Title

Anomaly Detection in IoT-Based PIR Occupancy Sensors to Improve Building Energy Efficiency

Authors

Samani, Ehsan
Khaledian, Parviz
Aligholian, Armin
et al.

Publication Date

2020-02-01

Peer reviewed

Anomaly Detection in IoT-Based PIR Occupancy Sensors to Improve Building Energy Efficiency

Ehsan Samani†, Parviz Khaledian†, Armin Aligholian†,
Evangelos Papalexakis‡, Shawn Cun*, Masoud H. Nazari*§, and Hamed Mohsenian-Rad†
†Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA
‡Department of Computer Science and Engineering, University of California, Riverside, CA, USA
*Department of Electrical Engineering, California State University, Long Beach, CA, USA
§Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI, USA.

***Abstract*— In this paper, we study the real-world data streams from hundreds of digital passive infrared (PIR) occupancy sensors that are integrated into LED lighting fixtures in a recent Internet-of-Things (IoT) Building Energy Management System (BEMS) deployment in a large building in California. We first develop a data-driven method to detect anomalies in these data streams. We then use the results to enhance energy efficiency in the building and also open up opportunities to offer demand response services. In addition, we provide load forecasting for the lighting load in this building using a deep neural network architecture with high accuracy. We show that our approach can result in about 30% load reduction across lighting fixtures.**

Keywords: Internet-of-things, building energy efficiency, anomaly detection, load forecasting, demand response, deep learning.

I. Introduction

Energy demand in buildings currently accounts for 40% of the total U.S. energy consumption [1]. This calls for efforts to make buildings more energy-efficient. In this regard, smart buildings are receiving growing attention with the integration of building energy management systems (BEMS) and the proliferation of Internet-of-Things (IoT) [2]–[4].

An IoT-based BEMS may include hundreds of IoT devices, such as sensors, actuators, and communications nodes. These IoT devices monitor and control various load components, such as lighting, heating, ventilating, and air conditioning (HVAC), and plug-in loads. The IoT sensors produce a huge amount of data streams, which can provide new opportunities to enhance energy efficiency in buildings.

In this paper, we analyze the *real-world* data streams that come from a recent IoT-based BEMS deployment in a large-sacle 101,670 sqft academic building at California State University, Long Beach with over 1000 IoT devices, which provide high granular monitoring and control capabilities for lighting, plug-in loads, and HVAC loads. Specifically, we look into the data from hundreds of digital passive infrared (PIR) occupancy sensors that are integrated into each lighting fixture in this building [5]. All lighting fixtures have LED lights as well as integrated wireless communications capabilities. Note that, each room is equipped with tens of such IoT-based PIR occupancy sensors, which provide us with the occupancy status of each covered area within the room.

This work is supported in part by CEC grant EPIC-16-033, UCOP grant LFR-18-548175, and NSF grant 1462530. The corresponding author is Hamed Mohsenian-Rad. E-mail: hamed@ece.ucr.edu.

Our goal in this paper is to detect anomalies in such real-world data streams from PIR sensors and to subsequently use the results to enhance energy efficiency in the building and open up opportunities to offer demand response services.

A. Literature Review

There are few studies that have addressed the challenges related to anomaly detection in data streams from IoT devices in smart buildings. In [6] a new pattern-based anomaly classifier, the collective contextual anomaly detection using sliding window (CCAD-SW) is proposed to identify anomalous consumption patterns. In [7], [8], anomaly detection based on methods such as fuzzy linguistic description and nearest neighbor clustering is used to improve state-awareness and the understandability of BEMS data. In [9], a rule-based method is presented to detect energy inefficiencies in smart buildings. In [10], the design and implementation of a presence sensor platform is discussed that can be used for accurate occupancy detection at the level of individual rooms. There are also some papers, such as [11]–[13] that address the broad topic of energy efficiency issues in smart buildings, and some other papers, such as in [14]–[16], that address energy consumption prediction in smart buildings. All of the above papers are one way or another related to this study; however, none of the previous papers have addressed anomaly detection in IoT-based lighting-fixture-integrated PIR occupancy sensors; and application to energy saving and demand response. Moreover, most prior studies are not based on real-world data, as opposed to this paper that is fundamentally a data-driven study built upon large volume of real-world data points.

B. Summary of Contributions

The contributions in this paper are summarized as follows:

1) A two-step algorithm is proposed to find anomalies in occupancy data. In the first step, a factor showing the reliability level of each IoT lighting sensor is defined by using historical data. In the second step, real-time data is analyzed to find possible anomalies, which result in energy loss due to incorrect lighting system operation.
2) The application of the proposed two-step anomaly detection is presented for energy saving in smart buildings. Based on the forecasted amount of such energy saving

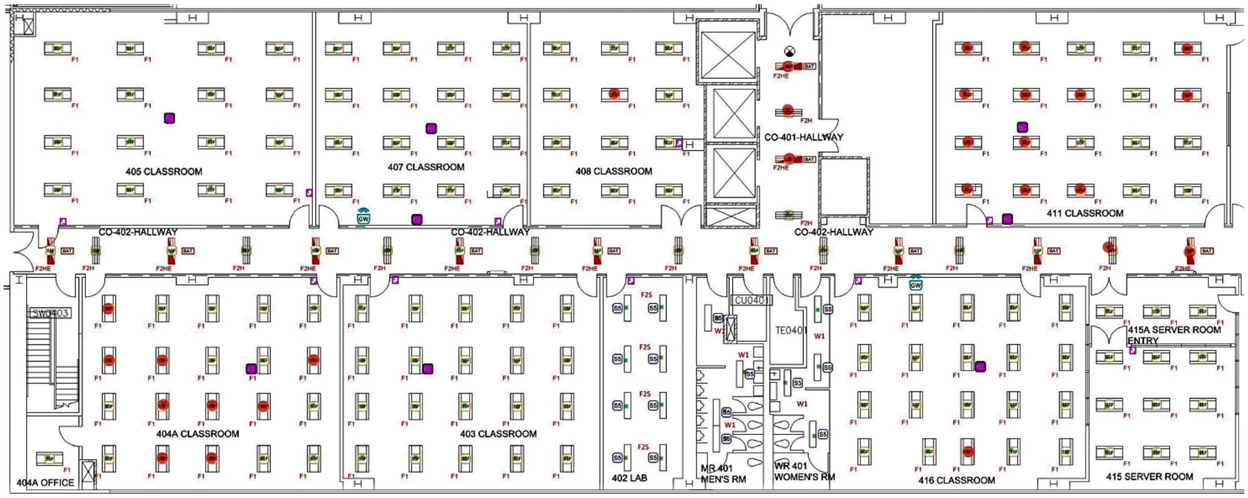

Fig. 1. The layout of the fourth floor at the test site and the locations of the lighting fixtures and their integrated PIR occupancy sensors [17].

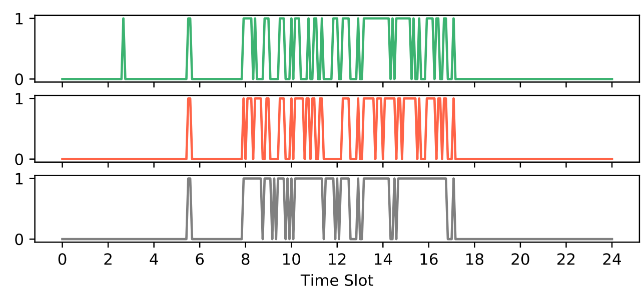

Fig. 2. An example daily output for three PIR sensors in room 408.

under the proposed method, the under-study building can efficiently participate in demand response programs.

3) The forecasting component is built upon a deep neural network architecture based on long short-term memory (LSTM). We show that this deep learning algorithm can forecast the power consumption with high accuracy.

4) The analysis in this paper is based on real-world data from hundreds of PIR sensors at the test site. A total of 2,088,170 data points are analyzed in this paper. Our estimated energy saving in the lighting system is 30%.

II. Problem Statement

The focus in this section is on explaining the problem we are facing in a real-world R&D project. The test site for this project is a large six-story academic building with hundreds of LED lighting fixtures. Each lighting fixture is equipped with a PIR occupancy sensor. As an example, the layout of the 4th floor and the location of the lighting fixtures are shown in Fig. 1. It should be noted that this building is equipped with three types of IoT devices, lighting, HVAC and plug-in load control; however, only the lighting system is the focus of this paper.

As an example regarding the type of data that is available for this study, Fig. 2 shows the daily output for one day for three PIR occupancy sensors in room 408. Note that, this room has 12 lighting fixtures, which provides 12 separate data streams, one for each PIR occupancy sensor. The reporting interval of each sensor is 5 minutes. Therefore, each sensor provides 288 data points per day. Each reading is either 1, indicating "Occupied", or 0, indicating "Not Occupied".

Currently, the lighting control system is set to work as follows: In each room, if any of the PIR sensors at any lighting fixture within that room detects occupancy, lights automatically turn on. Also, lights automatically turn off if none of the PIR sensors at any lighting fixture within that room detect occupancy for a duration of 5 minutes.

Given the above setup and the availability of the real-world data streams, we seek to answer the following questions: 1) Is the data coming from each PIR sensor at each lighting fixture reliable? For example, could it be that there is no one in a room, yet one or more of the PIR sensors incorrectly detect occupancy? 2) If the answer to the first question is Yes, then how can we detect such anomaly? 3) How should we take action on such finding, i.e., how should we incorporate such indication into the existing lighting control system? 4) How can we enhance the demand response capability of the lighting loads in this process? Note that, in general, anomalies may have different causes, such as sensor failure, improper setting of sensors, communications issues, or even cyber-attacks.

III. Proposed Methodology

A. Anomaly Detection

For each room, let m denote the number of PIR sensors. At each reading interval t, let $\delta[t,i]$ denote the reading of PIR sensor i, where $i = 1, \ldots, m$. The number of triggered sensors, i.e., those that return 1 as their output, is obtained as

$$N[t] = \sum_{i=1}^{m} \delta[t,i]. \tag{1}$$

First, we consider the output of the PIR sensors as *suspicious* if $N[t]$ is smaller than a certain threshold N_{th}. In particular, based on our experience in manually investigating the data streams in this project, we set the threshold to be $N_{\text{th}} = 2$, i.e., when only one or at most two PIR sensors detect occupancy. Other values could also be considered for this threshold.

Next, we define a reliability index for each PIR sensor i, as

$$S[i,t] = \sum_{\tau=0}^{T} b[i,t-\tau], \tag{2}$$

where

$$b[i,t] = \mathbb{I}\left(\delta[i,t] = 1 \mid N \leq N_{\text{th}}\right). \tag{3}$$

Note that, $b[i,t]$ indicates whether the PIR sensor i was a *cause* of the suspicious observation in the readings of the PIR sensors in this room at time interval t. As for the reliability index $S[i,t]$, it indicates whether the behavior observed from PIR sensor i under such suspicious condition was *persistent* or *momentary*. Note that, T is a parameter with respect to how far back in time we would like to check the operation of the sensor in order to obtain its reliability index. This parameter is set based on the knowledge of the expert operator of the building and also the available history record of sensors, such as the information on whether there was any maintenance or if the sensor was replaced or calibrated recently.

By keeping track of $S[i,t]$, it shows the reliability of each sensor, and whether the suspicious observation $N \leq N_{\text{th}}$ should indeed be declared as anomaly. In this regard, next, we define two thresholds with respect to $S[i,t]$, namely $S_{\text{th,min}}$ and $S_{\text{th,max}}$. Specifically, on one hand, if $S[i,t] > S_{\text{th,max}}$, then the suspicious observation $N \leq N_{\text{th}}$ was *persistent* to be caused by sensor i, suggesting that it is likely an anomaly and the room is likely not occupied. On the other hand, if $S[i,t] < S_{\text{th,min}}$, then the suspicious observation $N \leq N_{\text{th}}$ was *momentary*, suggesting that it is likely an unusual but valid occupancy pattern and the room is indeed occupied. As for the third case, where the following inequalities hold: $S_{\text{th,min}} \leq S[i,t] \leq S_{\text{th,max}}$, then we must check time interval t before we make a final conclusion, as we explain next.

Suppose $\mathcal{E}$ denotes the time slots during which the room is empty in normal circumstances. For example, for a classroom, we can check the class schedule, and set $\mathcal{E}$ to include the time slots from mid-night till 4:00 AM; or any other time frame(s). If $t \in \mathcal{E}$, when we know that the building is normally empty, thus, the observation is likely to be an anomaly; otherwise, it is treated as an unusual but valid occupancy. Parameters $S_{\text{th,min}}$ and $S_{\text{th,max}}$ can be set based on experiments and historical data. We set $S_{\text{th,min}} = 1000$ and $S_{\text{th,max}} = 2000$.

The outline of the proposed anomaly detection method is shown in Fig. 3. It takes in real-time data and detects possible anomalies, going through the steps that we explained above.

B. Potential for Energy Saving

Recall from Section II that the lighting control system is set in a way that triggering even one of the PIR occupancy sensors in a room results in turning on the lights in that room. In this regard, the proposed anomaly detection method can detect faulty or highly sensitive sensors which cause an anomaly for the lighting system in the room, i.e., unnecessarily turning on the lights. Therefore, having accurate occupancy data can help in saving energy in different parts of the building. Note that, since the purpose of this paper is to enhance energy saving, we only address faults that cause unneeded energy usage.

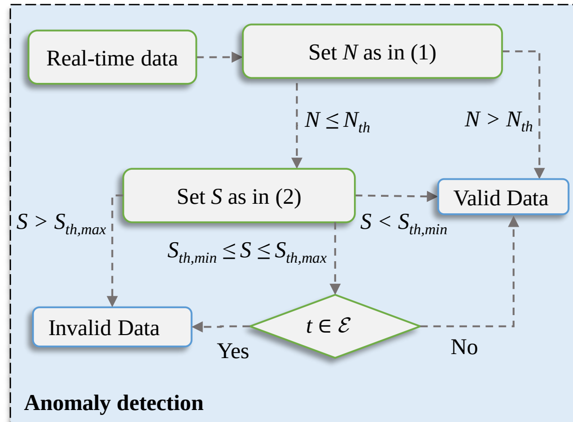

Fig. 3. Outline of the proposed anomaly detection method.

Since there is major overlap in the coverage areas among the PIR sensors in each room, there is redundancy in detecting occupancy. In fact, in our experiments, it has never happened that someone enters the room and none of the sensors pick it up. Therefore, a potential fault to miss occupancy is not a practical concern in this study and we do not address it.

Specifically, in the real-time operation of the BEMS, if an occupancy data is considered as anomaly based on the proposed method, the lights in the room should *not* be turned on. This can result in a major amount of energy saving. Consider a faulty or highly sensitive sensor in a room. This single faulty sensor can turn on the lights in a room all day and night, regardless of the operation of the rest of the occupancy sensors in that room. In fact, without considering the proposed method, certain components of smart buildings, such as certain rooms, may even result in more energy loss than conventional buildings. The difference in energy consumption between utilizing our proposed method and operating the system as is, i.e., ignoring the possible anomalies, is the amount of the energy that can saved when our method is implemented.

C. Demand Response and Energy Forecasting

The proposed anomaly detection method can be used also to create new capacities for the lighting control system to participate in demand response programs. The key is to adjust the parameters of the algorithm, i.e., N_{th}, $S_{\text{th,min}}$, $S_{\text{th,max}}$, T, and $\mathcal{E}$, under demand response operating conditions. That is, while the parameters can be set conservatively for energy saving during normal operating conditions; they can be set rather aggressively during demand response events. For example, we may set $N_{\text{th}} = 2$ during normal operating conditions, so that we check for anomaly if fewer than two PIR sensors are triggered. During a demand response event, we may change this to $N_{\text{th}} = 3$, so that we turn off the lights more aggressively when we observe potential faulty sensors. This opens up addition load reduction capacities that can we used during the demand response events. Such adjustments can be done also based on the location of the sensor, such as whether it is close to a door or a window, again based on the knowledge and experience of an expert operator for the understudy building.

Consider a "Basic Plan", which does *not* use the proposed anomaly detection method. Based on the actual historical data

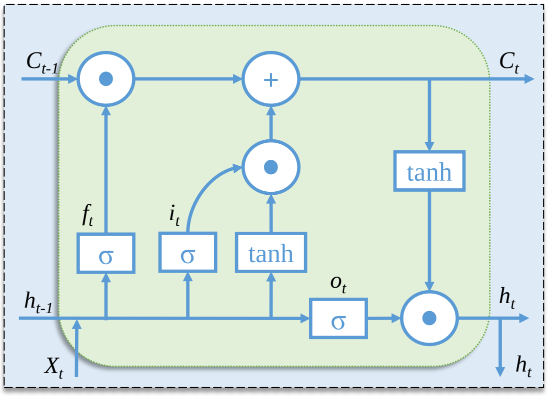

Fig. 4. Architecture of an LSTM cell that is used for load forecasting.

from the lighting fixtures, we train an LSTM model, which is a Recurrent Neural Network (RNN) [18], see Fig. 4. The trained LSTM model can predict the day-ahead energy usage of the lighting fixtures based on historical data. Next, consider a "Demand Response Plan" which uses the proposed anomaly detection method for any given choice of parameters. This time, we apply the anomaly detection method to historical data and train a second LSTM model to predict the day-ahead energy usage when the anomaly detection method is utilized.

Given the two prediction models, the *difference* between the Basic Plan and the Demand Response Plan is calculated to obtain the overall *predicted demand response capacity*. Such prediction is then reported to the demand response aggregator, as the amount of load reduction that we expect to be able to provide, in case a demand response event occurs.

Once the demand response mode is activated, the building operation is set such that the lights in each room do not turn on when *both* of the following conditions happens: 1) an anomaly is detected for the PIR occupancy sensors in that room; and 2) a demand response event occurs. In other words, we utilize the potential for energy reduction only during the demand response events. Again, the two prediction models are used in order to estimate the amount of available energy reduction, i.e., the demand response capacity, which is needed in order to participate in most practical demand response programs [19].

IV. Case Studies

In this section, we evaluate the performance of the proposed anomaly detection method using real-world data. It should be noted that while the data is real, the calculation of the energy saving is done numerically. The data contains the motion detection output of each sensor and the energy usage of its fixture. The analyzed historical data is for 45 days. We focus on the 4th floor of this building and it should be mentioned that corridors are excluded in this analysis due to safety.

A. Anomaly Detection

Based on the proposed anomaly detection method, in the first step, we determine the reliability level for each sensor in each room. As an example, Fig. 5 and Fig. 6 depict the number of single and double detections, respectively, in room 408. As we can see, sensor 72A902 with 1313 single detections

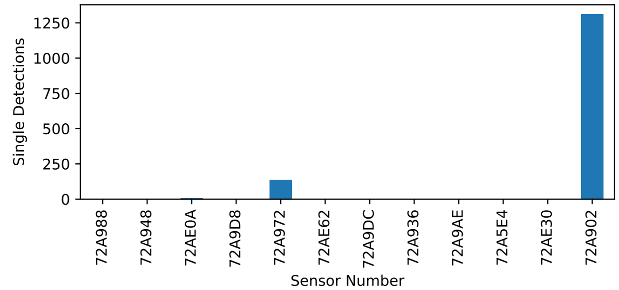

Fig. 5. Number of single detections for each sensor in room 408.

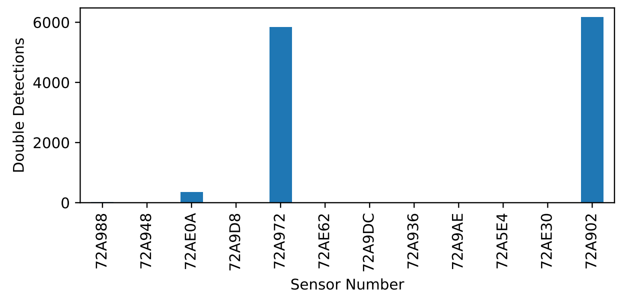

Fig. 6. Number of double detections for each sensor in room 408.

and 6169 double detections and also, sensor 72A972 with 138 single detections and 5845 double detections over 12970 time slots are two unreliable sensors in this room. Any suspicious detection with $N \leq 2$ corresponding to these two sensors will be considered as anomaly and the reported data is invalid. Note that, these two sensors are not close to each other.

B. Energy Usage Forecasting

Based on what explained in the previous section, valid and invalid occupancy detections will be separated in the historical data. For the invalid occupancy data, as we determined that room as unoccupied, the energy consumption of that room at that time will be considered as zero. Accordingly, there are two time series for energy consumption. One with modifying the data based on the proposed anomaly detection method and the other one without utilizing our proposed method. The amount of energy usage *with* and *without* employing the proposed method in the under-study period is 1.4626 MWh and 2.0633 MWh, respectively. Therefore, by utilizing this method energy usage reduces by about 30%, which is a significant amount.

In the forecasting part, we utilize LSTM to train a model for the power consumption time series. As mentioned before, this is a day-ahead forecasting and in order to forecast the power consumption of each time step, the model utilize the data for the days in the previous week. Note that as we are working on an academic building, only weekdays are taken into consideration. In order to train the model, the historical data is split into training and testing data sets. The first 80% of the data is used for training and the last 20% for testing the model. Fig. 7 shows the forecasted and the actual power consumption in the test data by utilizing the proposed method. The accuracy of this prediction model for the training and testing datasets is 85% and 84%, respectively.

Fig. 8 shows the forecasted and actual power consumption in the test data for the unmodified data. The accuracy of

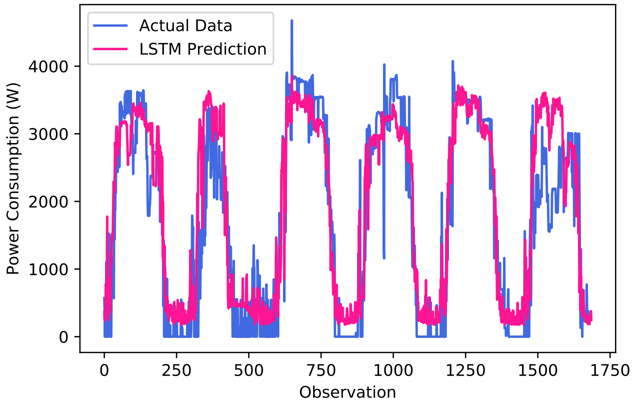

Fig. 7. Actual and forecasted load with utilizing our proposed method.

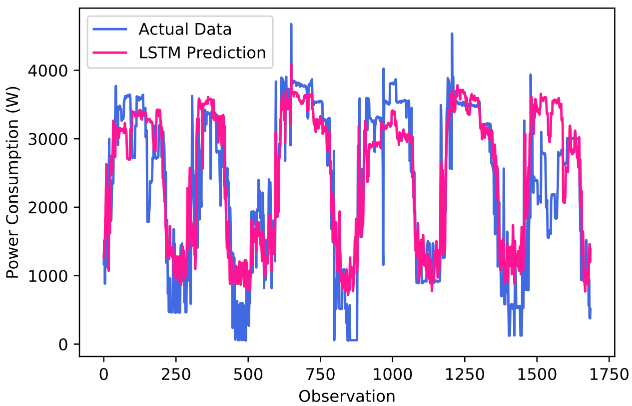

Fig. 8. Actual and forecasted load without utilizing our proposed method.

this model for the training and testing datasets is 73% and 72%, respectively. By comparing Fig. 7 and Fig. 8 , we can identify a biased power consumption in Fig. 8. This constant consumption is because of malfunction occupancy sensors, which result in unnecessary power consumption even at nights. Utilizing the proposed day-ahead forecasting model, we can forecast next day's power consumption and see how much it can save energy. Based on the day-ahead forecasted amount of energy saving, buildings can efficiently participate in demand response programs by having different override plans. Each plan will be constructed based on the proposed anomaly detection method by changing different parameters.

V. Conclusions

This paper established an anomaly detection method for occupancy data from PIR sensors in IoT-based lighting systems, with application to building energy efficiency. First, based on historical data, suspicious sensors were identified at each room or zone. This identification was based on occupancy detections which were out of normal expectation. Next, real-time occupancy data were analyzed to distinguish between valid and invalid data. We analyzed the lighting system in a large academic building in California, which is equipped with such IoT-based network of PIR sensors. Our analysis shows that utilizing our proposed method can reduce energy consumption by about 30% in this building. By utilizing LSTM as a deep neural network architecture, the day-ahead energy consumption was forecasted so that the identified energy consumption reduction can be used to offer demand response. The study in this paper can be extended in various directions. In particular, our analysis can be done also on the HVAC and plug-in load controllers and the effect of anomalies in occupancy data on these systems can be investigated.

References

[1] X. Cao, X. Dai, and J. Liu, "Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade," *Energy and buildings*, vol. 128, pp. 198–213, 2016.

[2] R. Jia, B. Jin, M. Jin, Y. Zhou, I. C. Konstantakopoulos, H. Zou, J. Kim, D. Li, W. Gu, R. Arghandeh, *et al.*, "Design automation for smart building systems," *Proceedings of the IEEE*, vol. 106, no. 9, pp. 1680–1699, 2018.

[3] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE communications surveys & tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[4] D. Amaxilatis, O. Akrivopoulos, G. Mylonas, and I. Chatzigiannakis, "An iot-based solution for monitoring a fleet of educational buildings focusing on energy efficiency," *Sensors*, vol. 17, no. 10, p. 2296, 2017.

[5] 2019. [Online]. Available: http://info.enlightedinc.com/rs/000-IKN-871/images/CSSU5E_SS_93-02053-01_Rev01_0428.pdf

[6] D. B. Araya, K. Grolinger, H. F. ElYamany, M. A. Capretz, and G. Bitsuamlak, "An ensemble learning framework for anomaly detection in building energy consumption," *Energy and Buildings*, vol. 144, pp. 191–206, 2017.

[7] D. Wijayasekara, O. Linda, M. Manic, and C. Rieger, "Mining building energy management system data using fuzzy anomaly detection and linguistic descriptions," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 3, pp. 1829–1840, 2014.

[8] O. Linda, D. Wijayasekara, M. Manic, and C. Rieger, "Computational intelligence based anomaly detection for building energy management systems," in *Proc. of the International Symposium on Resilient Control Systems*, Salt Lake City, UT, Aug. 2012.

[9] M. Peña, F. Biscarri, J. I. Guerrero, I. Monedero, and C. León, "Rule-based system to detect energy efficiency anomalies in smart buildings, a data mining approach," *Expert Systems with Applications*, vol. 56, pp. 242–255, 2016.

[10] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, and T. Weng, "Occupancy-driven energy management for smart building automation," in *Prof. of the ACM workshop on embedded sensing systems for energy-efficiency in building*, 2010.

[11] A. Allouhi, Y. El Fouih, T. Kousksou, A. Jamil, Y. Zeraouli, and Y. Mourad, "Energy consumption and efficiency in buildings: current status and future trends," *Journal of Cleaner production*, vol. 109, pp. 118–130, 2015.

[12] S. Ahmadi-Karvigh, A. Ghahramani, B. Becerik-Gerber, and L. Soibelman, "Real-time activity recognition for energy efficiency in buildings," *Applied energy*, vol. 211, pp. 146–160, 2018.

[13] B. Tan, Y. Yavuz, E. N. Otay, and E. Çamlıbel, "Optimal selection of energy efficiency measures for energy sustainability of existing buildings," *Computers & Operations Research*, vol. 66, pp. 258–271, 2016.

[14] P. A. Gonzalez and J. M. Zamarreno, "Prediction of hourly energy consumption in buildings based on a feedback artificial neural network," *Energy and buildings*, vol. 37, no. 6, pp. 595–601, 2005.

[15] J. Yang, H. Rivard, and R. Zmeureanu, "On-line building energy prediction using adaptive artificial neural networks," *Energy and buildings*, vol. 37, no. 12, pp. 1250–1259, 2005.

[16] M. Castelli, L. Trujillo, L. Vanneschi, and A. Popovič, "Prediction of energy performance of residential buildings: A genetic programming approach," *Energy and Buildings*, vol. 102, pp. 67–74, 2015.

[17] D. H. Tran, M. H. Nazari, A. Sadeghi-Mobarakeh, and H. Mohsenian-Rad, "Smart building design: a framework for optimal placement of smart sensors and actuators," in *Proc. of IEEE PES ISGT*, 2019.

[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[19] H. Mohsenian-Rad and A. Leon-Garcia, "Optimal residential load control with price prediction in real-time electricity pricing environments," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 120–133, 2010.