A Non-Intrusive Occupancy Monitoring System for Demand Driven HVAC Operations

Zheng YANG[1], Nan LI[2], Burcin BECERIK-GERBER[3], Michael OROSZ[4]

[1]Ph. D. Student, Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA 90089; PH (213) 740-0578; email: zhengyan@usc.edu
[2]Ph. D. Student, Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA 90089; PH (213) 740-0578; email: nanl@usc.edu
[3]Assistant Professor, A.M.ASCE, Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA 90089; PH (213) 740-4383; Fax: (213) 744-1426; email: becerik@usc.edu
[4]PhD, Project Leader, Information Sciences Institute, University of Southern California, Marina del Rey, CA 90292; PH (310) 448-8266; email: mdorosz@isi.edu

ABSTRACT

In the U.S., 40% of the energy consumption is from buildings, approximately 48% of which is consumed by heating, ventilation, and air conditioning (HVAC) systems. Implementing demand driven HVAC operations is a way to reduce HVAC related energy consumption, and ultimately to achieve sustainable building operations and maintenance. This relies on the availability of occupancy information, which determines peak/off-hour modes and impacts cooling/heating loads of HVAC systems. This research proposes an occupancy monitoring system that is built on a combination of non-intrusive sensors that can detect indoor temperature, humidity, CO_2 concentration, door status, light, sound and motion. The effectiveness of each sensor in occupancy estimation is evaluated. The sensor data is communicated wirelessly, and processed in real time using a back-propagation (BP) artificial neural network (ANN) algorithm. Field tests are carried out in a lab space that is shared by up to 9 people for 15 consecutive days. The test results report an overall detection rate of over 90%, which indicates the ability of the proposed system to monitor the occupancy information of multi-occupancy spaces in real time in support of demand driven HVAC operations.

Keywords: Energy consumption; HVAC; Demand driven; Occupancy monitoring; Non-intrusive sensor

INTRODUCTION

Energy conservation is becoming an increasingly important topic in the contemporary society due to the rising energy demand and diminishing energy resources. This has attracted wide attention from the government, industry, academia, and public, and resulted in an emerging consensus that the energy needs should be addressed in more sustainable ways. In the U.S., buildings account for 40% of total

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

energy consumption, 48% of which is consumed by heating, ventilation, and air conditioning (HVAC) systems (DOE 2011). Given the fact that in the U.S., existing facilities represent over 97% of the existing building stock (Shelley and Roessner 2004) and that buildings are generally in operation for 30 to 50 years, there is huge potential of energy savings through improving the operations of HVAC systems in existing buildings.

Traditional HVAC system operations assume the ventilation and conditioning demand is at the peak value, and rely on temperature and humidity as the only inputs in adjusting the operations, which often results in waste of HVAC related energy consumption (Agarwal et al. 2010). Even with improved HVAC systems that run at different capacities at different times of the day, e.g. minimum capacity at off hours, energy can still be wasted e.g. by over cooling unoccupied spaces. As a solution to such inefficiencies with HVAC related energy consumption, the idea of demand driven HVAC operations has attracted considerable attention from the academia, and given rise to active research on this topic. Previous research has proven that the application of demand driven HVAC operations, which replaces the assumption that the ventilation and conditioning demand is at the peak value with the actual demand based on real-time sensing of the environment, could save up to 56% of HVAC related energy consumption (Sun et al. 2011). The actual ventilation and conditioning demands depend on various factors that should be input into HVAC systems for energy-efficient operations, among which fine-grained occupancy information is a key input (Agarwal et al. 2010). Occupancy information enables timely reaction to changing ventilation and conditioning demands, and minimizes energy consumption without compromising the occupant comfort.

Due to the importance of the occupancy information, a number of occupancy detection systems have been proposed in previous research, which reported consequent HVAC energy savings between 10% and 56% based on simulations (Erickson et al. 2009; Erickson and Cerpa 2010; Sun et al. 2011; Tachwali et al. 2007). However, these occupancy detection systems have certain limitations with respect to accuracy, cost, intrusiveness, and privacy, and therefore bear considerable potentials for improvement. Melfi et al. (2011) defined a successful occupancy detection system as one that can provide accuracy with high resolution. Generally, a higher resolution means a more specific sensing space, a more defined occupancy, and quicker information retrieval. In addition to these criteria, the authors believe that cost efficiency and non-intrusiveness are also essential attributes of a successful occupancy detection system. The former ensures the system is affordable, and the latter ensures the convenience of deployment. Following these criteria, this paper proposes a system that has the following features: (1) low cost. The system is built on a number of off-the-shelf low-cost sensors, most of which are as inexpensive as several dollars; (2) high-resolution. The system is configured to count the number of occupants at the room level, with a sample rate of one reading per minute, leading to a short system latency; (3) accurate. To improve the accuracy, the proposed system includes a comprehensive list of ambient sensors to account for various factors that may be associated with the occupancy information; and (4) non-intrusive. The proposed system does not require the installation of intrusive infrastructure in

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

buildings, nor require occupants to wear or carry any sensors. Moreover, it does not identify and record occupants' identities, and therefore does not cause any privacy issues. This study is part of a DOE sponsored project, Building Level Energy Management System (BLEMS), which has the objective of optimizing building energy efficiency.

LITERATURE REVIEW

Potential benefits of energy savings by implementing demand driven HVAC operations have motivated considerable research efforts in providing an effective occupancy detection system. The system is expected to provide high-resolution and accurate occupancy information to drive the operation of HVAC systems. CO_2 sensors have been widely used for this purpose (Leephakpreeda et al. 2001; Sun et al. 2011), as a larger occupancy in a space usually results in higher CO_2 concentrations. However, it usually takes some time for the CO_2 concentration to build up, and the CO_2 concentration is affected by not only occupancy but also other factors such as the weather. Such limitations indicate that the CO_2 sensor based systems are low-resolution, and unable to provide accurate and real-time occupancy information. Researchers have also proposed various video based systems (Benezeth et al. 2011; Erickson et al. 2009; Wang et al. 2010), which detect the occupancy in a monitored space by using image-processing techniques. However, these video based systems generally suffer from the requirement for line of sight in the monitored spaces, which compromises the applicability of these systems especially in heavily-partitioned spaces. Moreover, the use of video cameras usually requires large image storage space, and can cause privacy concerns among users.

To overcome these limitations, researchers have proposed to use a combination of various ambient sensors. Agarwal et al. (2010) used a magnetic reed switch door sensor and a PIR sensor for occupancy detection, which could report the actual occupancy most of the time. However, their occupancy detection algorithm was only applicable for single-occupancy offices, and was built on an assumption that occupants always keep their doors open when they are in the offices or being somewhere nearby. Meyn et al. (2009) used measurements from cameras, PIRs, and CO_2 sensors, as well as historical data of building utilization, to estimate the building occupancy level. The estimation was done by solving a receding-horizon convex optimization problem. The reported accuracy was 89%. Limitations include that the system was not able to estimate the number of occupants at the room level, and that the error tended to accrue over time. Henze et al. (2006) proposed an occupancy detection system that comprised of three PIRs and one telephone sensor for each room and relied on the belief networks algorithm. The system could detect if any occupant was present with an accuracy of 76%, but was not able to count the number of occupants. Dong et al. (2010) proposed a system that estimated the occupancy of a space by sensing the CO_2 concentration, acoustics, and motion in the space. Field tests carried out in two rooms yielded an accuracy of around 75%. Hailemariam et al. (2011) built an occupancy detection system for cubicles in offices using a combination of light sensors, motion sensors, CO_2 sensors, and sound sensors. The decision trees algorithm was used. An accuracy of 98.4% in detecting if a room was

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

occupied or unoccupied was achieved using the motion sensor alone, and a decline in the accuracy was reported when other sensors were integrated.

Based on the previous research, this study aims to make contribution by introduction an algorithm, i.e. error back propagation (BP) artificial neural network (ANN), in estimating the number of occupants. This algorithm, to the authors' best knowledge, has not been tested for this purpose before. This study also is also distinguished from previous research in that it introduces principal component analysis (PCA) to evaluate the effectiveness of different sensor data in occupancy estimation, and prioritizes the sensors. This can provide guidance in choosing proper sensors and building an effective occupancy monitoring system.

METHODOLOGY

Principal Component Analysis

The PCA is an approach to pre-process the data before applying it to ANN. It converts a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has as high a variance as possible, and each succeeding component in turn has the highest variance possible under the constraint that it can be uncorrelated with the preceding components. As the occupancy calculation may not rely on all principal variables equally, there is a need to analyze and prioritize these variables.

Artificial Neural Network

ANN is a highly complicated and large-scale nonlinear adaptive system for simulating human neural network. It consists of huge amounts of simple processing units, and is often used to do massively parallel data processing through continuously adjusting the relationships between inner variables. Some of the advantages of ANN include storage of distributed information, fault tolerance and strong ability of learning and association.

The specific model of ANN used in this study is error back propagation ANN. As one of the most widely used ANN models, the BP network is a typical multilayer feedforward neural network. It is known for its ability to transfer signal forward and transfer error backward, and can be infinitely approximated to an unknown continuous system. The BP network can learn and store large amounts of relations between input and output data without any predefined mathematical functions to describe them. The main application areas of BP ANN are in function approximation, pattern recognition and classification, and data compression. The topological structure of the BP neural network is shown in Figure 1, where X_1 , X_2 …, X_n are the input values of the BP neural network, Y_1 , Y_2 …, Y_m are the predicted values, and W_{ij} and W_{jk} are the network weights.

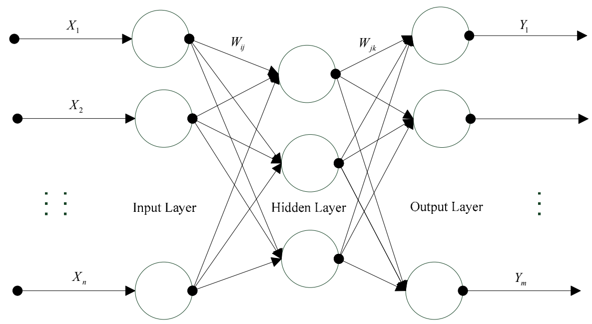

Figure 1: BP neural network topological structure (Rumelhart et al. 1986)

In a two-layer BP network, S excitation function is used in a hidden layer while a linear function is used in output layer. With a sufficient number of neurons in a hidden layer, the network can approximate to any arbitrary function.

TEST SETUP

A sensor node is built and used in the tests (Figure 2). The sensor node consists of an Arduino Black Widow stand-alone single-board microcontroller computer with integrated support for 802.11 WiFi. The sensor node includes the following sensors: a lighting sensor, a sound sensor, a motion sensor, a CO_2 sensor, a temperature sensor, a relative humidity sensor, a PIR sensor that detects objects as they pass through the door, and a door switch sensor that detects whether the door is open or closed. A script is written and uploaded to the sensor node using Arduino to configure the microcontroller to process the raw data. The processed data reported by the sensor node includes 13 variables, which can be categorized into three types: instant variables that show the instant output of a sensor at the time the data is queried, including *lighting, sound, motion, CO_2 concentration*, *temperature, relative humidity, reflector* (infrared), and *door status*; count variables that sum the number of times a sensor's output changes in the last minute, including *motion count, reflector count*, and *door count*; average variables that show the average value of a sensor's output over a certain period of time, including *sound average* (5 seconds) and *long sound average* (5 minutes). The data is automatically queried every one minute, time stamped, and stored in an SQL database.

Figure 2: Sensor node

Figure 3: Application used to collect occupancy ground truth

The sensor node is installed in a multi-occupancy lab in an educational building at the University of Southern California. The lab has an area of about 40 m^2,

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

and is shared by 5 PhD students. The lab hosts meetings at times, which can involve up to 10 attendees. In order to collect the ground truth occupancy information, a touch-screen Android phone is mounted close to the door (Figure 3), which displays a web application that is developed to collect the ground truth occupancy information. During the test period, all occupants are told to log in or out when they enter or leave the lab, by clicking "plus" or "minus" button of the web application on the phone screen. The data is then sent to and stored in the same database where the sensor data resides.

The sensor data was collected for 15 consecutive days, starting from 00:00 AM, Oct. 13th to 00:00 AM, Oct. 27th. At a one-minute sampling rate, after excluding all corrupted data points due to wireless connection breaks, a total of 20,156 data points were collected, of which the first 10,078 data points were used for training the BP network model, and the rest 10,078 data points were used for validation.

TEST RESULTS

Principal Component Analysis

All 13 variables reported by the sensor node are used for PCA processing. Each variable has 13 weights to indicate its influence on other twelve variables and itself. The numbers of the feature vector are averaged to a value for evaluating the influence on occupancy output. Further orthogonal linear transformation is done to convert the 13 variables to the relative optimized number of principal variables. Each variable's influential effect from weighting calculation is illustrated in Figure 4.

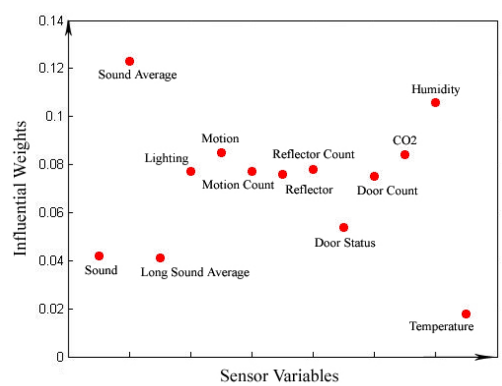

Figure 4: Variable influential weights

After the PCA processing, it is found that *sound average* has the highest priority followed by the *humidity*, *motion*, and CO_2 *concentration*. S*ound, long sound average, door status* and *temperature* have the lowest priority. After trying different number of variables in the BP ANN, it is found that using the first nine variables based on the influential weights yielded the highest accuracy. Therefore, s*ound, long sound average, door status* and *temperature* are excluded in the BP ANN process in order to increase the calculation speed and improve the quality of the results.

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

BP Single Hidden Layer Neural Network Analysis

Firstly, a one-hidden layer BP neural network is applied to calculate the relationships among the 9 input variables and the output occupancy information. The S function $f(net) = \frac{1}{1+e^{-net}} d$ is used as the activation function as it has the ability of non-linear amplifying coefficient and approximating complicated non-linear input-output relationships. BP network calculation exits the loop when the network output error decreases to a predefined threshold, set as 0.00004 in the test, or the number of iterations reaches the predefined maximum value, set as 100 in this test.

Two parameters are defined to evaluate the results. The first one is error covariance, which shows the standard deviation of the results. The error distance measures how much the estimated occupancy deviates from the actual occupancy. The second parameter is the error rate, which shows the accuracy of all validated data, i.e. the percentage of the data points where the estimated occupancy matches the actual occupancy. The concept of tolerance is also introduced that measures the tolerated error between estimated and actual occupancy. Tolerance is necessary in that for the purpose of driving HVAC systems, a small error can be acceptable, and the HVAC systems do not need to be adjusted every time the occupancy slightly changes. Due to the initialization random feature, five calculations were carried out and the best one was chosen for analysis. When tolerance=1, the result yielded an error covariance of 1.2817 and error rate of 9.18%, or an accuracy of 90.82%. When tolerance=0 and 2, the error rates were 37.59% and 2.86%, respectively. To better visualize the different between the estimated output (rounded) and the ground truth, they are both plotted in Figure 5.

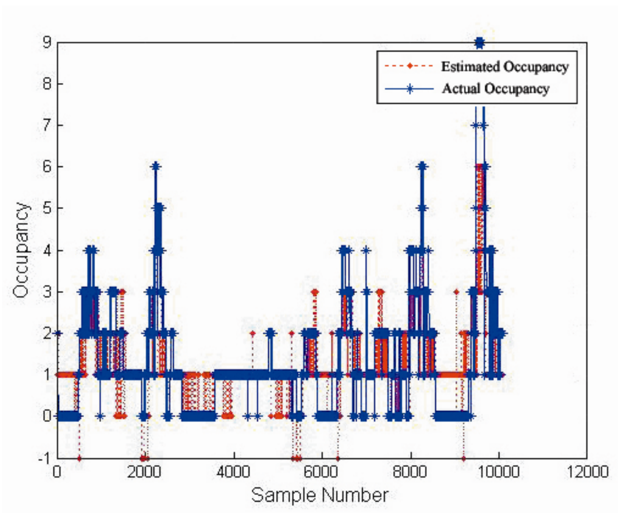

Figure 5: Estimated and actual occupancy with single hidden layer

BP Multi Hidden Layer Neural Network Analysis

Three hidden layers are included in the solution model as multiple hidden layers are more applicable for solving complicated non-linear relationship and

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

achieving more accurate result. However, inserting too many hidden layers may lead to huge computing load and increase calculation time significantly. Therefore, a balance is required and after a trial-and-error process, the solution with three hidden layers was found to be the most suitable for occupancy estimation in that it could generate satisfying result and maintain a relative high processing speed.

In the multi-hidden layer BP neural network, a higher accuracy was achieved. When tolerance=1, the error covariance was 1.1937 and error rate was 8.25%, i.e. estimation accuracy of 91.75%. When tolerance=0 and 2, the error rates were 35.23% and 2.03%, respectively. To better visualize the difference between the estimated output (rounded) and the ground truth data, they are both plotted in Figure 6.

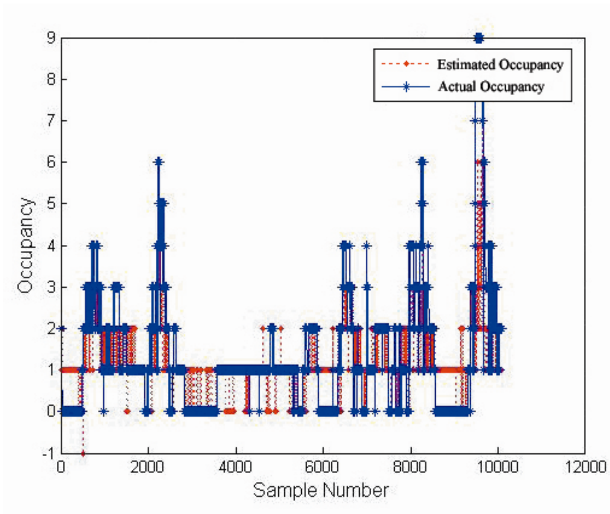

Figure 6: Estimated and actual occupancy with multiple hidden layer

DISCUSSION AND CONCLUSIONS

The PCA processing results show that *sound average, relative humidity, motion* and *CO_2 concentration* are the most influential variables in estimating occupancy at least for environments similar to the multi-occupancy lab used as the test bed in this study. Since the only difference between *sound, sound average and long sound average* is the sampling rate, the difference in their influential weights suggests that a proper sampling rate is important in making the best use of certain sensor data. It's interesting that relative humidity has a high priority, while temperature has so low a priority that it is discarded by the PCA processing. One possible explanation lies in the fact that the temperature in the test bed is dominated by the HVAC systems, which runs at the peak demand assumption and starts cooling the room as long as the room temperature reaches its set point. At the meanwhile, the humidity is less influenced by the HVAC system, and since the room has no direct access to the outdoor environment and is therefore not susceptible to the weather, occupancy is the only remaining factor that can influence the humidity. Motion and CO_2 concentrations are found to be influential. This is consistent with previous research that uses motion and CO_2 sensors as major indicators of occupancy. The analysis of the influence of each sensor identifies the sensors that are most closely

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

related to the occupancy, and the ones that are less related and negligible. This helps to accelerate the BP network processing, and improves the quality of the occupancy estimation. More importantly, knowing which sensor data is more influential provides suggestions for estimation.

The test result shows that the proposed system can estimate the exact number of occupants (tolerance=0) 62.41% of the time. This accuracy bears potential for improvement, and is partly because the output of BP network after the PCA processing is given in a decimal format, which needs to be rounded to compare with the ground truth. The rounding process causes additional errors. However, when the tolerance increases, the accuracy increases rapidly to over 90%, as the error caused by the rounding process is all offset by the tolerance. The associated practical hint is that, when used for demand driven HVAC operations, a certain level of error is fairly acceptable, as the HVAC systems do not need to be so sensitive that they respond to any slight changes in occupancy. Instead, adding or subtracting one or two occupants in a room shouldn't cause significant changes in HVAC operations, unless the room switches form unoccupied to occupied or vice versa. Therefore, it can be concluded that the proposed system can provide sufficiently accurate estimates of occupancy, under reasonable tolerance, to support demand driven HVAC operations. As previous research has proven that implementing demand driven HVAC operations could save up to 56% of HVAC related energy consumption (Sun et al. 2011), or 27% of total building energy consumption, the use of the proposed system is expected to make an important contribution to building energy savings.

Furthermore, the proposed system is low-cost and high-resolution. The sensor node prototype cost about $230 USD, and will be even lower if mass-produced. As respect to resolution, the proposed system can provide the occupancy information at the room level, and indicate the exact number of occupants, which can all be done instantly upon users' request.

ACKNOWLEDGEMENT

Authors would like to acknowledge the support of the Department of Energy through the grant # DE-EE0004019. Any opinions and findings presented in this paper are those of the authors and do not necessarily reflect the views of the Department of Energy.

REFERENCES

Agarwal, Y., Balaji, B., Gupta, R., Lyles, J., Wei, M., Weng, T. (2010). "Occupancy-driven energy management for smart building automation." *Proc., 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings, BuildSys'10*, November 2, 2010, Association for Computing Machinery, 1-6.

Alpaydın, E. (2010). Introduction to Machine Learning 2nd edition Ed., MIT Press, Cambridge, Mass.

Benezeth, Y., Laurent, H., Emile, B., Rosenberger, C. (2011). "Towards a Sensor for Detecting Human Presence and Characterizing Activity." *Energy Build.*, 43(2-3), 305-314.

Downloaded from ascelibrary.org by Cardiff University on 04/27/22. Copyright ASCE. For personal use only; all rights reserved.

DOE. (2011). "Building energy data book." <http://buildingsdatabook.eren.doe.gov/> (Nov. 10, 2011).

Dong, B., Andrews, B., Khee, P. L., Hoynck, M., Zhang, R., Chiou, Y.S., Benitez, D. (2010). "An Information Technology Enabled Sustainability Test-Bed (ITEST) for Occupancy Detection through an Environmental Sensing Network." *Energy Build*., 42(7), 1038-46.

Erickson, V. L., and Cerpa, A. E. (2010). "Occupancy based demand response HVAC control strategy." *Proc., 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings, BuildSys'10*, November 2, 2010, Association for Computing Machinery, 7-12.

Erickson, V. L., Lin, Y., Kamthe, A., Brahme, R., Surana, A., Cerpa, A. E., Sohn, M. D., Narayanan, S. (2009). "Energy efficient building environment control strategies using real-time occupancy measurements." *Proc., 1st ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings, BUILDSYS 2009, in Conjunction with ACM SenSys 2009*, November 3, 2009 , Association for Computing Machinery, 19-24.

Henze, G. P., Dodier, R. H., Tiller, D. K., Guo, X. (2006). "Building Occupancy Detection through Sensor Belief Networks." *Energy Build*., 38(9), 1033-43.

Leephakpreeda, T., Thitipatanapong, R., Grittiyachot, T., Yungchareon, V. (2001). "Occupancy-Based Control of Indoor Air Ventilation: A Theoretical and Experimental Study." *Science Asia*, 27(4), 279-84.

Melfi, R., Rosenblum, B., Nordman, B., Christensen, K. (2011). "Measuring building occupancy using existing network infrastructure." *Proc.*, *2011 International Green Computing Conference, IGCC 2011*, July 25-28, 2011, IEEE Computer Society.

Meyn, S., Surana, A., Lin, Y., Oggianu, S. M., Narayanan, S., Frewen, T. A. (2009). "A sensor-utility-network method for estimation of occupancy in buildings." *Proc., 48th IEEE Conference on Decision and Control Held Jointly with 2009 28th Chinese Control Conference, CDC/CCC 2009*, December 15-18, 2009, IEEE, 1494-1500.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). "Learning representations by back-propagating errors". *Nature,* 323 (6088): 533–536.

Shelley, I., and Roessner, H. (2004). "RFID tagging." <http://www.global-id-magazine.com/0904/140904.pdf> (Dec. 10, 2009).

Sun, Z., Wang, S., Ma, Z. (2011). "In-Situ Implementation and Validation of a CO_2-Based Adaptive Demand-Controlled Ventilation Strategy in a Multi-Zone Office Building." *Build. Environ*., 46(1), 124-133.

Tachwali, Y., Refai, H., Fagan, J. E. (2007). "Minimizing HVAC energy consumption using a wireless sensor network." *Proc.*, *33rd Annual Conference of the IEEE Industrial Electronics*, IEEE, 439-44.

Wang, H., Jia, Q., Song, C., Yuan, R., Guan, X. (2010). "Estimation of occupancy level in indoor environment based on heterogeneous information fusion." *Proc., 49th IEEE Conference on Decision and Control, CDC 2010*, December 15-17, 2010, IEEE, 5086-5091.