Managing Ventilation Systems for Improving User Comfort in Smart Buildings using Reinforcement Learning Agents

Jiawei Zhu[1], Fabrice Lauri[1], Abderrafiaa Koukam[1], and Vincent Hilaire[1]

IRTES-SET, UTBM, 90010 Belfort cedex, France
(jiawei.zhu, fabrice.lauri, abder.koukam, vincent.hilaire)@utbm.fr

Abstract. With the fast development of information technology and increasingly prominent environmental problems, building comfort and energy management become the major tasks for an intelligent residential building system. This paper identifies the system requirements of Smart Buildings, analyzes the problems that need to be solved and how Reinforcement Learning is suitable for dealing with them. It also proposes to represent parts of Smart Buildings as Cyber-Physical Systems. Although the global goal is to model and manage a complex and whole system of a Smart Building, since the work is in progress, in this paper we mainly focus on how Reinforcement Learning technique is good at controlling subsystems, specifically the Ventilation System. The experimental results show the advantages of our system compared with the widely used baselines: On/Off control and PI control approaches.

Keywords: energy, smart buildings, reinforcement learning, multi-agent system, cyber-physical system.

1 Introduction

According to United Nations Environment Programme [1], buildings use about 40% of global energy, 25% of global water, 40% of global resources, and they emit approximately 1/3 of Green House Gas (GHG) emissions. With the development of human society, environmental issues have drawn more and more attention. In this background, buildings can offer a great potential for achieving significant GHG emission reductions in different countries. Furthermore, energy consumption in buildings can be reduced by using advanced technologies and management. On the other hand, people spend greater part of their time in buildings. As the quality of life in home is increasingly considered as of paramount importance, many people constantly seek to improve comfort in their living spaces. Meanwhile, the popularization of the concept of home office makes the productivity in smart buildings economically significant. How to manage buildings in a proper way to improve energy efficiency and comfort level while reducing pollution at the same time is therefore a subject of uttermost importance.

Corresponding to the increasing demands for environment, comfort, energy, and productivity, advanced methods are applied for improving comfort conditions in smart buildings thanks to the dramatically rapid development of information technologies. Widespread utilization of computing devices, powerful but low cost sensors and actuators, and ubiquitous networks make the intelligent control more easily come true. Actually the implementation of smart buildings involves controls of different subsystems and devices. Hence itself is a system of systems.

Based on this context, Cyber-Physical System (CPS) can be used to model this complex system, which is integrations of computation, networking and physical processes, in which embedded computers and networks monitor and control the physical processes with feedback loops where physical processes affect computations and vice versa [2]. CPSs integrate the dynamics of the physical processes with those of the software and networking, providing abstractions and modelling, design, and analysis techniques for the integrated whole. Modelling and controlling smart buildings as CPSs can bring many advantages: different subsystems such as heating, ventilation and air-conditioning (HVAC) can communicate with other electrical devices to form an intelligent whole; more information can be integrated and shared, for example, real-time and forecasting local weather data from observatories can be used through networks to assist HVAC system to make better decisions or even to help power distributors to balance loads; the system can be more robust and the cost can be reduced by separating sensors and actuators from traditional electrical devices.

In this work, we try to reformulate smart buildings as CPSs and capitalise on Reinforcement Learning (RL) and Multi-Agent techniques to control the whole system. Due to the work being in progress, in this paper we mainly focus on modeling and controlling subsystems, specifically the ventilation system. Our contributions is threefold: firstly we identify the system requirements for smart buildings; then inspired from [3] we propose the method to model ventilation system as CPS; finally RL is proposed to control the ventilation system, its performance is analyzed and compared with PI control and On/Off control. The rest of this paper is organized as follows: Section 2 presents the related work. Section 3 analyzes system requirements for smart buildings. Section 4 focuses on RL and its feasibility. Section 5 investigates ventilation system and models it as CPS. Experimental results are presented and analyzed in Section 6. Finally we conclude in Section 7.

2 Related Work

Research is increasing in the emerging field of smart buildings. Kleissl *et al.* [4] regard smart buildings as CPS, and by examining different buildings and their energy use in detail they point out opportunities available to improve energy efficiency operation through various strategies from lighting to computing. However, the requirements for developing smart buildings and the architecture of the system have not been analysed and proposed. In order to improve comfort in buildings, authors of [5] propose an adaptive smart home system named CASAS, which utilizes machine learning techniques to discover patterns in occupant's daily activities and to generate automation policies that mimic these patterns. Although the user's explicit or implicit wishes can be adapted, the energy consumption of the building has not been taken into account. Actually energy efficiency is one of the key factors that need to be considered when designing smart buildings, since more comfort usually comes at the expense of higher energy consumption. Therefore these two conflicting points should be carefully balanced. In [3] and [6], thermal comfort and indoor air quality in buildings are improved separately by intelligent control methods while less energy is used compared with conventional controllers. However, a smart building is a system of systems, and only individual subsystems being well controlled is not enough. Hence, in our work we undertake the analysis and design from a global view, while implementing the system by a bottom-up approach.

Both to make subsystems have the ability to take intelligent decisions, and the global system learn good strategies to schedule and coordinate these subsystems, RL brings advantageous properties. Up to now, RL has been successfully used on a wide range of problems. Peters *et al.* [7] propose an intelligent decentralized control mechanism, which is able to operate in different Smart Electricity Markets, by using autonomous broker agents. These agents can accommodate arbitrary economic signals and learn efficiently over the large state spaces resulting from the signals with function approximation. After learning, they are capable of deriving long-term, profit-maximizing policies. Li *et al.* [8] present an improved MAXQ [9] method to minimize electricity costs on the premise of satisfying the power balance and generation limit of units in a microgrid and the proposed multi-agent architecture is beneficial to handle the problem of "curse of dimensionality" and speed up learning in the unknown large-scale world. Other works focusing on robotic and traffic light control can be found in [10,11] and [12,13] respectively.

3 System Requirements for Smart Buildings

Actually a Smart building is a system of systems. It requires to think about different subsystems and devices as an integrated whole that has a global objective. Different functional parts of this whole can be modelled as agents so that the individual objectives and the global objective can be reached by cooperating, coordinating, and negotiating among these agents. Therefore, the smart building can be regarded as a multi-agent problem. On the other hand, improving the comfort of a single building is important whereas reducing energy consumption so as to reduce electric bill, balancing power distribution, and shifting peak load are also vital. So the desired strategy is to reasonably balance comfort and energy consumption not only in one building but also among different buildings of a district. In order to model this complex system, we first need to analyse the requirements for smart buildings.

3.1 Multi-Authority and Multi-Level

A smart building comprises various agents, such as air-conditioning agent, ventilation agent, water-heating agent and so on. This forms multi-authority of lower level. If there does not exist communication between agents, each of them can be considered as selfish, that is each of them try to maximize its individual goal regardless of the states of other agents and the global objective. From a higher viewpoint, each smart building can be regarded as an independent agent and some of these agents constitute a smart district, in which smart buildings can improve their comfort, balance total grid load and reduce energy consumption through communication, cooperation and coordination. This is multi-level requirement for smart buildings.

3.2 Multi-Objective

Authorities of different levels want to reach their own objectives and global objectives of higher level. For instance, thermal comfort, indoor air quality, and visual comfort are three basic factors which determine the comfort conditions in buildings [14], and relative devices treat improving these comforts as their individual objective, while they also need to consider global objective: reducing energy consumption and peak load shifting. Individual objective and global objective are often conflicting: improving comfort often means consuming more energy.

3.3 Heterogeneity

Heterogeneity in smart buildings mainly comes from three aspects. First, the devices in buildings are diverse with different control strategies. How to integrate them together and have a proper management is important. Moreover, multi-level structure, that has been presented in Section 3.1, brings difficulties in designing the system. In addition, different occupants have different user preferences, including comfort definition, device type, and their physical activities in buildings. Electrical devices in buildings are heterogeneous. In general, they can be divided into two categories: power consuming devices and power producing devices. However, in [15] these two categories are unified by a new word called prosumer which means either producing or consuming. By convention, a positive prosumption represents a production and a negative one a consumption. But in order to analyse device properties we still use former notions.

Power Consuming Devices In this category, devices consume power when they are working and they are divided as negotiable and non-negotiable. Negotiable devices are these who can reduce working power, called power-negotiable (*e.g.* intelligent air-conditioning, intelligent ventilation system), so as to reduce comfort within a range that occupants could accept; who can postpone scheming start-time, named time-negotiable (*e.g.* washing machine), in order to shift peak load; who can both reduce working power and postpone scheming start-time, called power-time-negotiable (*e.g.* water heating system, storage system), to provide a flexible service. Non-negotiable devices are inflexible, that means when people turn them on, they always consume power as required (*e.g.* daylight lamp, TV, computer).

Power Producing Devices In smart buildings, there often exist power producing devices to make full use of green energy and help decrease the load on main grids. Some of these devices can provide constant power like fuel cells, micro turbines, and storage systems, while the others can merely offer variable outputs, which strongly depend on weather conditions, such as photovoltaic panels and wind turbines. At this time, storage systems are required for these devices to play a role as buffers, which can achieve constant outputs to protect the micro-grid.

3.4 Scalability

The desired system should be scalable. In a single building, devices often plug in and out, and in a district, new buildings may participate. This requires the architecture we design have the ability of scalability and the decision-making algorithms need to be decentralized.

3.5 Incremental Change

Although smart buildings can bring numerous benefits, traditional buildings with traditional devices already exit. Hence any changes introduced in the future should be reasonably gradual so as not to disturb and damage the working system and its service. This requires the designing system can tolerate and integrate the exiting traditional devices.

4 Reinforcement Learning

Compared with traditional control, CPSs enable consider more inputs to better realize the dynamic of the physical world so as to support decision making. For example, nowadays most air-conditionings use On/Off and PI control. On/Off control regulates temperature by using a compressor that is periodically either working at maximum capacity or switched off entirely, whereas PI control has a variable-frequency drive that incorporates an adjustable electrical inverter to control the speed of the motor and thus the compressor and cooling output. The inputs for On/Off and PI control are only indoor temperature. However, for CPSs they can capitalize on more inputs like occupant number, since bodies are also heat source that can increase indoor temperature. In order to benefit from CPSs, it seems that traditional control methods, which are often straightforward, are not enough.

Hence, we advocate the use of RL to control smart buildings, which offers a suitable set of techniques to address these challenges. The classical reinforcement learning framework is based on Markov Decision Processes (MDPs). An MDP can be depicted by a tuple (X, U, f, ρ, γ). X is the set of states it can perceive, U is the set of possible actions it can perform in these states, f is the state transition function, ρ is the reward function that evaluates the immediate effect of an action, and γ is the discount factor. The goal of RL is to find an optimal policy, $h : X \to U$, that maximizes the return from any initial state x_0: $R^h(x_0) = \sum_{k=0}^{\infty} \gamma^k \rho(x_k, h(x_k))$, where $\gamma \in [0, 1)$ and k is discrete time step. The discount factor can be interpreted intuitively as a measure of how "far-sighted" the controller is on its rewards, or as a way of taking into account increasing uncertainty about future rewards [16]. In order to characterize policies, state-action value function (Q-function) is used, $Q^h : X \times U \to \mathbb{R}$. After finding an optimal Q-function $Q^*(x, u)$, optimal policy $h^*(x)$ can be obtained greedily by $h^*(x) \in \arg\max_u Q^*(x, u)$. In this work, the model-free online algorithm Q-learning [17] is used to update the Q-function. The choice of this algorithm is motivated by the fact that no explicit model of the dynamics of a smart building is available, due to the great amount of involved devices. Moreover, some inputs like the number of occupants presented in the room are random. Hence, considering f as a stochastic function is more realistic. Q-learning can work as a sample-based algorithm to deal with stochastic approximation procedure. The Q-function is updated online at every new sample of the form $(x_k, u_k, x_{k+1}, r_{k+1})$, using the following equation: $Q_{k+1}(x_k, u_k) \leftarrow Q_k(x_k, u_k) + \alpha_k[r_{k+1} + \gamma \max_{u'} Q_k(x_{k+1}, u') - Q_k(x_k, u_k)]$. In subsystems, RL can be utilized as intelligent controller to control devices such as heating and ventilating systems. In global system, RL can optimize the coordination of subsystems.

5 Ventilation Controlling Subsystem

In this section, we mainly focus on the ventilation subsystem. A thermal subsystem has been investigated in our previous work [18].

The ventilation controlling system is used to improve indoor air quality. In most cases, people can obtain a good indoor air quality by simply opening windows. However, in some situations, we need mechanical ventilating devices to exchange indoor air, for example, when there are many visitors in the room, when outdoor air speed is close to zero, when there is no window in the room, and for people who live in modern skyscrapers in which opening windows will raise the possibility of hidden danger and hence is often forbidden.

The indoor air quality is mainly decided by CO_2 concentration. People generate CO_2 and consume oxygen, at a rate that depends primarily on their body size and their level of physical activity[19]. The rate of oxygen consumption in L/s of a person is given as

$$V_{O_2} = \frac{0.00276 A_D M}{(0.23 RQ + 0.77)} \tag{1}$$

where RQ is the respiratory quotient (the relative volumetric rates of carbon dioxide produced to oxygen consumed). M is the level of physical activity or the metabolic rate per unit of surface area in met (1 met = 58.2 W/m^2). A_D is the DuBois surface area in m^2, which can be estimated by the following equation

$$A_D = 0.203 H^{0.725} W^{0.425} \tag{2}$$

where H is the body height in meter and W is the body mass in kg. For an average size adult, A_D is about $1.8 m^2$.

The value of RQ depends on diet, the level of physical activity and the physical condition of the person. It is equal to 0.83 for an average size adult engaged in light or sedentary activities (about 1 met), and increases to a value of about 1 for heavy physical activity (about 5 met). The carbon dioxide generation rate in L/s of an individual is

$$V_{CO_2} = V_{O_2} \times RQ \tag{3}$$

Steady state CO_2 concentration can be determined for a given ventilation rate based on a single zone mass balance analysis. Assuming that in a room there are N adults and the room is equipped with an electric fan. The mass balance of CO_2 in the room can be expressed as follows:

$$V \frac{dC}{dt} \times 10^{-6} = G \times 10^{-3} + Q \times (C_{out} - C) \times 10^{-6} \tag{4}$$

where V is building volume in m^3, C is indoor CO_2 concentration in $ppm(v)$, C_{out} is outdoor CO_2 concentration in $ppm(v)$, t is time in second, G is indoor CO_2 generation rate in L/s, and Q is ventilation rate in m^3/s. Generally an acceptable value of indoor CO_2 concentration varies from 600 $ppm(v)$ to 1000 $ppm(v)$, and 800 $ppm(v)$ is set as a reasonable setpoint for a good indoor air quality.

Figure 1 depicts the architecture of the CPS for ventilation control. In this figure, there are three parts: Physical World, Network, and Cyber World. Physical World contains Sensor Domain and Actuator Domain. Different indoor and outdoor parameters, such as air speed, CO_2 concentration, occupant number, metabolic rate, etc., can be observed by different sensors and these data are transmitted to Cyber World though the network. In Cyber World, we use RL because it can deal with the random appearance of occupants. With this algorithm, the captured data are used to decide the current state and the reward of last time. RL can make good decisions automatically by trial and error without the need of a specific model of the problem. Based on this technique multiple devices in Physical World are controlled by the actuating signals from Cyber World.

6 Experiments

The goal of this work is to control the mechanical ventilation system to keep the CO_2 concentration at the setpoint while reduce energy consumption. Specifically, according to the present information of occupants' number and indoor

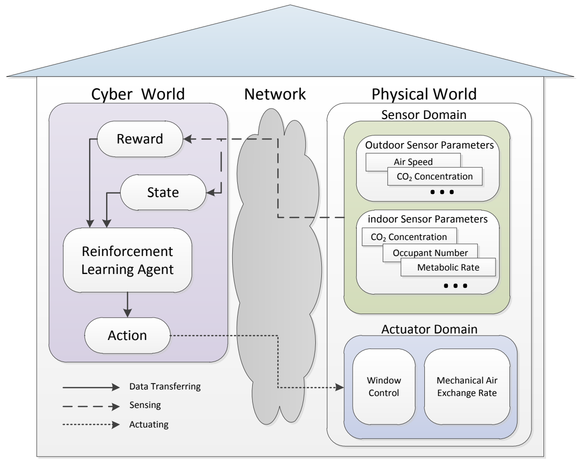
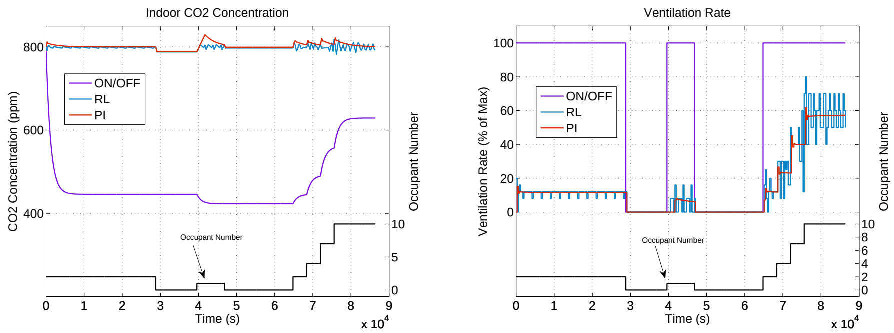


Fig. 1. Cyber-Physical System for ventilation control

CO_2 concentration, the system should adjust the electric fan's speed to control the ventilation rate. We assume the maximum number of occupants is 10 with average size and doing light or sedentary activities in a house of $100m^2 \times 3m$. The CO_2 concentration in $[750, 1200]$ is discretized into 450 states, plus 2 over boundary states. Therefore the total number of state is $452 \times 11 = 4972$. The mechanical ventilation system used has a maximum ventilation rate of $0.25m^3/s$ with power of 40 W and can take 13 actions: $\{0,4,8,12,16,20,25,30,40,50,60,70,80,100\}\%$ of the maximum ventilation rate. The reward function is $r_{k+1} = e^{-\frac{(C_k-800)^2}{20000}} \times 8 - 8$, where C_k is the indoor CO_2 concentration at time step k. Due to the slow variation property of CO_2 concentration, the time step is set to 300 seconds.

Figure 2(a) compares the CO_2 concentration variations within one day by three different control methods. The black line is the occupant number change during this period of time. The result indicates that the CO_2 concentration can be unnecessarily reduced far below 800 *ppm* by On/Off control, which simply turns on the fan with maximum power if any occupants are detected in the room while turns off if not. Although it can provide continuously fresh air flow, it consumes much more energy than the others. PI control (proportional gain: 0.003, integral gain: 0.000001) can keep the CO_2 concentration at the setpoint smoothly, except for every change of number of occupant's presence, which causes the overshoot of the concentration. RL method can offer the best comfort, even though the CO_2 concentration has small vibration, that is caused by the discrete definition of actions and occurs often in RL applications. For residential buildings, this slight fluctuation will not affect inhabitants' comfort. The comparison of ventilation rates is presented in Figure 2(b). It reflects that when there are more occupants in the room, the quicker dynamic of physical environment makes



(a) Indoor CO_2 Concentration (b) Ventilation Rate

Fig. 2. Experimental Results

it more challenging for RL to control. The total energy spent is 0.1379 kWh/day by PI control, 0.1403 kWh/day by RL, and 0.6401 kWh/day by On/Off control. Compared with On/Off control, RL can save 78.08% energy, and compared with PI control, although RL use 1.74% more energy, it is not only able to maintain good indoor air quality but also more suitable and feasible for implementing CPSs in smart buildings.

7 Conclusion

In this paper, we identified the system requirements for smart buildings, including multi-authority and multi-level, multi-objective, heterogeneity, scalability, and incremental change. Then we presented the framework of Reinforcement Learning and explained why RL is suitable to resolve most of the smart building challenges. After that, a subsystem, specifically a ventilation system, was investigated and modeled by a CPS approach. The experimental results revealed that Q-Learning, a model-free online RL technique, is more adaptable and feasible than conventional PI and On/Off approaches for managing a ventilation system in a smart building to improve comfort level while reduce energy consumption.

In the future, function approximation will be utilized, since there are more information (input variables) available for CPSs and often these input variables are continuous, so it will not be applicable to discretize them anymore. In addition, various RL techniques will be compared and analyzed to find their applicability for smart building management. After that, different subsystems in smart buildings will be integrated together based on Multi-agent System approach.

References

1. UNEP. http://www.unep.org/sbci/AboutSBCI/Background.asp

2. Lee, E.A.: Cyber physical systems: Design challenges. In: 11th IEEE International Symposium on Object Oriented Real-Time Distributed Computing (ISORC), IEEE (2008) 363–369
3. Cheng, Z., Shein, W.W., Tan, Y., Lim, A.: Energy efficient thermal comfort control for cyber-physical home system. In: IEEE International Conference on Smart Grid Communications (SmartGridComm). (Oct 2013) 797–802
4. Kleissl, J., Agarwal, Y.: Cyber-physical energy systems: Focus on smart buildings. In: 47th ACM/IEEE Design Automation Conference (DAC). (June 2010) 749–754
5. Rashidi, P., Cook, D.: Keeping the resident in the loop: Adapting the smart home to the user. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans **39**(5) (Sept 2009) 949–959
6. Wang, Z., Wang, L.: Intelligent control of ventilation system for energy-efficient buildings with co2 predictive model. IEEE Transactions on Smart Grid **4**(2) (June 2013) 686–693
7. Peters, M., Ketter, W., Saar-Tsechansky, M., Collins, J.: Autonomous data-driven decision-making in smart electricity markets. In: Proceedings of the 2012 European Conference on Machine Learning and Knowledge Discovery in Databases. (2012) 132–147
8. Li, F.D., Wu, M., He, Y., Chen, X.: Optimal control in microgrid using multi-agent reinforcement learning. ISA Transactions **51**(6) (2012) 743 – 751
9. Dietterich, T.G.: Hierarchical reinforcement learning with the maxq value function decomposition. Journal of Artificial Intelligence Research **13**(1) (nov 2000) 227–303
10. Smart, W., Kaelbling, L.: Effective reinforcement learning for mobile robots. In: IEEE International Conference on Robotics and Automation. Volume 4. (2002) 3404–3410
11. Mataric, M.: Reinforcement learning in the multi-robot domain. Autonomous Robots **4**(1) (1997) 73–83
12. Abdulhai, B., Pringle, R., Karakoulas, G.J.: Reinforcement learning for true adaptive traffic signal control. Journal of Transportation Engineering **129**(3) (2003) 278–285
13. Wiering, M.: Multi-agent reinforcement learning for traffic light control. In: ICML. (2000) 1151–1158
14. Wang, Z., Yang, R., Wang, L.: Multi-agent intelligent controller design for smart and sustainable buildings. In: 4th Annual IEEE Systems Conference. (April 2010) 277–282
15. Frey, S., Diaconescu, A., Menga, D., Demeure, I.: A holonic control architecture for a heterogeneous multi-objective smart micro-grid. In: IEEE 7th International Conference on Self-Adaptive and Self-Organizing Systems (SASO). (Sept 2013) 21–30
16. Busoniu, L., Babuska, R., De Schutter, B., Ernst, D.: Reinforcement learning and dynamic programming using function approximators. CRC Press (2010)
17. Watkins, C.J., Dayan, P.: Q-learning. Machine learning **8**(3-4) (1992) 279–292
18. Zhu, J., Lauri, F., Koukam, A., Hilaire, V., Simoes, M.G.: Improving thermal comfort in residential buildings using artificial immune system. In: IEEE 10th International Conference on Ubiquitous Intelligence and Computing(UIC). (Dec 2013) 194–200
19. Emmerich, S.J., Persily, A.K.: State-of-the-art review of co2 demand controlled ventilation technology and application. US Department of Commerce, Technology Administration, National Institute of Standards and Technology (2001)