Distributed Reinforcement Learning for Building Energy Optimization

[Introduction]

Heating, ventilation, and air-conditioning (HVAC) systems are ubiquitous and are one of the major components responsible for energy consumption in a typical building system. In light of the imminent energy crisis, there is an increasing demand to revisit the building systems and save as much energy as possible. In this regard, the scope of the thesis is to explore opportunities for data-driven optimization of HVAC systems. Traditionally, optimization in HVAC systems has relied on offline optimization requiring domain expertise to schedule a set of optimal controls. Thus, data-driven optimization approaches such as reinforcement learning (RL) are more appealing due to their ability to adapt online and model and solve complex problems.

[Main Text]

A primary objective of the thesis is to carry out exploratory data analysis experiments to quantify the savings potential and expose the optimization space for RL. The main objective is to explore distributed reinforcement learning via multi-agent RL strategies (MARL) strategies and compare and contrast the pros and cons of MARL with single-agent RL. The work benchmarks two of the popular contemporary MARL strategies, centralized training and decentralized execution, and value-mixing approaches, along with proposing two novel MARL enhancements in HVAC systems: a linear value-mixing strategy (inspired by Q-function mixing, QMIX) and turn-based games, that attempt to alleviate some of the problems of multi-agent credit assignment and non-stationarity surrounding MARL.

The experimental results include the learning performance of various RL strategies and the performance benchmarks against the closed-loop controller under realistic conditions. The experimental results reveal that the RL strategies perform significantly better than the closed-loop controller (with a few exceptions), achieving power savings of up to 15% on yearly simulations with live weather profiles. The results also highlight the tradeoffs between optimality and sampling efficiency, further corroborating the prejudice about MARL, where the single-agent RL performs better in terms of optimality, while the MARL approach displays faster learning.

Overall, the work was an attempt to employ a non-linear control strategy such as RL in HVAC systems. Although the results seem very optimistic, there are several aspects to implementing the RL learnings real world. A primary concern in RL is safety, future work may involve more rigorous RL problem formulation in honoring the physical constraints and exploring constrained RL optimization.