Predicting office's ambient parameters

Vito Janko
Jožef Stefan Institute
Jožef Stefan International Postgraduate School
Jamova 39, 1000 Ljubljana
vito.janko@ijs.si

Andreas Stensby
Jožef Stefan Institute
Jamova 39, 1000 Ljubljana
andreas.stensbye@hotmail.com

Mitja Luštrek
Jožef Stefan Institute
Jožef Stefan International Postgraduate School
Jamova 39, 1000 Ljubljana
mitja.lustrek@ijs.si

ABSTRACT

Bad environmental conditions in the office can negatively affect the workplace productivity. In the presented work we measure three ambient parameters - CO_2, temperature and humidity - asses their quality and predict their likely future values. To do so, we first heuristically determine the state of the office (are the windows open, air conditioner active etc.) and then try to mathematically model the parameter's future behavior. Based on the current and predicted state of ambient parameters, we can send a recommendation on how to best improve them. Experimental evaluation shows, that our models outperform the related work in terms of prediction accuracy.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

Keywords

CO_2, temperature, humidity, modeling, recommendations

1. INTRODUCTION

Good work environment is essential for keeping work productivity. In this paper, we are focusing on three office's ambient parameters: CO_2, temperature and humidity. The quality of these parameters is often hard for humans to objectively detect, especially if they are changing slowly. However, it has been shown [4, 5] that when their quality drops below certain thresholds, the work productivity in the office is negatively affected.

In this paper we present an intelligent system that is able to measure these parameters and estimate their future values. In the case of CO_2 and temperature, a simple mathematical model is used for prediction, in the case of humidity a machine learning model is used instead. Furthermore, it is able to asses the quality of these parameters, simulate several possible actions, that an user can take, and then recommend the one leading to the best working conditions. The system is meant to be used in offices without automatic ambient control and is a part of the larger "Fit4Work" project [2] that is focused on helping to raise the well-being of office workers. It requires no prior knowledge or manual input of office properties, yet it is able to adapt to them over time.

The ambient parameters were measured using the Netatmo commercial device [1]. The same device is expected to be used by the end users of the system, although it can be replaced with any similar device with the same functionality. This device has an indoor and an outdoor unit, both capable of measuring the CO_2, temperature and humidity, and sending the data to a web server. For easier testing and validation of our method we also had sensors that monitor if the windows are opened and closed and an application where we manually labeled the number of people in the office, the air conditioner state, heater state and humidifier state. As for the time of writing this paper we collected roughly two years of data for three different offices in our department. Data is continuously sent to a web server, where it is analyzed as described in Section 2. If a recommendation is deemed necessary, it is sent to a mobile device via a push notification.

The paper was inspired by another work [3], and proposes a different solution to the same problem. The proposed solution makes heavier use of mathematical modeling and produces more accurate predictions about the ambient parameter's future values.

2. METHODOLOGY

The goals of this paper are three-fold. First, to predict the state of the office: are the windows open, is air conditioner turned on, etc. This not only allows us to predict the probable changes in the ambient parameters, but also to make sensible recommendations: no need to recommend opening of the windows if they are already open. Second, to predict the future behavior of the following three parameters - CO_2, temperature and humidity. We are interested in predicting values up to 30 minutes in advance. Our data was measured every 10 minutes, so this corresponds to 3 data points. Behavior should be predicted for the current state and also for the cases where some office parameters changes. Finally, we use a combination of the previous two points to form recommendations to the user on actions that improve the work environment.

While the physical phenomena of temperature, humidity and CO_2 was already heavily studied in the past, the challenge we face here is in not knowing any attributes of the office where this system would be used: how big the office, how good the thermal insulation, the surface of the windows, etc. Using standard formulas for predicting the ambient parameters can therefore be infeasible, given how many unknowns they contain. In our approach we tried to simplify the models to simple versions with only a few unknowns. We use data recorded in the target office in the last two weeks (exact number of days may vary based on office usage) to estimate these unknowns, and then use them for real-time predictions in the following day.

2.1 Virtual sensors

Virtual sensors refer to values that are not directly measured. Instead, their value is derived from the measured data and then later used to help derive some other value. In our setting, there are five virtual sensors that affect the ambient parameters: the windows state, air conditioner state, heater state, humidifier state and number of people in the office.

"The number of people in the office" is calculated from raising CO_2 levels, and the humidifier state is tied to humidity data, so those two will be explored in the corresponding Sections 2.2 and 2.4. The remaining three can be determined with simple heuristics as described below.

2.1.1 Windows

Windows were modeled in a binary fashion: they are either open or they are closed. In a real office there might be many windows, some of them open, some closed, some perhaps half-open at any time; but lacking any knowledge about the window quantity or size, predicting their state more accurately is almost impossible.

An effect of opening the window is reflected on all three ambient parameters, but only in the case of CO_2 is the effect consistent. Whenever a window is opened, CO_2 falls drastically, whenever it is closed it starts to rise again. This allows us to make a simple heuristic: a.) if the CO_2 is falling faster then some threshold, window was opened; b.) if CO_2 keeps increasing, the window is closed; c.) if neither of those is happening, assume the last known state. Thresholds can be determined by looking at the data history and find such values that would generate predictions, where windows is opened/closed few times a day, as would realistically be the case.

This approach could be improved by correlating changes in CO_2 to those in temperature and humidity, but the described simple heuristic appeared to work well in practice.

2.1.2 Air conditioner

Again we assume binary outcome - the air conditioner is either on or off - additionally we assume that the temperature set on it is constant, or at least is changing infrequently. The distinguishing pattern of air conditioning is one of temperature inside decreasing while the temperature outside is higher then inside. Since the temperature naturally tries to equalize itself with its surroundings and since all other factors (people, computers, etc..) only serve to warm the office, it is reasonable to conclude that such a temperature drop was caused by the air conditioner. After a while of the air conditioner working, the temperature will converge to value that can be stored for later predictions. If the temperature starts rising again, the air conditioner is assumed to be turned off.

2.1.3 Heater

The same assumptions and methods are used here as with the air conditioner, except in reverse: the heater is on if the inside temperature rises significantly more then expected from the outside temperature, etc.

2.2 CO_2 predictions

We start by modeling CO_2, as it the most "well-behaved" of the three ambient parameters, and we describe the process in depth. We later use a similar methodology for temperature modeling. Intuitively, CO_2 level inside the office is increasing linearly with respect to the number of people present, but at the same time it tries to equalize itself with the outside CO_2 level. The bigger the difference between outside and inside, the faster it moves from one side to another. If window is opened, the same happens, only to a significantly larger degree. This can be encapsulated in the following equation.

$$C_{n+1} = C_n + \alpha(C_{out} - C_n) + \beta p \tag{1}$$

$C_n = CO_2$ inside at timestep n
$C_{out} = CO_2$ outside
$p =$ the number of people in the room
$\alpha =$ the coefficient of diffusion speed (between 0 and 1) - small for closed windows, big for open ones
$\beta =$ how much a single person raises CO_2 in a given time unit

Using all the labeled data, the α and β are mostly trivial to compute using linear regression. Using them results in an almost perfect match between the predicted and real values. In Figure 1 we plot a scenario where we know the initial CO_2 level and all future windows states and all future numbers of people, and we are able to predict CO_2 level two days in advance. This strongly signifies that the model captures the real-life behavior of CO_2, and it is only a matter of determining the correct coefficients.

Calculating the coefficients for a given office without the labeled data, however, is a challenging task as the above formula has 5 unknowns - α when windows are closed, α when windows are opened, window state, β and number of people p. Furthermore these coefficients can behave very similarly: CO_2 level in a room with many people and open window can be close to CO_2 level in a room with closed windows and few people. The first improvement is to combine the two variables β and p into one - γ, as we never need those two individually and are only interested in their product. This shortens the formula to:

$$C_{n+1} = C_n + \alpha(C_{out} - C_n) + \gamma \tag{2}$$

This formula can be rewritten in an analytical way (Equation 3) so it can predict an arbitrary time step instead of only steps of integer size (10 minutes). A simple explanation of this formula goes as follows - CO_2 always converges to a value L. The number of people in the office dictates this limit, while the value α dictates how fast we approach this limit. The inverse of this formula will also be useful and can be trivially computed using some basic algebra.

$$C_n = \begin{cases} \gamma^n, & \text{if } \alpha = 0 \\ L + (C_0 - L)(1 - \alpha)^n, & \text{otherwise} \end{cases} \tag{3}$$

$$L = C_{out} + \frac{\gamma}{\alpha}$$

Determining the window state is described in Section 2.1.1. If we know the α value for the current window state, the γ value becomes the only unknown in the formula and can be determined with a simple linear regression, using last three data points. Since γ correlates with the number of people in the office, it must be recalculated for every prediction. The α value on the other hand is dependent on the office heat insulation level, office size and windows size, and is therefore a constant. We can therefore estimate the α value by trying different values on the past two weeks of data and then select the one that has the lowest error rate when predicting - this is possible since when predicting on the past data, we already know what CO_2 value will be reached.

2.3 Temperature predictions

We used the same base formula - Equation 3 - for the inside temperature prediction. This model however, has to be made more complex because of two factors.

First, the temperature does not converge towards the outside one, but goes towards some function of the outside temperature instead. For example, even if the outside temperature is below zero, the temperature in the office never went below 10 degrees, even without heating. There are several reasons for this behavior, including the heat of the building itself, and the fact that building is warming and cooling at different rates than the exterior when the external temperature changes. This is dealt by calculating a function from last two weeks of data that models the expected inside temperature as a linear function of the outside temperature. The calculation is made during rest days, when no one is in the office, reducing the noise in the data. This calculated value then replaces the value C_{out} in the Equation 3.

Second, we have to account for both air conditioning and heating. The detection of their state is described in Sections 2.1.2 and 2.1.3. In the same section it is also described how to collect the limiting temperature value these devices generate. If either device is on, the corresponding limiting value replaces L in Equation 3. Improvement of this rather simplistic modeling of the devices is subject to future work. A prediction example is plotted in Figure 2.

2.4 Humidity predictions

Humidity was not changing much in our data, and when it did, there was no obvious pattern. So instead of plugging the data into the same equation, we used a classical machine learning approach. The last few humidity and temperature measurements, together with the window state are fed into a machine learning model, and a prediction for future humidity is given. Again the training of the model is made on the previous two weeks. If it turns out that the prediction underestimated the humidity in the office, the humidifier is determined to be active. If the classifier overestimates the humidity and humidifier was considered active, it is considered inactive from then on.

2.5 Recommendation system

Each ambient parameter has predefined quality ranges - good, medium and bad. For example: "good" CO_2 is under 500 ppm, "bad" over 800 ppm and "medium" in between. The ideal case is to have all three parameters in the "good" quality range. This, however, is not always possible as improving one parameter may damage another - opening the window may improve the CO_2, but it may reduce the temperature quality. The priority of the system is to have the minimum number of "bad" parameters. If all the parameters are "medium" or above, the maximum number of "good" parameters is prioritized.

A possible action is a change in one of the devices/windows that exist in office. In the current version, all the devices are assumed to be binary (air conditioner is either on or off, windows are opened or closed, etc.). The list of all possible actions is generated based on the current assumed state of the office. If the windows are assumed opened, "open the window" action will be omitted. Some hand-selected actions may appear in pairs, as they are commonly done simultaneously: turn on the air conditioner and close the windows for example. A default action "do nothing" is also included on the list.

Each action effect is simulated over the period of 30 minutes. The action that results in the best state after that time interval is selected. If the action has a higher score than the default action of doing nothing, it is recommended to the user.

While not fully implemented yet, there are two areas with possible improvements that are currently worked on. One is to try to make the recommendations more time-specific. Instead of "open the window", we could recommend "open the window for 7 minutes, then close again". This can be done by first determining all the relevant time frames - times where a parameter shifts from one quality range to another. All the possible actions can then be tested against every relevant time frame. This multiplies the number of combinations checked, but the total number is still reasonably low. Second is to predetermine which actions are even sensible, given the context. If the only problem is the temperature inside being too cold and it is also cold outside, then the sensible options are only to close the window or to turn on the heater. This is being implemented by an ontology that contains facts about some ambient parameters, configured in a way that is able to search for relevant actions given current state.

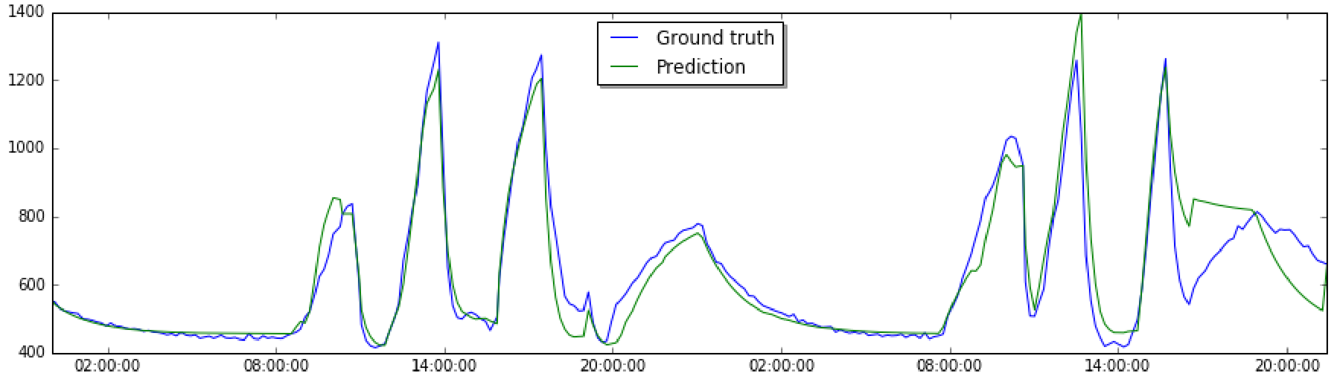

Figure 1: CO_2 prediction and ground truth, predicting values for the next two days, supposing that we have perfect information about the current and future office state.

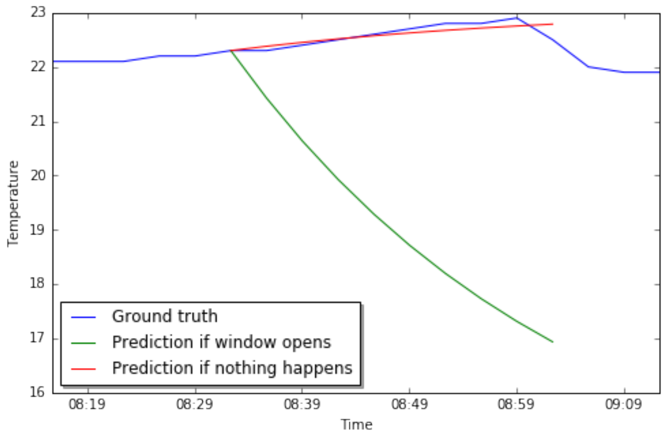

Figure 2: Temperature prediction and ground truth. We predict what happens if no action is done, against what happens if window is opened. The prediction starts in the past so we can compare it to the actual measurements.

3. RESULTS

As results we list (Table 1) the mean absolute error when predicting a parameter 30 minutes in advance, during a three month period. The test are made to be comparable with those in paper by Frešer et al. [3]. We show that our predictions for CO_2 and temperature display lower error then the before-mentioned work. Their humidity measurements were better, probably because of better selection of features in their model.

Table 1: Mean absolute error

Parameter	Our error	Error reported by Frešer [3]
CO_2 [ppm]	43	79
Temperature [°C]	0.36	0.50
Humidity [%]	1.2	0.74

4. CONCLUSIONS

In this work we model three ambient parameters in the office. For two of them, we show a simple mathematical model, that predicts their future behavior. For those two we get more accurate predictions than those in the related work. This is probably a consequence of using a physically-inspired formula. For humidity we use a machine learning model, that while showing promising results, still has room for improvement. We also predict the state of devices and windows in the office, although the accuracy of this prediction has not yet been directly tested. Furthermore we presented a recommendation system that we plan to test with multiple real offices in the future.

5. REFERENCES

[1] "netatmo. 2016. netatmo. (2016)". "https://www.netatmo.com". "Accessed: 2017-09-03".

[2] B. Cvetković, M. Gjoreski, M. Frešer, M. Kosiedowski, and M. Luštrek. Monitoring and management of physical, mental and environmental stress at work.

[3] M. Frešer, A. Gradišek, B. Cvetković, and M. Luštrek. An intelligent system to improve thc parameters at the workplace. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 61–64. ACM, 2016.

[4] L. Lan, P. Wargocki, and Z. Lian. Optimal thermal environment improves performance of office work. *REHVA European HVAC Journal*, 2:12–17, 2012.

[5] D. P. Wyon. The effects of indoor air quality on performance and productivity. *Indoor air*, 14(s7):92–101, 2004.