Data-driven Reinforcement Learning for Smart Controllers in Large Building Facilities

B. Schubnel, R. Carillo, A. Hutter, P.-J. Alet

Today, buildings consume almost half of the primary energy in Switzerland, 30% of which is used for heating, air conditioning and domestic hot water [1]. Making smart controllers that optimize energy consumption and comfort is part of the way to reduce this primary energy consumption. Within the European project THERMOSS, CSEM is building the next generation of industrial smart building controllers using a combination of simulation, reduced modelling and reinforcement learning techniques.

CSEM has a longstanding experience in designing smart controllers for building automation, in particular using model predictive control (MPC). Within the THERMOSS project, we used the recent advances in deep learning to simulate and design a smart controller for a large office building.

This building, located on the EPFL campus, has 100 rooms of various sizes and orientations, a centralized and fluid-based heating/cooling unit, and two centralized ventilation units that each provide air supply to half of the rooms. In each room valves control the supply of heating/cooling fluid, and blinds can reduce solar gain. The diversity in room configuration (types of valves, number of blinds per rooms, variable or constant air supply) makes the modelling and controlling problem particularly cumbersome. Creating with existing tools (e.g., EnergyPlus) a realistic simulation of the whole building dynamics would require a lot of efforts. On the other hand, it is not possible to try new control strategies directly on the building site so creating a realistic model is a necessity.

With an objective to build smart controllers that are sufficiently general and robust to be embedded in industrial applications, we have investigated the following approach to modelling the building.

First we have constructed a simple simulation of the building using the DIMOSIM suite developed by CSTB, a partner in the THERMOSS project. In line with our replicability objective, we tried and minimized the hand-crafted feature work. In particular, we did not try to fine-tune building parameters, shapes and devices. Instead of modelling each room in the building, we have modelled only a few room classes using the relevant attributes: size, orientation, floor location and types of inputs to control the rooms (valves, blinds). Our principal constraint was to have the same inputs and outputs for every room in the simulated model and in reality.

We have then trained a neural network to mimic the dynamics of the simulated building, and retrained the neural network with real data from the site (four months of data, sampled every ten minutes) to have a better fit to reality. We have used careful retraining techniques to keep as much as possible the physics learnt with the simulation and to avoid overfitting the model to the narrow band of real observed data and setpoints.

The main advantage of our approach is to remove the process of fine-tuning of parameters and architectural design. Instead this part is automatically done by the neural network via retraining on real data. Moreover, we can produce different models by varying the length of the retraining episodes and the learning rates, fitting more or less closely the reality. These different models can be used to make the building controller robust.

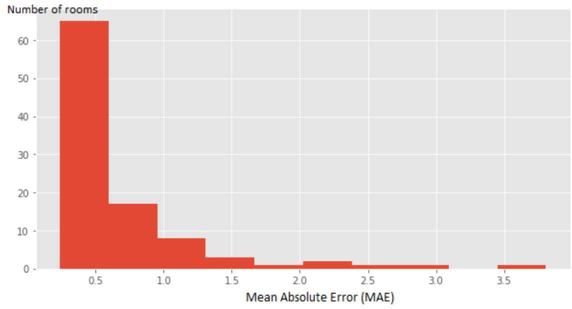

Figure 1: Distribution of mean absolute error in the room temperatures predicted by the model (trained on a separate range of dates) in June. Some errors are high (>1.5 °C) because of sensor problems in the rooms.

In a second step, we are developing a controller using state of the art techniques for continuous and discrete action space control. In every room, we have a local controller that controls the blinds (and/or the valves) and a global controller that controls ventilation and heating/cooling supply. The global controller works in continuous action space, whereas the local controllers have discrete action spaces.

Our controllers are modelled with neural networks and trained with the most recent reinforcement learning techniques (in particular DQN [2], DDPG [3] and PPO [4]) on our offline models of the building. We are currently investigating which of these algorithms are the most robust and need less fine-tuning of hyper-parameters. One of these algorithm, PPO, has recently proven to be easily transferable to real situations for complex robotics control tasks and has been used to construct collaborative agents in video games. It can be used for continuous and/or discrete action spaces but is less data efficient than the other two algorithms. Work is ongoing to address these issues for the THERMOSS project.

This work was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 16.0106.

[1] http://www.bfe.admin.ch/themen/00507/00607/index.html?lang=en

[2] V. Mnih, *et al*., “Playing Atari with Deep Reinforcement Learning,” presented at NIPS Deep Learning Workshop, 2013.

[3] T. P. Lillicrap, *et al*., “Continuous control with deep reinforcement learning,” arXiv:1509.02971 [cs, stat] (2015).

[4] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, “Proximal Policy Optimization Algorithms,” arXiv:1707.06347 [cs] (2017).