
A Reinforcement Learning based HVAC Control for Thermal Comfort and Energy Efficiency in Office Buildings

Zhiang Zhang

Siliang Lu

Chenlu Zhang

1 Overview of project idea

1.1 Importance

Occupants' comfort and energy efficiency are the most important evaluation criteria for the modern office buildings. Among all comfort metrics, thermal comfort is ranked by building occupants as the most important one and has more impact on the overall comfort.[1] However, conventional control methods of the heating, ventilation, and air conditioning (HVAC) systems often leads to uncomfortable thermal environment and high energy consumption. The project proposes to control the HVAC systems by employing existing deep reinforcement learning algorithms, with the goal of improving occupants' thermal comfort and HVAC energy efficiency.

1.2 Challenges

Building is a complicated engineering system because of its hardly-predictable and slow thermal behavior as well as dynamic boundary conditions (e.g. weather variation). It becomes even more convoluted when occupants' thermal comfort is involved since thermal comfort is subjective, varies person-by-person and is affected by both environment conditions (e.g. indoor air temperature, air velocity, etc.) and occupant conditions (e.g. clothing insulation, metabolic rate, etc.)[2]. An effective control algorithm of HVAC systems should be able to deal with disturbances from weather variation, unclear building thermal behavior and changing occupants' thermal sensations to deliver a comfortable indoor environment.

In addition, office's thermal environment is non-uniform but HVAC can only be controlled at the room level. For example, people sitting near the window may experience higher air temperature than people sitting in the core area of the room in summer, but there is only one HVAC air diffuser that can be controlled for the room. An ideal control algorithm should be able find an optimal control decision that minimize the thermal discomfort for all people in the room.

In conclusion, the challenges in term of reinforcement learning algorithm implementation are:

- State and reward design: the state should be designed carefully to represent the building and occupant which are both slow and highly non-linear systems. Also, reward should be able to fulfill the goal of minimum thermal discomfort for all occupants in the same room and minimum energy consumptions.
- Possibly large state spaces: depending on the state design, the state is expected to be large since it should includes all important factors that affect the thermal comfort and building environment. Most of state variables are inherently continuous variables. Moreover, since thermal comfort is location and individual dependent, state space may increase very quickly with number of occupants in the room.

1.3 Simulator

EnergyPlus [4] will be the simulator for the project to simulate the building thermal dynamics and HVAC energy consumption. EnergyPlus is a widely used building simulation tool developed by the Department of Energy and has been validated against the U.S building simulation standards, such as ANSI/ASHRAE Standard 140-2011.

Predicted Mean Vote (PMV) [2] model will be used to simulate the occupants' thermal comfort. PMV is an empirical model that builds relationship between humans' thermal comfort and measurable parameters such as air temperature, metabolic rate, etc.

1.4 Performance Metrics

The HVAC control algorithm using reinforcement learning will be evaluated regarding energy efficiency of HVAC systems and thermal comfort level of the occupants. More specifically, through the EnergyPlus simulation, the metrics are energy consumption of the HVAC systems (in kWh) and the PMV value of the occupants.

2 Literature Review

The application of reinforcement learning to building HVAC control is not yet well explored in building research. Liu and Henze [5, 6] implemented a Q-learning algorithm with Neural Network as the Q function approximation method to control the global building temperature setpoint and active thermal storage charge/discharge rate. The objective is to reduce the HVAC operation cost during the cooling season. Through a simulation study based on one level of a building, it was found that reinforcement learning can achieve 22% cost saving compared to the conventional rule-based control. However, the authors mentioned that the learning period for the reinforcement learning is too long to be realistic for the real-life application. A similar application of reinforcement learning was reported in [7], where authors use fitted Q iteration method to find the optimal action for the heater's on/off. The objective is also cost saving of HVAC operation. The algorithm was tested in a simulation model of a room, and it was found after 20 days of learning, the algorithm can reach 90% of the performance of a MPC algorithm (viewed as the "true" optimal solution). Dalamagkidis [8] implemented a linear reinforcement learning controller (using linear functions to approximate the value function) based on TD learning algorithm to control the heat pump, air ventilation system and window opening. The authors found the reinforcement learning controller can only achieve similar energy efficiency and thermal comfort performance with the conventional controllers, but takes significant time to learn. It can be found that there is no agreement on the performance of the reinforcement learning controller in the HVAC systems. In this project, the team plans to build the work based on the above mentioned researches, and try to evaluate the feasibility of the reinforcement learning control for the building HVAC systems.

3 Milestone

By the second milestone:

- Determine the suitable state, reward and action design;
- Build the test-bed environment based on EnergyPlus;
- Design and test a relatively simpler reinforcement learning algorithm (e.g. Q-learning) with simpler representation of the state/reward/action (e.g. discrete state, only 1 occupant in the room).

By the end of the project:

- Build and test a sophisticated deep reinforcement learning algorithm for PMV-based HVAC control with sophisticated state/reward/action representation (e.g. continuous state) and compare the result with the previous one.
- Build a model that multiple occupants (each occupant with different clothing insulation, metabolic rate, etc.) are in an office room at the same time. Also, artificial stochasticity

will be added to the thermal comfort of each occupant. Under such simulators, test the performances of both sophisticated and simple algorithms.

4 Computation resources

1) GPU computing resource for reinforcement learning; 2) Computer clusters for EnergyPlus simulation.

References

- [1] Frontczak, Monika, and Pawel Wargocki. 2011. "Literature Survey on How Different Factors Influence Human Comfort in Indoor Environments." *Building and Environment* 46 (4). Elsevier Ltd: 922–37. doi:10.1016/j.buildenv.2010.10.021.
- [2] FANGER, P O. 1970. *Thermal comfort. Analysis and applications in environmental engineering*. Copenhagen: Danish Technical Press.
- [3] de Dear, Richard, and Gail Brager. 1998. "Developing an Adaptive Model of Thermal Comfort and Preference." *ASHRAE Winter Meeting* (San Francisco, CA, USA: ASHRAE).
- [4] Department of Energy. "EnergyPlus." Software application. <https://energyplus.net>
- [5] Liu, Simeng, and Henze, Gregor P. 2007. "Evaluation of Reinforcement Learning for Optimal Control of Building Active and Passive Thermal Storage Inventory." *Journal of Solar Energy Engineering*, Vol. 129, May 2007.
- [6] Liu, Simeng, and Henze, Gregor P. 2006. "Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory Part 2: Results and analysis." *Energy and Buildings* 38 (2006) 148-161.
- [7] Costanza, G.T., Iacovella, S., et al. 2016. "Experimental analysis of data-driven control for a building heating system." *Sustainable Energy, Grids and Networks* 2016(6):81-90.
- [8] Dalamagkidis, K., Kolokotsa, D., Kalaitzakis, K., Stavrakakis, G.S. 2007. "Reinforcement learning for energy conservation and comfort in buildings." *Building and Environment* 42(2007) 2686-2698.