Intelligent Energy Management Systems in Office Buildings

Carlos Gomes*1, Carlos A. Silva2

1Instituto Superior Técnico, Universidade de Lisboa
2IN+, Instituto Superior Técnico, Universidade de Lisboa, WS Energia Chair, IN+, Instituto Superior Técnico, Universidade de Lisboa

Abstract

Current energy management systems in office buildings are not flexible enough to minimize power consumption while maximizing comfort to users. In fact, they tackle this problem by considering general comfort functions that often neither do they describe individual preferences nor the best energy efficient scenario.

The aim of this thesis is to design a power management system for office buildings that can recognize the user’s preferences using learning algorithms and adjust the air conditioning systems to maximize user comfort, and at the same time minimize the energy consumption. This was validated in the offices on the 2nd floor of the Instituto Superior Técnico - Tagus Park campus, equipped with an automation system.

Different control and learning algorithms were implemented, using the Matlab software, for lighting and air conditioner and a suitable profile for each user was seek by analyzing the user’s needs to regulate the office’s conditions before he felt comfortable. The developed algorithms proved to be adequate to estimate the comfort temperature in the tested offices. The obtained results show the convergence of the inside temperatures to the comfort threshold upon the application of the learning control to the air conditioning systems.

Keywords: building automation, energy, thermal comfort, control, learning, management

1. Introduction

Buildings are the sector that consumes more energy[1][2]. Currently it is estimated that the amount of energy consumed by the European Union’s buildings is around 40 to 45% of the total final energy consumption of these countries. The office equipment operation is responsible for 40% of the electricity consumed in an office building. As a result, energy use in these sectors in the EU accounts for approximately 50% of the greenhouse effect[3].

There is a global consensus that energy consumption in the world is high and there is great potential to increase its efficiency. According to U. White[4], with today’s technology, the world could operate with the same level of functionality and comfort using 30% less energy. It thus becomes essential to seek ways to optimize the use of energy in these spaces and minimize equipment operation’s cost, making it automatic and user-friendly.

In this thesis, it is proposed the creation of a learning algorithm with the ability to analyze human behavior in space occupation, in search of the optimal levels of comfort in a physical location, and play these individual conditions while optimizing the energy consumption of the building.

The hypotheses to be tested are:

• If the algorithm is able to discover a comfort temperature for a certain user;
• If the algorithm can satisfy the user's needs over a certain period of time (day, week, month, etc.);
• The time taken for the algorithm to have credible levels of ideal comfort, so it may generate them without the user’s intervention in the system. Therefore, autonomously from an enhanced learning process.

2. Bibliography review

2.1 Building Energy Management System (BEMS)

A BEMS is a control system installed in buildings that oversees mechanical and electrical building equipment such as ventilation, lighting, power systems, fire systems, and security systems. These systems enhance the performance of the building and ease of
operation over its life-cycle [5]. The primary goal for a higher performing building is to minimize the long-term costs of facility ownership to owners, occupants and the environment. In a higher performing building all components of the building are integrated in order to work together. This improves operational performance, increases occupant comfort and satisfaction and provides the owner with systems, technologies and tools to manage and minimize energy consumption.

According to the Smart Building Institute [5] a BEMS represents a building that:

1. Provides actionable information regarding the performance of building systems and facilities;
2. Proactively monitors and detects errors or deficiencies in building systems;
3. Integrates systems to an enterprise business level for real-time reporting and management utilization of operations, energy and occupant comfort;
4. Incorporates the tools, technologies, resources and practices to contribute to energy conservation and environmental sustainability.

However, current BEMS generally operate according to fixed schedules and maximum design occupancy assumptions. While temperature, airflow, and lighting set points for most commercial facilities can be controlled digitally through the centralized BEMS, these set points must still be determined and scheduled manually [6]. Typically, operational settings are dictated according to assumed occupied and unoccupied periods of the day and do not consider when buildings are only partially occupied.

There is a great potential to reduce building energy consumption by tailoring system control to actual occupancy loads. Addressing this demand, various occupant detection solutions have been implemented and tested to understand actual occupancy loads as they compare to maximum occupancy design assumptions. Occupant detection technologies include motion sensors, cameras, infrared sensors, RFID (radio frequency identification), WSN (wireless sensor networks) and CO\textsubscript{2} sensors. Theoretical energy simulations have found that the input of real-time occupancy information can reduce HVAC energy consumption by 10–20% and can reduce lighting energy consumption by up to 30% [6].

2.2 Comfort

Thermal comfort is the condition of mind that expresses satisfaction with the thermal environment and is assessed by subjective evaluation (ANSI / ASHRAE Standard 55). Maintaining this standard of thermal comfort for occupants of buildings or other enclosures is one of the most important goals of HVAC (heating, ventilation, and air conditioning) design engineers [7].

Thermal neutrality is maintained when the heat generated by human metabolism is allowed to dissipate, thus maintaining thermal equilibrium with the surroundings. The main factors that influence thermal comfort are those that determine heat gain and loss, namely metabolic rate, clothing insulation, air temperature, mean radiant temperature, air speed and relative humidity. Psychological parameters such as individual expectations also affect thermal comfort [8].

Satisfaction with the thermal environment is important for its own sake and because it influences productivity and health. Office workers who are satisfied with their thermal environment are more productive [9]. The combination of high temperature and high relative humidity serves to reduce thermal comfort and indoor air quality [10].

Since there are large variations from person to person in terms of physiological and psychological satisfaction, it is hard to find an optimal temperature for everyone in a given space. Laboratory and field data have been collected to define conditions that will be found comfortable for a specified percentage of occupants.

Regarding the factors that influence thermal comfort, there are:

1. Metabolism - People have different metabolisms that may fluctuate due to the level of activity and environmental conditions [11]. The ASHRAE Standard 55-2010 defines the metabolic rate as the level of conversion of chemical energy into heat and mechanical work and metabolic activity within the body, typically expressed in terms of unit area of the total body surface.
2. Clothing - The amount of insulation used by a person has a substantial impact on the thermal comfort, since it influences the heat loss and hence the heat balance. Insulating layers of clothing prevent heat loss and may help to keep a person warm or lead to overheating. Generally, the thicker the garment, the greater the insulating capacity. Depending on the type of material which the garment is made, the flow of air and relative humidity can reduce the insulation capacity of the material [12].

3. Air Temperature - The temperature is the average temperature of air around the occupant, with respect to location and time. The air temperature is measured with a dry-bulb thermometer.

4. Mean radiant temperature - the radiant temperature is related to the amount of radiation heat transferred from a surface, and which depends on the material's ability to absorb or emit heat, its emissivity. The mean radiation temperature depends on the temperature and emissivity values of the surrounding surfaces, as well as the shape factor.

5. Air speed - The air velocity is defined as the rate of airflow at a point, without regard to its direction. According to ASHRAE Standard 55, air speed is the average air velocity that the body is exposed with respect to location and time.

6. Relative humidity is the ratio of the amount of water vapor in the air to the amount of water vapor that the air could hold at the specific temperature and pressure. While the human body has sensors within the skin that are fairly efficient at feeling heat and cold, relative humidity (RH) is detected indirectly. Sweating is an effective heat loss mechanism that relies on evaporation from the skin. However, at high RH, the air has close to the maximum water vapor that it can hold, so evaporation, and therefore heat loss, is decreased. On the other hand, very dry environments (RH < 20-30%) are also uncomfortable because of their effect on the mucous membranes. The recommended level of indoor humidity is in the range of 30-60% in air conditioned buildings, but new standards such as the adaptive model allow lower and higher humidity, depending on the other factors involved in thermal comfort [13].

2.3 Comfort Measurement Model

The model used in this thesis to measure the thermal comfort is the PMV / PPD, developed by Fanger [10], who used heat balance equations and empirical studies of skin temperature to define comfort. It measures the opinion of thermal sensation on a scale of seven points, from cold (-3) to hot (+3). Zero is the optimal value, and represents thermal neutrality, and the comfort zone is defined by all combinations of the six parameters for the PMV to be within the recommended limits (-0.5 <PMV <0.5).

The ASHRAE Standard 55-2010 uses the PMV model to define the requirements for the thermal conditions inside. This pattern requires that at least 80% of occupants are satisfied.

3. Case study description

The case study presented within this article is a group of 14 office spaces in the building of the IST Taguspark. The offices’ control uses KNX home automation. This system has a lighting and temperature controller in the cabinet and allows the user to control the lighting and the HVAC system.

The data measured by the sensors are possible to be read on the system interface or by URL links. This system is usually controlled with a panel interface where the level of lighting can also be controlled. The user does not define the HVAC velocity but, instead, sets a desired temperature.

The system installed to date is very simplified; the user must set a comfort temperature (temperature set point) and the air conditioner will turn on depending on the temperature difference between the set point and the instant temperature inside the room. As the temperature difference decreases, the velocity of the air conditioner also decreases. Just as shown in the block diagram of the figure 1.
4. Methodology

Reinforcement learning (RL) is a powerful unsupervised learning scheme, which has been widely studied by researchers such as Kaelbling, Littman, Moore, Barto, Watkins, Mahadevan, all of them authors of many articles and books on the subject that have been quoted in this text. RL is the name given to a group of methods to deal with the problems when the learning agent needs to find the optimal behavior by interacting with an unknown environment, usually involving a delay in reward [14].

In RL there are elements of high importance [15], such as:

1. Policy – defines the learning agent's way of behaving at a given time. A policy is a mapping from perceived states of the environment to actions to be taken when in those states.

2. Reward Function – defines the goal in a reinforcement learning problem. It maps each perceived state (or state-action pair) of the environment to a single number, a reward, indicating the intrinsic desirability of that state. A reinforcement learning agent's sole objective is to maximize the total reward it receives in the long run. The reward function defines the good and the bad events for the agent.

3. Value function – specifies what is good in the long run. The value of a state is the total amount of reward an agent can expect to accumulate over the future, starting from that state.

4. Model of the environment - given a state and action, the model might predict the resultant next state and next reward. Models are used for planning, which means any way of deciding on a course of action by considering possible future situations before they are actually experienced.

5. Agent – the learner with the decision maker

These last two elements interact continually, the agent selecting new actions and the environment responding to those actions and presenting new situations to the agent.

The agent and environment interact at each of a sequence of discrete time steps, $t = 0,1,2,3 \ldots$. At each time step $t$, the agent receives some representation of the environment's state, $s_t \in S$, where $S$ is the set of possible states, and on that basis selects an action, $a_t \in A(s_t)$, where $A(s_t)$ is the set of actions available in state $s_t$. One time step later, in part as a consequence of its action, the agent receives a numerical reward $r_{t+1} \in R$, and finds itself in a new state, $s_{t+1}$. Figure 2 diagrams the agent-environment interaction.
At each time step, the agent implements a mapping from states to probabilities of selecting each possible action. This mapping is called the agent's policy and is denoted $\pi_t$, where $\pi_t(s, a)$ is the probability that $a_t = a$ if $s_t = s$. Reinforcement learning methods specify how the agent changes its policy as a result of its experience. The agent's goal, generally speaking, is to maximize the total amount of reward it receives over the long run.

The mapping is done by a matrix where the rows represent the states and the actions column. In the tables below there are 2 examples, table 1 represents an example standard array of an agent without anything learned yet, here for one state (temperature), there are four options for the AC's velocity which means each one has a 25% of happening. Table 2 represents a final array of that same state after the learning process.

Table 1 – Standard array of an agent without anything learnt

<table>
<thead>
<tr>
<th>$T[\degree C]$</th>
<th>AC = 0</th>
<th>AC = 1</th>
<th>AC = 2</th>
<th>AC = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Table 2 - Final array of that same state after the learning process

<table>
<thead>
<tr>
<th>$T[\degree C]$</th>
<th>AC = 0</th>
<th>AC = 1</th>
<th>AC = 2</th>
<th>AC = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The goal of the agent, generally speaking, is to maximize the total amount of reward it receives over time by the equation that generates the value of each state.

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha [R_{t+1} + \gamma \max_{a'} Q_t(s_{t+1}, a') - Q_t(s_t, a_t)]$$  \hspace{1cm} (Eq. 1)

In this thesis, the policy will be the probability distribution of turning on a certain HVAC velocity. The actions will be the orders calling a certain velocity air conditioning. The states are real-time temperatures of a cabinet.

5. Models

In this thesis it was created an algorithm constituted by two parts: one capable of finding an ideal comfort temperature for the user in the cabinet, requiring his intervention while this algorithm runs, and the second part with the capability of controlling the AC system on its own and at the same time learning what fan speed to use depending on the temperature difference depending on the comfort temperature and the real-time temperature of the room.

When the program starts the user is presented with an interface capable of controlling the HVAC system and the lights (ON-OFF). This interface is always the same no matter which algorithm is running but tells the user which one is running.

To find out the ideal comfort temperature the learning algorithm used was the following:
1. Initializing probability distribution, array (temperatures, ΔAC);
2. While there are no probabilities below 0.9;
3. Read room temperature;
4. Checks if the user interfere with the AC velocity, calculates ΔAC:
   \[ \Delta AC = (\text{Last AC velocity}) - (\text{actual AC velocity}) \]
5. Reward given depending on the difference of the velocities:
   \[ \text{reward} = 3 - \Delta AC \]
6. Rearrange of the probability distribution:
7. Wait 5 minutes:
8. Repeat points 3 to 7 until point 2 is not valid.

Following this algorithm, after a certain time we are able to say what the user’s ideal comfort temperature is because the probability s of the matrix will, in column ΔAC = 0, be higher than 0.85 (limit value for to be considered validated). The line of the matrix where this succeeds, is the temperature that we were looking for.

For the automatic control of the AC system, with no need for the user’s intervention, the learning algorithm used was the following:

1. Initializing probability distribution, array (temperature, AC);
2. Read room temperature – state \( t \);
3. Turns on air conditioning in a random velocity depending on the probabilities;
4. Waits 5 minutes;
5. Read room temperature – state \( t+1 \);
6. Reward given as function of the person’s state of comfort:
   \[ PMV = T(\text{atual}) - T(\text{comforto}) \]
   \[ \text{reward} = 3 - |PMV| \]
7. Rearranging probabilities in state \( t \):
   \[ \text{matrix}(t, AC) = \text{matrix}(t, AC) + \frac{\text{reward}}{10} \]
   \[ \text{matrix}(t,:) = \text{matrix}(t,: \sim AC) - \frac{\text{reward}}{30} \]

6. Results

During the trial period it was possible to test the algorithm in some offices and observe different types of learning, mainly due to external environmental factors such as precipitation and solar exposure of the cabinet, and to the opening and closing the doors.

6.1. Algorithm with fixed Comfort Temperature

In office 2N-14.16, it is observed that the 1st mode was active up to 19 timelines, i.e., during 1h35, when it learned a comfort temperature of 26°C (Figure 3).

![Figure 3 – Trial at office 2N-14.16](image)

From the 20th iteration, the system started to be controlled automatically, and it was observed the fan speeds being turned on randomly every 5 minutes, because of all the probabilities were the same.

From iteration 30 to 53, i.e., after 2:30 application has been powered on, we note that the office’s temperature begins to be constant around the value 26.5°C, with a difference of
0.5°C for comfort temperature, initially learned. However, it still exists a high randomness at fan speeds of the AC until the 45th iteration, 3h45 after. One can still note a constant value of the temperature and fan speed of the AC to the iteration 53, 4h25 after the algorithm has been initiated.

From iteration 53 onwards, there is a similar behavior described between iterations 20-53, perhaps initiated by an external element, such as a window or a door being opened.

In office 2N-14.24 (Figure 4), the application was turned on for 4h35, and 2nd was activated in iteration 18. The 1st mode was on for 1h30 and learned a comfort temperature of 25°C.

![Figure 4 - Trial at office 2N-14.16](image)

We can see a satisfactory performance in terms of temperature difference between the comfort temperature and the office’s real time temperature. Although as it was already expected, the appearance of the AC’s random velocities in the iterations of the 2nd mode (18-25), it is visible, an almost constant temperature around the comfort temperature (24.6°C - 25.4°C). In the last 50 minutes there has been a sharp rise in cabinet temperature, probably due to the opening of the door to the corridor because the corridor is constantly at higher temperatures than the offices powered by the AC.

The office 2N-14.28 is different from the previous spaces because it is an inner enclosure and is shared by more than one user. In certain simulations there was an encounter with unexpected results, it was seen that the room temperature, sometimes, converged to a different temperature from the comfort temperature learned (+/- 1°C), as it can be seen in the test for the office 2N-14.28 (Figure 5). It was then concluded that the comfort temperature of a certain user, even after learned, could change throughout the day, and if the algorithm had this event into account, it could operate with a better performance than the one observed. These changes could be motivated by the person's needs, or simply because there was another person in the room.

![Figure 5 - Test to the office 2N-14.28 where there is a convergence of the room temperature to 1°C below the comfort temperature learned](image)
Therefore, knowing that the algorithm consists of two separate modes (comfort temperature and automatic control), and considering that the automatic control functions on the basis of the previously learned comfort temperature, it would be a lot better if the comfort temperature could change over the learning time of the automatic control. This would only happen if the algorithm felt the need to it.

The following change was made in the automatic control algorithm: if the user gives at least three opposite orders to the system, the algorithm checks which was the change, and the comfort temperature will change. If the third order is to switch the AC fan speed at level 2 or 3, the comfort temperature decreases 1°C, since it is assumed that the occupant was too hot. If the fan speed is switched to 0 or 1, the comfort temperature increases 1°C, since the user was cold.

6.2. Algorithm with Adjustable Comfort Temperature

With the new version of the algorithm, there was a better performance compared with the previous, and a greater capacity of the system to regulate a temperature much closer to user’s comfort temperature.

The switch to automatic control can be seen in Figure 6 at iteration 13, 1h05 after the application has been activated, with a learned comfort temperature of 24°C, encountering right away a quick change of the comfort temperature to 26°C. This abrupt change is justified by previous orders given by the user to continue to have the AC on fan speed 0. As is shown, six straight orders to trigger the fan speed at 0, generated a 2°C increase of the comfort temperature. In the remaining time until the close of the application, there is a good reading of office’s temperature, never exceeding the difference +/- 0.3°C in relation to the comfort temperature.

![Figure 6 – Test to the office 2N-14.16 which shows a change of the comfort temperature of 2 °C during the automatic control](image)

In the simulation shown in figure 7, there was a behavior similar to the previous one, featuring a comfort temperature of 26 °C, with a rapid decrease of 2°C, 50 minutes after the automatic control was triggered. This change was justified by the orders given by the user by switching the fan speed of the AC repeatedly during the iterations 24-30.
7. Conclusions and future work

The implemented program was developed with a simple interface interpretation by informing users about the main parameters of office’s state in real time.

The tests carried out showed that the user’s comfort level could be defined, whereas the system showed good performances while achieving temperature stabilization very close to the learned comfort temperature, and this taking into account that the fan speed of AC was triggered by probabilities and never by direct orders.

It was also concluded that an initial learning comfort temperature was often premature, since the occupant had their needs change over the day, therefore the calculation of the user’s comfort temperature was extend to the automatic control.

It is recommended to do longer studies, and testing in several seasons and variations in climate, since this work was only tested during the summer.

Another significant improvement would be implementing the program depending on other climatic variables such as outdoor temperature, or the level of sun exposure the day. This would have the aim to create profiles for each user, even taking account the time of day, and thus make the program capable of recognizing the desire of the occupant at any day of the year. Further to this, the incorporation of a cost function that also regards energy cost and consumption would be a step towards the development of an intelligent BMES that could maximize both the comfort of the occupants and also energy and economic efficiency.

References


