12th IFAC/IEEE Workshop on Discrete Event Systems
Cachan, France. May 14-16, 2014

Sample Path Sharing in Policy Improvement for Indoor Air Temperature Control ⋆

Di Wu and Qing-Shan Jia *

* *Center For Intelligent and Networked Systems (CFINS), Department of Automation, Tsinghua University, Beijing, 100084, China (e-mail: woody10074026@gmail.com, jiaqs@tsinghua.edu.cn)*

Abstract: Stochastic simulation is a commonly used tool in building energy management and indoor air temperature control. We consider an HVAC system, where the speed of the fan is controlled to minimize a weighted sum of the energy cost and the thermal comfort of the occupants. The problem is formulated as a Markov decision process. Simulation-based policy improvement (SBPI) is used to improve from a given base policy. An important issue is to allocate the limited computing budget among action candidates to find the best action with the highest probability. We make two major contributions in this paper. First, we use event-based policies to handle the large state space. Second, we develop a sample path sharing method and combine it with equal allocation, successive rejects and optimal computing budget allocation. Numerical results show that, through sample path sharing, these methods can output policies with better performance than using them alone.

Keywords: Discrete event dynamic systems, simulation-based optimization, computing budget allocation, sample-based optimization

1. INTRODUCTION

Buildings are responsible for 40% of the energy consumption in developed countries (Djuric et al., 2007), and around 30% in China (Shui et al., 2009). In particular, heating, ventilation, and air conditioning (HVAC) systems account for nearly 40% among all the energy consumption in buildings (Sun et al., 2013). It is therefore of great practical interest to optimize the HVACs to achieve a reasonable tradeoff between energy reduction and indoor comfort.

Due to the complicated dynamics in HVACs, the system performance is usually evaluated by simulation models instead of closed-form expressions. Since it is generally difficult, if not impossible, to obtain exact solutions to such optimization problems, many studies have been conducted to solve simulation-based policy optimization approximately. A brief review will be provided in section 2. Among these methods, simulation-based policy improvement (SBPI) is an important candidate. The basic idea of SBPI is to estimate the Q-factors, which are the total cost for using an action candidate for the current state and following a base policy in the rest of the stages. The action that can minimize these Q-factors is selected. When the system evolves to the next stage, similar procedure can be applied.

⋆ This work was supported in part by the National Science Foundation of China under grants (Nos. 60704008, 60736027, 61174072, 61222302, 90924001, 91224008 and U1301254), the Specialized Research Fund for the Doctoral Program of Higher Education (No. 20070003110), the Tsinghua National Laboratory for Information Science and Technology (TNLIST) Cross-discipline Foundation, and the National 111 International Collaboration Project (No. B06002).

It has been shown that if the aforementioned Q-factors can be accurately evaluated, then SBPI guarantees to output a policy that is no worse than the base policy (Bertsekas and Castanon, 1999). However, in many practical problems, these Q-factors can only be evaluated using stochastic simulations. Jia (2012) showed that SBPI cannot guarantee an improvement using estimated Q-factors. Therefore it is of great practical interest to allocate the computing budget to improve the probability for correctly selecting the best action for each state.

In this paper, we apply SBPI to the indoor air temperature control in buildings, and make the following contributions. First, in order to handle the large state space, we define certain events that trigger control actions for HVACs and focus on these event-based policies. Second, we apply sample path sharing (Jia and Wu, 2012; Wu et al., 2014) to efficiently share the sample path among the simulation of different actions. We then use numerical results to show that this sample path sharing techniques can be easily combined with equal allocation, successive rejects, and optimal computing budget allocation (OCBA), and improves the corresponding performance.

The rest of this paper is organized as follows. We briefly review the related works in section 2, mathematically formulate the indoor air temperature control problem in section 3, provide our methodology in section 4, present the numerical results in section 5, and briefly conclude in section 6.

2. LITERATURE REVIEW

Simulation-based optimization methods have received increasing attention in HVAC optimization. An early overview and comparison of some of the methods is available in Wetter and Wright (2004). In addition, complete simulation-based sequential quadratic programming (Sun and Reddy, 2005), robust evolutionary algorithm (Fong et al., 2009), genetic algorithm and Artificial Neural Network (Magnier and Haghighat, 2010) and many other methods have also been studied in HVAC optimization problems.

Markov decision process (MDP) provides a general framework for many control, decision-making and optimization problems. While traditional policy iteration and value iteration algorithms are usually computationally infeasible, SBPI, or simulation rollout (Bertsekas and Castanon, 1999) can be adopted to improve policies for large-scale MDPs. Jia et al. (2012) applied SBPI (i.e., stochastic simulation-based rollout) to energy management in commercial office buildings.

Due to the large amount of simulation used in SBPI, it is of great practical interest to study how to allocate the computing budget so that the best action is selected with high probability. Many computing budget procedures have been studied in simulation-based optimization. Chen et al. (2011) provided an excellent recent review. In this paper, we consider three allocation methods: equal allocation, successive rejects (Audibert et al., 2010) and OCBA (Chen et al., 2000).

3. PROBLEM FORMULATION

In this section, we formulate the problem as an MDP. Consider a typical HVAC system as shown in Fig. 1. The combination of the fan and the coil is referred to as fan coil unit (FCU). In summer, the chiller supplies cold water to the coil, where the flow rate of the water is controlled by a valve. By changing the speed of the fan, we can control the rate of heat exchange between inlet air and chilling water. We focus on the control of FCU, because it plays a central role in indoor thermal comfort. In particular, we intend to attain a tradeoff between energy cost and thermal comfort by controlling the speed of the fan, whereas the valve is assumed to maintain open throughout the control process.

3.1 System States and Actions

Similar to Sun et al. (2013), we define the system state as a combination of indoor air temperature T_a, indoor air humidity H_a, wall temperature T_w, outdoor air temperature T_o and the number of occupants, P, i.e.,

$$S = [T_a, H_a, T_w, T_o, P]^T, \tag{1}$$

where the superscript T stands for transpose. The action, defined as the speed of the fan, is denoted by V_{fan}. The action space is discretized into four actions, i.e. $\mathcal{V} = \{1, 2, 3, 4\}$, where action 1 corresponds to turning off the fan, while actions $2, 3$ and 4 correspond to setting the fan at low, medium and high speed, respectively.

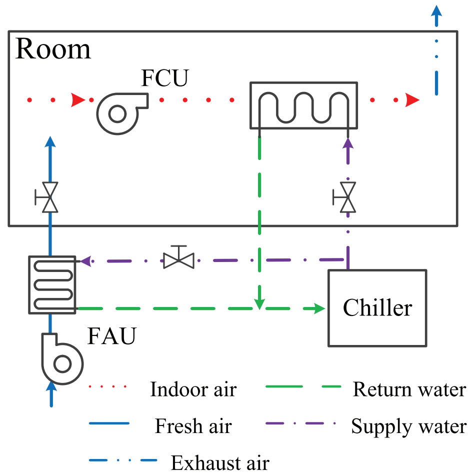

Fig. 1. A typical HVAC system.

3.2 System Dynamics

Based on the definition of system states, we describe the system dynamics using the models in Sun et al. (2013). During discretization, two time scales are used: a small time scale (1 minute) for simulating system dynamics, and a large time scale (10 minute) for decision making. We use superscript (t) to denote the time index of the system dynamics, and superscript (k) to denote decision epochs.

According to energy conservation rule, the dynamics of T_a can be given by the following:

$$\begin{aligned} m_a T_a^{(t+1)} = {} & m_a T_a^{(t)} + \Delta t[P^{(t)} Q_g + h_{gs} A_{gs}(T_o^{(t)} - T_a^{(t)}) \\ & + h_{w,in} A_w (T_w^{(t)} - T_a^{(t)}) + Q_{base}]/C_p \\ & + \Delta t(G_{FCU}^{(t)} T_{FCU}^{(t)} + G_{nv}^{(t)} T_o^{(t)}) \\ & - \Delta t(G_{FCU}^{(t)} + G_{nv}^{(t)}) T_a^{(t)}, \end{aligned} \tag{2}$$

where m_a is the mass of indoor air, Q_g is the heat generated per occupant, h_{gs} is the heat transfer coefficient between indoor and outdoor air through the glass curtain wall facing outside, A_{gs} is the area of the glass curtain wall, $h_{w,in}$ is the heat convection coefficient between interior walls and indoor air, A_w is the area of interior walls, Q_{base} is the base cooling load, C_p is the air specific heat, T_{FCU} is the temperature of FCU's outlet air, and G_{nv} is the flow rate of natural ventilation, which is assumed to be invariable for simplicity.

Similarly, the dynamics of H_a can be derived as follows:

$$\begin{aligned} m_a H_a^{(t+1)} = {} & m_a H_a^{(t)} + \Delta t P^{(t)} H_g + \Delta t(G_{FCU}^{(t)} H_{FCU}^{(t)} \\ & + G_{nv}^{(t)} H_o^{(t)}) - \Delta t(G_{FCU}^{(t)} + G_{nv}^{(t)}) H_a^{(t)}, \end{aligned} \tag{3}$$

where H_g is the humidity generated by each occupant, H_{FCU} is the humidity of FCU's outlet air, and H_o is the humidity of outdoor air. Here we assume that H_o varies only slightly over time and can be approximated by a constant. In addition, without any dehumidification, H_{FCU} can be approximated by H_a, because FCU uses indoor air as its inlet air supply.

Suppose the room can hold 3 occupants at most, i.e., $P^{(t)} \in \{0, 1, 2, 3\}$. The occupants' movement can be