Occupancy Estimation in a room using $C0_2$ Sensor

Joanne Assaf Bou Saba[1*] Mateus Moreira[1*] Davi Jose Marques Vieira[1*]

[1] Univ. Grenoble Alpes, Grenoble INP, Ense[3], 38000 Grenoble, France

*joanne.assaf-bou-saba@grenoble-inp.org
*mateus.moreira@grenoble-inp.org
*davi-jose.marques-vieira@grenoble-inp.org

Abstract. The concept of smart house has gained increasing attention in the past years. Information technology tools have been adopted in modern buildings such as wireless sensor networks, data analytics, IoT i.e. The internet of things and machine learning concepts. By integrating the concept of smart buildings, we can achieve a perfect energy efficiency as well as electrical devices monitoring. In this experiment, we are going to investigate the occupancy in a room and estimate the number of people using a CO2 sensor network. The concentration of CO2 in a room is a good indicator of the number of people in this room. Data collection is going to be performed and we will be able to visualize all the data coming from the sensors using Grafana. We are going to visualize and analyze this data in order to correctly address this problem. Finally we will make a conclusion about this analysis.

Keywords CO_2 sensor, occupancy, data

1. Introduction

The measurement of the CO_2 concentration is a very important indicator to be able to predict the number of occupants in a room. In new smart buildings, this information is very important since this information is relevant for maintaining indoor quality standards and ventilation to be able to regulate HVAC systems and the ventilation rate based on the number of occupants.

This paper is organized as follows. First we are going to read the data from Grafana that will represent the CO_2 concentration variation in the different rooms of the house. After that, we will try to find a specific correlation between the concentration of CO_2 and the number of

occupants in a room and this correlation can be used later for further studies on ventilation systems control .

2. Data Visualisation Via Grafana

Grafana is an openSource platform tailored for monitoring, analysis and data visualization. It comes with a web server allowing access to it via an http API. under the Apache 2.02. Grafana is able to generate charts and dashboards from time series databases such as Graphite or InfluxDB. This platform is also an essential tool to create alerts.

In our case, we have CO_2 sensors installed in different rooms of the house. and through grafana we are able to access in real time the variation of the concentration. As mentioned before this information is very important since it indicated the number of people present in that room, and using this information, we can control the ventilation required and control HVAC systems.

After accessing the data from Grafana, we exported the data in a csv format. We used this csv file to generate a dataframe using python. Why dataframe ? A DataFrame is a python data structure similar to a database. Python generates powerful data analysis tools that are very powerful for data analytics when applied to real time data and provides us with go-to libraries for data manipulation, in order to conduct analysis that can help in decision making situations.

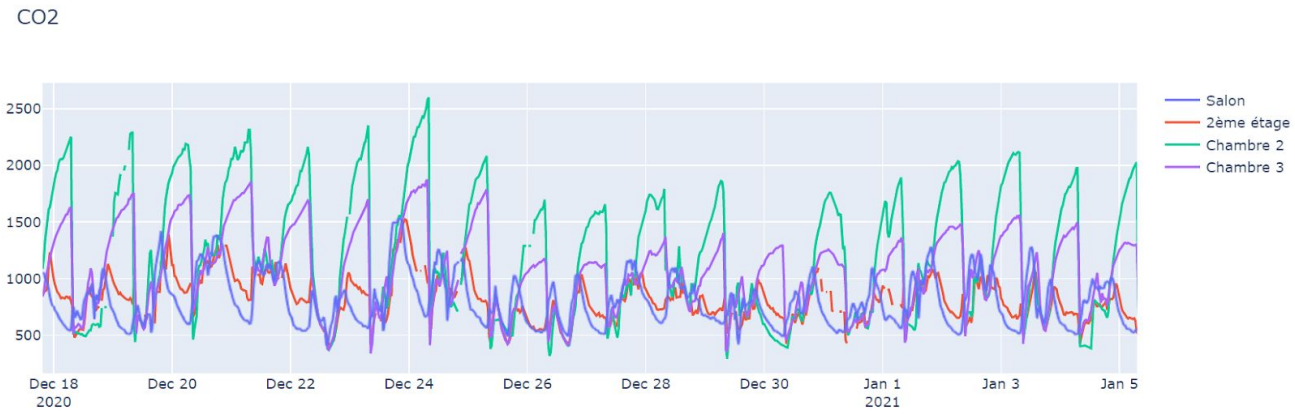

figure 1 : Real time variation of CO_2 concentration in different rooms of the house

The data collected in our case, uses a time step of 30 minutes. The data collected from the sensors at each time step, will enable us to observe the variation of CO_2 and therefore conclude about the number of people present in that room, since people exhale CO_2, the higher the concentration, the higher the number of people are present.

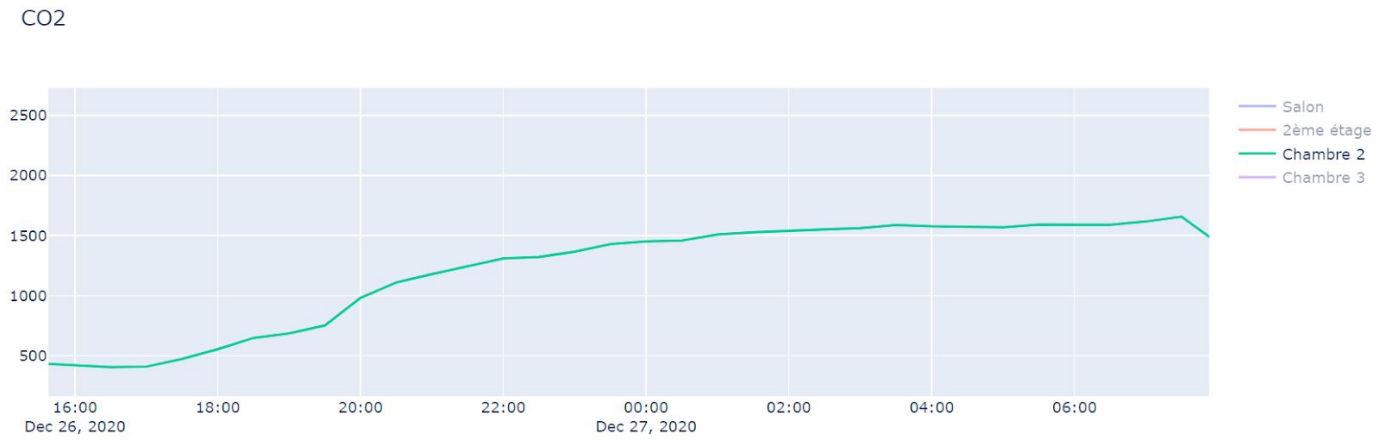

figure 2 : hourly variation of CO_2 concentration in Chambre 2

3. Correlation between CO_2 concentration and number of people

At first, it was important for us to understand the mathematical reasoning between the concentration of CO_2. Very much past studies worked on CO_2 data concentrations obtained with specific sensors in conjunction with some specific mathematical models aiming to detect occupancy in a room.

In our analysis, we are going to take a reference to previous studies and information to provide specific results about occupancy.

In all standards, CO_2 concentrations are expressed in ppm (parts per million). The reference value of a CO_2 concentration ranges between 390 ppm and the variation above this standard may provide information about the occupancy in the room.

In order to be able to present a correlation mathematical formula between the number of occupants in a room and the evolution of the CO_2 concentration over time we are going to define some standards values.

The CO_2 generation rate value per person can be estimated around $1.2\ m^3/h$ provided from literature[1] and given the concentration ofCO_2 in the air exhaled by a person which is $400\ ppm$ we can write the following formula.

$$Number\ of\ occupants\ = \frac{\frac{\Delta CO_2}{\Delta T} \times V + Q_{supply} \times (CO_2 - CO_{2\,air})}{Q_{person}\ CO_{2\,person}}$$

- ΔCO_2 : Variation of concentration of CCO_2 for each time step
- ΔT : Time step
- V : Volume of the room ($V_{chambre\ 2} = 3.2 \times 2.77 \times 2.5\ m^3$ and $V_{chambre\ 3} = 2.53 \times 2.6 \times 2.5\ m^3$
- CO_2 : Concentration of Co2 at time t
- $CO_{2\,air}$: Concentration of at the ambient air (390 ppm)
- $CO_{2\,person}$: Concentration of Co2 exhaled by a person in the air (400 ppm)
- Q_{supply} : the ventilation rate in a room (in our case it is 0)

We implemented this mathematical formula that includes a direct correlation between the number of occupants and the variation rate of Co2 into our python code in order to perform operations on our data and calculate the number of occupants present in the room.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7411428/#bib25

4. Data Processing

It's important to notice that the number of people in the 'Salon' and in the '2ème étage' was not analysed because theses places are an open space with many variables that would be difficult to implement in the code, besides that, the formula has a better result when used to study the concentration of CO_2 in small places.

[1] "Measurement of CO2 concentration for ... - NCBI - NIH." 7 Aug. 2020, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7411428/.

Once the number of people in the Chambres 2 and 3 had been computed with the raw data, it was evident that some imperfections should be corrected. It was visible the presence of some outliers, since at some moments the number of people inside these small rooms grew up a lot. We removed them using the Interquartile Range Method (between the 10th and the 90th percentile values)

After these processes, there were still few moments left when the calculations indicated a negative number of people.This may be due to the fact that for some moments there may be a flow of air out of the room, decreasing the concentration of CO2 sharply. Depending on the airflow, the calculated number of people may distance themselves from reality.

This highlights something that can be improved in this project, to take into consideration, besides motion sensors, doors and windows in these spaces so that, at least, we can ensure that if there are people or not and if the number of people obtained is reliable or not.

5. Results

After that all the data was treated, it's clear to see that the results are not 100% correct, because this data comes from a household where a family of 5 lives and the maximum number of people in the 'Chambre 2' between december 18th and january 16th was 7 people and in the 'Chambre 3' was 6 people.

figure 3 : Real time occupancy in Chambre 2 and Chambre 3 of the house

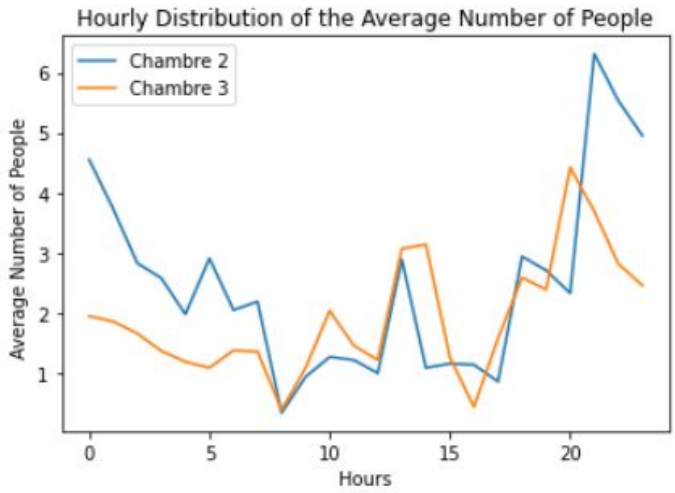

figure 4 : Average hourly occupancy in Chambre 2 and Chambre 3 of the house

In the decision making process that involves controlling some installation in a house, it is important to perform some data analysis that can give us the hourly distribution of the average number of people. Knowing that only 5 people live in that house, we have some imperfections and outliers data that are not very logical. This could be due to some technical imperfections in the sensors or some insufficient data needed to evaluate the number of occupants.

6. Conclusion

In this report, we showed the steps and the method to determine the number of people in a room using CO_2 data. The method used is basically based on the concentration of CO_2 in the room. However, it's important to highlight that other features may be important to this analysis, like if the windows and doors are open, causing natural ventilation and decreasing the concentration of the CO_2 in the rooms.

The study made for this project is still very simple, but it could be further analysed with Machine Learning techniques to try to discover the most important variables that can be affected by the occupation other than the CO_2 concentration. Given the different sensors present in a room, we are able to perform machine learning techniques based on the data collected by sensors and try to learn our model in order to predict occupancy and monitor

HVAC and ventilation systems that are very important in indoor spaces. With that further analysis, the occupancy in the rooms would probably be more accurately.