

## Research Article

# A Comfort-Aware Energy Efficient HVAC System Based on the Subspace Identification Method

O. Tsakiridis,<sup>1</sup> D. Sklavounos,<sup>2</sup> E. Zervas,<sup>1</sup> and J. Stonham<sup>2</sup>

<sup>1</sup>Department of Electronics, TEI of Athens, Egaleo, 12210 Athens, Greece

<sup>2</sup>Department of Engineering and Design, Brunel University, Kingston Lane, Middlesex UB8 3PH, UK

Correspondence should be addressed to O. Tsakiridis; [odytsak@teiath.gr](mailto:odytsak@teiath.gr)

Received 12 October 2015; Revised 12 January 2016; Accepted 8 February 2016

Academic Editor: Aleksander Zidansek

Copyright © 2016 O. Tsakiridis et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A proactive heating method is presented aiming at reducing the energy consumption in a HVAC system while maintaining the thermal comfort of the occupants. The proposed technique fuses time predictions for the zones' temperatures, based on a deterministic subspace identification method, and zones' occupancy predictions, based on a mobility model, in a decision scheme that is capable of regulating the balance between the total energy consumed and the total discomfort cost. Simulation results for various occupation-mobility models demonstrate the efficiency of the proposed technique.

## 1. Introduction

As the control and limitation of the energy consumption remain a field with an exceptional technological and economical interest, areas that have been considered as high energy consumers comprise very challenging research issues. According to the US Energy Information Administration from 2013 through 2040 the electricity consumption in the commercial and residential sectors will be increasing by 0.5% and 0.8% per year [1]. It is well established and widely accepted through research studies that the main energy consumers in the commercial and residential buildings are the Heat Ventilation and Air Conditioning (HVAC) systems, as well as the lighting systems.

Buildings contributed a 41% (or 40 quadrillion btu) to the total US energy consumption in 2014. On an average, about 43% of the energy consumption in a commercial and residential building is due to HVAC systems [2]. Therefore, due to the high rate of energy consumption, the necessity of the demand-driven control in the HVAC systems has become inevitable. In modern buildings several sophisticated systems have been applied aiming at providing this type of control in the HVAC systems. The state-of-the-art technology of the HVAC control considers the occupancy of the zones a very

important parameter, playing a key role in methods aiming at reducing energy consumption. The detection and prediction of the zones' occupancy are a very challenging research field and several techniques have been proposed based on historic statistical data as well as on probabilistic models.

Another equally important factor taken into account in advanced HVAC control systems is the thermal comfort of the occupants. The objective of modern HVAC control systems is to reduce energy consumption without compromising the comfort of the occupants. Wireless sensor networks, equipped with temperature, humidity, and occupancy detection sensor nodes, are nowadays the basic platform to build automated HVAC control systems. A number of methods aiming to maintain the thermal comfort while saving energy have been proposed. Some of them are described in Section 2. Towards this direction, a novel technique is proposed in this paper, which aims at balancing the comfort and energy costs in a multizone system. The decisions on heating the zones or not may be taken either centrally or in a distributed manner by wireless sensor nodes scattered in the multizone system. In any case temperature and zone occupancy information must be exchanged between a node, responsible for a zone, and its neighboring nodes. The decision process itself relies on two kinds of predictions: (a) temperature-time predictions for

the zones and (b) the zones' occupancy profile. The emphasis in this paper is on the zones' temperature predictions and to this end a deterministic subspace identification method is used for modelling the thermal dynamics of each zone. That is, each zone is modelled by a simple state-space model capable of producing accurate predictions based on the surrounding temperatures, the heating power of the zone, and the current state that summarizes the temperature history of the zone. For the zones' occupancy predictions we consider a semi-Markov model, where occupants (moving as a swarm) stay in a zone for a random period of time and then move to adjacent zones with given probabilities. What is needed by the decision process is the distribution of the first entrance time to unoccupied zones. Aiming at a proactive action, the proposed method periodically computes the risk of activating the heater or not and decides in favor of the action that produces the smaller risk. The computation of the risks relies on the relative weights of the energy and discomfort costs so that the balance between the total energy consumed and the total discomfort cost may be regulated.

The structure of the paper is as follows: In Section 2, prior and related work of the field of the demand-driven HVAC systems that utilize occupancy and prediction methods is presented. In Section 3, the proposed comfort-aware energy efficient mechanism is described, along with the subspace identification method, used for obtaining temperature-time predictions, and a discussion on the distribution of the first entrance time to unoccupied zones. Section 4 contains simulation results and assesses the effectiveness of the proposed algorithm. Finally, Section 5 summarizes the paper and proposes research directions for future work.

## 2. Prior Work

The necessity of demand-driven HVAC systems, for energy efficient solutions, has orientated the researchers towards occupancy-based activated systems. In multizone spaces the reactive and proactive activation of the heating/cooling of the zones contributes to significant energy savings and improves the thermal comfort for the occupants. Several research works with valuable results of energy savings have utilized the occupancy detection and prediction, in order to control the HVAC system appropriately. For the occupancy detection several types of sensors are used ( $\text{CO}_2$ , motion, etc.), while, for the prediction, combinational systems are usually applied, utilizing mathematical predictive models (e.g., Markov chains) alongside with detected real data.

The authors in [3] proposed an automatic thermostat control system which is based on an occupancy prediction scheme that predicts the destination and arrival time of the occupants in the air-conditioned areas, in order to provide a comfortable environment. For the occupants' mobility prediction the cell tower information of the mobile telephony system is utilized and the arrival time prediction is based on historical patterns and route classification. For the destination prediction of locations very close to each other (intracell), the time-aided order-Markov predictor is used. The authors in [4] proposed a closed-loop system for optimally controlling HVAC systems in buildings, based on actual

occupancy levels named the Power-Efficient Occupancy-Based Energy Management (POEM) System. In order to accurately detect occupants' transition, they deployed a wireless network comprising two parts: the occupancy estimation system (OPTNet) consisting of 22 camera nodes and a passive infrared (PIR) sensors system (BONet). By fusing the sensing data from the WSN (OPTNet and BONet) with the output of an occupancy transition model in a particle filter, more accurate estimation of the current occupancy in each room is achieved. Then, according to the current occupancy in each room and the predicted one from the transition model, a control schedule of the HVAC system takes over the preheat of the areas to the target temperature. In [5] the authors used real world data gathered from a wireless network of 16 smart cameras called Smart Camera Occupancy Position Estimation System (SCOPES) and in this way, they developed occupancy models. Three types of Markov chain (MC) occupancy models were tested and these are the single MC, the closest distance, and finally the blended (BMC). The authors concluded that BMC is the most efficient and thus they embodied it as the occupancy prediction method in their proposed "OBSERVE" algorithm, which is a temperature control strategy for HVAC systems.

The authors in [6] have developed an integrated system called SENTINEL that is a control system for HVAC systems utilizing occupancy information. For the occupancy detection and localisation, the system utilizes the existing Wi-Fi network and the clients' smart phones. The occupancy localization mechanism is based on the access point (AP) communication with a client's smart phone, so that if an occupant's phone sends packets to an AP then he is located within the range of the AP. By classifying the building areas into two main categories, namely, the personal and the shared spaces, the proposed mechanism activates the HVAC system when an owner of a personal space (e.g., office) has been detected within an area where his/her office is located or when occupants are detected within shared spaces. In [7] an integrated heating and cooling control system of a building is presented aiming to reduce energy consumption. The occupancy behavior prediction as well as the weather forecast, as inputs to a virtual (software based) building model, determines the control of the HVAC system. The occupancy detection technique utilizes Gaussian Mixture Models (GMM) for the categorization of selected features, yielding the highest information gain according to the different number of occupants. This categorization was used for observation to a Hidden Markov Model for the estimation of the number of occupants. A semi-Markov model was developed based on patterns comprised by sensory data of  $\text{CO}_2$ , acoustics, motion, and light changes, to estimate the duration of occupants in the space. The work in [8] proposes a model predictive control (MPC) technique aiming to reduce energy consumption in a HVAC system while maintaining comfortable environment for the occupants. The occupancy predictive model is based on the two-state Markov chain, with the states modelling the occupied and unoccupied condition of the areas. The authors in [9] propose a feedback control algorithm for a variable air volume (VAV) HVAC system for full actuated (zones consisting of one room) and underactuated (zones consisting

of more than one room) zones. The proposed algorithm is called MOBSua (Measured Occupancy-Based Setback for underactuated zones) and it utilizes real-time occupancy data, through a WSN, for optimum energy efficiency and thermal comfort of the occupants. Moreover, the algorithm can be applied on conventional control systems with no need of occupancy information and it is scalable to arbitrary sized buildings.

### 3. Comfort-Aware Energy Efficient Mechanism

We consider a multizone HVAC system consisting of  $Z$  zones  $z_i$ ,  $i = 1, 2, \dots, Z$ . Each zone is equipped with a wireless sensor node capable of conveying the zone's temperature and occupancy information to its neighbors or to a central processing unit. The occupancy information may be very simple, that is, binary information indicating the presence or absence of individuals in the zone, or more advanced like the number of occupants in the zone. The objective is to minimize the total energy and discomfort cost defined as

$$\begin{aligned} \mathcal{E} &= \mathcal{E}_{\text{energy}} + \mathcal{E}_{\text{discomfort}} \\ &= \sum_{i=1}^Z \int_{-\infty}^{\infty} W_i I(H_i = \text{on}) dt \\ &\quad + \sum_{i=1}^Z \int_{-\infty}^{\infty} C_i I(T_i < T_c) dt. \end{aligned} \quad (1)$$

Note that only the heating problem is addressed in this paper. Thus, the case of spending energy for cooling the zones and keeping the temperature  $T_i$  within a certain comfort zone is not treated for simplicity. In (1)  $I(A)$  denotes the indicator function which takes the value 1 or 0 depending on whether condition  $A$  is true or false.  $W_i$  is the power of a heater that covers zone  $z_i$  and  $H_i$  is the state of the heater; that is, the heater is on or off. Thus, the first term of (1) is the total energy consumed by the multizone system. For the discomfort cost we define the cost per unit time  $C_i$  and the target comfort threshold  $T_c$ . As long as the zone's temperature  $T_i$  is higher than  $T_c$  the occupants do not feel discomfort. The parameter  $C_i$  may depend on the zone, the number of the occupants (e.g., the total system's discomfort cost is proportional to the number of people experiencing discomfort), and the difference of the zone's temperature  $T_i$  and the comfort threshold  $T_c$ . It is worth noting that there is a tradeoff between energy savings and discomfort cost. We may preheat all zones to the comfort level thus rendering the discomfort cost equal to zero. This policy is inefficient due to the large amount of energy consumed. In the other extreme we could act reactively by heating a zone only upon detecting occupants in it. In this case the energy cost would be as small as possible but the discomfort cost could increase dramatically. In the sequel we will describe a method that acts proactively by taking periodically decisions on whether to heat a zone or not. If the entrance to a zone is delayed then we postpone heating until the next decision epoch. If, on the contrary, it is anticipated that a zone will be occupied before the zone's temperature reaches the comfort level, then we preheat the zone.

The decision of whether or not the heater of an unoccupied zone  $z_i$  should be turned on depends on (a) the current state of the zone, (b) the relative value of the energy cost ( $W_i$ ) and discomfort cost ( $C_i$ ) per unit time, (c) the temperatures of the surrounding zones, and (d) an estimate of the time that the zone will become occupied. This decision may be taken either centrally from the central processing unit that gathers information by all zones or in a distributed manner by each zone's node after collecting the relative information. There is also the possibility of dividing the decision process functionality between the central unit and the wireless sensor network nodes. For example, the prediction of the first entrance times to unoccupied zones may be taken by the central unit that is aware of the location of people in the multizone system, and then the predicted values may be communicated back to the zones' nodes for the final decision. The specific implementation of the decision process is irrelevant to this paper and it will not be further analyzed.

We assume that time is discretized  $t = mT_s$ ,  $T_s$  is the sampling period, and that each node takes its decisions periodically every  $D$  sampling periods. Note that there is no need for the nodes to be synchronized to each other. Let  $Y_i(t)$  denote the random variable that models the remaining time from the current time instant  $t$  until entrance to the unoccupied zone  $z_i$ . It is obvious that the random variables  $Y_k(t)$ , for the various unoccupied zones  $z_k$ , do not follow the same distribution. We further assume that the node is capable of making predictions for the time it takes to exceed the comfort threshold  $T_c$ . To this end, let  $S_i^{\text{on}}(t)$  denote the time period that is required to exceed  $T_c$  if the heater is turned on immediately. Similarly, we define  $S_i^{\text{off}}(t)$  as the predicted time to reach the comfort threshold if the heater remains off for a period  $D$  and then, at the next decision epoch, it is turned on. The relation of the aforementioned parameters is shown in Figure 1.

Suppose now that at time  $t$  the node of the unoccupied zone  $z_i$  has to take a decision whether to turn the zone's heater on or remain in the off state. At the decision epoch  $t$  the risk to turn the heater on is

$$\begin{aligned} R_i^{\text{on}}(t) &= W_i DP \{S_i^{\text{off}}(t) \leq Y_i(t)\} \\ &\quad + W_i (Y_i(t) - S_i^{\text{on}}(t)) P \{S_i^{\text{on}}(t) < Y_i(t) \leq S_i^{\text{off}}(t)\}. \end{aligned} \quad (2)$$

That is, if  $S_i^{\text{off}}(t) \leq Y_i(t)$ , there is plenty of time to reach the comfort threshold even if we start heating at the next decision epoch and therefore we consume unnecessarily  $W_i D$  units of energy (this is the case depicted in Figure 1). Similarly, if  $S_i^{\text{on}}(t) < Y_i(t) \leq S_i^{\text{off}}(t)$  we waste  $W_i (Y_i(t) - S_i^{\text{on}}(t))$  units of energy. Note that there is no risk of turning on the heater if the entrance to the zone happens sooner than it is predicted. In a similar fashion we calculate the risk of keeping the heater off. In this case

$$\begin{aligned} R_i^{\text{off}}(t) &= C_i DP \{Y_i(t) \leq S_i^{\text{on}}(t)\} \\ &\quad + C_i (S_i^{\text{off}}(t) - Y_i(t)) P \{S_i^{\text{on}}(t) < Y_i(t) \leq S_i^{\text{off}}(t)\}. \end{aligned} \quad (3)$$

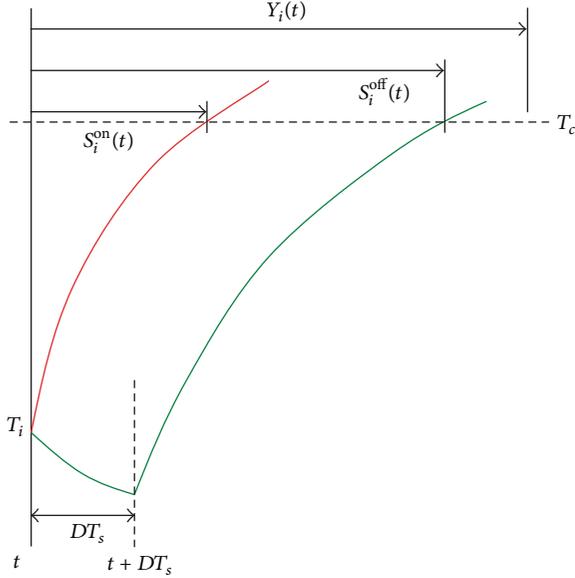


FIGURE 1: Predicted time intervals related to the decision process.

Using (2) and (3) the decision criterion takes the form

$$R_i^{\text{on}}(t) \underset{\text{off}}{\leq} R_i^{\text{off}}(t). \quad (4)$$

Consider for the moment that there is no randomness on  $Y_i(t)$ , that is, strict time scheduled occupation of zones. Typical examples are the classrooms in a university campus. In this case, criterion (4) takes a very simple form since one of the involved probabilities is equal to one and the rest assume the value of zero. Thus, if  $Y_i(t) \leq S_i^{\text{on}}(t)$ , then the heater is turned on, if  $S_i^{\text{off}}(t) < Y_i(t)$ , the heater remains off, and if  $S_i^{\text{on}}(t) < Y_i(t) \leq S_i^{\text{off}}(t)$  the heater will be turned on/off depending on the values of the relative energy and discomfort cost. The balance between these two costs is regulated by the weights  $W_i$  and  $C_i$ .

If  $Y_i(t)$  is random, then criterion (4) needs to be slightly modified as the value of  $Y_i(t)$  is unknown. In this case we replace  $Y_i(t)$  with its conditional expected value given that  $S_i^{\text{on}}(t) < Y_i(t) \leq S_i^{\text{off}}(t)$ . Therefore, in criterion (4), we substitute  $Y_i(t)$  by  $E[Y_i(t) | S_i^{\text{on}}(t) < Y_i(t) \leq S_i^{\text{off}}(t)]$ .

From the previous discussion, it becomes clear that the decision process needs the estimates  $S_i^{\text{on}}(t)$ ,  $S_i^{\text{off}}(t)$  and the distribution of the process  $Y_i(t)$ . The former is the subject of the next subsection, whereas a discussion on the distribution of  $Y_i(t)$  is left for Section 3.2.

**3.1. Temperature-Time Predictions Based on SID.** It is conceivable that in order to make predictions about a zone's temperature evolution we need a model able to capture the thermal dynamics of the tagged zone. To this end we resort to the deterministic subspace identification method. Each zone

is treated separately and is represented by a discrete time, linear, time invariant, state-space model. That is,

$$x_{k+1} = Ax_k + Bu_k + w_k, \quad (5)$$

$$y_k = Cx_k + Du_k + v_k, \quad (6)$$

where  $y_k$  is the zone temperature at time  $k$  and is in general an  $\ell$ -dimensional vector if  $\ell$  temperature sensors scattered in the zone have been deployed. Assuming a uniform zone temperature we use only one sensor node and thus  $y_k$  is scalar.  $v_k$  represents the measurement noise due to imperfections of the sensor.  $u_k$  is a vector of input measurements and its dimensionality varies from zone to zone. This vector contains the measurable output of sources that influence zone's temperature. Such sources are the power of a heater located in the zone, the air temperature at various positions outside the zone, and so forth. The vector  $x_k$  is the state vector of the process at the discrete time  $k$ . The states have a conceptual relevance only and they are not assigned physical interpretation. Of course a similarity transform can convert the states to physical meaningful ones. The unmeasurable vector signal  $w_k$  represents noise due to the presence of humans in the zone, lamps' switching on and off, and so forth. For the rest of the paper we neglect the effects of state and output noise and thus we deal with a deterministic identification problem, which is the computation of the matrices  $A$ ,  $B$ ,  $C$ , and  $D$  from the given input-output data. The no noise assumption makes sense since we need temperature-time predictions for the empty zones only which are "free" from occupants and other disturbances.

In order to estimate the state-space matrices  $A$ ,  $B$ ,  $C$ , and  $D$  of each zone, we need a training phase during which all the nodes in the structure report their relative measurements to a central computing system. These measurements consist of the temperature of the zone and the current power of a heater, if the sensor node is in charge of a heated zone, or only the temperature if the sensor node monitors an area immaterial to the decision process but relevant to other zones, that is, the exterior of a building. The training phase should be performed only once but under controlled conditions, for example, no extra sources of heating and closed windows. Moreover, the length of this period should be quite large in order to exhibit several variations in the input signals to excite the modes of the system. In the simulation part we considered a training period of 24 h, that is, 86400 samples with sampling period 1 sec. The sampling period of 1 sec is too small for all practical reasons. However, as the dominant factor of performance degradation is the estimation of the residual arrival times we consider an ideal temperature prediction model (high accuracy sensors and temperature noiseless zones) and we access only the uncertainty of the occupation model.

After receiving all the measurements, the central system organizes them to input-output data for each zone. For example, if a zone is named **A**, neighbors zones are named **B**, **C**, and **D** and the exterior to the building zone is named **E**, then the temperature of zone **A** is the output of the system to be identified, whereas the temperatures of zones **B**, **C**, **D**, and **E** as well as the power profile of the heater of zone **A** are the input data for the system. A deterministic subspace

identification process, described in the following paragraphs, is then run for each zone. The relevant matrices  $A$ ,  $B$ ,  $C$ , and  $D$  of the state-space model of each zone are communicated back to the sensor node responsible for the monitoring of the zone.

Following the notation and the derivation in [10] we define the block Hankel matrices

$$U_p = U_{0|i-1} = \begin{pmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_j \\ \vdots & \vdots & \vdots & \vdots \\ u_{i-1} & u_i & \cdots & u_{i+j-2} \end{pmatrix}, \quad (7)$$

$$U_f = U_{i|2i-1} = \begin{pmatrix} u_i & u_{i+1} & \cdots & u_{i+j-1} \\ u_{i+1} & u_{i+2} & \cdots & u_{i+j} \\ \vdots & \vdots & \vdots & \vdots \\ u_{2i-1} & u_{2i} & \cdots & u_{2i+j-2} \end{pmatrix}$$

which represent the “past” and the “future” input with respect to the present time instant  $i$ . Stacking  $U_p$  on top of  $U_f$  we have the block Hankel matrix

$$U_{0|2i-1} = \begin{pmatrix} U_{0|i-1} \\ U_{i|2i-1} \end{pmatrix} = \begin{pmatrix} U_p \\ U_f \end{pmatrix} \quad (8)$$

which can also be partitioned as

$$U_{0|2i-1} = \begin{pmatrix} U_{0|i} \\ U_{i+1|2i-1} \end{pmatrix} = \begin{pmatrix} U_p^+ \\ U_f^- \end{pmatrix} \quad (9)$$

by moving the “present” time  $i$  one step ahead to time  $i + 1$ . Similarly, we define the output block Hankel matrices  $Y_{0|2i-1}$ ,  $Y_p$ ,  $Y_f$ ,  $Y_p^+$ ,  $Y_f^-$  and the input-output data matrices

$$W_p = W_{0|i-1} = \begin{pmatrix} U_{0|i-1} \\ Y_{0|i-1} \end{pmatrix} = \begin{pmatrix} U_p \\ Y_p \end{pmatrix}, \quad (10)$$

$$W_p^+ = \begin{pmatrix} U_p^+ \\ Y_p^+ \end{pmatrix}.$$

The state sequence  $X_m$  is defined as

$$X_m = (x_m \ x_{m+1} \ \cdots \ x_{m+j-1}) \quad (11)$$

and the “past” and “future” state sequences are  $X_p = X_0$  and  $X_f = X_i$ , respectively. With the state-space model (13), (15) we associate the extended observability matrix  $\Gamma_i$  and the reversed extended controllability matrix  $\Delta_i$ , where

$$\Gamma_i = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{i-1} \end{pmatrix}, \quad (12)$$

$$\Delta_i = (A^{i-1}B \ A^{i-2}B \ \cdots \ AB \ B).$$

The subspace identification method relies on determining the state sequence  $X_f$  and the extended observability matrix  $\Gamma_i$  directly from the input-output data  $u_k$ ,  $y_k$  and based on them, in a next step, to extract the matrices  $A$ ,  $B$ ,  $C$ , and  $D$ . To this end we define the oblique projection of the row space of  $Y_f$  along the row space of  $U_f$  on the row space of  $W_p$ :

$$\mathcal{O}_i = Y_f /_{U_f} W_p = [Y_f /_{U_f}^\dagger], [W_p /_{U_f}^\dagger]^\dagger W_p, \quad (13)$$

where  $\dagger$  denotes the Moore-Penrose pseudoinverse of a matrix and  $A/B^\dagger$  is a shorthand for the projection of the row space of matrix  $A$  onto the orthogonal complement of the row space of matrix  $B$ ; that is,

$$A/B^\dagger = I - AB^T (BB^T)^\dagger B. \quad (14)$$

As it is proved in [10] the matrix  $\mathcal{O}_i$  is equal to the product of the extended observability matrix and the “future” state vector

$$\mathcal{O}_i = \Gamma_i X_f. \quad (15)$$

Having in our disposal the matrix  $\mathcal{O}_i$  we use its singular value decomposition

$$\mathcal{O}_i = (U_1 \ U_2) \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = U_1 S_1 V_1^T \quad (16)$$

and we identify the extended observation matrix  $\Gamma_i$  and the state vector  $X_f$  as

$$\Gamma_i = U_1 S_1^{1/2} T, \quad (17)$$

$$X_f = T^{-1} S_1^{1/2} V_1^T = \Gamma_i^\dagger \mathcal{O}_i,$$

where  $T$  is an  $n \times n$  arbitrary nonsingular matrix representing a similarity transformation.

One method [10] to extract the matrices  $A$ ,  $B$ ,  $C$ , and  $D$  uses in addition the oblique projection

$$\mathcal{O}_{i-1} = Y_f^- /_{U_f^-} W_p^\dagger = \Gamma_{i-1} X_{i+1}, \quad (18)$$

where  $\Gamma_{i-1}$  is obtained from  $\Gamma_i$  by deleting the last  $\ell$  rows ( $\ell = 1$  in our case). Similar to the previous derivation we have

$$X_{i+1} = \Gamma_{i-1}^\dagger \mathcal{O}_{i-1} \quad (19)$$

and the matrices  $A$ ,  $B$ ,  $C$ , and  $D$  are obtained by solving (in least squares fashion) the system

$$\begin{pmatrix} X_{i+1} \\ Y_{i|i} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} X_i \\ U_{i|i} \end{pmatrix}. \quad (20)$$

After completion of the training phase the state-space model matrices  $A$ ,  $B$ ,  $C$ , and  $D$  for each zone are communicated back to the node responsible for the zone. Upon reception of the matrices the nodes enter a “convergence” phase; that is, starting from an initial state  $x_0$ , they update

the state (5) for a predetermined period of time, using the measurements of the surrounding zones and the zone's heater status. Note that the training and the convergence phase need to be executed only once prior to the normal operation of the nodes. After the convergence phase the nodes enter the "decision" phase. During this phase the nodes update the state of the model (as in the convergence phase) and periodically they take decisions on whether to turn the zone's heater on or not. Suppose that at the decision epoch  $t$  the state of the tagged unoccupied node is  $x_t$ . The node bases its decision on the estimates  $S_i^{\text{on}}(t)$  and  $S_i^{\text{off}}(t)$ . Recall that  $S_i^{\text{on}}(t)$  is the time needed to reach the comfort level  $T_c$  if the heater is on. That is,

$$S_i^{\text{on}}(t) = \min \ell : y_{t+\ell} \geq T_c. \quad (21)$$

The node "freezes" the input-vector  $u_k$  to its current value  $u_t$  (more recent temperatures of the surrounding zones and the tagged zone's heater is on) and starting from state  $x_t$  repeats (5), (6) until  $y_{t+k}$  exceeds the value  $T_c$ . Note that

$$y_{t+k} = CA^k x_t + CA^{k-1} Bu_t + CA^{k-2} Bu_{t+1} + \dots + CABu_{t+k-2} + CBu_{t+k-1} + Du_{t+k} \quad (22)$$

and for  $u_t = u_{t+1} = \dots = u_{t+k}$  we obtain

$$\begin{aligned} y_{t+k} &= CA^k x_t + C(I + A + \dots + A^{k-1}) Bu_t + Du_t \\ &= CA^k x_t + C(I - A)^{-1} (I - A^k) Bu_t + Du_t. \end{aligned} \quad (23)$$

If we set  $y_{t+k} = T_c$  and

$$w = T_c - C(I - A)^{-1} Bu_t - Du_t, \quad (24)$$

then (23) takes the form

$$\begin{aligned} w &= CA^k x_t - C(I - A)^{-1} A^k Bu_t \\ &= CV\Lambda^k V^{-1} x_t - C(I - A)^{-1} V\Lambda^k V^{-1} Bu_t, \end{aligned} \quad (25)$$

where we have considered the diagonalization of the matrix  $A$  as  $A = V\Lambda V^{-1}$ , with  $\Lambda$  being the diagonal matrix of the eigenvalues and  $V$  being the matrix of the corresponding eigenvectors. Numerical methods can be used to solve (25) for  $k$ . In a similar manner we compute  $S_i^{\text{off}}(t)$ . In this case, however, we first repeat (5)  $D$  times with input-vector  $u_k$  reflecting the fact that the heater is off, and then starting from the new state  $x_{t+D}$  we estimate the time needed to reach  $T_c$  if the heater is on.

**3.2. On the Distribution of  $Y_n(t)$ .** The random variable  $Y_n(t)$  expresses the remaining time from the current time instant  $t$  until entrance to zone  $z_n$ . That is, we assume that occupants move around as a swarm, visiting zones either in a predetermined order or in a random fashion and finally they end up in zone  $z_n$ . Each time the occupants visit a zone  $z_\ell$  they

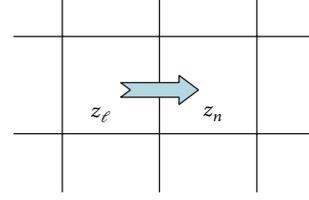


FIGURE 2: Movement of occupants from zone  $z_\ell$  to the adjacent unoccupied zone  $z_n$ .

stay in it for a nonzero time period which is itself a random variable with distribution  $F_\ell(x)$ . We use the variable  $t$  to designate calendar time, the time elapsed since the process started, and  $\tau$  to designate the time elapsed since entry to a zone. The zone occupation time distribution  $F_\ell(x)$  may be given parametrically either a priori or inferred from examination of historic data by suitable fitting of the distribution's parameters. We may use the empirical distribution function (e.d.f.) instead which is the nonparametric estimate of  $F_\ell(x)$ . That is, if a sample of  $N$  occupation periods of zone  $z_\ell$ ,  $\{x_1, x_2, \dots, x_N\}$  is available, then

$$\widehat{F}_\ell(x) = \frac{1}{N} \sum_{m=1}^N I(x_m \in [0, x]), \quad (26)$$

where  $I(\cdot)$  is the indicator function.

Let us consider first the simple case of two adjusting zones  $z_\ell, z_n$  with zone  $z_\ell$  being occupied and  $z_n$  unoccupied as it is shown in Figure 2. We assume that the occupants remain to zone  $z_\ell$  for a random period of time  $X_\ell$  which is distributed according to  $F_\ell(x)$  and then upon departure they enter zone  $z_n$ . In this case  $Y_n(t)$  expresses the remaining waiting time to zone  $z_\ell$  and follows the residual life distribution which is defined by

$$\begin{aligned} \overline{R}_{Y_n(t)}(y) &= 1 - R_{Y_n(t)}(y) = P\{X_\ell > \tau + y \mid X_\ell > \tau\} \\ &= \frac{\overline{F}_\ell(\tau + y)}{\overline{F}_\ell(\tau)}. \end{aligned} \quad (27)$$

Next we consider the case that occupants are currently in a zone (call it  $z_0$ ) and upon departure they visit in cascade zones  $z_1, z_2, \dots, z_n$  as it is shown in Figure 3. Then

$$Y_n(t) = Y_0 + X_1 + X_2 + \dots + X_{n-1}, \quad (28)$$

where  $Y_0$  is the remaining time in zone  $z_0$  and  $X_i$  is the occupation period of zone  $z_i$ . It is well known that the distribution of  $Y_n(t)$  is the convolution of the distributions of the random variables  $Y_0, X_1, \dots, X_{n-1}$ . That is

$$\begin{aligned} F_{Y_n(t)}(y) &= P\{Y_n(t) \leq y\} \\ &= R_0(y) * F_1(y) * F_2(y) * \dots * F_{n-1}(y) \end{aligned} \quad (29)$$

with  $R_0(y) = 1 - \overline{F}_0(\tau + y)/\overline{F}_0(\tau)$ . A more compact form is obtained by considering a transform of the distributions, that

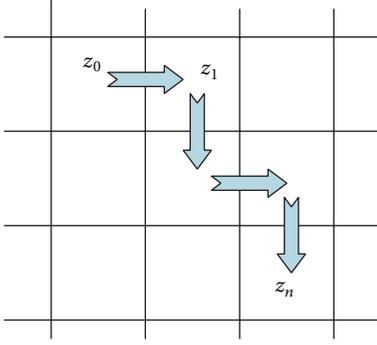


FIGURE 3: Movement of occupants in cascade from zone  $z_0$  to zone  $z_n$ .

is, the Laplace transform, the moment generating function, or the characteristic function. Using the moment generating function (mgf) which for a distribution  $F_X(x)$  of a nonnegative random variable  $X$  is defined by

$$F_X(s) = \int_0^{\infty} e^{sx} dF_X(x), \quad (30)$$

(29) takes the form

$$F_{Y_n(t)}(s) = R_0(s) \cdot F_1(s) \cdot F_2(s) \cdot \dots \cdot F_{n-1}(s). \quad (31)$$

In some cases (unfortunately few) (26) can be inverted analytically in order to calculate the distribution  $F_{Y_n(t)}(y)$ . For example, if the occupation periods  $X_i$ ,  $i = 0, \dots, n-1$ , are exponentially distributed with corresponding parameters  $\lambda_i$ , then  $Y_0$  is also exponentially distributed with parameter  $\lambda_0$  due to the memoryless property of the exponential distribution, and the probability density function of  $Y_n(t)$  is (assuming that the parameters  $\lambda_i$  are all distinct)

$$f_{Y_n(t)}(x) = \sum_{i=0}^{n-1} \frac{\lambda_0 \cdots \lambda_{n-1}}{\prod_{j=0, j \neq i}^{n-1} (\lambda_j - \lambda_i)} \exp(-x\lambda_i). \quad (32)$$

(The expression for the general case, of nondistinct parameters, is slightly more involved but still exists.) If a closed form for the density  $f_{Y_n(t)}(x)$  does not exist or it is difficult to be obtained, we resort to the saddlepoint approximation [11] to invert  $F_{Y_n(t)}(s)$ . To this end, for a random variable  $X$  with cumulative distribution function (cdf)  $F_X(x)$ , we define the cumulant generating function (cgf)  $K_X(s)$  as the logarithm of the corresponding mgf  $F_X(s)$ ; that is,

$$K_X(s) = \log F_X(s). \quad (33)$$

Then, the saddlepoint approximation of the probability density function (pdf)  $f_X(x)$  is

$$f_X(x) \approx \left( \frac{1}{2\pi K_X''(\widehat{s}(x))} \right)^{1/2} \exp\{K_X(\widehat{s}(x)) - \widehat{s}(x)x\}, \quad (34)$$

where  $\widehat{s}(x)$  is the solution to the saddlepoint equation

$$K_X'(s) = x. \quad (35)$$

Estimation (34) relies on numerical computations based on cgf  $K_X(s)$ , which in turn depends on  $F_X(s)$ . Note that we may use the empirical mgf  $\widehat{F}_X(s)$  instead, so that a nonparametric estimation of the pdf  $f_X(x)$  is possible.

The more general case is to model the movement of the occupants in the multizone system as a semi-Markov process. The states of the process are the zones of the system and at the transition times they form an embedded Markov chain with transition probabilities  $p_{ij}$ . Given a transition from zone  $z_i$  to zone  $z_j$ , the occupation time in  $z_i$  has distribution function  $F_{ij}(x)$ . The matrix  $\mathbf{Q}(x)$  with elements  $Q_{ij}(x) = p_{ij}F_{ij}(x)$  is the so called semi-Markov kernel. The occupation time distribution in zone  $z_i$  independent of the state to which a transition is made is  $F_i(x) = \sum_j Q_{ij}(x)$ . Consider now the case that at time  $t$  occupants exist in zone  $z_i$  and the node of zone  $z_n$  has to make a decision on whether to start heating the zone or not. For the decision process the distribution of time until entrance to zone  $z_n$ ,  $Y_n(t)$ , is needed.  $Y_n(t)$  is the sum of two terms. The first term is the residual occupation time at zone  $z_i$  and the second term is the time to reach  $z_n$  upon departure from zone  $z_i$ . The distribution  $R_i(x)$  of the first term is given by (27), whereas for the second term we have to consider the various paths taken upon departure from state  $z_i$ . Thus, if zone  $z_i$  neighbors  $N_i$  zones, namely,  $z_k$ ,  $k = 1, \dots, N_i$ , then the time to reach zone  $z_n$  upon departure is distributed according to

$$G_{in}(x) = \sum_{k=1}^{N_i} p_{ik} H_{kn}(x), \quad (36)$$

where  $H_{kn}(x)$  is the distribution of the first passage from zone  $z_k$  to zone  $z_n$ . Pyke [12] and Mason [13, 14] provided solutions for the first passage distribution in a semi-Markov process. For example, Pyke's formula states that  $H(s)$ , the elementwise transform of the matrix  $H(x) = \{H_{ij}(x)\}$ , is given by

$$H(s) = \mathcal{T}(s) [I - \mathcal{T}(s)]^{-1} \left\{ \left[ [I - \mathcal{T}(s)]^{-1} \right]_d \right\}^{-1}, \quad (37)$$

where  $\mathcal{T}(s)$  is the transmittance matrix with the transforms  $\mathcal{T}_{ij}(s)$  of the distributions  $F_{ij}(x)$ . Notation  $M_d$  denotes the matrix that is formed by the diagonal elements of  $M$ . A simplified version of (37) is obtained if we let  $\Delta(s) = \det(I - \mathcal{T}(s))$  and  $\Delta_{ij} = (-1)^{i+j} \det(I - \mathcal{T}(s))_{ij}$ , the  $ij$ th cofactor of  $I - \mathcal{T}(s)$ . Using this notation, for a semi-Markov process of  $n$  states, we get

$$H_{1n}(s) = \frac{\Delta_{n1}(s)}{\Delta_{nn}(s)}. \quad (38)$$

By relabeling the states  $1, \dots, n-1$  of the process we can find the transform of the first passage from any zone to zone  $z_n$ .

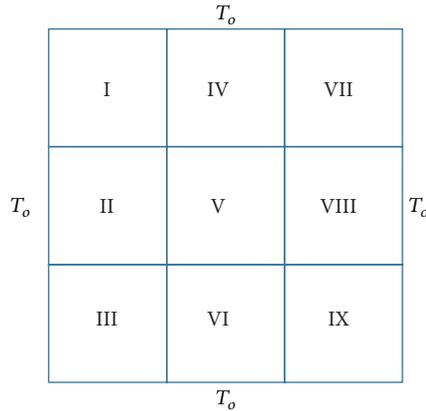


FIGURE 4: Multizone model.

In summary,

$$F_{Y_n(t)}(s) = R_i(s) \cdot G_{in}(s) = R_i(s) \cdot \left( \sum_{k=1}^{N_i} p_{ik} H_{kn}(s) \right), \quad (39)$$

with  $H_{kn}(s)$  computed using (38). Next, the saddlepoint approximation technique can be used to calculate the density  $f_{Y_n(t)}(x)$ .

We have to mention that the process of estimating the first passage distributions needs to be performed only prior to the decision process. Thus, given the topology of the multizone system, we posit a parametric family for the occupation time distributions and we estimate the distributions' parameters from sample data. Then, the transition probabilities are estimated based on the data and the occupation mgf  $F_{ij}(s)$  is obtained. Having in our disposal the transforms  $F_{ij}(s)$  we can form the transmittance matrix  $\mathcal{T}(s)$  and solve for the first passage transform  $H_{ij}(s)$  using (38). Finally, numerical methods may be used to invert these transforms in order to obtain the necessary probability density functions.

#### 4. Simulation Results

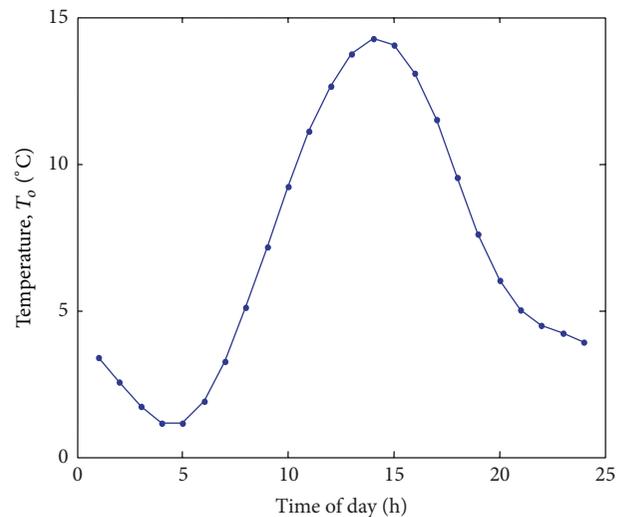
We simulated a multizone system consisting of a squared arrangement of rooms (zones) where each room (zone) is equipped with a wireless sensor node as shown in Figure 4. We assume that heat transfer takes place between room's air and the walls as well as between room's air and the roof. Furthermore, we neglect the effects of ground on the room temperature. North and south walls have the same effect on the room temperature and we assume the same for the east and west walls. According to these assumptions, there is a symmetry in the dynamics of the zones. For example, zones I, III, VII, and IX exhibit the same behavior. Note that, for the multizone system of Figure 4 several state variables, such as wall temperatures, are common to the individual zone systems. In the simulations we treat the multizone system as a single system with 42 states and nine outputs (the zones' temperatures). The parameter values for this lumped capacity zone model were taken from [15] (see also [16]).

TABLE 1: Training and testing parameters for outer temperature model and heat gain.

Phase	Walter's model			Heat gain
	$T_{\max}$	$T_{\min}$	$T_m$	$q(t)$
Training	15	2	7	600
Testing	16	1	6	700

TABLE 2: Training and testing zones' target temperatures.

Zone #	Target temperatures								
	I	II	III	IV	V	VI	VII	VIII	IX
Training	16	18	16	15	14	13	12	16	14
Testing	17	20	17	17	17	14	18	19	21

FIGURE 5: Empirical estimation of  $T_o$ .

In Figure 4,  $T_o$  denotes the outside temperature which is assumed to be uniform with no loss of generality. For simulation purposes daily outside temperature variations are obtained using Walters' model [17]. For this model, the maximum temperature of day  $T_{\max}$ , the minimum temperature of day  $T_{\min}$ , and the mean of the 24 hourly temperatures  $T_m$  need to be provided in order to estimate  $T_o$ . As an example, if  $T_{\max} = 15^\circ\text{C}$ ,  $T_{\min} = 2^\circ\text{C}$ , and  $T_m = 7^\circ\text{C}$ , the daily variations of temperature in Figure 5 are obtained, which shows that temperature is higher at 2:00–3:00 pm.

**4.1. Training Phase.** For the deterministic subsystem identification we used 86400 samples with sampling period  $T_s = 1$  sec. For the outside temperature we used Walter's model with parameters given in the first row of Table 1 and a heater gain equal to 600 W. We set the zones' target temperatures  $T_{tg}$  to those of Table 2 (first row) with a margin equal to  $T_{mg} = 1^\circ\text{C}$ . That is, for a specific zone the heater is on until temperature  $T_{tg} + T_{mg}$  is reached. After this point the heater is turned off until temperature hits the lower threshold  $T_{tg} - T_{mg}$  and the whole process is repeated.

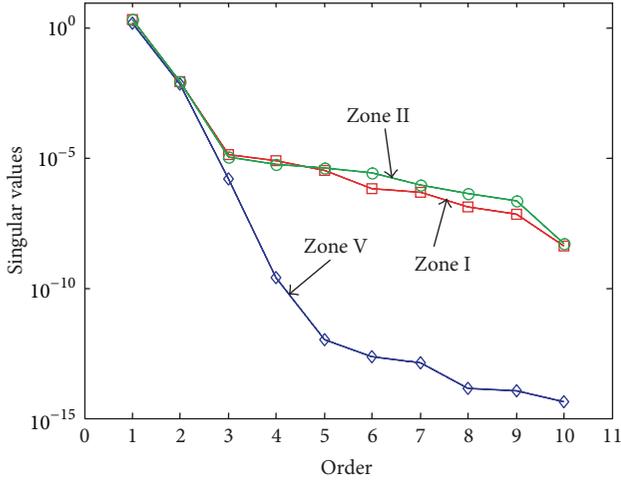


FIGURE 6: Singular values of zones I, II, and V.

Having collected the data over the period of 24 hours, the measurements of the zones' temperature, the outer temperature, and the heat gains of each zone are organized into input-output data for the subsystem identification process. For example, for zone I, the outer temperature, the heating gain of zone I, and the temperatures of zones II and IV are the input data to the identification process whereas the temperature of zone I itself is the output data. Note that the number of input signals differs from zone to zone. Zones I, III, VII, and IX use 4 input signals, whereas the rest of the zones use 5 input signals. Based on the input-output data we identify the matrices  $A$ ,  $B$ ,  $C$ , and  $D$  for each zone as described in Section 3. During this process we have to decide on the order of the subsystems. This is achieved by looking at the singular values of the SVD decomposition in (20) and deciding on the number of dominant ones. Figure 6 depicts the singular values for zones I, II, and V. All subsystems exhibit similar behavior regarding the profile of their singular values and therefore we set the order of all subsystems equal to 2.

Next we run a test on the obtained state-space models. We set the outer temperature parameters as in the second row of Table 1 and the target temperatures as in the second row of Table 2. The initial state of the subsystem model was set to  $[-10 \ 0]$ , whereas the elements of the 42-state vector of the multizone system were set equal to 3 (the initial output temperature). Figure 7 shows the evolution of the actual and predicted temperatures for zones I and IV. For zone I the heater was on (heat gain equal to 700) whereas the heater for zone IV was off. As it is observed, after 5000 samples (approximately 80 minutes period) the state of the subsystems has converged to one that produces almost the same output as the original system. After this point the WSN nodes can enter in the "decision" phase.

We should notice at this point that the "original" simulated system (42-state-space model) is also linear and thus the use of linear subspace identification may be questionable for more complex and possibly nonlinear systems. As the simulation results indicate, although the dynamics of each

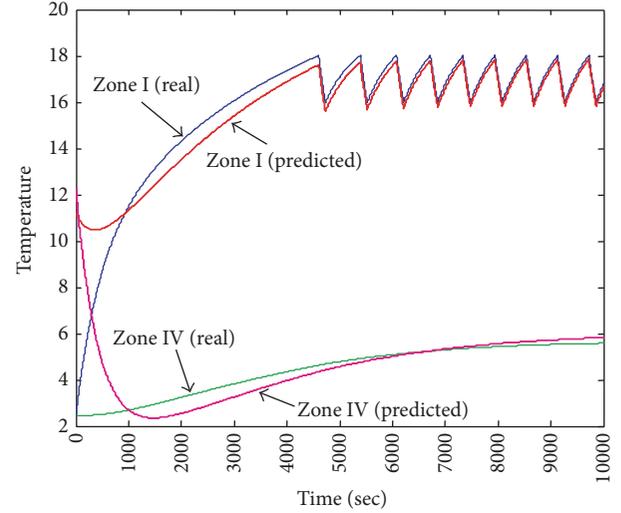


FIGURE 7: Real and predicted temperature of zones I and IV.

zone are more complex (including the roof temperature, e.g.), a low dimensionality subsystem of order 2 can capture its behavior. For more complex systems we can choose the order of the identified subsystem to be high enough. For nonlinear systems the identification of a time varying system is possible by using a recursive update of the model.

Next we simulate the scenario of Figure 2. We assume that occupants enter zone I (at time instant 5000) and then after a random period of time they enter zone IV, where they remain for 3600 sec. Zone I is heated since the beginning of the process (heat gain 900 W) until the occupants leave the zone to enter zone IV. Zone IV uses a heater with gain 1200 W, which is turned on (or off) according to the decisions of its node. The target temperatures for zones I and IV are  $17^\circ\text{C}$  and  $16^\circ\text{C}$ , respectively. For the first set of experiments, the occupation period of zone I is Gamma distributed, that is,

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad (40)$$

with the shape parameter  $\alpha = 5$  and  $\beta$  such that the expected value of  $X$ ,  $E[X] = \alpha/\beta$ , is 3600 and 7200. Decisions were taken every  $D = 67$  seconds. Figure 8 shows the total comfort cost achieved for various values of the comfort weight  $C_i$ . The results are averages over 200 runs of the simulation. The case of  $C_i = 0$  is the full reactive case where heating of zone IV is postponed until occupants enter the zone. On the other extreme, for high values of  $C_i$ , the decision process acts proactively by spending energy in order to reduce the discomfort level.

Figure 9 shows the total energy cost in kWh for the two cases of Figure 8. As it is observed, the higher the value of  $C_i$  the more the energy consumed since then heating of zone IV starts earlier. The difference between the two curves of Figure 9 is justified by the difference of the means of the occupation period,  $E[X] = 3600$  and  $E[X] = 7200$ , for the two cases.

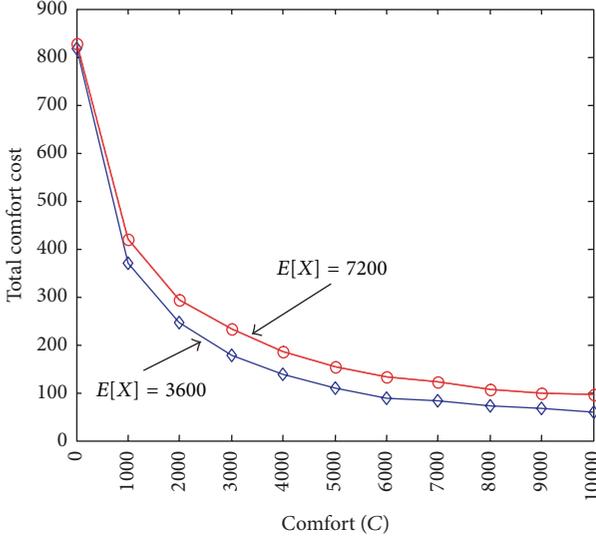


FIGURE 8: Total comfort cost versus  $C_i$ . The occupation time of zone I is Gamma distributed.

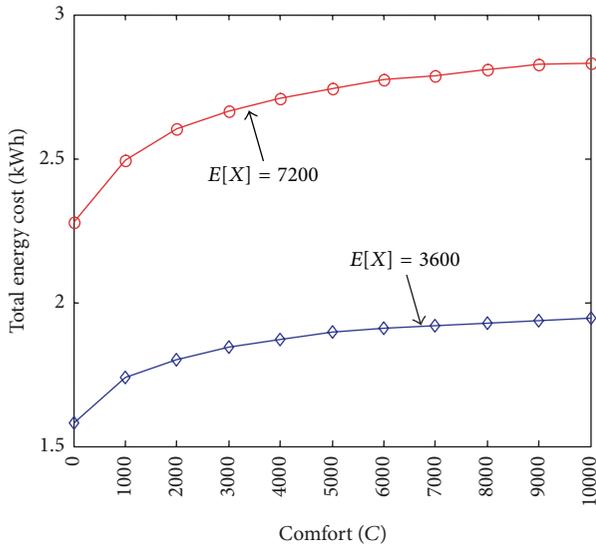


FIGURE 9: Total energy cost versus  $C_i$ . The occupation time of zone I is Gamma distributed.

In Figure 10 we plot the temperature profiles for zones I, IV and for two different values of  $C$ . As it is observed the value of  $C = 0$  corresponds to the reactive case; that is the heater of zone IV remains off until occupants are detected in the zone. On the other extreme a large value of  $C$  (10000) will force the heater of zone IV to be turned on as soon as possible in an effort to reduce the discomfort penalty.

The scenario of Figure 2 was also simulated for Pareto distributed occupation times of zone I. The pdf for Pareto distribution is

$$f_X(x) = \begin{cases} \frac{\alpha x_m^\alpha}{x^{\alpha+1}}, & x \geq x_m \\ 0, & x < x_m \end{cases} \quad (41)$$

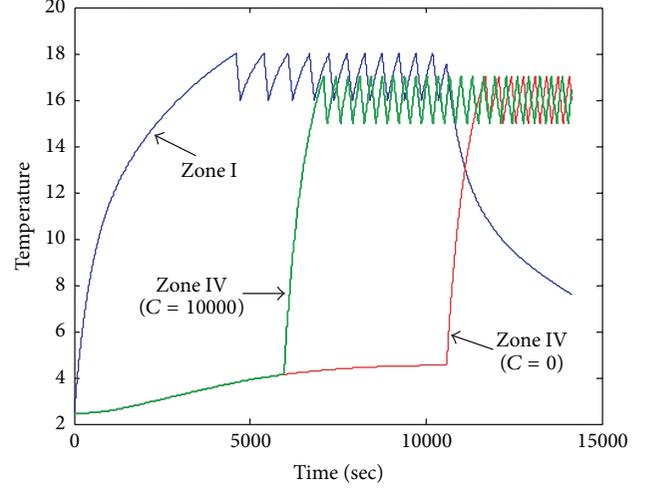


FIGURE 10: Temperature versus time for  $C = 10000$  and  $C = 0$ .

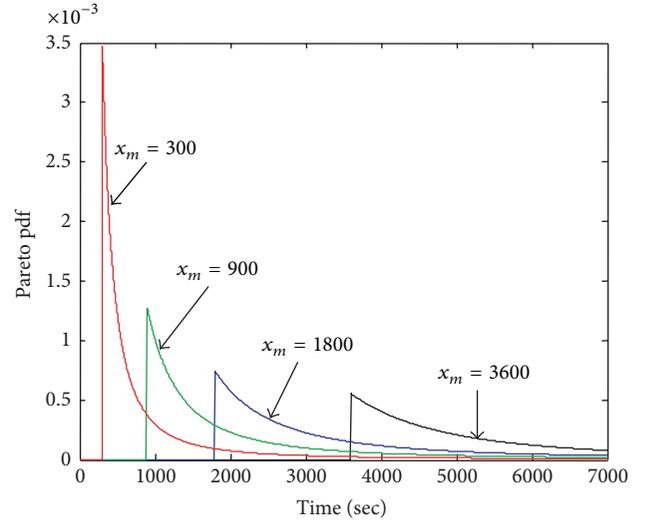


FIGURE 11: Pareto pdf for various values of  $x_m$ .

with expected value

$$E[X] = \begin{cases} \frac{\alpha x_m}{\alpha - 1}, & \alpha > 1 \\ \infty, & \alpha \leq 1. \end{cases} \quad (42)$$

Figure 11 shows the Pareto pdf for four values of  $x_m$  ( $x_m = 300, 900, 1800, 3600$ ) and  $\alpha$  suitably chosen so that the mean value of the occupation period is 7200 in all cases.

Figures 12 and 13 show the total comfort cost and the total energy consumed, respectively, for various values of the comfort gain  $C_i$ . As it is observed, a 50% improvement on the comfort cost (compared to the reactive case) can be achieved for a value of  $C$  equal to 1000. However, there is a performance floor (this is evident for the case of  $x_m = 300$ ), which is due to the nature of the Pareto distribution. Most of the samples for the occupation time period are concentrated close to  $x_m$

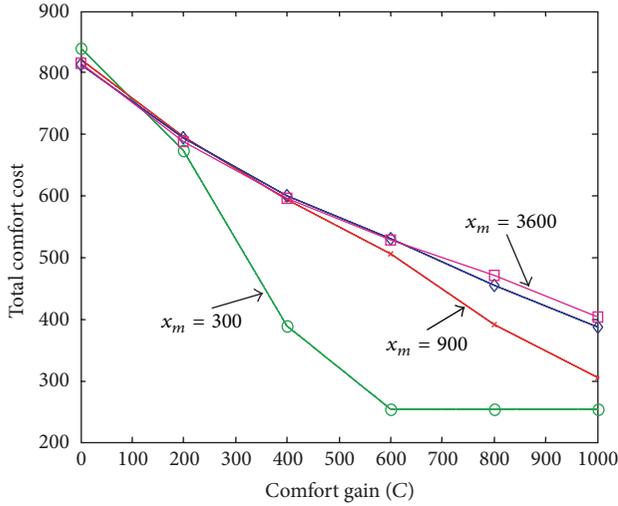


FIGURE 12: Total comfort cost for Pareto distributed occupation period.

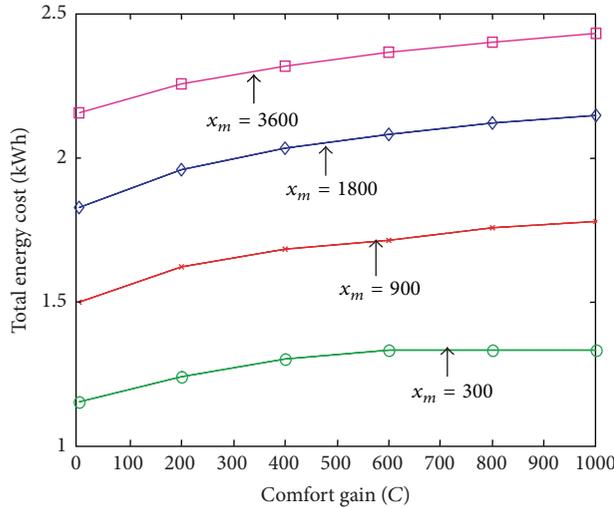


FIGURE 13: Total energy cost for Pareto distributed occupation period.

and therefore there is not enough time to preheat the next visiting zone, which in turn implies high discomfort costs. Moreover, the heavy tail of the Pareto distribution causes some extremely high values of the occupation period which result in an increase of the total consumed energy, as it can be observed clearly from Figure 13.

Next, we simulate the scenario of Figure 3. The occupants move in cascade from zone to zone as it is depicted in Figure 14. The target temperatures for the zones are given in the first row of Table 3, whereas the heat gain of each zone is given in the second row of Table 3. For the occupancy time of each zone we considered an exponential random variable with mean provided in the last row of Table 3. Note that this assumption is the least favorable for adjacent zones due to the memoryless property of the exponential distribution.

TABLE 3: Parameters used in cascade movement scenario.

Zone #	I	IV	V	VIII	IX	VII	III	II
Target °C	17	16	17	17	16	14	17	18
Power (KW)	0.9	1.2	0.9	0.9	0.9	1.2	0.9	1.2
Mean time $\times 10^2$	36	27	18	40	54	45	30	20

TABLE 4: Energy and comfort costs for constant occupancy times.

	Comfort weight C	500	1000	1500	2000
$D = 67$	Comfort cost	2615	72	67	67
	Energy cost	7.75	9.07	9.73	10.20
$D = 167$	Comfort cost	2856	72	71	64
	Energy cost	7.79	9.11	9.77	10.25

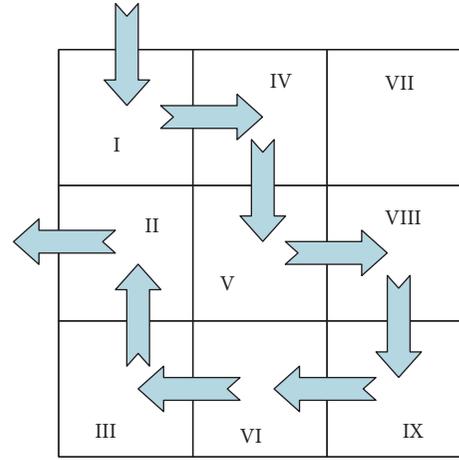


FIGURE 14: Movement of occupants in cascade (zones I, IV, V, VIII, IX, VI, III, and II).

The total comfort and energy costs for this scenario are plotted in Figure 15 for various values of the comfort weight  $C$ . The results are averages over 100 runs. The two curves for each cost correspond to  $D = 67$  and  $D = 167$ , respectively. As it is observed, considerable comfort gains are achieved for a moderate increase of the consumed energy. In Figure 15 we also show the costs for a “fixed” proactive scenario in which the heating of a zone starts as soon as occupants are detected to its neighbor zone. For example, when occupants enter zone IV, the heater of zone V is turned on, and so on.

Table 4 shows the comfort and energy costs obtained when there is no randomness on the occupancy times for various values of the comfort weight  $C$  and for two different values of  $D$  ( $D = 67$  and  $167$ ). The occupancy times of the zones were set to the mean values provided in the last row of Table 3. As it is observed, extremely low comfort costs can be achieved in this case which implies that the right modelling of the occupancy times is of paramount importance to the process.

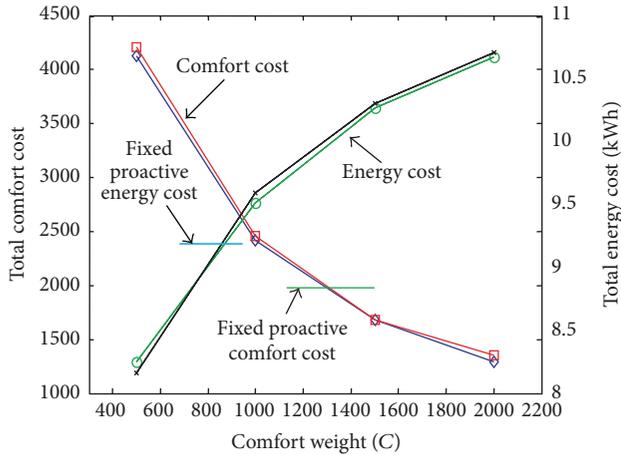


FIGURE 15: Total energy and comfort costs (cascade movement scenario).

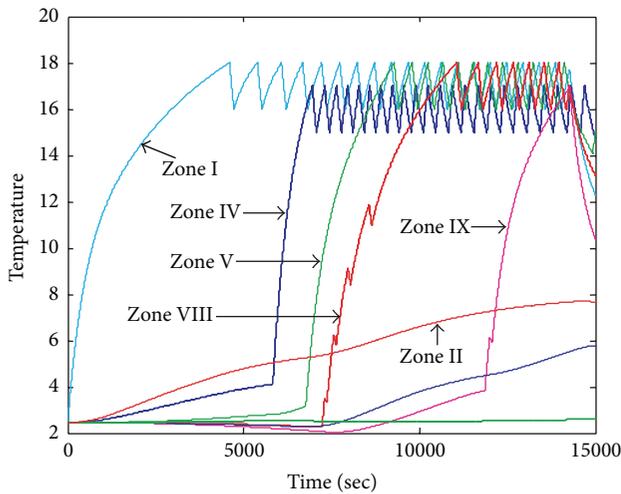


FIGURE 16: Temperature profiles of zones (cascade movement scenario).

Finally, Figure 16 shows a sample of the zones' temperature profile. Note that the temperature of zone II, although zone II is the last zone visited, starts increasing (at a slower rate) much earlier. This is because zone II neighbors zones I and V which are heated in earlier stages of the process.

## 5. Conclusions and Future Work

A method that fuses temperature and occupancy predictions in an effort to balance the thermal comfort condition and the consumed energy in a multizone HVAC system was presented. Temperature predictions are based on a subspace identification technique used to model the thermal dynamics of each zone independently. This technique is quite robust and accurate predictions are possible after a convergence period which may last 1-2 hours. With regard to the occupancy predictions the decision process makes use of the distribution of the first passage time from an occupied zone

to an unoccupied one. Simulation results for different first passage distributions demonstrated the effectiveness of the proposed technique.

Several extensions and modifications of the proposed method are possible. The comfort parameter  $C_i$  may depend on the zone, the number of occupants, and the current status of the zones. Moreover, this parameter may be time variable, that is, different values may be used for day and night hours. Regarding the decision process itself, there is no need that it be executed at regular time epochs. If the energy consumed by the wireless sensor nodes is an issue, we may take decisions at irregular time epochs, depending on the occupancy predictions of the zones. The computation of the risks, used by the decision process, may take into consideration additional parameters that affect energy consumption and/or the thermal comfort of the occupants. For example, open windows situation may be easily detected and incorporated suitably in the decision process.

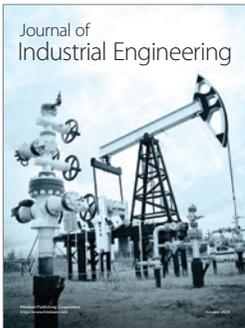
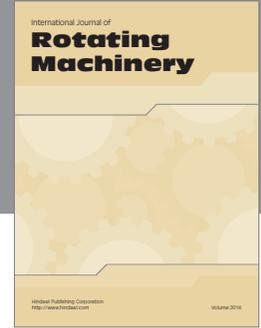
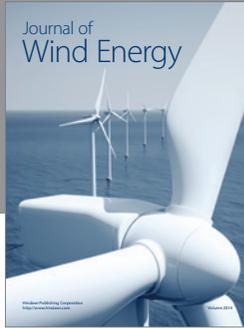
## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] US Energy Information Administration, "Annual Energy Outlook 2015," DOE/EIA-0383, April 2015, <http://www.eia.gov/>.
- [2] U.S. Energy Information Administration, June 2015, <http://www.eia.gov>.
- [3] S. Lee, Y. Chon, Y. Kim, R. Ha, and H. Cha, "Occupancy prediction algorithms for thermostat control systems using mobile devices," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1332–1340, 2013.
- [4] V. L. Erickson, S. Achleitner, and A. E. Cerpa, "POEM: power-efficient occupancy-based energy management system," in *Proceedings of the 12th International Conference on Information Processing in Sensor Networks (IPSN '13)*, pp. 203–216, ACM, Philadelphia, Pa, USA, April 2013.
- [5] V. L. Erickson, M. Á. Carreira-Perpiñán, and A. E. Cerpa, "OBSERVE: occupancy-based system for efficient reduction of HVAC energy," in *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN '11)*, pp. 258–269, Chicago, Ill, USA, April 2011.
- [6] B. Balaji, J. Xuy, A. Nwokafor, R. Guptay, and Y. Agarwal, "Sentinel: occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings," in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (SenSys '13)*, Rome, Italy, November 2013.
- [7] B. Dong and K. P. Lam, "A real-time model predictive control for building heating and cooling systems based on the occupancy behavior pattern detection and local weather forecasting," *Building Simulation*, vol. 7, no. 1, pp. 89–106, 2014.
- [8] J. R. Dobbs and B. M. Hency, "Predictive HVAC control using a Markov occupancy model," in *Proceedings of the American Control Conference (ACC '14)*, pp. 1057–1062, Portland, Ore, USA, June 2014.
- [9] J. Brooks, S. Kumar, S. Goyal, R. Subramany, and P. Barooah, "Energy-efficient control of under-actuated HVAC zones in commercial buildings," *Energy and Buildings*, vol. 93, pp. 160–168, 2015.

- [10] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems, Theory-Implementation-Applications*, Kluwer Academic, New York, NY, USA, 1996.
- [11] H. E. Daniels, "Saddlepoint approximations in statistics," *Annals of Mathematical Statistics*, vol. 25, pp. 631–650, 1954.
- [12] R. Pyke, "Markov renewal processes with finitely many states," *Annals of Mathematical Statistics*, vol. 32, pp. 1243–1259, 1961.
- [13] S. J. Mason, "Feedback theory-some properties of signal flow graphs," *Proceedings of Institute of Radio Engineers*, vol. 41, no. 9, pp. 1144–1156, 1953.
- [14] S. J. Mason, "Feedback theory—further properties of signal flow graphs," *Proceedings of the IRE*, vol. 44, no. 7, pp. 920–926, 1956.
- [15] B. Tashtoush, M. Molhim, and M. Al-Rousan, "Dynamic model of an HVAC system for control analysis," *Energy*, vol. 30, no. 10, pp. 1729–1745, 2005.
- [16] D. Sklavounos, E. Zervas, O. Tsakiridis, and J. Stonham, "A subspace identification method for detecting abnormal behavior in HVAC systems," *Journal of Energy*, vol. 2015, Article ID 693749, 12 pages, 2015.
- [17] A. Walter, "Notes on the utilization of records from third order climatological stations for agricultural purposes," *Agricultural Meteorology*, vol. 4, no. 2, pp. 137–143, 1967.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

