2010 American Control Conference
Marriott Waterfront, Baltimore, MD, USA
June 30-July 02, 2010

An Integrated Approach to Occupancy Modeling and Estimation in Commercial Buildings

Chenda Liao and Prabir Barooah

Abstract— **The problem of real-time estimation of occupancy in a commercial building (number of people in various zones at every time instant) is relevant to a number of emerging applications, such as green buildings that achieve high energy efficiency through feedback control. Due to the high deployment cost and large errors that people counting sensors suffer from, measuring occupancy throughout the building accurately from sensors alone is not feasible. Fusing sensor data with model predictions is essential. Due to the highly uncertain nature of occupancy dynamics, modeling and estimation of occupancy is a challenging problem. This paper makes two contributions toward addressing these challenges. We develop an agent-based model to simulate the behavior of all the occupants of a building, and extract reduced-order graphical models from Monte-Carlo simulations of the agent-based model. The agent-based model is validated with sensor data for the special case of one room and one occupant. Noisy measurements from a few sensors are fused with the graphical model predictions using the classical LMV estimator to estimate room-level occupancy in the building. Simulations illustrate the effectiveness of the proposed method.**

I. Introduction

In the United States, buildings are responsible for 38% of CO_2 emissions, 71% of electricity consumption, and 39% of energy use [1]. Improving energy efficiency of buildings is therefore becoming a national priority. Feedback control based on real-time sensing and estimation is expected to play a significant role in achieving the high-degree of energy efficiency. One of the key variables that needs to be monitored to achieve high energy-efficiency is occupancy, which refers to the number of people. The number of people in a zone affects the zone's cooling load and ventilation load, which determines the amount of conditioned air to be delivered to that zone to maintain thermal comfort and air quality. Besides energy efficiency, occupancy information in a large commercial buildings is valuable during evacuation of fire, terrorist attacks, earthquake and other emergencies.

Due to the highly uncertain nature of occupancy and wide fluctuations seen over multiple time-scales, occupancy cannot be predicted ahead of time based on expected building use, and has to be monitored in real-time. Several kinds of sensors currently can provide information on occupancy, such as video cameras equipped with people-counting software [2], [3], optical tripwires and PIR (pyroelectric infrared) motion sensors that count the number of people crossing a particular area, CO_2 sensors that measure the concentration of CO_2, etc. However, sensing alone is not enough for occupancy measurement, because each of these sensors have large uncertainty [4]. Optical tripwires, PIR sensors and video cameras suffer from false counts, while CO_2 sensors have large delay (order of 10-15 minutes [5]), calibration drift and uncertainty in the relationship between the number of people and CO_2 concentration. As a result, simply estimating the total number of people in a building from various sensor data is a challenging problem; sophisticated methods are required to estimate total building occupancy [4], [5]. In addition, deploying enough sensors to cover all areas of the building to measure zone-level occupancy will lead to prohibitively large deployment and maintenance costs and unacceptable level of intrusion into building occupants' privacy. Therefore, real-time occupancy estimation with high spatial resolution requires fusion of sensor data with model predictions. In fact, Meyn et. al. [5] reports that a large improvement in building occupancy results when prior information, such as knowledge of meetings at a specific room, is used, compared to using sensor information alone.

Constructing mathematical models of occupancy dynamics in a building that are appropriate for real time estimation is a challenging problem, because of the high uncertainty of people movement that governs occupancy evolution. The occupancy models suggested in [6] rely on the correlation between occupancy and lighting and equipment load. To estimate zone-level occupancy with this method, sensors at each zone will be required. The paper [7] proposes a probabilistic model to predict and simulate occupancy in a single person office, where vacant and occupied intervals are modeled as exponential distributions. The results of the study were mixed. A stochastic model of occupancy in UK households was proposed in [8] that can be used to generate statistically representative sample paths of household occupancy.

Another line of research on occupancy modeling that has seen a recent spurt in activity is through agent-based models. Agent-based models to mimic individual behavior has been in development over the last 40 years; see the review article [9] and reference therein. Many of these models are used to simulate behavior of individuals in emergency situations such as fire [10], or to study the patterns in pedestrian traffic [11]. However, there has been little work on agent-based modeling for building occupancy in normal, day-to-day, operations. A notable exception is the work by Page et. al. [12]; which proposes an agent-based model for a single person in a single room and validates the prediction of the model from collected sensor data over a long period of time. However, extending the model proposed in [12] to

C. Liao and P. Barooah are with the Dept. of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL. {cdliao, pbarooah}@ufl.edu

the case of a building with many rooms occupied by a large number of individuals is not trivial.

A weakness of agent-based-models is that even if accurate models are constructed, they suffer high-degree of complexity that makes them unsuitable for real-time data fusion. We adopt a two-tiered strategy to address the modeling challenge. (i) First, we develop a novel agent-based stochastic model which can be used to simulate the behavior of an arbitrary number of individual occupants (agents) of a building with arbitrary number of zones. The information needed to construct these models can be obtained from surveys, collected sensor data, or a combination of the two. Simulations of these models yield realizations (time series data) of zone-level occupancy. (ii) Next, we extract so-called graphical models of zone-level occupancy from Monte-Carlo simulations of the agent-based model. Graphical models are widely used in spatial statistics, image analysis and bioinformatics, and are convenient for representing statistical relationships between variables (see [13], [14] and references therein). The usefulness of graphical models for occupancy estimation is that they extract a reduced-order statistical model of room-level occupancy (means and correlations as a function of time) from the stochastic agent-level behavior. A key advantage of a graphical model is that it identifies correlations among room-level occupancy that arise out of complex behavior of individual occupants. This information helps predict occupancy in locations without sensors based on measured sensor data in other locations. Finally, the extracted graphical model is used with the classical LMV (Linear Minimum Variance) estimator [15] to compute the number of people in all the zones of the building from the noisy measurements of occupancy from a few sensors. We assume that each sensor measures a noise-corrupted value of the occupancy in a zone in which it is located.

The primary contribution of this paper is a novel agent-based model for occupancy simulation in multi-zone multi-occupant buildings and its preliminary validation from experimental data for the special case of single-room single-occupant. The experimental data is borrowed from [12]. We calibrate the agent-based model for a building in the University of Florida campus based on preliminary survey data, extract graphical models and perform occupancy estimates. We present performance evaluation of estimation method through simulations, which show promising results.

The paper is organized as follows. Section II describes the model of agent-based model of building occupancy that we have developed, and preliminary model validation from sensor data. Section III describes identification of covariance graph models from agent-based models. Section IV describes estimation of occupancy using sensor data and covariance graph models. Simulation studies to evaluate performance of the proposed method is described in Section V. The paper concludes with a discussion on future work in Section VI.

II. AGENT-BASED MODEL OF INDIVIDUAL OCCUPANTS

Although agent-based models have a long history of application in many disciplines, to the best of our knowledge,

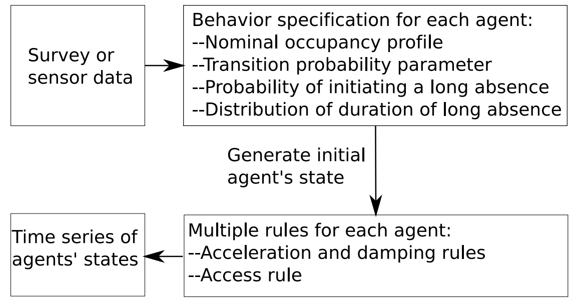


Fig. 1. The process of constructing the agent-based MARM model and generating occupancy time-series using it.

the paper by Page *et. al.* [12] is the only one that focuses on developing agent-based models for simulating the occupancy profile in a building during normal operation over long time intervals. We will refer to the model in [12] as the "Page model" in the sequel. The Page model considers a single person who nominally occupies a single room. The motion between the room and "outside" is modeled as a time-inhomogeneous Markov chain. Extending the model to multiple rooms/zones that are peopled by a large number of individuals is not trivial. The reason is that since accurate motion model of an individual necessarily has to be time-inhomogeneous, it is cumbersome to specify the time-varying Markov transition probability matrices for each individual. Therefore in this paper we propose an alternate agent-based model that is easily scalable to large number of zones and individuals. The model is inspired by that in [12], and shares several key attributes.

Consider a building with n zones that is nominally occupied by m individuals. The model is described in terms of $n+1$ *nodes* that are indexed as $j = 1, \dots, n, n+1$ ($n+1$-th node refers to the "outside of the building"). The individuals (*agents*) are indexed as $i = 1, \dots, m$. Time is measured by a discrete time index k. In this paper, each increment in the time counter corresponds to a 15-minute interval. We model the motion of agents for one week, assuming that the statistics of agents' behavior does not change among weeks, so that $k = 1, \dots, 672$. The model consists of deciding on the *state* of each agent at each time instant, where state refers to the node that the agent occupies. The model, named *Mixed Agent-based Rules Model* (MARM), consists of a number of modules for each agent i, the decision on the state of an agent being modified by each successive module, the decision by the last module being the final state. We now describe these modules: (see Figure 1)

A. Description of the MARM model

- Denote by $P_{i,j}(k)$ the probability of agent i occupying node j at time k. A *nominal occupancy profile*, $\{P_i(k), k = 1, \dots, K\}$ is first specified for every agent i, where $P_i(k) = [P_{i,1}(k), \dots, P_{i,n+1}(k)]^T$, and $K = 672$ for the current investigation. During simulation, each occupant's state is initialized according to its nominal occupancy profile with the help of a pseudo-random number generator. The nominal behavior could be obtained by conducting a survey, or collected by occupancy sensors for a long term as done in [12].

- Occupants in the hallway tend to leave the hallway quickly, while occupants of the other zones tend to stay in where they are. An acceleration rule and a damping rule are used to mimic this behavior, which is based on a pre-specified *transition probability parameter* p_h and p_r. The implementation of acceleration rule is as follows. Suppose the k-th state of an agent is "hallway" and his/her tentative $(k+1)$-th state is also "hallway". With probability p_h, the tentative $(k+1)$-th state will be reassigned by using nominal occupancy profile in order to force that agent to transit out of the "hallway" state as soon as possible. The damping rule is similar: if the tentative $(k+1)$-th state is different from the k state (except hallway), with probability p_r, the $(k+1)$-th state will be reassigned back to the value of the k-th state.
- To simulate long absence, we borrow the method used in [12]. Long absence corresponds to an absence of more than one days due to vacation, sickness, business trips and conference, but not weekends. The probability of initiating a long absence and the distribution of duration of long absence for each agent is specified a-priori, which are obtained from survey or sensor data. During simulation, a long absence for an agent is randomly initiated at time k according to the specified probability using a random number generator, while the distribution of duration of long absence decides the duration of the long absence once it is initiated.
- Finally, each agent has associated access profile that specifies which rooms he/she has access to. The access rule is used to ensure that agents do not occupy nodes without access, which is the last step in agents' state generation. Some nodes that have a maximum occupancy like classrooms and restrooms also fit this rule.

B. Preliminary model validation with experimental data

We perform a preliminary validation of the proposed model by simulating the special case of 1 room and 1 occupant ($m = n = 1$), and comparing the statistics predicted by MARM model with processed occupancy sensor data, as well as that by Page model. The processed sensor data consists of binary state information (occupied/unoccupied) collected at 15 minute intervals, which has been provided to us by the authors of [12]. The raw sensor data was collected using a motion sensor for years in a room that was used by only one person. We estimate a nominal occupancy profile as well as long absence probabilities from the collected sensor data, as done in [12]. Due to lack of space, the interested reader is referred to [12] for details on the data collection and processing. These profiles are used as inputs to both the proposed MARM model and Page model. Figure 2 shows the nominal occupancy profile extracted from the collected data and used in the simulations for model validation. Monte-Carlo simulations of our model and Page model are conducted, and the resulting time series are used to estimate the CDFs (cumulative distribution functions) of the following random

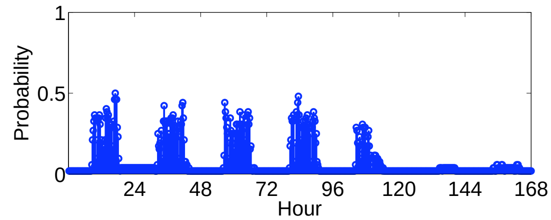


Fig. 2. Nominal occupancy profile of the single occupant in the one-room building extracted from sensor data reported in [12].

variables: first arrival time, last departure time, total duration of daily presence, length of continuous presence, number of daily changes and probability of presence. The statistics of these variables are also estimated from the repeated segments of one-week-long processed sensor data.

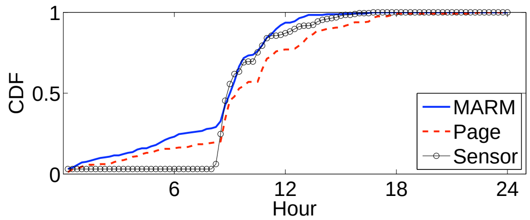
Figure 3(a) compares the CDF of the first arrival time predicted by the proposed model, and Page model and sensor data. The figure shows that the proposed model predicts the late arrivals better than it does the early ones, whereas the Page model predicts the early first arrivals more accurately than late first arrivals. Figure 3(b) shows the CDF of the total duration of daily presence, which refers to the total number of hours the occupant stays in the room in a 24-hour period. Both MARM and Page models essentially track the monitored data with equal amount of error. Figure 3(c) shows the CDF of the length of continuous presence in a day. The figure shows that MARM model predicts the length of continuous presence better than the Page model does. The interested reader is referred to [16] for comparison of the remaining variables, which are omitted here due to space limitation.

Based on this comparison, we conclude that occupancy statistics predicted by the MARM model sufficiently matches that estimated from sensor data, and the proposed MARM model has the same level of accuracy as that of the Page model. Efforts to gather data to validate the MARM model in the multi-zone multi-person case is ongoing.

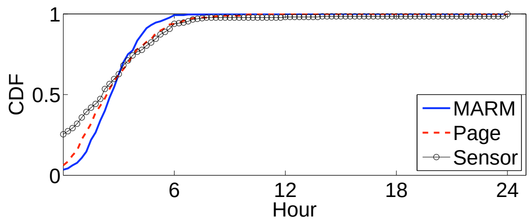
III. Covariance Graph Model Identification

The agent-based model described in the previous section is useful to simulate room-level occupancy profiles in buildings, and it predicts the statistics of occupancy evolution. However, it is too complex for real-time estimation with sensor data. A more compact representation of the occupancy dynamics is required for real time estimation. Our choice of such a compact modeling paradigm is *Covariance Graph Model* [17]. These graphical models are attractive for real-time estimation since they can compactly represent marginal dependencies among the occupancy of various nodes. In addition, they can be identified from time series data collected from sensors or agent-based model. Graphical models have been widely used in spatial statistics, image analysis and bioinformatics; see [13], [14] and references therein.

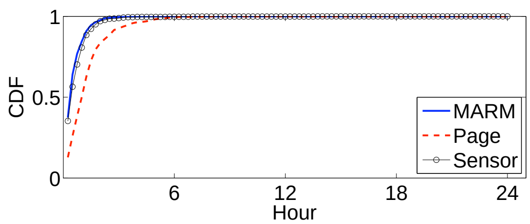
A covariance graph model of an n-variate distribution is specified in terms of the mean μ and covariance matrix $\Sigma = \{\sigma_{ij}\}$. It is called a graphical model since the structure of the covariance matrix defines a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where



(a) CDF of first arrival time



(b) CDF of total daily presence



(c) CDF of length of continuous presence

Fig. 3. Comparison the statistics predicted by the MARM model, Page model and sensor data.

$\mathcal{V} = \{1, \ldots, n\}$ is the node set and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the edge set with the property that $(i,j) \notin \mathcal{E} \Rightarrow \sigma_{i,j} = 0$ [17]. For occupancy modeling, the random vector whose distribution we are interested is the occupancy vector $X = [X_1, \ldots, X_n]^T$, where X_i is the occupancy (number of people) of node i, $i = 1, \ldots, n$. Note that the "outside" node is not part of the graph. Since occupancy statistics vary with time, the model varies with time as well, so that we have a sequence of covariance graph models $(\mu(k), \Sigma(k))$, $k = 1, \ldots, K$.

To describe model identification, we first consider the case when the model does not change with time. The identification of a covariance graph model (μ, Σ) from samples of the random vector X consists of two steps: (i) model selection and (ii) parameter estimation. Model selection refers to choosing the structure of the graph $\mathcal{G}$ (or equivalently, the sparsity pattern of Σ), while parameter estimation refers to choosing the values of those entries of Σ that have been decided to be non-zero in the model selection step. The goal of this two-step identification is to estimate the possibly sparsest graph structure that can still explain the first and second order statistics of the data. We follow the methods proposed in [17], [18] to carry out the model selection and parameter estimation steps. The first step is the computation of the sample covariance matrix from N samples of data. Assuming we conduct N Monte-Carlo experiments, for every time k, we compute the sample mean

$$\bar{X}(k) = \textstyle\sum_{j=1}^{N} X^{(j)}(k), \qquad (1)$$

where $X^{(j)}(k)$ is $X(k)$ observed in the j-th simulation. Similarly, we compute the sample covariance of the state at time k as

$$W(k) = \frac{1}{N}\sum_{j=1}^{N}(X^{(j)}(k) - \bar{X}(k))(X^{(j)}(k) - \bar{X}(k))^T. \quad (2)$$

Model selection is based on hypotheses testing on all edges (i.e., all entries of $W(k)$) at an overall confidence level determined by a designed parameter α [17]. Assuming that the data comes from the true graph model $\mathcal{G}$, this method leads to an estimated graph model $\hat{\mathcal{G}}_\alpha$ with the following confidence level $\lim_{n\to\infty} \inf\ P(\hat{\mathcal{G}}_\alpha = \mathcal{G}) \geq 1 - \alpha$. This means, for fixed α, the correct model is selected with probability at least $1 - \alpha$ for large sample size. Once the structure of the graph model is chosen based on model selection, an iterative conditional fitting algorithm based on maximum likelihood estimation is used for estimating the values of the non-zero entries of Σ. The interested reader is referred to [18] for the details. Although rigorous results on identification of graphical models require the assumption that the underlying distribution is multi-variate Gaussian, applications of these models to non-Gaussian data is common [19].

Using multiple time-series of occupancy in the n-room building obtained from Monte Carlo simulations of the agent-based model, we estimate 672 covariance graph models $(\mu(k), \Sigma(k))$, $k = 1, \ldots, 672$, one for each 15-minute interval. There is one difficulty in applying methods for graphical model identification to this application in a straightforward manner. The sample covariance matrix W in (2) is required to be non-singular for these methods to be applicable [17], [18]. At certain time instants, especially at night, the probabilities of certain rooms being occupied are very small. In this case, we may get zero rows and columns in the sample covariance matrix due to finite number of Monte-Carlo simulations. In order to use proposed model identification method, we eliminate these rows and columns and construct reduced sample covariance matrix $W(k)_r$. The reduced covariance matrix $\Sigma(k)_r$ can be obtained by using the techniques described above, and then $\Sigma(k)$ can be regained by plugging in the zero rows and columns back to $\Sigma(k)_r$.

IV. Occupancy Estimation

In this paper we consider sensors that can measure a function of the occupancy at a location directly. The total number of sensors in the building is denoted by N_s; and $s_i \in \mathcal{V}$ is used to denote the node of the graph where the i-th sensor is located (excluding node $n+1$, the "outside"). Note that this means s_i can be any zone of the building except the "outside". We assume a measurement of the form $y_{s_i}(k) = \eta(x_{s_i}(k) + \epsilon_i(k))$ is available from the s_i-th sensor, where ϵ is a measurement noise and $\eta(\cdot)$ is a (typically) nonlinear function of its argument. For example, video cameras with people counting software fall into this category. The measurement vector at time k is denoted by $Y(k) = [y_1(k), \ldots, y_{N_s}(k)]^T$. The estimation problem is to estimate the state vector $X(k)$ from $Y(k)$, using the graphical model information $(\mu_x(k), \Sigma_x(k))$, for $k = 1, \ldots, K$.

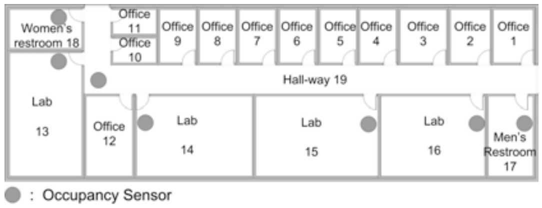


Fig. 4. The floor plan of the 3rd floor of MAE-B in the University of Florida campus, in which estimation was carried out.

The LMV (Linear Minimum Variance) estimator of a random vector X in terms of another Y is given by [15]

$$\hat{X}(k) = \mu_X + \Sigma_{xy}\Sigma_{yy}^{-1}(Y(k) - \mu_Y), \tag{3}$$

where $\Sigma_{xy} = cov(X, Y)$ and $\Sigma_{yy} = cov(Y, Y)$. In the linear sensing model ($\eta(x) = x$), $Y(k) = CX(k) + \epsilon(k)$, where C is a $N_s \times n$ matrix of 0's and 1's, and $\epsilon(k) = [\epsilon_1(k), \ldots, \epsilon_{N_s}(k)]^T$. Assuming $\epsilon(k)$ is zero-mean and $\mathrm{E}[\epsilon(k)\epsilon(\kappa)^T] = R\delta(k - \kappa)$, the LMV estimate in (3) reduces to

$$\hat{X}(k) = \bar{X}(k) + \Sigma(k)C^T(C\Sigma(k)C^T + R(k))^{-1} \times (Y(k) - C\bar{X}(k)), \tag{4}$$

where the invertibility of $(C\Sigma C^T + R)$ is assured by the positive-definiteness of R and positive semi-definiteness of $C\Sigma C^T$. When a sensor measures the occupancy of a room with an additive white noise, the estimate of occupanices in all the zones can be computed by using the classical LMV estimator as described above. The mean and covariance information needed to compute the estimates are provided by the covariance graph model. When the sensing model is nonlinear, the LMV estimate does not have a simple expression as it does in the linear case.

V. Performance evaluation via simulation

In this section we describe the results of performance evaluation of the proposed methodology carried out for a building shown in Figure 4. This is the simplified floor plan of the 3rd floor of the MAE-B building in the University of Florida campus. The floor has 19 nodes (12 professor's offices, 4 labs, 2 restrooms and 1 hallway). Forty-five people work in this floor in a typical work day.

A. Modeling

An informal survey was conducted to obtain information on schedules of occupants (such as first arrival and last departure time, frequency and during of absence, etc.) as well as information on who occupies which room, etc. Based on this information, we generated a nominal occupancy profile, the probability of initiating a long absence, and the distribution of long absence, for each occupant. The transition probability parameters were fixed at 0.5, which was the value used in validation of a single room building model. With these inputs, Monte-Carlo simulations of occupancy are carried out which yields stochastic occupancy histories for each room over a one-week period.

A total of 672 covariance graph models $(\mu(k), \Sigma(k))$, $k = 1, \ldots, 672$ (one week at 15-minute intervals) are identified from the time series data obtained from 1000 Monte-Carlo simulations of the agent-based model, using the method described in Section III. The value of α for hypothesis testing was chosen as 0.1. The model selection step in model identification reduces model complexity by almost an order of magnitude. If the covariance Σ were simply taken as the sample covariance, skipping the model selection part, the number of edges in the graph would have been between 153 and 171. In contrast, the number of edges in the identified model was between 0 and 18. More details are available in [16].

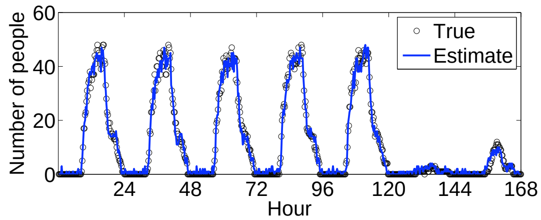
B. Estimation

Estimation was performed from simulated sensor data that was generated with the help of the agent-based model. We used a total of 7 sensors, whose locations are shown in Figure 4. In the simulations reported here, we obtain a single time trace from the agent-based model for one week, and then add random noise to the profile so generated, which is taken as the "true" occupancy. The purpose of adding noise is to simulate people who occupy the building for short periods and who are not taken into account in the agent-based model, e.g., visitors to the nominal occupants, repairmen and janitorial staff, etc. The variance of the random disturbance added was 0.16. Noisy sensor measurements are generated using a nonlinear sensor model from the true occupancy data:

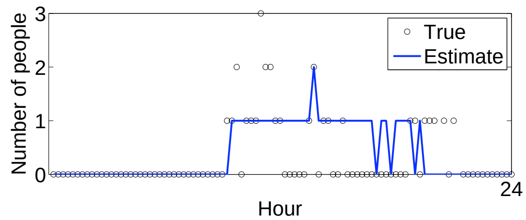
$$Y_i(k) = round(x_{s_i}(k) + \epsilon_i(k)), \tag{5}$$

where $\epsilon_i(k)$ is a zero mean Gaussian noise with standard deviation 0.5. Estimate of the occupancy state $\hat{X}(k)$ is computed at every time k using the LMV estimator (4). Note that the data is generated using a nonlinear sensor model though the estimator is constructed under the assumption of linearity of the sensing model. Thus, the computed estimates are no longer linear optimal estimates of the states.

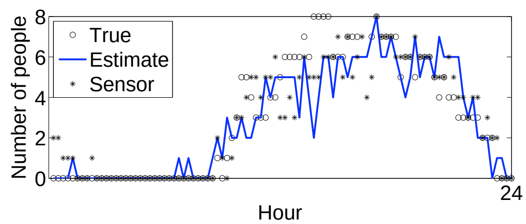
Figure 5(a) shows the estimate of total occupancy in the building and the true total occupancy. Note that the true total occupancy cannot be inferred from the sensors alone since the sensors monitor less than half of the rooms. The figure shows that the total occupancy estimate matches quite well with the true value, with a mean error of 0.1 and a standard deviation of 2.1 (computed from a single one-week time-series). Figure 5(b) shows the estimated and true occupancy of room 4 (which is an office with a nominal occupancy of 1 without sensor) for a 24-hour interval (Monday). The figure shows that even without a sensor, the error between the estimated occupancy and the true value is small, with a mean of 0.05 and a standard deviation of 0.59 (computed from a single one-week time-series). Figure 5(c) shows the estimated occupancy, the sensor measurements and the true occupancy in room 16 for a 24-hour interval (Monday). This room is a lab with a nominal occupancy of 7 that has a sensor installed in it. The mean and standard deviation of estimation error are 0.03 and 1.1, which are computed in the same fashion. The estimator provides estimates that, on an average, are only slightly more accurate than that by the sensor.



(a) Estimated and true occupancy of the whole floor with extra uncertainty.



(b) Estimated and true occupancy of room 4 (no sensor) with extra uncertainty.



(c) Estimated and true occupancy of room 16 with extra uncertainty.

Fig. 5. Estimation result with extra uncertainty

VI. Summary and Future Work

In this paper we presented an integrated approach to modeling and real-time estimation of occupancy in a large commercial building. The modeling portion consists of two stages, agent-based model to simulate behaviors of individual occupants and covariance graph models to extract relevant statistical information to aid in real time estimation. Limited and noisy sensor data from occupancy sensors are then fused with the predictions of the graphical model using the classical LMV estimator to estimate the number of people in every zone of the building, even in those without sensors. The agent-based model is validated for a special scenario of one room and one occupant using collected sensor data that was reported in [12]. Simulations suggest that the modeling and estimation technique proposed in this paper can produce accurate estimates of building-wise as well as zone-wise occupancy.

The present work is a first step in addressing a very challenging problem of modeling and estimating occupancy in commercial buildings. There are numerous avenues of future work. Validation of the agent-based model for the multi-room and multi-occupant case is currently under way. The next step would be to test experimentally the accuracy of the occupancy estimates in a real building. Apart from experimental validation, new approaches are needed in model identification and estimation techniques. Instead of identifying distinct graphical models at every time instant, a more appropriate approach would utilize the slow rate of change of occupancy dynamics to identify a sequence of graphical models with minimum number of changes of the model from one time step to the next. Once such a framework is in place, estimation can be posed in a filtering framework, which is more natural for the targeted application than performing a sequence of static estimates.

Acknowledgment

The authors wish to thank Prof. Robinson and F. Haldi for providing us with the data reported in [12].

References

[1] U. S. D. of Energy (DOE), "The 2008 building energy databook," Washington, D.C., 2008.

[2] T.-H. Chen, T.-Y. Chen, and Z.-X. Chen, "An intelligent people-flow counting method for passing through a gate," in *IEEE Conference on Robotics, Automation and Mechatronics*, June 2006, pp. 1–6.

[3] K. Terada, D. Yoshida, S. Oe, and J. Yamaguchi, "A method of counting the passing people by using the stereo images," in *International conference on image processing*, 1999.

[4] J. Hutchins, A. Ihler, and P. Smyth, "Modeling count data from multiple sensors: a building occupancy model," in *IEEE International Workshop on Computational Advances in MultiSensor Adaptive Processing*, December 2007.

[5] S. Meyn, A. Surana, Y. Lin, S. M. Oggianu, S. Narayanan, and T. A. Frewen, "A sensor-utility-network method for estimation of occupancy distribution in buildings," in *48th IEEE Conference on Decision and Control*, December 16-18 2009.

[6] B. Abushakra and D. E. Claridge, "Modeling office building occupancy in hourly data-driven and detailed energy simulation programs," in *ASHRAE Transactions*, July 2008.

[7] D. Wang, C. C. Federspiel, and F. Rubinstein, "Modeling occupancy in single person offices," *Energy and buildings*, vol. 37, no. 2, pp. 121–126, 2005.

[8] I. Richardson, M. Thomson, and D. Infield, "A high-resolution domestic building occupancy model for energy demand simulations," *Energy and Buildings*, vol. 40, no. 8, pp. 1560 – 1566, 2008.

[9] D. Helbing, "Traffic and related self-driven many-particle systems," *Review of Modern Physics*, vol. 73, pp. 1067–1141, 2001.

[10] S. M. V. Gwynne and E. D. Kuligowski, "Simulating a building as a people movement system," *Journal of Fire Sciences*, 2009.

[11] D. Helbing, "Self-organisation phenomena in pedestrian crowds," pp. 569–577, 1997.

[12] J. Page, D. Robinson, N. Morel, and J.-L. Scartezzini, "A generalised stochastic model for the simulation of occupant presence," *Energy and Buildings*, vol. 40, pp. 83 – 98, 2008.

[13] S. L. Lauritzen, *Graphical Models*. Oxford: Clarendon Press, 1996.

[14] H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applications*, ser. Monographs on Statistics and Applied Probability. London: Chapman and Hall, 2005, vol. 104.

[15] I. Rhodes, "A tutorial introduction to estimation and filtering," *IEEE Transactions on Automatic Control*, vol. 16, pp. 688– 706, December 1971.

[16] C. Liao and P. Barooah, *An Integrated Approach to Occupancy Modeling and Estimation in Commercial Buildings(Technical Report)*, Feb 2010. [Online]. Available: http://plaza.ufl.edu/cdliao/

[17] M. Drton and M. Perlman, "Model selection for gaussian concentration graphs," *Biometrika 91*, pp. 591–602, 2004.

[18] S. Chaudhuri, M. Drton, and T. S. Richardson, "Estimation of a covariance matrix with zeros," *Biometrika*, pp. 199–216, 2007.

[19] C. Borgelt and R. Kruse, *Graphical Models: methods for data analysis and mining*. Wiley, 2002.