number.

The corresponding observer can be obtained by the general result from the previous section

$$\hat{u}_t(x,t) = -b\hat{u}_x(x,t) + \begin{pmatrix} b_X & 0 \end{pmatrix}\begin{pmatrix} \hat{X}(t) \\ \hat{V}(t) \end{pmatrix} + \mathbf{r}(x)\begin{pmatrix} L_1 \\ L_2 \end{pmatrix}(u(1,t) - \hat{u}(1,t)) \tag{19}$$

$$\hat{u}_t(0,t) = U(t) \tag{20}$$

$$\begin{pmatrix} \dot{\hat{X}}(t) \\ \dot{\hat{V}}(t) \end{pmatrix} = \begin{pmatrix} -a & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \hat{X}(t) \\ \hat{V}(t) \end{pmatrix} + \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}(u(1,t) - \hat{u}(1,t)) \tag{21}$$

where $\mathbf{r}(x) = \begin{pmatrix} \pi_1(x) & \pi_2(x) \end{pmatrix}$, and

$$\pi_1(x) = \frac{b_X}{a}\left(e^{\frac{a}{b}x} - 1\right) \tag{22}$$

$$\pi_2(x) = \frac{b_X}{ba}x + \frac{b_X}{a^2}\left(1 - e^{\frac{a}{b}x}\right). \tag{23}$$

Note that the conditions of Corollary 1 are satisfied in this case when $b_X \neq 0$. The corresponding occupancy detection algorithm is shown in Algorithm 1.[1] As CO_2 is one of the most important indoor air parameters for occupants' health, and the installment of CO_2 sensors can be found in many HVAC systems, while our framework is expandable to incorporate several sensors in a fusion paradigm, we will base our experiments and analysis using CO_2 alone [36].

Algorithm 1: SbP for Occupancy Detection

function SENSINGBYPROXY(X^R,X^S, $Param$)
  Inputs: X^R: measurements at air return of size $1 \times T$
  X^S: measurements at air supply of size $1 \times T$
  $Param$: hyperparameters
    1) Model specification as in Table III: convection coefficient b, source term coefficient b_X, time constant of human effect a, human emission rate V^H, equilibrium concentration in air U_e
    2) Control parameters: L_1, L_2 as in (21) and (19).
    3) Spatial resolution d_s, temporal resolution d_t
    4) Smoothing window for median filter: w
  Initialization:
  $\hat{u}, \hat{X}, \hat{V} \leftarrow U_e\mathbf{1}(d_s, Td_t), \mathbf{0}(1, Td_t), \mathbf{0}(1, Td_t)$
  $x^R, x^S, \tau \leftarrow kron(X^R, \mathbf{1}(1, d_t)), kron(X^S, \mathbf{1}(1, d_t)), \frac{1}{d_t}$
  $r(n) \leftarrow L_1\frac{b_X}{a}(e^{a/b} - 1) + L_2\left(\frac{b_X}{ba} + \frac{b_X}{a^2}(1 - e^{a/b})\right)$
  Main program:
  for $t \in \{1, \ldots, Td_t\}$ **do**
    $\hat{u}(0,t) \leftarrow x^S(t)$ ▷ Eq.(20)
    for $n \in \{1, \ldots, d_s\}$ **do** ▷ PDE updates
      $\hat{u}_x(n,t) \leftarrow \left(\hat{u}(n,t) - \hat{u}(n-1,t)\right)d_s$ ▷ spatial
      $\hat{u}(n,t+1) \leftarrow \hat{u}(n,t) + \tau\Big(-b\hat{u}_x(n,t) + b_X\hat{X}(t) + r(n)\left(x^R(t) - \hat{u}(d_s,t)\right)\Big)$ ▷ Eq.(19) updates
    end for
    $\hat{X}(t+1) \leftarrow \hat{X}(t) + \tau\Big(-a\hat{X}(t) + \hat{V}(t) + L_1\left(x^R(t) - \hat{u}(d_s,t)\right)\Big)$ ▷ updates by (21)
    $\hat{V}(t+1) \leftarrow \hat{V}(t) + \tau L_2\left(x^R(t) - \hat{u}(d_s,t)\right)$
  end for
  Outputs: $y^{\text{Occupants}} \leftarrow \lfloor \frac{\text{median}(\hat{V}, w)}{V^H} + \frac{1}{2} \rfloor$ ▷ round of signal after median filter with window size w
end function

III. Experimental Design

A. Hardwares

As our approach is not particularly demanding of the accuracy of the proxy measurements, we employ the low-cost K30 CO_2 sensor [39], shown in Fig. 2, as the main module in our sensor platform. We implemented a local data storage solution with SD card, and plan to integrate a wireless transmission module in the long run to directly deposit data in our database. The sensor is capable of measuring CO_2 concentrations from 0 to 5000 ppm at a frequency of 1 Hz with an accuracy of ± 30 ppm, or $\pm 3\%$ of measured value, which is considered sufficient for the purpose of occupancy detection.

Sensor calibration is performed by the baseline method. We leave the sensors in a well-ventilated room with outdoor supply air for a few hours. The systematic offset ξ_i is given as

$$\xi_i = \frac{1}{T_{\text{cal}}}\sum_{t=1}^{T_{\text{cal}}} y_t - x_{\text{outdoor}} \tag{24}$$

where T_{cal} is the length of the calibration period, y_t is the sensor reading at time t, and x_{outdoor} is the outdoor CO_2 concentration, usually at 400 ppm. The offset ξ_i is subtracted from sensor i under *the well mixed assumption*, which states that "at steady state, the air in the room is well mixed, with the CO_2 concentration the same as the fresh air from the air supply vent."

B. Testbed Deployment

We implemented the experiments in a typical conference room, shown in Fig. 2, located in the Cory Hall on the University of California, Berkeley campus, whose occupancy is on-

[1]The code can be accessed at: http://people.eecs.berkeley.edu/∼jinming/

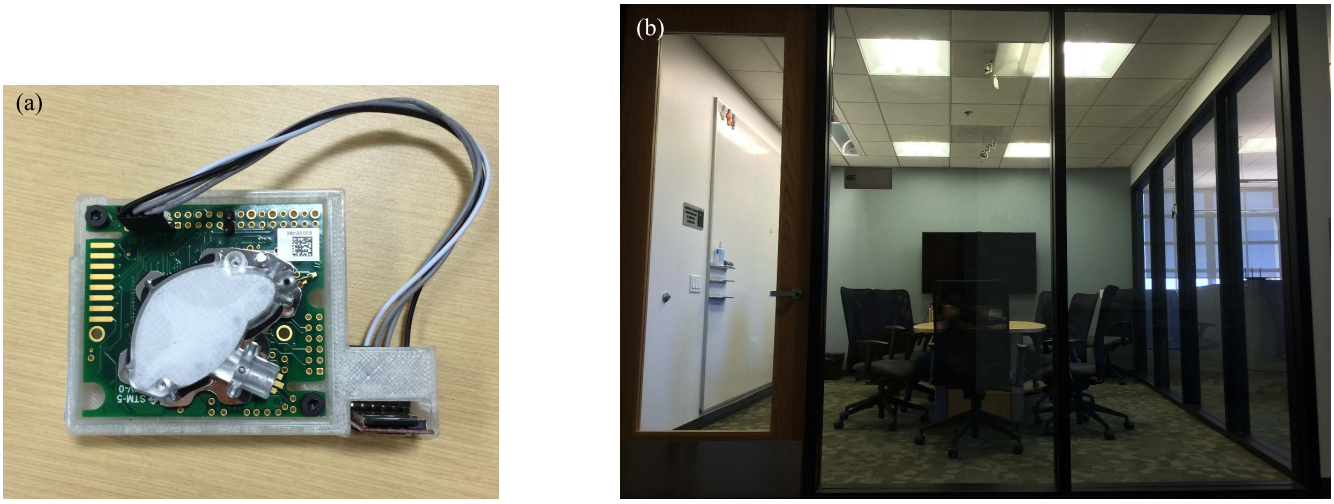

Fig. 2. (a) CO_2 sensor up close. (b) Testbed is a conference room of size $14 \times 10 \times 9$ ft^3, equipped with a full ventilation system including an air return vent and air supply vent, as illustrated in Fig. 1.

demand and not regular. The room bears close resemblance to other typical indoor spaces, with a ventilation system including air supply and air return vents on the ceiling. The sensors are placed on both the vents, in addition to the blackboard on the sidewall.

C. Experiments

Two types of experiments are performed, namely, CO_2 pump and occupants experiments, with different focuses.

For the CO_2 pump experiments, an outlet placed ∼20 cm above the desk injects beverage-grade CO_2 through a 200 W personal heater to emulate warm human breaths. The experiment is designed with two purposes. First, we want to examine the spatial dependence of the CO_2 concentration in the room. Second, we can collect data to identify the parameters of the model whose output matches the measured data, under different frequencies of excitation. Hence, we conducted experiments with the pump alternating between ON and OFF states, with the length of a full period of 30 min (A), 1 h (B), 3 h (C), and 10 h (D), whose results are detailed in Section IV.

For the occupants' experiments, the purpose is to validate SbP in a real setting. Hence, we performed both the controlled experiments (E) and field measurements (F, G). Our excitation procedure for the controlled experiments consists of adding or removing one of two participants of the experiment, and noting the time that the occupancy changes. The subjects are graduate students with a similar physique. The door is closed during the experiment, while the participants are engaged in normal activity such as working on their computers and talking to each other. The field measurements require much less commitment from the occupants, who are using the conference room for meetings or group study. The occupancy schedules for E, F, and G are demonstrated in Figs. 5 and 8.

IV. Results and Discussion

A. Experimental Results and Data Analysis

As described in the section of experimental design, we performed two groups of experiments, namely, one with

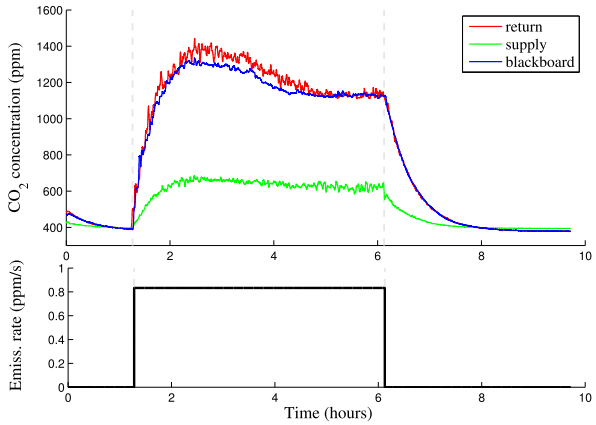

Fig. 3. CO_2 pump experiment D. The measured CO_2 concentration from different locations inside the room for a 5-h CO_2 release from a pump are shown.

CO_2 pump and the other with varying number of occupants. Based on the measurements, we make qualitative and quantitative analysis as a preparation.

1) CO_2 Pump Experiments (Hypothesis): When the CO_2 is injected for a long time with constant emission rate, the system reaches steady state.

The steady-state characterization experiment is conducted, when the pump is turned ON for five consecutive hours. Fig. 3 illustrates the measurements from the supply vent, return vent, and blackboard.

The rate of CO_2 concentration starts to decrease after a few hours, and reaches a plateau in the last hour. The steady-state concentration settles at around 1200 ppm as a result of mixing of fresh incoming air and CO_2 release.

Hypotheses: When the CO_2 is released periodically, the measurement exhibits periodic patterns according to the PDE–ODE system. Further, besides the transient behavior due to changes in the ventilation rate, the CO_2 concentrations from different points in the room react the same, albeit with different magnitudes.

Both the short period and long period excitation experiments are performed, with the periods of 30 min (15 min ON,

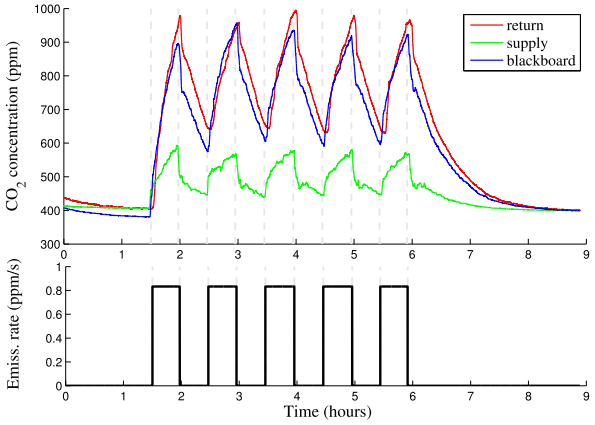

Fig. 4. CO_2 pump experiment B. Short-term excitation with period of 1 h. Measurements at return (red), supply (green) vents, and blackboard (blue) are shown.

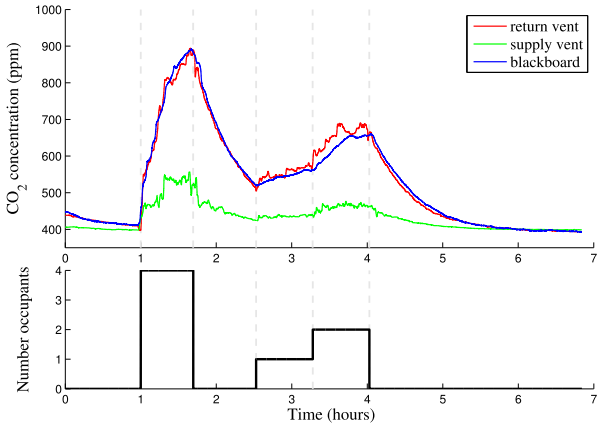

Fig. 5. Occupants experiment F. Field measurements during project discussion. Top: Proxy measurements. Bottom: Corresponding occupancy.

TABLE I

CROSS-CORRELATION VALUE OF CO_2 MEASUREMENTS AT DIFFERENT LOCATIONS FOR EXPERIMENTS A, B, AND C. THE MEAN AND STANDARD DEVIATION (IN PARENTHESIS) ARE SHOWN

Location	Cross-correlation
Return-Supply	.9277 (.0607)
Return-Blackboard	.9645 (.0316)
Supply-Blackboard	.9541 (.0292)

15 min OFF, same for the following), 1 h, and 3 h, as shown in Figs. 4 and 7.

As can be seen the CO_2 concentrations at all the sensed locations are responsive to the periodic injection, though the measurement at the air supply vent has a smaller magnitude compared with the blackboard and the air return vent. While the CO_2 accumulates from the start of the injection, the first-order derivative decreases as the room reaches a higher CO_2 concentration.

To quantitatively evaluate the spatial dependencies of sensors in the room, we now derive the cross correlation between measurements from three different locations for the CO_2 pump experiments. The definition of the cross correlation $r_{y_1y_2}$ between two signals y_1, y_2 that is employed here is given as

$$r_{y_1y_2} = \frac{\sum_{k=1}^{T}(y_1(k)-\bar{y}_1)(y_2(k)-\bar{y}_2)}{\sqrt{\sum_{k=1}^{T}(y_1(k)-\bar{y}_1)^2(y_2(k)-\bar{y}_2)^2}} \tag{25}$$

where $\bar{y}_1$ and $\bar{y}_2$ are the sample mean of y_1 and y_2, respectively. The cross correlation is a measure of the degree of linear dependence between two signals, and hence, it is a meaningful measure for comparing the measurements from different locations inside the room. The values of the cross correlations are shown in Table I. One can observe that the cross correlation between return and blackboard measurements is high, whereas the cross correlations that involve supply measurements are lower. This implies that the signals have a high degree of linear dependence [note that when $y_1(k) = c_1y_2(k) + c_2$, for all k, the cross correlation is one] on each other, although the correlation with the supply measurements is lower due to the ventilation operation. Note that the cross correlation between any two locations is derived as the average cross correlation obtained from the measurements of the experiments.

2) Occupants Experiments: As SbP aims at accurately infer occupancy through proxy, in addition to CO_2 pump experiments, occupants experiments are necessary for validation.

As described in the experimental design, we perform strictly controlled and field experiments. The former implements a designed schedule of occupancy, and requires the occupants to sit in designated chairs and remain in the room during the experiment, while allowing them to be engaged in normal activities such as using computers and chatting. The latter is taken during daily events and requires much less commitment from the occupants.

The following shows results from several such experiments, which substantially cover the usage of the conference room, and can be easily extended to other areas in the building. The field measurements are shown in Figs. 5 and 8 (right), and the strictly controlled experiment is illustrated in Fig. 8 (left). Note that to avoid a significant overlap between the graphs of this section and those of the simulation section, we arbitrarily decide which graphs show the blackboard measurement and the others show the simulated return, as long as the evidence is sufficient for the argument.

Similar to the pump experiment, CO_2 concentration increases almost immediately at the start of occupancy, and the concentration level and rate have a clear correspondence to the number of occupants in the room. The possibility of relating proxy measurements, namely, CO_2 concentration, to latent factors, namely, occupancy, lays the foundation for SbP.

Although the system is responsive to the change of occupancy, the time it takes to accumulate or deplete CO_2 to the corresponding stationary value is fairly long. From vacancy to a high-level occupancy, the measurement slowly sweeps across several intermediate levels. The difficulty of most distribution-based classification methods is illustrated in Fig. 6, where the significant overlapping of regions and misplacement of modes corresponding to different levels of occupancy will lead to confusion for standard machine learning algorithms. By modeling the temporal and spatial

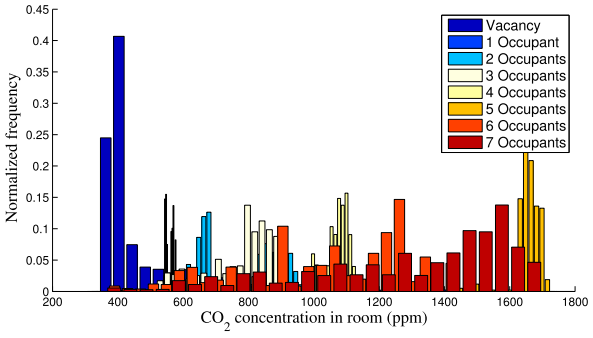

Fig. 6. Empirical distribution of CO_2 concentration for all occupants experiments corresponding to different occupancy.

TABLE II

PHYSICAL PARAMETERS OF PROXY MODEL USED IN ALL THE CO_2 PUMP EXPERIMENTS (A, B, C, D)

Physical parameter	Symbol	Value
Convection coefficient ($\frac{1}{100s}$)	b	2.5
Source coefficient ($\frac{1}{100s}$)	b_X	1.00
Time constant of human effect ($100s$)	$\frac{1}{a}$	16.67
Pump emission rate (ppm/sec)	ρ_{pump}	0.833
Equilibrium concentration in air (ppm)	U_e	400

dynamics of the system, as we demonstrate in the subsequent sections, we can develop an inference method that is both robust to noise and responsive to change of occupancy.

B. Simulation With Proxy Model

This section applies the model as described by (15)–(18) which links the location-specific proxy measurements to latent CO_2 emission factors to the CO_2 pump experiments and occupants measurements. In particular, we are concerned with the reproduction of the return vent measurements $u(1, t)$, i.e., the output of the system, given the supply vent measurements $U(t)$ and emission rate $V(t)$.

The results are illustrated in Fig. 7, where two experiments from CO_2 pump measurements are arbitrarily shown as the results are very similar. The set of parameters for the group of CO_2 pump experiments is determined by visual evaluation of the matching of simulation to the air return measurements, which is listed in Table II. The process of parameter evaluation is actually very simple, given the derived equation for stationary distribution

$$u^{\text{stationary}} = U_{\text{stationary}} + \frac{b_X V}{ab} \tag{26}$$

according to the model (15)–(18), where V is the fixed emission rate.

The stability of the CO_2 system can be seen in the good matching of all the air return vent measurements. There are, nevertheless, occasionally overmatching and undermatching, especially around the peak and valleys, which might be caused by the fluctuation of ventilation rates. The mismatch, even though not frequent, might introduce bias in our emission rate and occupancy estimations as we show in the following section. It is, therefore, recommended to examine the cause of the mismatch in actual building operations and periodically calibrate the model in order for sensing SbP to make the most

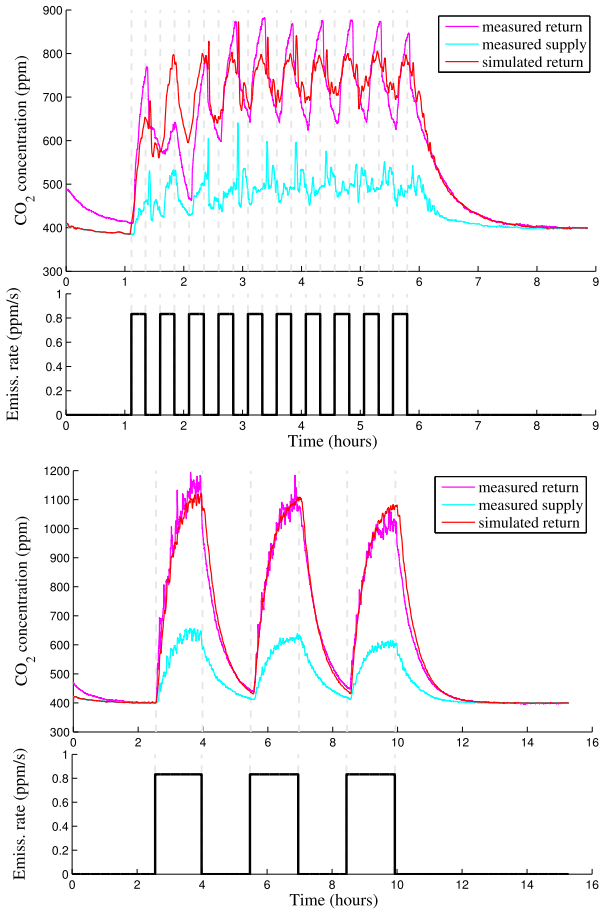

Fig. 7. Proxy model simulation with CO_2 pump experiment A with 30 min (top) and experiment B with 1 h (bottom) periodic excitation. Measurements at supply (green), return (blue), and simulated return (red) vents are shown.

reliable inference. It is also possible to design an automatic calibrator for each distributed sensor system.

Based on our experiences in the CO_2 pump experiments, we designed occupants controlled and field experiments to collect occupancy ground truth and validate our model in practice, as shown in Fig. 8.

In actual building usage, especially conference rooms and common areas, the occupancy is often irregular, as exemplified by the experimental profiles. The simulation of proxy measurements, therefore, is direct estimation of the effects of the irregular change of latent factors. The closeness of simulation matching to actual proxy measurements, as can be seen, is a clear indication of the accuracy of the model, and also ensures reliable inference of latent factors. The spatial and temporal simulation is illustrated in Fig. 9.

As a general remark, our proxy model is extremely simple and parsimonious with parameters. The set of parameters, including the convection coefficient b, the source coefficient b_X, time constant of human effects $(1/a)$, in addition to the human emission rate V and CO_2 concentration of fresh air U_e which are standard fixed parameters, are shared among all the experiments in the same group of CO_2 pump and occupants experiments, with relatively small difference between different groups due to the extent of emulation by the pump to human breathing. This makes our model extremely easy to train and

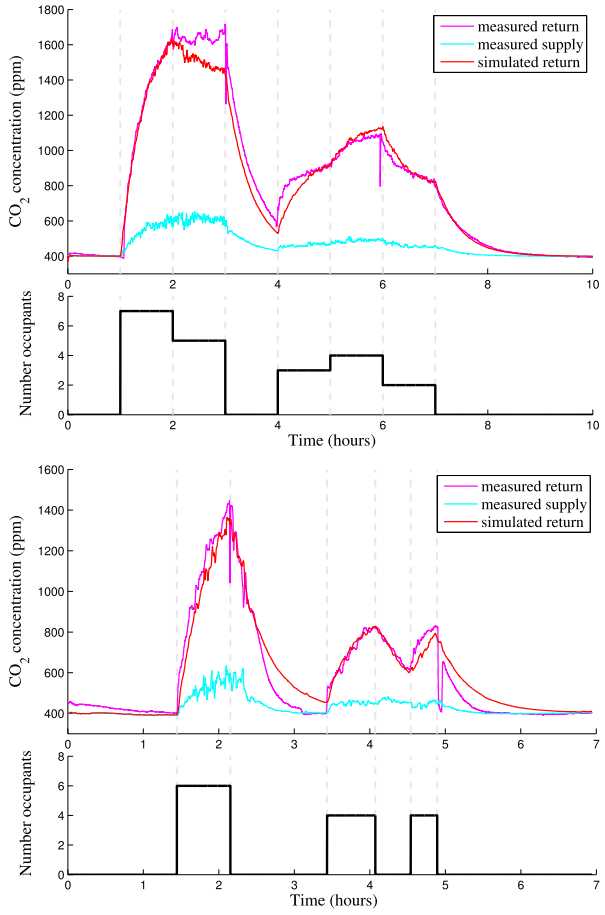

Fig. 8. Proxy model simulation with occupants experiments E (top) and G (bottom). The proxy measurements at return (blue), supply (green), and simulated return (red) vents are demonstrated.

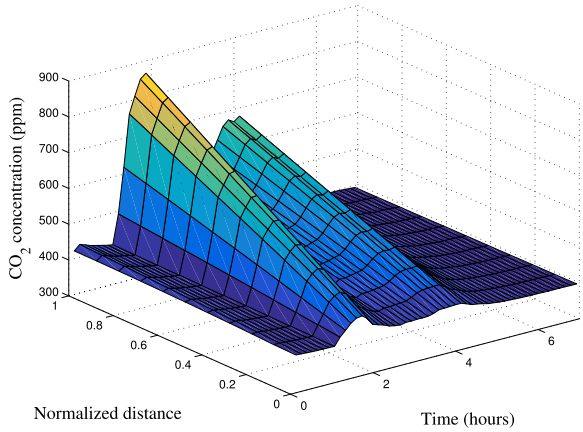

Fig. 9. Spatial and temporal dynamics of CO_2 concentration as represented as the states in the proxy model.

employ in practice. The additional advantage of parsimonious model relies on its stability and robustness by avoiding the potential overfitting problem. As we demonstrate next, the SbP approach substantially outperforms other popular methods and yet remains physically meaningful.

C. Proxy Inference of Occupancy

The observer model as described by (19)–(21) and Algorithm 1 are applied in this section to infer the CO_2 emission rate and occupancy based on proxy at return and supply vents.

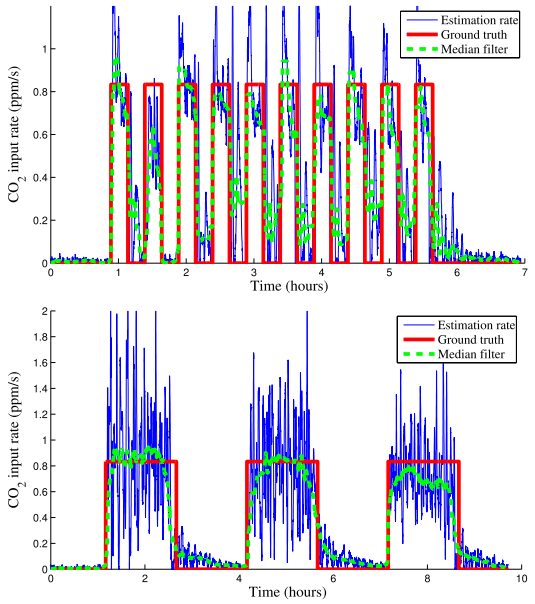

Fig. 10. Sensing the latent CO_2 emission rate by proxy for CO_2 pump experiments A (top) and C (bottom). The estimated emission rate (blue), median filtered rate (green), and ground truth (red) are given.

SbP distinguishes from other machine learning methods that assume independence of samples by implicitly considering time-autocorrelation of the latent emission rate. The advantage as a result is to have smooth state trajectory after simple signal processing, where we employed median filter directly on the estimated emission rate, $\hat{V}$, with a window of 8 min for experiment A, 20 min for B and C, and 25 min for all the occupant experiments. The median filter is a useful denoising method in signal processing, which is often preferred to mean filter to preserve relevant details and sharp transitions in the trace, as we will demonstrate next. Fig. 10 is plotted for the CO_2 pump experiment with periods of 30 min and 3 h, respectively.

Contrary to the common belief that CO_2-based methods are slow in response, SbP exhibits fast response to the change of occupancy. The previous argument is based on the fact that it takes time to accumulate CO_2 to a level that can be detected, and this accumulation time is fairly long as we observed in the experiments. During the accumulation, the concentration value sweeps across the stationary values for lower occupancy when several people enter the room, which accounts for the significant overlap in Fig. 6.

SbP, however, tackles this issue by modeling the dynamics of the measurements based on our model, which implicitly considers the increasing rate and stationary values to infer the actual occupancy. As a result, SbP is immediately responsive to changes of occupancy even when the transition is fairly frequent in the case of Fig. 10 (left), which is not possible with other methods since the concentration remains at a relatively high level even when the pump is turned OFF. The parameters chosen for the estimation are $L_1 = 2$, $L_2 = 0.02$, and the other physical parameters of the model are shown in Table II.

TABLE III

PHYSICAL PARAMETERS OF PROXY MODEL USED IN ALL THE OCCUPANTS EXPERIMENTS (E, F, G)

Physical parameter	Symbol	Value
Convection coefficient ($\frac{1}{100s}$)	b	2.5
Source coefficient ($\frac{1}{100s}$)	b_X	1.50
Time constant of human effect ($100s$)	$\frac{1}{a}$	16.67
Human emission rate (ppm/sec)	V^H	0.183
Equilibrium concentration in air (ppm)	U_e	400

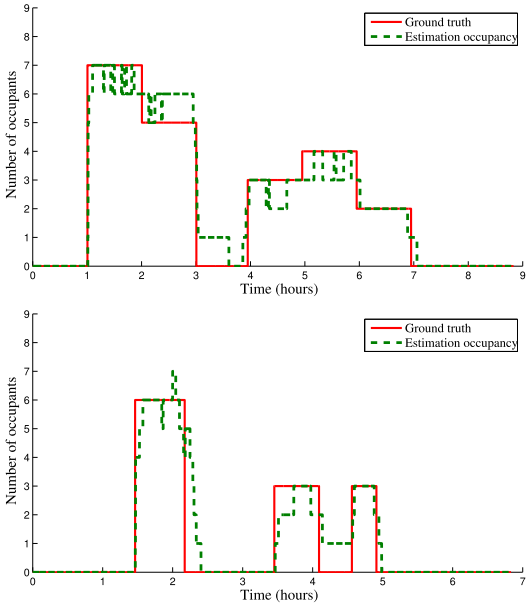

Fig. 11. Occupancy detection by SbP for experiments E (top) and G (bottom). The response times from vacancy to occupancy and *vice versa* are about 10 s and 5–10 min, respectively.

In the case of occupants estimation, the task is more difficult due to the following reasons. First, humans are not uniform in physique, so the emission rate must vary for different occupants. Second, the positions of the people sitting in the room are arbitrary, which might question our assumption that the human emission has a uniform effect on measurements on the ceiling regardless of positions of sources. Also, the ventilation rate, opening and closing of doors, and different activities might all introduce additional noise to the measurements. Nevertheless, regardless of these factors, Fig. 11 shows that SbP is reasonably robust to these influences, where we plot the estimated number of occupants together with the ground truth.

The fast transition behavior exhibited in the CO_2 pump experiments is also observed for the occupants experiments, even without any sensors to explicitly sense the exits or entry of people as in other methods such as PFs [40] or Markov models [2], [21]. The occupancy inference is accurate without explicitly specifying the transition rates of the occupancy model. For all these inferences, the parameters chosen are the same, namely, $L_1 = 2$, $L_2 = 0.02$, and physical parameters from Table III.

To compare with other models, we employ the root mean-squared-error (RMSE) with units of fractional people,

TABLE IV

COMPARISON OF RMSE OF ESTIMATION WITH OTHER MODELS IN OCCUPANTS EXPERIMENTS

	Exp. E	Exp. F	Exp. G	Mixed
Naïve Bayes	1.3080	0.7454	1.7457	1.3555
Bayes Net	1.2345	0.6555	1.5406	1.2061
Logistic regression	1.0796	0.6109	1.8414	1.4736
Multi-Layer Perceptron	0.9686	0.5672	1.6221	1.2321
RBF Network	1.0837	0.6760	1.6496	1.3341
Seq. Min. Opt. (SMO)	1.2326	0.6185	1.8803	1.6118
AdaBoostM1	1.6415	0.7053	2.2257	2.3927
Sensing by Proxy	**0.5644**	**0.5156**	**0.7331**	**0.6044**

given as

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{k=1}^{T}(\phi(k) - \hat{\phi}(k))^2} \tag{27}$$

where $\phi(k)$ is the ground truth occupancy at time k and $\hat{\phi}(k)$ is the estimated occupancy at time k given as

$$\hat{\phi}(k) = \left\lfloor \frac{\bar{\hat{V}}(k)}{\rho_{\text{human}}} + \frac{1}{2} \right\rfloor \tag{28}$$

where $\bar{\hat{V}}(k)$ is the median-filtered estimated emission rate at time k, $\rho_{\text{human}} = 0.183$ ppm/s is the average sedentary person emission rate, and $\lfloor x \rfloor$ is the floor operation to obtain the largest integer smaller than x.

The comparison of SbP with other methods is shown in Table IV. As all the other models require substantial training phase, the data are split into training and testing sets and the RMSE is computed by tenfold cross-validation. The algorithms take the measurements from the air supply and air return vents as features, where the corresponding labels are the number of occupants. The outputs for each testing point are the number of occupants obtained by classification, which are compared against the ground truth. For standardization purposes, we employ the Weka Machine Learning Toolkit [41] for the implementation of these algorithms. No time dynamic models such as PFs are learned for comparison, as it requires additional sensors to measure transitions and extra knowledge of transitional probabilities, which require substantial learning data and might not be reliable for the case of nonstationary activities in practice.

Although the parameters of our proxy model are shared across all the cases, the training for other models might be significantly different for each experiment, which differ by scale and time. Therefore, we decide to separate the RMSE for each experiment, as shown in the first three columns of Table IV, which might make it obvious which model is consistently better even with different data sets. In the last column, we combine all the occupants' experiments data and test each model. Note that as it is possible for other models to yield different outputs due to different training, SbP will output the same value given the chosen parameters, which is desirable as it is less susceptible to training noise.

SbP, as can be seen, delivers standout performance in all the testings, while the second best (underlined) positions are shared between multilayer perception (MLP) and Bayes Net,

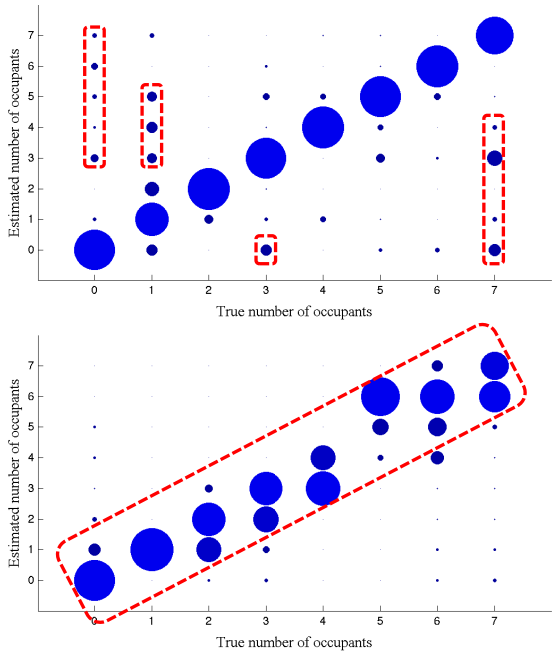

Fig. 12. Visualization of confusion matrix for Bayes Net (top) and SbP (bottom), where the position of circles represents the true number of occupants (x-axis) and estimated number of occupants (y-axis), and the size indicates the percentage normalized for each column. Bayes Net introduces nonnegligible misclassification with large deviation (red rectangle), whereas SbP makes inference within the ± 1 occupant region.

whole error metric almost doubles that of SbP in the mixed data set case. By ignoring the dynamics of CO_2 concentration, these algorithms are confused by the overlapping concentration region as shown in Fig. 6, especially during CO_2 accumulation and depletion period.

Close examination of the confusion matrix for our model and the second best model, in the mixed data set case, the Bayes Net, as visualized in Fig. 12, reveals an additional advantage of SbP. In the illustration, the size of the bubble represents the percentage of data classified as $\hat{\phi}$ (y-axis) for ground truth ϕ (x-axis), normalized for the sample size corresponding to ϕ. Bayes Net has a straight diagonal pattern, but it is undermined by the nonnegligible points far off the diagonal, representing misclassification error with large magnitude. On the contrary, SbP, though not possessing the straight diagonal pattern as in Bayes Net, is fairly clean of points far off the diagonal region. The point mass is also concentrated in the narrowband of subdiagonals, which indicate that the estimation is within an error of one person. This is clearly preferred in practice, as large estimation deviation might defeat the purpose of energy saving.

V. Projected Ventilation Impact

A. Experiments

The impact on ventilation saving and occupants comforts is analyzed based on the ASHRAE 62.1 standards for office buildings. As energy saving may vary among different buildings [34], which differ in materials, ventilation, heating and cooling efficiency, as well as outside temperature and humidity, we choose to focus on the minimum required ventilation, V_{bz}, to provide indoor air quality that is acceptable to occupants and minimizes adverse heal effects, given as

$$V_{bz} = R_p \cdot P_z + R_a \cdot A_z \tag{29}$$

where R_p and R_a are the outdoor airflow rate required per person and per unit area, respectively, A_z is the zone floor area, and P_z is the zone occupancy, whose values are listed in Table V.

TABLE V
Simulation Parameters According to ASHRAE for Minimum Ventilation Rates in Office Buildings

Simulation parameters	Value
A_z: Zone floor area (ft^2)	140
R_p: Outdoor airflow rate per person (CFM/person)	5
R_a: Outdoor airflow rate per area (CFM/ft^2)	0.06

For comparison, we include two conventional strategies recommended by ASHRAE 62.1, namely, conservative control that assumes maximum occupancy from 8 A.M. to 11 P.M. and 0 occupancy otherwise, and a fixed control that assumes maximum occupancy from 8 A.M. to 8 P.M. and 0 occupancy otherwise [34]. We also implemented the demand-control ventilation based on Bayes Net, SbP, MLP, and Logistic Regression, where R_p is replaced by $\hat{R}_p$ in (29) estimated according to the corresponding confusion matrices as visualized in Fig. 12, given the simulated number of occupants in each hour.

The simulation is based on two typical schedules, namely, under light usage and heavy usage, where the number of occupants in each hour is distributed according to a multinomial distribution. The top plot in Fig. 13 is the distribution of hourly occupancy (we omit the plots from the heavy schedule due to similarity and page limitation). The daily total ventilation, in cubic feet, is shown as the cyan boxplots in the middle plot. To ensure occupancy comforts, we want the estimated ventilation to be no lower than the required ventilation for the true occupancy less one (1), we also count the daily number of violations of this rule for each method, as shown in the cyan boxplots of the bottom plot. As the daily number of violations is worrisome for all the demand-control schemes except for SbP, we manually adjust the estimated occupancy for these methods by offsetting it with a margin of occupancy until we can keep the violations in an acceptable range, and redo the simulations, as shown in the magenta bars in these plots. Table VI is a summary of the simulation in the light usage case, where Sims A and B correspond to the methods without and with adjustments to meet the acceptable violation range.

Although conventional methods have negligible daily violations, they consume double the size of ventilation air compared to demand-control schemes. However, the saving of ventilation for demand-control, except for SbP, is at the expense of more violations, which deteriorates occupants' comforts. This is ameliorated by manual adjustments (Sim B), though it brings up the ventilation almost at the level of conventional strategies. SbP, nevertheless, represents 55% for the daily ventilation usage, respectively, compared with the fixed strategy, with a similar violation profile.

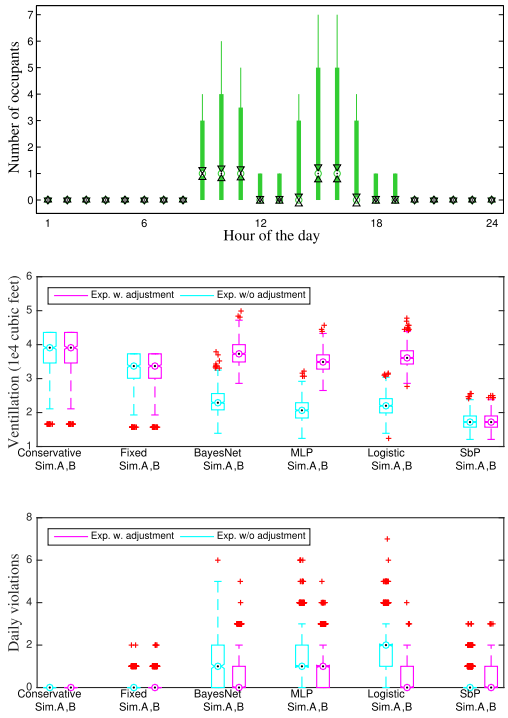

Fig. 13. Top: daily occupancy distribution for 1000 simulations. Middle: outside air capacity (cubic feet) for ventilation controlled by different schemes, without (Sim A, cyan) and with (Sim B, magenta) manual adjustments. Bottom: daily violation profiles for different methods in Sims A and B. The main body of the each box represents the first and third quartiles. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the "+" symbol.

TABLE VI

PROJECTED DAILY VENTILATION AND NUMBER OF VIOLATIONS FOR DIFFERENT STRATEGIES, WITHOUT (SIM A) AND WITH (SIM B) MANUAL ADJUSTMENTS FOR ACCEPTABLE VIOLATION PROFILES

	Ventilation (10^4cubic feet)		**Number of Violations**	
	Sim.A	**Sim.B**	**Sim.A**	**Sim.B**
Conservative strategy	3.8619	3.8750	0	0
Fixed strategy	3.3314	3.3419	0.0920	0.0950
Bayes Net	2.3412	3.7570	1.2960	0.6550
Multi-layer Perceptron	2.0925	3.5007	1.5480	0.9180
Logistic Regression	2.2055	3.6217	1.7350	0.3580
Sensing by Proxy	1.7473	1.7417	0.2640	0.2570

B. Limitation of Study

Although SbP demonstrates improved performance in both field and simulation experiments, a clear limitation of the study is that only the ceiling supply and return ventilation have been examined, and it remains to see how SbP applies to other types of air circulation, like the underfloor air distribution system. A viable mitigation strategy is to place sensors on the walls, as is shown in Figs. 3 and 5, where the blackboard CO_2 exhibited similar trends as the return vent in response to human presence.

The projected savings of 55% in the energy impact study are derived in simplified simulation studies based on ASHRAE 62.1 for only ventilation; if the combined heating/cooling and ventilation system is used in the building, both thermal and air quality performance need to be investigated to make sure the indoor environment meets the requirement of both ASHRAE 55 and 62.1. The numbers are subject to change for different locations and seasons in the undergoing field test.

VI. CONCLUSION

This paper describes an occupancy detection algorithm using environmental parameters based on the SbP methodology, which explores the spatial and temporal features of the system with constitutive models. Controlled field experiments are conducted in a typical indoor space to show that the proposed model can reproduce the CO_2 measurements given the latent emission rates. It is demonstrated that SbP can reliably detect the number of occupants based on "proxy" observations with RMSE of 0.6044 (factional person), compared with 1.2061 (factional person) of the best alternative machine learning algorithm. Investigation of the confusion matrices reveals that the estimation by SbP is within one occupant of the ground truth with high probability, while the estimation by Bayes Net sometimes has large deviations. Results from the projected ventilation analysis show that SbP can potentially save 55% of total ventilation compared with the traditional fixed schedule ventilation strategy, while at the same time maintain a reasonably comfort profile for the occupants. By successfully identifying the proxy candidates in the problem, SbP can also be applied to other tasks, such as indoor pollutants source identification, while requiring minimal capital investments.

ACKNOWLEDGMENT

BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore.

REFERENCES

[1] J. McQuade, "A system approach to high performance buildings," United Technol. Corp., Hartford, CT, USA, Tech. Rep. ADA559156, 2009.

[2] V. L. Erickson, M. Á. Carreira-Perpiñán, and A. E. Cerpa, "Occupancy modeling and prediction for building energy management," *ACM Trans. Sensor Netw.*, vol. 10, no. 3, Apr. 2014, Art. no. 42.

[3] J. Scott *et al.*, "PreHeat: Controlling home heating using occupancy prediction," in *Proc. 13th Int. Conf. Ubiquitous Comput.*, 2011, pp. 281–290.

[4] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, and T. Weng, "Occupancy-driven energy management for smart building automation," in *Proc. 2nd ACM Workshop Embedded Sens. Syst. Energy-Efficiency Building*, 2010, pp. 1–6.

[5] G. Diraco, A. Leone, and P. Siciliano, "People occupancy detection and profiling with 3-D depth sensors for building energy management," *Energy Buildings*, vol. 92, pp. 246–266, Apr. 2015.

[6] A. L. Pisello, M. Bobker, and F. Cotana, "A building energy efficiency optimization method by evaluating the effective thermal zones occupancy," *Energies*, vol. 5, no. 12, pp. 5257–5278, 2012.

[7] T. Hong, H. Sun, Y. Chen, S. C. Taylor-Lange, and D. Yan, "An occupant behavior modeling tool for co-simulation," *Energy Buildings*, vol. 117, pp. 272–281, Apr. 2016.

[8] R. Jia, R. Dong, S. Shankar Sastry, and C. J. Spanos. (Jul. 2016). "Privacy-enhanced architecture for occupancy-based HVAC control." [Online]. Available: https://arxiv.org/abs/1607.03140

[9] C.-C. Huang and S.-J. Wang, "A Bayesian hierarchical framework for multitarget labeling and correspondence with ghost suppression over multicamera surveillance system," *IEEE Trans. Autom. Sci. Eng.*, vol. 9, no. 1, pp. 16–30, Jan. 2012.

[10] S. Meyn, A. Surana, Y. Lin, S. M. Oggianu, S. Narayanan, and T. A. Frewen, "A sensor-utility-network method for estimation of occupancy in buildings," in *Proc. 48th IEEE Conf. Decision Control, 28th Chin. Control Conf. (CDC/CCC)*, Dec. 2009, pp. 1494–1500.

[11] A. Beltran, V. L. Erickson, and A. E. Cerpa, "Thermosense: Occupancy thermal based sensing for hvac control," in *Proc. 5th ACM Workshop Embedded Syst. Energy-Efficient Buildings*, 2013, pp. 1–8.

[12] F.-M. Chang, F.-L. Lian, and C.-C. Chou, "Integration of modified inverse observation model and multiple hypothesis tracking for detecting and tracking humans," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 1, pp. 160–170, Jan. 2016.

[13] P. Dames and V. Kumar, "Autonomous localization of an unknown number of targets without data association using teams of mobile sensors," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 3, pp. 850–864, Jul. 2015.

[14] A. Corna, L. Fontana, A. A. Nacci, and D. Sciuto, "Occupancy detection via iBeacon on Android devices for smart building management," in *Proc. Design, Autom. Test Eur. Conf. Exhibit.*, 2015, pp. 629–632.

[15] A. A. Adamopoulou, A. M. Tryferidis, and D. K. Tzovaras, "A context-aware method for building occupancy prediction," *Energy Buildings*, vol. 110, pp. 229–244, Jan. 2016.

[16] T. Labeodan, K. Aduda, W. Zeiler, and F. Hoving, "Experimental evaluation of the performance of chair sensors in an office space for occupancy detection and occupancy-driven control," *Energy Buildings*, vol. 111, pp. 195–206, Jan. 2016.

[17] A. Nag and S. C. Mukhopadhyay, "Occupancy detection at smart home using real-time dynamic thresholding of flexiforce sensor," *IEEE Sensors J.*, vol. 15, no. 8, pp. 4457–4463, Aug. 2015.

[18] W. Kleiminger, C. Beckel, T. Staake, and S. Santini, "Occupancy detection from electricity consumption data," in *Proc. 5th ACM Workshop Embedded Syst. Energy-Efficient Buildings*, 2013, pp. 1–8.

[19] M. Jin, R. Jia, Z. Kang, I. C. Konstantakopoulos, and C. J. Spanos, "PresenceSense: Zero-training algorithm for individual presence detection based on power monitoring," in *Proc. 1st ACM Conf. Embedded Syst. Energy-Efficient Buildings*, 2014, pp. 1–10.

[20] K. P. Lam *et al.*, "Occupancy detection through an extensive environmental sensor network in an open-plan office building," in *Proc. 11th Int. IBPSA Building Simulation*, Jul. 2009, pp. 1452–1459.

[21] B. Dong *et al.*, "An information technology enabled sustainability testbed (ITEST) for occupancy detection through an environmental sensing network," *Energy Buildings*, vol. 42, no. 7, pp. 1038–1046, Jul. 2010.

[22] D. Calì, P. Matthes, K. Huchtemann, R. Streblow, and D. Müller, "CO_2 based occupancy detection algorithm: Experimental analysis and validation for office and residential buildings," *Building Environ.*, vol. 86, pp. 39–49, Apr. 2015.

[23] S. Zikos, A. Tsolakis, D. Meskos, A. Tryferidis, and D. Tzovaras, "Conditional random fields—Based approach for real-time building occupancy estimation with multi-sensory networks," *Autom. Construct.*, vol. 68, pp. 128–145, Aug. 2016.

[24] A. Ebadat, G. Bottegal, D. Varagnolo, B. Wahlberg, and K. H. Johansson, "Regularized deconvolution-based approaches for estimating room occupancies," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 4, pp. 1157–1168, Oct. 2015.

[25] M. Jin, N. Bekiaris-Liberis, K. Weekly, C. Spanos, and A. M. Bayen, "Sensing by proxy: Occupancy detection based on indoor CO_2 concentration," in *Proc. 9th Int. Conf. Mobile Ubiquitous Comput., Syst., Services Technol. (UBICOMM)*, Jul. 2015, pp. 1–10.

[26] M. Jin, H. Zou, K. Weekly, R. Jia, A. M. Bayen, and C. J. Spanos, "Environmental sensing by wearable device for indoor activity and location estimation," in *Proc. 40th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2014, pp. 5369–5375.

[27] R. H. Dodier, G. P. Henze, D. K. Tiller, and X. Guo, "Building occupancy detection through sensor belief networks," *Energy Buildings*, vol. 38, no. 9, pp. 1033–1043, Sep. 2006.

[28] S. Srirangarajan and D. Pesch, "Occupancy estimation using real and virtual sensors," in *Proc. 12th Int. Conf. Inf. Process. Sensor Netw.*, 2013, pp. 347–348.

[29] D. Yan *et al.*, "Occupant behavior modeling for building performance simulation: Current state and future challenges," *Energy Buildings*, vol. 107, pp. 264–278, Nov. 2015.

[30] J. Zhao, B. Lasternas, K. P. Lam, R. Yun, and V. Loftness, "Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining," *Energy Buildings*, vol. 82, pp. 341–355, Oct. 2014.

[31] Q. S. Jia, H. Wang, Y. Lei, Q. Zhao, and X. Guan, "A decentralized stay-time based occupant distribution estimation method for buildings," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 4, pp. 1482–1491, Oct. 2015.

[32] L. M. Candanedo and V. Feldheim, "Accurate occupancy detection of an office room from light, temperature, humidity and CO_2 measurements using statistical learning models," *Energy Buildings*, vol. 112, pp. 28–39, Jan. 2016.

[33] S. Goyal, P. Barooah, and T. Middelkoop, "Experimental study of occupancy-based control of HVAC zones," *Appl. Energy*, vol. 140, pp. 75–84, Feb. 2015.

[34] *Ventilation for Acceptable Indoor Air Quality*, ASHRAE Standard 62.1, American Society of Heating, Refrigeration and Air-Conditioning Engineers, Atlanta, GA, USA, 2004.

[35] A. V. Baughman, A. J. Gadgil, and W. W. Nazaroff, "Mixing of a point source pollutant by natural convection flow within a room," *Indoor Air*, vol. 4, no. 2, pp. 114–122, Jun. 1994.

[36] K. Weekly, N. Bekiaris-Liberis, M. Jin, and A. M. Bayen, "Modeling and estimation of the humans' effect on the CO_2 dynamics inside a conference room," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 5, pp. 1770–1781, Sep. 2015.

[37] A. Ebadat, G. Bottegal, D. Varagnolo, B. Wahlberg, and K. H. Johansson, "Estimation of building occupancy levels through environmental signals deconvolution," in *Proc. 5th ACM Workshop Embedded Syst. Energy-Efficient Buildings*, 2013, pp. 1–8.

[38] N. Bekiaris-Liberis and M. Krstic, "Lyapunov stability of linear predictor feedback for distributed input delays," *IEEE Trans. Autom. Control*, vol. 56, no. 3, pp. 655–660, Mar. 2011.

[39] *K-30 10, 000 ppm CO2 Sensor*, accessed on Nov. 1, 2016. [Online]. Available: http://www.co2meter.com/products/k-30-co2-sensor-module

[40] V. L. Erickson, S. Achleitner, and A. E. Cerpa, "POEM: Power-efficient occupancy-based energy management system," in *Proc. 12th Int. Conf. Inf. Process. Sensor Netw.*, Apr. 2013, pp. 203–216.

[41] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: An update," *ACM SIGKDD Explorations Newslett.*, vol. 11, no. 1, pp. 10–18, 2009.

Ming Jin received the B.Eng. (Hons.) degree in electronic and computer engineering from the Hong Kong University of Science and Technology, Hong Kong. He is currently pursuing the Ph.D. degree in electrical engineering and computer science with the University of California at Berkeley, Berkeley, CA, USA.

His current research interests include intelligent buildings, statistical learning, and control theory, with a focus on indoor positioning, occupancy detection, and social game for energy savings.

Mr. Jin was a recipient of the Best Paper Award at the 2015 International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, the Electronic and Computer Engineering Department Scholarship, the School of Engineering Scholarship, and the University Scholarship at the Hong Kong University of Science and Technology.

Nikolaos Bekiaris-Liberis is currently a Post-Doctoral Researcher with the Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece. He has co-authored the SIAM book *Nonlinear Control under Nonconstant Delays*. His interests are in delay systems, distributed parameter systems, nonlinear control, and their applications.

Dr. Bekiaris-Liberis was a Finalist for the Student Best Paper Award at the 2010 ASME Dynamic Systems and Control Conference and the 2013 IEEE Conference on Decision and Control. He was a recipient of the Chancellor's Dissertation Medal in Engineering from the University of California at San Diego, San Diego, in 2014, and the Best Paper Award at the 2015 International Conference on Mobile Ubiquitous Computing, Systems, Services, and Technologies.

Kevin Weekly received the B.S. degree in electrical engineering and computer sciences from the University of Texas at Dallas, Richardson, TX, USA, and the Ph.D. degree in electrical engineering and computer sciences from the University of California at Berkeley, Berkeley, CA, USA.

He is currently a Research Scientist with Fitbit, Inc., San Francisco, CA, USA. His current research interests include distributed autonomous robotics, wireless sensor networks, and estimation and control for embedded systems, with particular applications to indoor environmental sensing to reduce the energy use of office buildings and environmental sensing of water networks.

Costas J. Spanos received the master's and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 1981 and 1985, respectively.

After graduation, he worked for three years with the Advanced Computer-Aided Design Group, Digital Equipment Corporation, Maynard, MA, USA. In 1988, he joined the University of California at Berkeley, Berkeley, CA, USA, where he was the Department Chair of Electrical Engineering and Computer Sciences, the Associate Dean for Research, and the CEO of the Berkeley Educational Alliance for Research in Singapore. He is currently the Director of CITRIS and the Banatao Institute, University of California, Berkeley. His current research interests include the application of statistical analysis in the design and fabrication of integrated circuits, and the development and deployment of novel sensors and computer-aided techniques in semiconductor manufacturing.

Prof. Spanos was an elected Fellow of the Institute of Electrical and Electronic Engineers for contributions and leadership in semiconductor manufacturing in 2000 and the Andrew S. Grove Distinguished Professor with the Department of Electrical Engineering and Computer Sciences in 2009.

Alexandre M. Bayen received the Engineering degree from École Polytechnique, Palaiseau, France, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, USA.

He was a Visiting Researcher with the NASA Ames Research Center, Mountain View, CA, USA, from 2000 to 2003, and the Director of the Autonomous Navigation Laboratory with the Laboratoire de Recherches Balistiques et Aerodynamiques, Ministere de la Defense, Vernon, France, where he holds the rank of Major. He is currently a Chancellor Professor of Electrical Engineering and Computer Science and Civil and Environmental Engineering and the Director of the Institute for Transportation Studies with the University of California at Berkeley, Berkeley, CA, USA. He has authored two books and over 150 articles in peer-reviewed journals and conferences.

Dr. Bayen was a recipient of the CAREER Award from the National Science Foundation, the Best of ITS Award for Best Innovative Practice (2008 ITS World Congress), the TRANNY Award (California Transportation Foundation), the PECASE Award (The White House), the Huber Prize (ASCE), the Ruberti Prize (IEEE), and the Okawa Research Grant Award.