



# Killing two birds with one stone: improving building energy efficiency and occupant comfort

Fateh Boulmaiz, Amr Alzouhri Alyafi, Stephane Ploix, Patrick Reignier

## ► To cite this version:

Fateh Boulmaiz, Amr Alzouhri Alyafi, Stephane Ploix, Patrick Reignier. Killing two birds with one stone: improving building energy efficiency and occupant comfort. 2nd International Conference on Energy and AI, Aug 2021, Londres, United Kingdom. hal-03313725

**HAL Id: hal-03313725**

**<https://hal.science/hal-03313725>**

Submitted on 6 Aug 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# KILLING TWO BIRDS WITH ONE STONE: IMPROVING BUILDING ENERGY EFFICIENCY AND OCCUPANT COMFORT

(Full paper)

Fateh Boulmaiz<sup>1</sup>, Amr Alzouhri Alyafi<sup>1</sup>, Stephane Ploix<sup>1</sup>, Patrick Reignier<sup>1</sup>

<sup>1</sup> Uni. Grenoble Alpes, CNRS, Grenoble INP, LIG, Grenoble, France

## ABSTRACT

Faced with the challenges posed by global warming and the depletion of fossil fuels, all the actors in society (individuals, industries, communities, and governments) are led to find ways to reduce their energy consumption. As residential buildings contribute significantly to the overall energy consumption, a particular effort is requested from buildings occupants to reduce their energy consumption without having to compensate on their quality of life. The energy behaviour of buildings is directly influenced among others by the behaviour of their occupants. This paper describes a new approach based on historical data to sensitize occupants for their actions' impact on building energy comportment without the need to develop a physical model thanks to case-based reasoning (CBR). We report on an experimental evaluation of the approach based on real-world data collected from an office building including two years of historical data.

**Keywords:** Case-based reasoning, energy management system, genetic algorithm.

## 1 INTRODUCTION

Energy is vital in our modern life but people do not pay attention to how and how much they consume it. Energy is almost present in all activities directly or indirectly. In its reference scenario, the American Energy Information Administration (EIA) estimates that world energy consumption could strongly increase in the next decades. The EIA<sup>1</sup> projects a rise of 28% of the global energy demand from 2015 to 2040. The residential and office building sector is considered the most energy-intensive, accounting for 42% of final energy consumption in France [1]. For that, it is important to optimize the usage of the energy and help the users to reduce the waste of energy in general. This can be applied especially in buildings, where most of our activities are done. Indeed, occupants spend up to 90% of their time inside these buildings [2] and a significant part of this energy is used by the heating and cooling systems to satisfy the thermal comfort of the building occupants.

Various factors can influence the energy behaviour of a building that can be classified into four categories as design, building materials and construction techniques, environmental characteristics, and occupant behaviour. Doors and windows positions, orientation of the building, and aperture-to-wall ratio are examples of the design parameters. Quality of materials used to insulate walls and ceilings, and glazing thermal performance are mentioned as the second category. As environment parameters, weather characteristics, and position of the building about the site surroundings, can be cited. Occupant behaviour is defined by the presence and the interaction of the occupant regarding his environment by performing actions (or not) to change the indoor environment. As actions, windows/doors opening/closing and adjusting the HVAC (heating, ventilation, and air conditioning) system settings can be mentioned. The occupant behaviour and its impact on the building energy performance are extensively studied in the literature (e.g., see [3], [4]) from different aspects: psychological, cultural, economic, physiological, and time. These studies are unanimously in agreement that occupant behaviour has a substantial impact on building energy consumption. Therefore, guiding the occupant by intelligently utilizing his actions can bring an improvement in his comfort (thermal and air quality) at the same energy consumption cost if not less.

Attempts have been made to attend occupants to elaborate an optimized energy consumption plan. Traditionally, smart thermostats are used. The latter can only display the current temperature and they might give some general energy advice. Moreover, the cooperation and interaction between occupants and these systems are limited to define only the setpoints and some states. In Rule-based EMS [5], the occupant can interact more with the system than with smart thermostats by setting different information about the context, the home state, and also personalize some scenarios. However, they do not provide the ability to simulate or assess the evolution of

---

<sup>1</sup> EIA, «International energy outlook» 2017

the environment. For example, they can not predict how the occupant's comfort would change when weather variables change. In the perspective of overcoming these issues, knowledge Model-based energy management systems [6] are proposed, which simulate and predict buildings behaviour while integrating occupant interactions with buildings. They can be either white-box models (relying on physical knowledge); black-box models (based on statistical estimation methods) or a hybrid of the latter two. Knowledge modelling approaches, although offering undeniable advantages in terms of accuracy, are very difficult and still an ongoing scientific problem [7]. They are generally hard to build and maintain. Designing energy models requires a lot of expert time, important knowledge about the building characteristics, and profound physical understanding. Once designed, when there is a change in the environment like adding an electric heater or other appliances, it is necessary to re-calibrate or even modify the model. At the same time, while the occupants are unable to correlate and understand the functioning of their habitat, due to these complexities, they are able to determine their comfort and intentions. Unfortunately, the modelling of human behaviour is difficult to estimate and still a very challenging task because of the stochastic nature and randomness of people due to the lack of information about cognitive processes. To overcome these difficulties, we propose an approach allowing constructing a data-based model without the need neither for profound physical knowledge to generate a knowledge model nor for modelling occupant's behaviour. Our working hypothesis that will subsequently guide our methodology can be formulated as follows: historical data holds historical information about the environment as well as about the occupant's behaviour. Thus, that inferred information could be used to construct a model to predict building behaviour from environment features and occupant behaviour. The goal is to maximize the occupants comfort while minimizing the energy consumption thanks to the CBR methodology.

In the CBR approach, problems are solved by adapting solutions from previous similar cases held in a cases base. According to [8], CBR involves four processes: 1) *retrieve* the most similar cases; 2) *reuse* of the retrieved cases, and if necessary, after adapting them to address the problem; 3) *revise* the proposed solution if required; and 4) *retain* the problem description and the proposed solution as a new case used to solve future problems.

## 2 METHOD

Various phenomena can impact occupant comfort, they are modelled by variables that can be categorized at least as *context* variables which model phenomena that cannot be controlled like the weather; *action* variables which can be controlled by the occupant like opening windows/door and *effect* variables that represent consequences of the other variables like indoor temperature and indoor CO<sub>2</sub> concentration.

Days are first of all automatically grouped into similar days, i.e., days where the same action will have the same consequences (opening a window in winter does not have the same impact as opening a window in summer for instance). These similar days are then used to extract an action plan that has previously led to good thermal comfort or air quality. To achieve this goal, we proceed in four steps:

- 1) grouping variables with the same nature and normalize them.
- 2) using a Genetic Algorithm for variables weighting to determine the importance of context and action variables in the resulting effects.
- 3) extracting days having similar context variables as the current day (based on weather forecast).
- 4) obtaining the best set of actions and recommending them to occupants.

### 2.1 VARIABLES WEIGHTING

The global objective is to find the best set of weights for context and action variables, which satisfies the condition: if a context or action variable is more important for effect variables, it should have a higher weight and if two days have a small context-action weighted distance  $Dist_{CA}$ , they should have a small effect distance  $Dist_E$  and vice versa. In the following, we propose an approach based on a genetic algorithm for estimating the weights.

1) set  $i = 0$ .

2) create an initial random weight vector  $w_i$  for context and action variables.

3) cluster the days, which are close, using a k-means algorithm based on the weighted context-action distance

$$\text{defined by: } Dist_{CA}(D1, D2) = \left( \sum_{i=1}^{n_C} [w_C[i] * (C_{D1}[i] - C_{D2}[i])]^2 + \sum_{i=1}^{n_A} [w_A[i] * (A_{D1}[i] - A_{D2}[i])]^2 \right)^{1/2}$$

where:  $C_{D1}[i]$  is the context variable  $i$  of the day  $D1$  -  $w_C[i]$ : is the weight of variable  $i$ .

4) the objective function to optimize consists of two stages. On the one side, minimizing the number of clusters to maximize the diversity of actions in each cluster (to able to propose at the end a pertinent set of actions to the user for the current day). On the other side, reducing the average distance defined in equation  $Dist_E$  within the effects in each cluster to meet our starting hypothesis (two similar days should have the same consequences

from the same actions). The algorithm holds if either the number of generations without fitness improvement is performed in a population, reaches a predefined value or if the number of generations reaches a specified maximum number of evolutions. Otherwise, go to Step 5.

$\text{Dist}_E(D1, D2) = ((E_{D1}[i] - E_{D2}[i])^2)^{1/2}$ , where:  $E_{D1}[i]$  is the effect variable  $i$  of the day  $D1$ .

5) adjust  $w_i$  to  $w_{i+1}$  (through operations of crossover and mutation) considering the objective function and iterate from step 3 until reaching a stable state.

## 2.2 EXTRACT SIMILAR DAYS

Considering the morning of a new day, the estimation of the similarity with previous days in the database is based on the context variables only (using weather forecast and scheduled occupancy) since action variables are unknown (that's what we want to determine) and, consequently, effect variables are either unknown. At this stage, KNN-algorithm is applied for predicting similar context days based on weighted distance introduced by:

$\text{Dist}_C(D1, D2) = (\sum_{i=1}^n (w_C[i] * C_{D1}[i] - w_C[i] * C_{D2}[i])^2)^{1/2}$ . The objective is to find the set of similar days of the new day  $D$  that satisfies the condition: all days in this set are similar to  $D$  based on the context variables and satisfy the *sensitive distance* for effect variables (the difference in the effect variables will not be perceived by the occupant. E.g., the difference of indoor temperature  $< 2^\circ\text{C}$  and difference of  $\text{CO}_2$  concentration  $< 400$  ppm). Though, there is no fixed number for similar context days. If the number of similar context days is too big, there might be several non-similar days in the cluster. On the other hand, if there is a small number of similar context days, the diversity of the set of actions may not be enough to be able to recommend actions for occupants. To address this issue, we define TC - a threshold context distance for two days  $D1, D2$  while guaranteeing the sensitive effect distance between them. However, it is difficult to determine precisely the TC. If TC is too low, it is not always possible to obtain some similar days satisfying the above condition (recall). But if TC is too high, it might have some incorrect days in the result (precision). To find the best compromise between accuracy and obtaining the maximum number of similar days based on their effects for a day  $D$ , a classification method with the F1-score is proposed to determine the TC.

## 2.3 EXTRACT THE BEST SET OF ACTIONS

As the aim is to strengthen the occupant's best comfort, the plan is to find the day having the best satisfaction from the set of similar context days. Then, extract the set of actions from this day to recommend it to the occupants. From the set of similar days based on context, the best day  $D^*$  to get the set of actions is defined by:  $D^* = \text{argmax } S(D)$ , where  $S(D)$  - satisfaction (comfort) of day  $D$  based on the effect variables.

# 3 EXPERIMENTAL EVALUATION

## 3.1 DATA SET

The paper focuses on a real case of study: an office, where 4 persons at work. It is equipped with 27 sensors collecting information about  $\text{CO}_2$  concentration, temperature, etc. The sensor data and weather conditions were recorded for a consecutive period starting from April 2015 and going till October 2016 at hourly intervals. Only working days were chosen (no weekends or closed days, as there are no occupants on these days). After filtering, 100 days are found for training the model. For the testing part, another period is selected: 22/06/2016-30/07/2016 (34 days). The test periods were in summer, so the heater during these periods was turned off. For the purpose of assessing the performance of the proposed approach, a physical model for the office was employed to simulate the effects of proposed actions (see sections 3.3 and 3.4).

## 3.2 VARIABLES WEIGHTING

The genetic algorithm was able to compute the following weights corresponding to the context and action variables. The algorithm has converged in 4784 seconds after 11 iterations. Experiments were conducted on a Mac laptop with Core-i5 (2,7 Ghz) processor and 8 Giga Bytes of memory.

$W = [0.211, 0.102, 0.081, 0.025, 0.041, 0.040, 0.087, 0.058, 0.051, 0.107, 0.039, 0.003, 0.051, 0.107, 0.039, 0.003]$

## 3.3 VALIDATING THE DATA MODEL

This step aims to validate the correctness of the hypothesis about the relations between context, action and effect variables. Since  $D1$  is the historical best similar day to day  $D$ , then the effects  $E^*$  (simulated effects of the day  $D$  when using recommended actions  $A^*$  from  $D1$ ) should be similar to the original effects  $E1$  of the day  $D1$ . This condition was used to validate the hypothesis. With each day  $D$  of 34 testing days, suppose  $\Delta T, \Delta \text{CO}_2$  are the difference of indoor temperature and  $\text{CO}_2$  concentration respectively between  $E^*$  and  $E1$ . They are computed by the maximum difference between each hour of two days respectively. The mean difference of temperature

$\Delta T$  is about  $1.56^{\circ}\text{C}$  and the mean  $\Delta\text{CO}_2$  is about 199 ppm, the maximum value of  $\Delta T$  is  $2^{\circ}\text{C}$  and the maximum value of  $\Delta\text{CO}_2$  is 400 ppm. It demonstrates that when applying the same set of actions to the days having similar context, they would obtain a similar effect and the experimental results agree with the hypothesis.

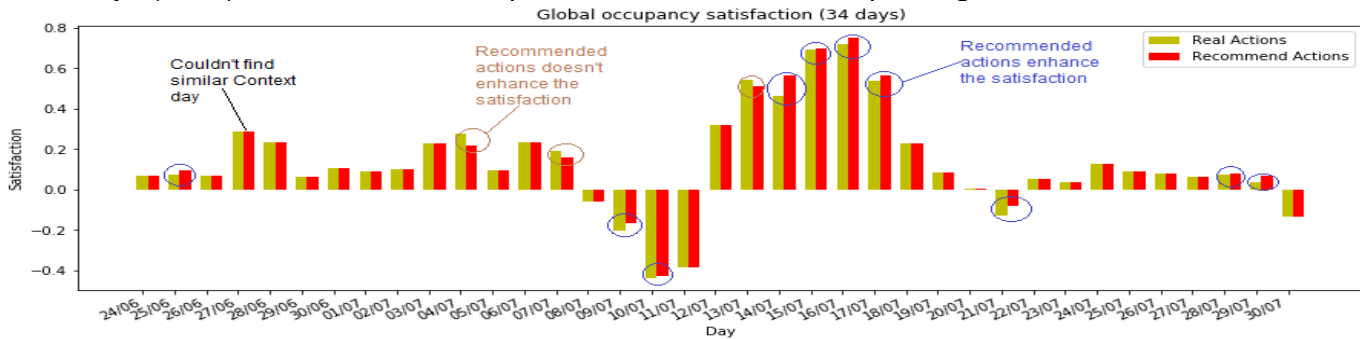
### 3.4 VALIDATING THE RECOMMENDED SET OF ACTIONS

This step aims to verify that the recommended set of actions would enhance the occupant's satisfaction. So, we compare simulated effects  $E^*$  of the day  $D$  when using recommended actions  $A^*$  to the real effect  $E$  with the real actions  $A$  of occupants on this day. For instance, the occupant satisfaction  $S$  is composed of thermal  $TS$  and air quality satisfaction  $CS$ . They are defined as follows:  $TS$  gets the highest value when the temperature is between  $21^{\circ}\text{C}$  to  $23^{\circ}\text{C}$  and  $CS$  is maximized when  $\text{CO}_2$  concentration is less than 500 ppm. We estimate the global satisfaction  $S$  as:  $S(t, c) = \alpha * TS(t) + (1 - \alpha) * CS(c)$ ;  $\alpha \in (0, 1)$  - the relative importance of  $TS$  to  $CS$ .

Suppose that  $S^*$  is the occupant's satisfaction obtained by the effect  $E^*$ , and  $S$  is the occupant's satisfaction derived by the effect  $E$ . The enhancement obtained when applying the recommended set of actions is given as:

$$H(S, S^*) = (S^* - S) / |S|$$

Figure below. Shows satisfactions with recommended and real actions for 34 testing days. Resultants show that we could recommend a set of actions for only 13 days on the 34 testing days (with  $T_c = 1.2$ ), this is due to the limitation of real data. Specifically, with 13 days, the recommended set of actions could enhance the satisfaction for 10 days (77 %). The model could improve from 8% to 18% occupant's global satisfaction.



## 4 CONCLUSION AND FUTURE WORKS

We propose a data-model technique to enhance occupant comfort based on the CBR approach. The proposed model predicts actions without the need for a knowledge model and depending only on the historical data. We are proposing a new approach for extracting similar days based on genetic algorithm, KNN, and F1-score function. A major strength of this technique lies in its capacity to determine the weights of variables in a fully automated way without the need for any domain expert assistance. Experiments on real data have shown the feasibility and effectiveness of this method.

This is a work in progress, research is underway to improve the process of the best set of actions estimation. Particularly, it can be helpful to look into how to use regression and extrapolation techniques to better adapt the retrieved solutions to fit well the problem to be solved.

## 5 REFERENCES

- [1] "Ministère de l'environnement de l'énergie et de la mer France", "Chiffres clés de l'énergie - Edition 2016," 2017.
- [2] YouGov, "Survey Modern indoor living can be bad for your health," 2018.
- [3] Si-Larb and al, "Occupant behaviour: a major issue for building energy performance," *Materials Science and Engineering*, 2019.
- [4] C. Jiayu, J. E. Taylor and W. Hsi-Hsien, "Modeling building occupant network energy consumption decision-making: The interplay between network structure and conservation," *Energy and Buildings*, vol. 47, pp. 515-524, 2012.
- [5] K. Tomoya, F. Naotaka, Y. Tomoki and T. Masahiko, "An Evaluation and Implementation of Rule-Based Home Energy Management System Using the Rete Algorithm," *The Scientific World Journal*, 2014.
- [6] H. Viot., "Modeling and instrumentation of a building and its systems," PhD report, Bordeaux University, 2016.
- [7] F. Amara, K. Agbossou, A. Cardenas, Y. Dubé and S. Kelouwani, "Comparison and Simulation of Building Thermal Models for Effective Energy Management," *Smart Grid and Renewable Energy*, vol. 6, pp. 95-112, 2015.
- [8] A. Aamodt and E. Plaza, "CBR: foundational issues, methodological variations and system approaches," *AI Communications*, vol. 7, no. 1, pp. 39-59, 1994.