

Yin, X., and Keoh, S. L. (2017) Personalized Ambience: An Integration of Learning Model and Intelligent Lighting Control. In: IEEE World Forum on Internet of Things (WF-IoT), Reston, VA, USA, 12-14 Dec 2016, pp. 666-671. ISBN 9781509041305 (doi:10.1109/WF-IoT.2016.7845398)

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/130025/

Deposited on: 12 October 2016

Enlighten – Research publications by members of the University of Glasgow http://eprints.gla.ac.uk

Personalized Ambience: An Integration of Learning Model and Intelligent Lighting Control

Xiangwei Yin and Sye Loong Keoh School of Computing Science, University of Glasgow Glasgow G12 8QQ, United Kingdom

Email: 2167861y@student.glasgow.ac.uk, SyeLoong.Keoh@glasgow.ac.uk

Abstract—The number of households and offices adopting automation system is on the rise. Although devices and actuators can be controlled through wireless transmission, they are mostly static with preset schedules, or at different times it requires human intervention. This paper presents a smart ambience system that analyzes the user's lighting habits, taking into account different environmental context variables and user needs in order to automatically learn about the user's preferences and automate the room ambience dynamically. Context information is obtained from Yahoo Weather and environmental data pertaining to the room is collected via Cubesensors to study the user's lighting habits. We employs a learning model known as the Reduced Error Prune Tree (REPTree) to analyze the users' preferences, and subsequently predicts the preferred lighting condition to be actuated in real time through Philips Hue. The system is able to ensure the user's comfort at all time by performing a closed feedback control loop which checks and maintains a suitable lighting ambience at optimal level.

I. Introduction

The rapid advancement in IoT technology offers multiple opportunities to improve our daily lives. Home and office automation used to be a science fiction which modelled an ideal standard of living. With technology advancement in IoT, 1.5 million homes in the United States of America have adopted home automation technology and this figure is expected to rise in 2016 [16], [17]. The upsurge in the number of home/office automation anticipates higher demand for diversified software that caters to the needs of different users. This has led to an emergence of different environments, mainly a healthier indoor environment or a sick building [15]. The differentiating factors include air ventilation measured in volatile organic compound (VOC), the temperature of the room measured in degree Celsius and the lighting condition measured in lux, and by performing continuous monitoring to ensure that these factors are at their optimal level, the user's comfort can be guaranteed.

Deploying a home/office automation system will raise the standard of living through creation of a conducive environment for users. Firstly, an ideal lighting in workplaces mitigates the risk of health issues such as blurry vision and dizziness due to insufficient light, and it effectively boosts the user's productivity [9]. Secondly, the use of home/office automation allows for an efficient use of energy [11], [7] as it provides a convenient way of controlling the connected devices, such as lighting, termostat, and air-conditioner. This resolves the issue of energy wastage due to unutilized appliances caused by users' negligence. Thirdly, the user's comfort level can

be further enhanced if the system is able to automatically recognize the user's habits, activities and personal preferences. This would allow the system to interact with the connected devices and actuators in the environment seamlessly without having to request input from the users.

However, the comfort level across different users varies and does not comply with the ideal scenario of a one-size-fits-all. This is mainly because the acceptance level of each user for a comfortable environment may vary with differences in each individual's preference and tolerance level. Therefore, this may cause a contradiction in the home/office automation software.

This paper proposes an integration of machine learning and IoT technology to realize personalized ambient intelligent system. It emphasizes on the aspect of room illumination which allows adjustment of lighting according to the user's preferences and the context information obtained from the surrounding in a dynamic manner. Insights of indoor environment status are collected by deploying wireless sensors (CubeSensors) [2] in the room, while other contributing factors such as day-lighting and weather conditions are collected through Yahoo Weather. The sensor readings of environment, and weather data are processed, evaluated by the home/office automation software using a machine learning approach, i.e. REPTree [22] in order to learn about the user's preferences and habits. The system can be trained to differentiate the lighting needs of users through profiling. With a feedback-control loop, the system performs monitoring and adjusts the room setting in order to achieve an optimal level deemed conducive for an individual user in real-time.

This paper is organized as follows: Section II presents the architectural overview of Personalized Ambience system, while Section III details the prototype implemented describing the REPTree learning model, and integration with IoT devices. In Section IV, we present an analysis of different learning models and the optimization of REPTree using pruning. Section V presents related work, and finally we conclude with future work in Section VI.

II. SYSTEM OVERVIEW

The overall architecture of the system is illustrated in Figure 1. The user's environment consists of several peripherals which include *Philips Hue light bulbs* [13], *CubeSensors* and a computer hosting the machine learning software. The system contains three main functions, namely *data collection, machine learning*, and *feedback control loop*. They are described in the following subsections:

978-1-5090-4130-5/16/\$31.000 ©2016 IEEE

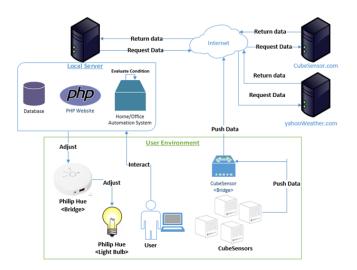


Fig. 1. Architectural Overview of Personalized Ambience

A. Data Collection

1) Cubesensor: is an indoor environment sensor system which consists of a bridge and multiple sensors (cubes). The bridge is used to connect (up to) 30 cubes simultaneously and relay sensor data to the cloud. Each cube contains several sensors which monitor the air quality, temperature, barometric pressure, humidity, noise, vibrations, and light. Cubesensors use facial emotions to convey the status of various rooms and scores are given to indicate the comfort level.

However, the collected sensor data are relayed onto *cube-sensor.com*, and the storage space provided is only limited to a period of 24 hours. This limitation causes a loss of potentially vital information, and more importantly the existing cubesensor portal only provides visualization for data that are captured but does not provide any data analysis. Therefore, we have implemented a local server to periodically poll *cubesensor.com* in order to extract all data that is relevant to the user and then store them in a local database for further analysis.

2) Weather Data: is known to affect the room ambience, as a hot sunny day will trigger the window blind to be drawn, while in a cloudy and rainny day, the users tend to switch on the lights for better illumination. Weather data such as weather description, temperature, wind temperature, wind direction, wind speed, time of sunrise and time of sunset are important metrics to determine the room ambience, so that artificial lighting can be illuminated, dimmed and adjusted to match the user's comfort needs.

The collected data are then fed into a learning model in our system, so that the system is able to learn about the user's preferences and habits at different weather conditions. Based on the learning model, we are able to perform an assessment of the environment and then predict a suitable comfort level to be actuated automatically without involving the users.

B. Machine Learning

1) Model: Machine learning allows the computer to identify complex user behaviour and resolve irregular decision-making [12], and this works well in the home/office automation context. We employs the learning model of decision tree [22]

in the proposed system. This model is constructed using a top-down approach, in which each test is specified as a form of a node in the tree structure with each branch of the node corresponding to a single value of the attribute known as the outcome of the test. The test is done in a recursive manner until the leaf node, otherwise known as the classification, is arrived. In particular, we have chosen the Very Fast Decision Tree (VFDT) [22], as it uses a white box model which can easily be interpreted and explained at each stage of the test (node). VFDT, also known as Reduced-Error Prune Tree (REPTree) is able to address missing values with C4.5's method [14] of fractional instance [10], thus making it highly efficient. The setback of the decision tree is that any small changes in the parent node may lead to a ripple effect. This will then lead to an undesirable reconstruction of the decision tree in addressing the changes made.

In the proposed Personalized Ambience system, we analyzes the users' behaviour and pattern using the this learning model to understand their needs for different types of conditions. Based on their needs, the model is used predict the amount of lighting (Lux) required by the users. Thus, the home/office environment is able to use the energy efficiently as compared to setting the light bulb at the maximum (by default).

2) Training: The proposed system uses unsupervised training as it allows the learning algorithm to further detect unseen patterns [21], [19]. The use of unsupervised learning is less trivial as compared to supervised learning because supervised learning requires a labeling process to tag the data [21], which can be very costly.

Five attributes are used to perform the training in this system, they includes: 'timestamp', light (Lux) obtained from *cubesensors*, 'weather conditions', time of 'sunset' and 'sunrise' obtained from Yahoo weather. The 'timestamp' refers to the time when the data was collected. The 'light' signifies the amount of lux required by the user. The 'weather condition' implies the weather type such as heavy rain shower, sunny, fair and cloudy. The 'sunrise' and 'sunset' act as a guideline for the system to understand the possible interferences caused by external factors such as sunlight.

These data are collected according to each user's preferences, thus creating a profile for each user.

C. Feedback Control Loop

One important aspect of the proposed Personalized Ambience system is to maintain the lighting in an acceptable state comfortable to the user. To achieve this, the system considers the time and the current lighting which are influenced by external factors such as day-light, weather conditions as well as the room lighting condition (implicitly, it is the user's preferences on the ambience in the room) to be the main factors affecting the user's comfort.

As shown in Figure 2, the *cubesensor* senses the lighting condition in the room, and external weather condition are gathered. These data are then fed into the system to calculate the desired lighting level to be set for the room based on the learned historical data. In this system, we used *Philips Hue* bulbs in order to achieve a feedback-control loop, where the

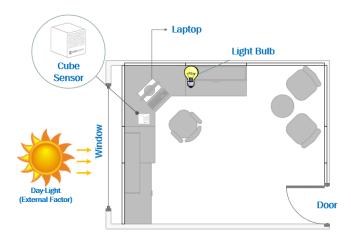


Fig. 2. Office Layout

desired lighting can be translated into a control message to the light bulb in real-time.

Let lux_x be the current room lighting which is affected by both the room illumination and the day-lighting (external factor). Let lux_o be the optimal lighting desired by the user at a particular weather condition, derived from the learning model. y is the adjustment to be made to the light bulb, such that its effect, \hat{y} in lux when added to the existing lighting condition lux_o can achieve the optimal lighting for the user.

$$lux_o = lux_x + \hat{y} \tag{1}$$

 lux_o is an output predicted by the system, and in order to achieve this optimal level, the system has to recursively change the value of y by continuously adjusting the light bulb until the optimal lighting in Lux is achieved. This step has to be executed dynamically in real-time, as the value of lux_x varies depending on the environmental factors, such as external lighting. Therefore, if the current room illumination exceeded the optimal lighting derived from the learning model due to external factors such as sunlight, the system will have to trigger the blind to be drawn in order to block the sunlight from affecting the room ambience.

D. Day vs. Night

Another important factor is that *day* and *night* context play an crucial role in the automated setting of room ambience, especially the lighting. During the day, the sunlight is perceived as a source of interference that will cause the learning model to be inaccurate. On the contrary, at night, external factors causing the interference on the lighting are non-existence, and this makes the learning for user's preferences and behaviour at night much simpler.

As shown in Figure 3, the decision process takes into account the causal factors before adjustments to the room ambient are made. The contributing factors include weather changes and the time of window blind drawn. Prior to any adjustment to the room illumination, an actuation strategy is derived to control various connected devices in the room so that the most optimal ambient for the user can be achieved.

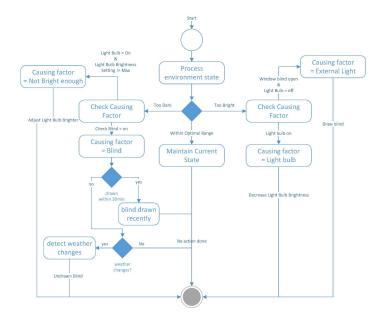


Fig. 3. Feedback Control Loop for Day Time Personalized Ambience

On the other hand, the night module is a simplified approach which only illustrates a filtering process. An adjustment of the lighting will be relayed to the *Phiilps Hue* light bulb when the system derives an outcome of either "too dark" or "too bright".

III. PROTOTYPE IMPLEMENTATION

A. Data Mining

Prior to learning about the user's habits and behaviour, we first collected contextual data from the *Cubesensor* deployed at the sites and crawled the corresponding weather information from Yahoo. Yahoo weather information was extracted in XML format, and then parsed using an API created by [4]. We also implemented a crawling service to indicate the intended location using "where on earth Id (WOEID)", and the data was then stored in a local PostgreSQL database.

As the *cubesensors* data were uploaded to *Cubesensors.com*, a data retrieval service was implemented to poll the server in order to retrieve the collected sensor data in JSON format in real time. The parsing of JSON was done using an API created by [18]. An access token was obtained from *Cubesensors.com*, so that the retrieval of sensor data can be done via OAUTH. A cron job was created and scheduled to run on OpenShift server on a per minute basis to retrieve both weather and sensor data. Similarly, an instance of this job was deployed on a local machine, serving as a hot redundancy in order to prevent a single point of failure.

The system was deployed at two locations, one at university office and the other at author's apartment, using two sets of *Cubesensors*. The purpose of deploying the cubesensors at both locations was to gather more data within a short timeframe. In addition, this approach also enables us to perform compatibility testing at both locations.

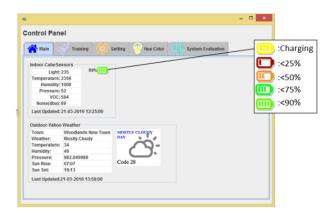


Fig. 4. Main GUI of the Personalized Ambience Application

The data obtained from the *cubesensors* is represented in a simple GUI as shown in Figure 4

B. Generation of Learning Model

The learning model was constructed using WEKA with three steps, namely (1) Data retrieval from local database, (2) Parsing of data into ARFF File, and (3) Building a regression model. The process of generating the ARFF was implemented into the system to accommodate the use of WEKA. The training dataset consisted of five attributes, namely a timestamp (Date and Time), light (Lux), sunrise (time-24hours), sunset (time-24hours) and weather description (String). These attributes have the potential of altering the amount of light (Lux) by analyzing the data collected.

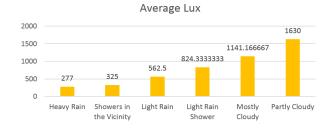


Fig. 5. Average Lux for cloudy/rainning/showers weather

As the weather condition is known to cause a drastic reduction of light, an experiment was done to investigate the lighting condition for different weather. This is especially important when the reported weather is cloudy or rainning. As shown in Figure 5, the raining condition is ranked among the lowest in a test conducted at a specific time (14:28:00).

By knowing the 'sunrise' and 'sunset' time, it aids the learning model to different between the contextual data for *day* and *night*. Once these five set of attributes data were collected, and then fed into the learning model for supervised training, a REPTree model is generated as shown in Figure 6. This model was used to predict the optimal lighting condition to be set.

C. Calibration of Philips Hue Light Bulbs

In order to adjust the night ambience effectively, it is important that the light readings (lux) obtained from the *cubesensors*

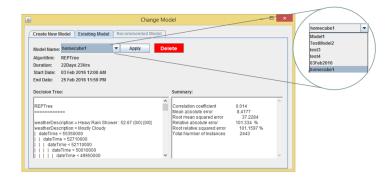


Fig. 6. An Example of Learning Model Generated

and *Philips Hue* light bulb were calibrated. This is to allow the system to recognize the relation between the amount of lux produced for the brightness setting of *Philips Hue* (Range from 0-254) in relation to the current lighting condition. The calibration was done by using a sampling approach to plot the regression line. The light bulb brightness was increased by 50 per iteration during the calibration process. Through this experiment, we derived a regression line of y = 5.632x + 397.2 as the best-fitted line as illustrated in Figure 7.

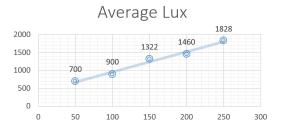


Fig. 7. Best-Fitted Line for Philips Hue calibration

D. Automation and Control

Invocation of lighting control and evaluation of the environment context were implemented using Java RMI. A recursive method was implemented to invoke the method of evaluating the users' environment status every five seconds, and if necessary the lighting condition will be adjusted accordingly based on the outcome predicted by the learning model. This process was used to maintain the comfort level of the users.

IV. EVALUATION AND RESULTS

A comparison of learning models was performed by creating a visualization of predicted light (Lux) across four different types of learning models as illustrated in Figure 8. The light (Lux) prediction of REPTree and M5P replicated a pattern similar to the actual user's behavior and preferences. Conversely, the Decision Stump was unable to predict accurately.

The detailed results of the comparison are also shown in Table I. We observed that REPTree outperformed M5P and DecisionStump in predicting the user's behaviour. Based on the error rates, the REPTree attained the lowest rates out of the three learning models.

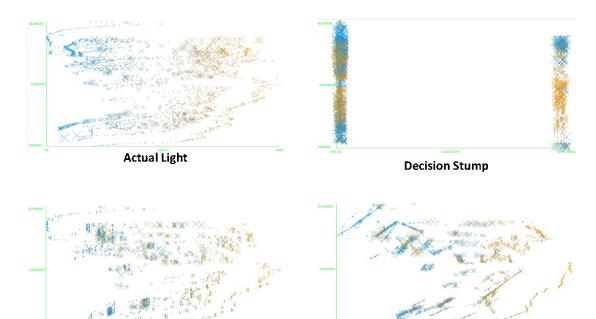


Fig. 8. Comparison of learning Model

WEKA Model	REPTree	M5P	DecisionStump
Attribute	5	2	5
Correlation coefficient	0.913	0.8871	0.325
MAE	124.2687	167.0409	415.354
RMSE	203.5367	231.7811	471.3571
RAE	27.6516 %	37.1691 %	92.4224 %
RRSE	40.8119 %	46.4753 %	94.5135 %
Time Taken To Build Model	0.1 seconds	1.21 seconds	0.01 seconds

REPTree

TABLE I. RESULTS OF USING DIFFERENT LEARNING MODELS

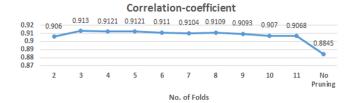
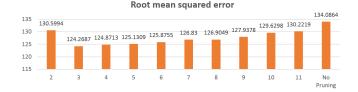


Fig. 9. Effect of Pruning

We also investigated the effect of pruning. It is observed that pruning reduces the size of the REPTree according to the number of defined folds. This approach optimizes the tree by examining the nodes and branches that can be removed without affecting the performance. The finding is illustrated in Figure 9 in which the REPTree with pruning shows an improvement of correlation coefficient compared to without pruning. The most efficient number of folds for pruning is 3 with the highest result of 0.913 and the lowest error rate of 124.2687 as shown in Figure 10.

V. RELATED WORK

In [8], the authors proposed a Smart Home Design for Disabled People based on Neural Network to improve the daily life of handicap users. It uses two types of neural network,



M₅P

Fig. 10. Pruning Effect on RMSE

namely Feed-Forward and Recurrent, for adaptive learning. These are used to predict the probability of an event causing a trigger that allows the software and the connected devices to act according to users' needs. The feed-forward neural network was used in the smart fire alarm detection system to derive a Boolean decision of a possible outbreak. The recurrent neural network is used to understand the users' habits and behaviors so as to predict the users' actions. Daily activities are recorded and processed by the recurrent neural network. This allows the recalculation of weights for each perceptron until a minimal error rate is achieved. In order to derive this prediction, the software monitors the previous actions made by the user to derive a probability of the next action. Unfortunately, it relies heavily on cameras for movement detection and a wireless microphone for users to input their desired actions.

In 2009, a smart home project proposed to use SVM to classify users' activities[5]. The objective was to monitor the health status of the elderly at home. The SVM method selected by the project to address the issue of undeterminable hyperplane is "One-Versus-One" as compared to "One-Versus-All". Similarly, an intrusion detection for smart home[1] was developed using SVM in 2015. The use of SVM was to prevent false positive due to pets' movements. The moving image was first processed with Background Subtraction (BGS) to generate

the Histogram Oriented Gradients (HOG). The use of high dimension vector SVM was shown in the project to identify the outcome of the HOG with a classification of the dataset into human or pet.

Smart building [6] was developed using fuzzy logic to monitor and control the energy usage of connected devices. The project used a proportional integral derivative (PID) controller and fuzzy based controller to determine the amount of potential energy saved. The fuzzy controller was used to evaluate the difference between a current condition and the optimal condition. This triggered an increment or decrement in the connected device to match the users' desired condition. Another work on energy efficient smart house[20] was built based on the idea of fuzzy logic as well. The fuzzy logic captures five parameters, namely price signal, battery storage, solar, humidity and temperature. The membership function depicts the degree of truth which were vaguely defined. The output generated by fuzzy logic was refined with centroid defuzzification technique to attain a value. This serves as an input for the intelligent lookup table to derive the state of energy consumption.

In 2014, the Very Fast Decision Tree (VFDT) was introduced to predict smart home lighting using sensors and user input [3]. The project's focus was to improvise the VFDT, namely VFDT++, which uses Kernel Density Estimation (KDE) and Laplace correction. It aimed to predict the lighting behavior which was affected by lifestyle and habit. A comparison of learning algorithm between VFDT, VFDT++, Naïve Bayes (NB) and Artificial Neural Network (ANN) was evaluated and accessed based on its capability of predicting the value of smart home lighting switch (On/Off).

VI. CONCLUSIONS AND FUTURE WORK

We have shown a novel integration of machine learning and IoT devices to provide personalized ambience, ensuring the user's comfort. The capability of learning the user's habit and behaviour is important to identify the different requirements and comfort needs of different users at different environmental conditions. The learned knowledge is then used to perform prediction and actuation of lighting control in real time. We have demonstrated the feasibility of such an integrated system to create a seamless IoT ecosystem, and the evaluation of the learning models has shown that REPTree is effective in predicting the users' preferences.

In the future, we plan to improve the system further by looking at the tracking of energy usage within the household/office. This can be incorporated into the existing system to allow users to understand the energy usage of the selected learning model/settings. By having an energy tracking system, the users would be able to monitor the monthly energy usage efficiently. This includes generating a monthly report to allow the users to take action against the affected appliances. In addition, an image processing technique can be implemented to recognize the users' activities to enable the system to accurate determine the desired ambience setting for the activities.

Finally, a layer of security should be added to protect the integrity and confidentiality of the system. Data transmission between the connected devices can be encrypted to prevent attackers from tampering with the data.

REFERENCES

- [1] D. Chowdhry, R. Paranjape, and P. Laforge. Smart home automation system for intrusion detection. In *Information Theory (CWIT), 2015 IEEE 14th Canadian Workshop on*, pages 75–78. IEEE, 2015.
- [2] CubeSensors. CubeSensors smart home appliances. https://cubesensors. com/. Accessed October 25, 2015.
- [3] I. Dinata and B. Hardian. Predicting Smart Home Lighting Behavior from Sensors and User Input using Very Fast Decision Tree with Kernel Density Estimation and Improved Laplace Correction. In IEEE International Conference on Advanced Computer Science and Information Systems (ICACSIS), pages 171–175, 2014.
- [4] F. D. Faveri. Yahoo Weather Java API [Online]. https://github.com/ fedy2/yahoo-weather-java-api, November 2013. Accessed Decmber 28, 2015
- [5] A. Fleury, N. Noury, and M. Vacher. Supervised classification of activities of daily living in health smart homes using svm. In *IEEE EMBC 2009*, pages 6099–6102, 2009.
- [6] Y. Y. Ghadi, M. Rasul, and M. M. K. Khan. Potential of saving energy using advanced fuzzy logic controllers in smart buildings in subtropical climates in australia. *Energy Procedia*, 61:290–293, 2014.
- [7] T. Hamm. Turn Off the Lights to Save Money...But How Much? [Online]. http://www.thesimpledollar.com/how-much-money-does-turning-off-the-lights-really-save/, August 2014. Accessed November 26, 2015.
- [8] A. Hussein, M. Adda, M. Atieh, and W. Fahs. Smart home design for disabled people based on neural networks. *Procedia Computer Science*, 37:117–126, 2014.
- [9] A. Isakovic. It just got too dark to work; didyounotice? [Online]. http://blog.cubesensors.com/2013/08/too-dark-to-work-office-lighting-tips/, August 2013. Accessed November 5, 2015.
- [10] S. Kalmegh. Analysis of weka data mining algorithm reptree, simple cart and random tree for classification of indian news. *International Journal of Innovative Science, Engineering & Technology (IJISET)*, 2(2):438–446, 2015.
- [11] O. of Energy Efficiency & Renewable Energy. When to turn off your lights [Online]. http://energy.gov/energysaver/when-turn-your-lights. Accessed November 26, 2015.
- [12] M. Pantic, A. Pentland, A. Nijholt, and T. S. Huang. Human computing and machine understanding of human behavior: a survey. In *Artifical Intelligence for Human Computing*, pages 47–71. Springer, 2007.
- [13] Philips. Philips philips hue personal wireless lighting. http://www2. meethue.com/. Accessed November 29, 2015.
- [14] J. R. Quinlan. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.
- [15] C. A. Redlich, J. Sparer, and M. R. Cullen. Sick-building syndrome. The Lancet, 349(9057):1013–1016, 1997.
- [16] A. Research. 1.5 Million Home Automation Systems Installed in the US This Year [Online]. https://www.abiresearch.com/press/ 15-million-home-automation-systems-installed-in-th/, November 2012. Accessed November 25, 2015.
- [17] T. M. Research. Global Home Automation Market will be worth US\$21.6 billion by 2020 [Online]. http://www.transparencymarketresearch.com/pressrelease/ home-automation-market.htm, September 2015. Accessed November 25, 2015.
- [18] J. Rogers. CubeSensors API Java [Online]. https://github.com/1337joe/ cubesensors-for-java, March 2015. Accessed Decmber 8, 2015.
- [19] R. Sathya and A. Abraham. Comparison of supervised and unsupervised learning algorithms for pattern classification. *Int J Adv Res Artificial Intell*, 2(2):34–38, 2013.
- [20] D. Shahgoshtasbi and M. Jamshid. Energy efficiency in a smart house with an intelligent neuro-fuzzy lookup table. In *System of Systems Engineering (SoSE)*, 2011 6th International Conference on, pages 288–292. IEEE, 2011.
- [21] S. Stenudd. Using machine learning in the adaptive control of a smart environment. *Utigivare, Vuorimiehentie*, 2010.
- [22] J. Su and H. Zhang. A fast decision tree learning algorithm. In AAAI, volume 6, pages 500–505, 2006.