

# Event-Based Optimization Within the Lagrangian Relaxation Framework for Energy Savings in HVAC Systems

Biao Sun, *Student Member, IEEE*, Peter B. Luh, *Fellow, IEEE*, Qing-Shan Jia, *Senior Member, IEEE*, and Bing Yan, *Student Member, IEEE*

**Abstract**—Optimizing HVAC operation becomes increasingly important because of the rising energy cost and comfort requirements. In this paper, an innovative event-based approach is developed within the **Lagrangian relaxation framework** to minimize an HVAC's day-ahead energy cost. To solve the HVAC optimization problem based on events is challenging since with time-dependent uncertainties in weather, cooling load, etc., the optimal policy is not stationary. The nonstationary policy space is extremely large, and it is time consuming to find the optimal policy. To overcome the challenge, **we develop an event-based approach to make the nonstationary optimal policy stationary in the planning horizon**. The key idea is to augment state variables to include the time-dependent variables that make the optimal policy nonstationary and then define events based on the extended state variables. In addition, we develop within the Lagrangian relaxation framework a Q-learning method where Q-factors are used to evaluate event-action pairs and to obtain the optimal policy. **Numerical results demonstrate that, as compared with time-based approaches, the event-based approach maintains similar levels of energy costs and human comfort, but reduces computational efforts significantly and has a much faster response to events.**

**Note to Practitioners**—Traditionally, the HVAC operation problem is solved by using time-based approaches where decisions are calculated and executed at each discrete time instant. These approaches usually have large computational requirements and slow responses to events. In this paper, an innovative event-based approach is developed so that decisions are calculated and executed only on an “as needed” basis to reduce computational requirements and have a fast response to events. The key of the approach is to aggregate the future information that affects day-ahead HVAC energy costs and augment the state to include the aggregated information. The events are then defined based on the augmented state and make the non-stationary optimal policy stationary in the planning horizon. Numerical results demonstrate significant computational time reduction and fast responses to events.

Manuscript received October 08, 2014; revised April 04, 2015; accepted May 27, 2015. Date of publication August 06, 2015; date of current version October 02, 2015. This paper was recommended for publication by Associate Editor L. H. Lee and Editor L. Shi upon evaluation of the reviewers' comments. The preliminary results of the work reported here were presented in the IEEE Conference on Automation Science and Engineering, August 2013 [13].

B. Sun and Q.-S. Jia are with the Center for Intelligent and Networked Systems (CFINS), TNLIS, Tsinghua University, Beijing 100084, China (e-mail: sun-b05@mails.tsinghua.edu.cn; jiaqs@tsinghua.edu.cn).

P. B. Luh is with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269-2157 USA, and also with CFINS, TNLIS, Tsinghua University, Beijing 100084, China (e-mail: Peter.Luh@uconn.edu).

B. Yan is with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269-2157 USA (e-mail: biy10001@engr.uconn.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TASE.2015.2455419

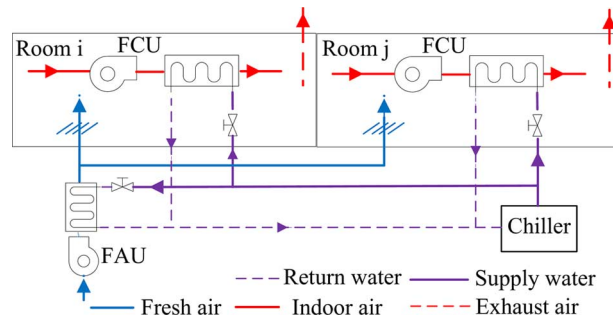


Fig. 1. HVAC system.

**Index Terms**—Event-based optimization, HVAC energy optimization, Lagrangian relaxation, Q-learning.

## I. INTRODUCTION

OPTIMIZING HVAC operations to minimize energy costs while satisfying human comfort requirements becomes increasingly important because of the rising energy cost and human comfort requirements [2]. As presented in Fig. 1, an HVAC usually includes four major parts: 1) fan coil units (FCUs), one for each room to cool and dehumidify indoor air; 2) a fresh air unit (FAU) shared by all rooms to cool, dehumidify, and provide fresh air; 3) chillers to produce chilled water to FCUs and the FAU; and 4) pumps and cooling towers that are not shown in the figure since this paper focuses on the control of terminal devices, i.e., FCUs and FAU.

There are a number of approaches to solve HVAC energy optimization problems in the literature, including model predictive control [18], [22], fuzzy control [6], [21], and genetic algorithms [7] for small buildings and dynamic programming within the Lagrangian relaxation framework [2] for large buildings. In these approaches, time is discretized, and the optimal decision is found for each state at each discrete time instant. In our recent paper [3], HVACs, lights, shading blinds, and windows were jointly controlled to minimize daily energy cost subject to time-of-day electricity prices. Rooms were coupled by sharing chillers with limited capacities. Because of the coupling, computational requirements increase exponentially as the number of rooms increases. To reduce computational requirements, Lagrangian relaxation (LR) was used to relax the chiller capacity constraints. Subproblems, each related to one room, were solved by using dynamic programming and the dual problem by a sub-gradient method.

In contrast to the above-mentioned “time-based” approaches, there is another type of approach which is based on events but has not been used for HVACs. In event-based approaches [10], [25], an action, i.e., control of HVAC devices, is triggered by an event which is defined as a set of state transitions. The objective is to find the optimal action for each event. There are two major advantages of event-based approaches. One is that, as compared with time-based approaches, event-based approaches generally have a much faster response to events. The other is that, for problems whose underlying process and optimal policy are stationary, the computational requirements can be reduced significantly since the number of events requiring actions is usually much smaller than that of states. For other problems whose underlying process is not stationary, the optimal event-based policy is nonstationary. The size of the non-stationary policy space is usually extremely large, and it is time consuming to obtain the optimal policy in the nonstationary policy space.

As for our HVAC problem, the optimal policy is nonstationary because the underlying process is nonstationary due to the nonstationary uncertainties in outside temperature and numbers of occupants, and also because the objective is a function of time-varying electricity prices. In this paper, an innovative event-based approach is developed to augment the states to make the nonstationary optimal policy stationary in the planning horizon. In this way, we only need to search for the optimal policy in the stationary policy space whose size is usually much smaller than that of nonstationary policy space. The computational requirements to solve the HVAC optimization problem can then be significantly reduced.

After reviewing detailed building and HVAC models in simulation software and optimization approaches in Section II, the problem formulation with simplified building and HVAC models is presented in Section III. The simplified models are the same as those in [3]. In [3], the problem focused on the integrated control of HVAC, lights, shading blinds, and windows. In this paper, only HVAC is considered in order to focus on the development of the event-based approach. The objective is to minimize HVAC energy costs of  $N$  (e.g., 24) hours ahead. Traditionally, an event is defined as a set of state transitions. Since rooms are coupled by sharing chillers with limited capacities, the definition of an event in one room needs to consider states of other rooms, causing large computational complexity. In addition, the nonstationarity of the optimal policy is another source that causes large computational complexity.

To reduce the computational requirement caused by the couplings among rooms, LR [2], [3] is introduced in Section III to relax the chiller capacity constraints. Similar to [3], the relaxed problem is then decomposed into subproblems, each related to one room. The objective of a subproblem is to find the optimal action for each event to minimize energy cost of a room. An event of a room no longer needs to be defined based on states of other rooms. A high-level dual problem is also formulated as in [3] to iteratively coordinate subproblem solutions so that the relaxed constraints can be satisfied. Since chiller capacity constraints are for each discrete time instant in the next  $N$  hours, Lagrangian multipliers should also be introduced for each discrete time instant. The dual problem therefore is still based on

time, and the multipliers need to be updated for all discrete time instants in the next  $N$  hours.

In Section IV, an event-based approach is developed to solve subproblems. If an event is defined as a set of state transitions in the traditional way, the optimal policy would be non-stationary due to the non-stationary underlying process. The nonstationarity of the optimal policy would cause large computational requirements. For the easy computation, our key idea is to augment the state to include these time-varying variables such as the multipliers and cooling load in the future and then define an event based on the augmented state. In the LR framework, these future variables are updated iteratively, causing the augmented state and the optimal action to be updated iteratively until converge. The optimal policy is assumed to be stationary in the planning horizon of the next  $N$  hours, i.e., the optimal action for an event is the same no matter when it happens in the planning horizon. The stationarity of the optimal policy then makes it easy for computation and the optimal policy in the planning horizon is obtained by iteratively updating a stationary policy.

In Section V, the dual problem is solved by the very recent surrogate LR method [23], [24]. In the method, the surrogate subgradient directions for updating multipliers are functions of cooling load of all discrete time instants in the next  $N$  hours. However, the subproblem solutions provide only the cooling load for the first event in the next  $N$  hours but not the rest of the cooling load. To overcome this challenge, the sample average approximation method is combined with the event-based approach to approximate the rest cooling load.

In Section VI, three examples are considered for buildings with different sizes. The testing results show that: 1) as compared with time-based approaches, the new approach maintains similar levels of energy costs and human comfort but reduces the computational time significantly and is scalable as the size of buildings increases and 2) the new approach also has a much faster response to events, because it acts immediately upon an event while the time-based approach waits till the next time instant and then acts.

Our approach opens up a new way to solve event-based problems with non-stationary optimal policies. In addition, if these problems are solved within the LR framework, our approach is also applicable by using the event-based method to solve subproblems and the surrogate Lagrangian relaxation method as well as sample average approximation to solve a dual problem.

## II. LITERATURE REVIEW

There are many studies on the HVAC energy optimization. In Section II-A, building and HVAC models and HVAC energy cost-optimization approaches are presented. In Section II-B, the event-based optimization approach is presented. Our recent work on the time-based optimization of building energy cost is presented in Section II-C.

### A. Building and HVAC Models and HVAC Energy Cost-Optimization Approaches

A typical approach for HVAC energy optimization is first to establish building and HVAC models to describe building behavior and calculate energy costs. An energy cost-optimization problem is then formulated, and optimization methods [2]–[8]

are developed to solve the problem to obtain the optimized policy.

Building and HVAC models that are commonly used are the simulation models **developed in building simulation software, such as DeST and EnergyPlus**. Although simulation models are reasonably accurate, they are usually complicated to establish since they have a large number of parameters and require a great deal of effort for model calibration. Also, they tend to have large computational requirements. Simulation models are therefore usually not suitable for searching for the optimal policy, but they are useful for evaluating optimized policies obtained by using some simplified models [17]–[20].

Black box models of buildings and HVACs are much more simplified than simulation models. They are usually established by using data mining techniques to find the relationships between inputs and outputs of buildings and HVACs. They are developed based on measured data without using physical knowledge of buildings and HVACs. Common techniques to establish black-box models [19], [20] include artificial neural network (ANN), linear regression, and Principal Component Analysis (PCA). The establishment of black-box models is usually easy and the prediction based on black-box models does not require large computational time, but black-box models might cause large prediction errors for working conditions that have not been included in the training data.

Based on building and HVAC models, optimization approaches are developed to solve energy cost optimization problem. An approach commonly used is the model predictive control (MPC) [18], [22]. It has been extensively studied to minimize building energy costs by taking advantages of building thermal mass, a variable energy price, and a water tank for storing chilled water, and so on. Fuzzy logic controllers are also in common use. They have been demonstrated by a large number of studies that fuzzy logic controllers are able to save significant energy costs [6], [21]. Another approach is to use genetic algorithms (GA) [7]. For example, a GA was applied in [7] to optimize indoor temperature set points in individual rooms and the fresh air flow rate shared by all rooms to save energy costs.

### B. Event-Based Optimization and State Aggregation

In some practical problems, actions are triggered by events. For example, take a problem in finance. A trader decides to buy or sell an asset when its price falls below or above a threshold. To solve this type of problems, the event-based optimization approach is developed and has been widely used in the field of finance, queue system, and networks [10]–[12] but not HVACs yet. The event-based approach first needs to define events. An event is a set of state transitions in two successive discrete time instants, and all state transitions in an event should share some common characteristics [16].

A performance potential or Q-factor is then used to evaluate the performance of each pair of event and action by how the problem cost is affected when an action is taken upon an event. These performance potentials or Q-factors are estimated from a sample path of the underlying system or historical data. For the sample-path-based method, the key idea is to estimate the

performance potentials of Q-factors for all states on a given sample path under a policy [16]. Based on performance potentials or Q-factors, the policy iteration or gradient-based methods are used to obtain the optimal policy.

The event-based optimization is usually developed for problems whose optimal policies are stationary, such as problems with steady-state distribution in infinite time. For finite-time problems whose optimal policies are not stationary, although some event-based methods have been developed, they need to run a large number of simulations to estimate the nonstationary potentials or Q-factors [11], [12]. The simulations usually require large computational efforts and take a long time.

In the HVAC energy cost optimization problem, the optimal policy is nonstationary because the underlying process and the objective are non-stationary due to the time-dependent variables such as cooling load and electricity prices. The potentials or Q-factors are nonstationary and difficult to obtain. The size of the nonstationary policy space is extremely large, and it is time consuming to find the optimal policy.

The event-based approach developed in this paper is based on our preliminary results in a conference paper [13]. Based on [13], the event-based approach is further improved with rigorous derivations to solve subproblems based on Q-factors. Numerical testing is strengthened with a more detailed comparison of the performance between the event-based approach and time-based approaches, and also with more insights on the results, including reduced computational time and a faster response time.

An event is a set of state transitions. For an event, if the state that is transitioned to can take any feasible value, then the event is equivalent to a state aggregation [25]. To solve Markov decision process (MDP) problems with state aggregations, a typical approach is to approximate the optimal value function by simple functions [28]. The challenge is that as the number of the aggregated states reduces, the computational requirements to MDP problems is reduced but the approximation errors increase. To have a good tradeoff between the computational time and the approximation errors, an ordinal optimization method [29] is developed to find good simple state aggregations with high probabilities.

### C. Our Recent Paper on Building Energy Cost Optimization

Rooms are coupled by sharing an HVAC system with limited capacities. The coupling makes it time consuming to search for the optimal policy. Decomposition and coordination methodologies, such as LR [2], [3], are effective in solving the building energy cost optimization problem with the coupling, especially for a building with a large number of rooms. In our recent paper [3], which focused on the integrated control of HVACs, lights, shading blinds, and natural ventilation, the chiller capacity constraints were relaxed by LR to obtain a dual problem and subproblems. Each subproblem was related to one room. A subproblem was to find the optimal decision variable of a room to minimize its energy cost. The dual problem was to coordinate subproblem solutions to satisfy the relaxed constraints. To obtain a near-optimal policy, the subproblems and the dual problem were solved iteratively by stochastic dynamic programming and surrogate subgradient method, respectively.

In the above-mentioned method as well as some other time-based methods, a day is discretized with a fixed time step and the problem is formulated and solved in each discretized time instant. One limitation of these time-based methods is that they will have large computational requirements if a finer time step is required to have a more accurate calculation and a faster response to the changes of occupancy, the changes of comfort requirements, etc.

### III. PROBLEM FORMULATION AND LR FRAMEWORK

The problem formulation for daily building energy cost optimization and the LR framework are presented in this section. The formulation is the same as that in [3], except that the control of lights, blinds, and natural ventilation is no longer considered. Device and room models are presented in Section III-A. The objective function and human comfort requirements are presented in Section III-B. The LR framework is presented in Section III-C.

#### A. HVAC and Room Models

The models of HVAC devices are established in [3] and briefly presented below. Assume there are  $I$  rooms in a building and a day is divided into  $K$  discrete time intervals of equal duration  $\Delta t$  (e.g., 10 min), with time index  $k$  ranging from 1 to  $K$ . The FAU fresh air flow rate equals the sum of fresh air flow rates to all rooms,  $\sum_{i=1}^I G_{fa,i}^k$ , where  $G_{fa,i}^k$  is the fresh air flow rate to room  $i$  at time  $k$ . The electric power of the FAU fan,  $P_{fan,FAU}^k$ , is nonlinear to the FAU air flow rate as

$$P_{fan,FAU}^k = P_{fan,FAU,Rated} \left[ \frac{\sum_{i=1}^I G_{fa,i}^k}{G_{a,FAU,Rated}} \right]^3 \quad (1)$$

where  $P_{fan,FAU,Rated}$  and  $G_{a,FAU,Rated}$  are the rated FAU fan power and air flow rate, respectively [3].

The chiller shared by FAU and FCUs has a limited capacity  $C_{HVAC}$ , and rooms are thus coupled by the chiller capacity constraints as

$$C_{FAU}^k + \sum_{i=1}^I C_{FCU,i}^k \leq C_{HVAC}, k = 1, \dots, K \quad (2)$$

where  $C_{FAU}$  and  $C_{FCU,i}$  can be calculated by [3, eq. (3)].

According to [3], decision variable at time  $k$  is

$$[T_{FAU}^k, v_1^k, v_2^k, \dots, v_I^k]^T \quad (3)$$

with

$$v_i^k = [T_{FCU,i}^k, G_{fa,i}^k, G_{a,FCU,i}^k]^T, i = 1, \dots, I \quad (4)$$

where  $T_{FAU}^k$  and  $T_{FCU,i}^k$  are outlet air temperatures of FAU and FCU in room  $i$ , respectively,  $G_{fa,i}^k$  is the fresh air flow rate to room  $i$  provided by the FAU, and  $G_{a,FCU,i}^k$  is the air flow rate of the FCU in room  $i$ .

Energy consumption of fans in FAU and FCUs is calculated by (1). Energy consumption of chillers, pumps, and cooling towers is calculated for simplicity based on a coefficient of performance (COP) which is defined as the ratio of cooling load of the FAU and FCUs to the electric power of chillers, pumps, and cooling towers [3].

Indoor air temperature, wall temperature, indoor humidity, and indoor CO<sub>2</sub> concentration are chosen as elements of the state variable. The state variable for room  $i$  at time  $k$  is

$$x_i^k = [T_{ai}^k, T_{wi}^k, H_i^k, CO2_i^k]^T. \quad (5)$$

Dynamic equations of the state are developed in our recent paper [3] based on the energy and mass conservations (See [3, eqs. (8)–(11)]). The major uncertainties that affect energy cost are outside temperatures and numbers of occupants which are modeled in [3].

#### B. Objective Function

An energy cost optimization problem is formulated when an event occurs. For example, an event could be the rising of indoor temperature above a threshold. Traditionally, an event is defined as a set of state transitions. Rigorous definition of events is presented later in Section IV-A. The objective of the problem is to find the current time's optimal decision to minimize the expected total costs of the HVAC for  $N$  hours ahead. Upon an event, the energy cost optimization problem is formulated as

$$\min J$$

with

$$J \equiv E \left\{ \Delta t \sum_{k=1}^K c^k \left[ \sum_{i=1}^I \left( \frac{C_{FCU,i}^k}{COP^k} + P_{fan,FCU,i}^k \right) + \frac{C_{FAU}^k}{COP^k} + P_{fan,FAU}^k \right] \right\} \quad (6)$$

where  $\Delta t$  is the time interval,  $K$  the number of time intervals in  $N$  hours ahead,  $c^k$  the electricity price at time  $k$ , and the expectation is over uncertain outside temperatures and numbers of occupants. The problem is subject to the chiller capacity constraints (2) and the human comfort requirements during occupied periods [3]

$$T_a \in [22^\circ\text{C}, 26^\circ\text{C}], H \in [40\%, 60\%], CO_2 < 900 \text{ ppm}. \quad (7)$$

In unoccupied periods, there are no comfort requirements.

#### C. LR Framework

In the problem formulated in (6), all rooms share chillers with limited capacities. To find the optimal action for an event in a room, the cooling load of other rooms needs to be considered. An event would therefore need to be defined as transitions of the combined states of all rooms. In that case, the number of events would increase exponentially as the number of rooms increases. In addition, the problem has a two-level structure. The low level focuses the control of devices in individual rooms and the high level coordinates all of the rooms to satisfy the chiller capacity constraints. In order to overcome the computational difficulty caused by the coupling in rooms, a possible method is to use an LR-based method [23], [24], [26] which is a decomposition and coordination approach.

For problems which are separable, the surrogate LR method can be directly used. The problem (6), however, is not separable [3] because: 1) the FAU is shared by all rooms and therefore its



decision variable the outlet air temperature  $T_{\text{FAU}}$  cannot be determined by any individual room; and 2) the FAU fan power  $P_{\text{fan,FAU}}$  is nonlinear to the sum of fresh air flow rates to all rooms as in (1). As in [3], the first inseparability is overcome by introducing new decision variables,  $T_{\text{FAU},i}^k$ ,  $i = 1, \dots, I$ , representing fresh air temperatures supplied by the FAU to  $I$  individual rooms at time  $k$ . At time  $k$ , the following requirements are induced as in [3]:

$$T_{\text{FAU},i}^k = T_{\text{FAU},i+1}^k, \quad i = 1, \dots, I, \quad k = 1, \dots, K. \quad (8)$$

The decision variable is therefore changed from (3) and (4) to

$$u^k = [u_1^k, u_2^k, \dots, u_I^k]^T \quad (9)$$

with

$$u_i^k = [G_{fa,i}^k, T_{\text{FAU},i}^k, G_{a,\text{FCU},i}^k, T_{\text{FCU},i}^k], \quad i = 1, \dots, I. \quad (10)$$

By using multipliers  $\lambda$  and  $\mu$  to relax the chiller capacity constraints (2) and the introduced constraints (8), respectively, the relaxed problem is to minimize the Lagrangian  $L$  as

$$\begin{aligned} \min_{u^k, k=1, K} L \text{ with } L \equiv & E \left\{ \Delta t \sum_{k=1}^K c^k \left[ \sum_{i=1}^I \left( \frac{C_{\text{FCU},i}^k}{\text{COP}^k} \right. \right. \right. \\ & \left. \left. + P_{\text{fan,FCU},i}^k \right) + \frac{C_{\text{FAU}}^k}{\text{COP}^k} + P_{\text{fan,FAU}}^k \right] \\ & + \lambda^k \left( C_{\text{FAU}}^k + \sum_{i=1}^I C_{\text{FCU},i}^k - C_{\text{HVAC}} \right) \\ & \left. + \mu_i^k (T_{\text{FAU},i}^k - T_{\text{FAU},i+1}^k) \right\}. \quad (11) \end{aligned}$$

The first term in the last line of (10) can then be separated to subproblems of individual rooms.

The second inseparability related to  $P_{\text{fan,FAU}}$  is overcome by the surrogate optimization framework [9] and the subproblem for room  $i$  is obtained as

$$\begin{aligned} \min_{u_i^k, k=1, K} L_i, \text{ with } L_i \equiv & E \left\{ \sum_{k=1}^K \left[ \left( \frac{c^k \Delta t}{\text{COP}^k} + \lambda^k \right) \right. \right. \\ & \cdot (C_{\text{FAU}}^k + C_{\text{FCU},i}^k) + c^k \Delta t (P_{\text{fan,FAU}}^k + P_{\text{fan,FCU},i}^k) \\ & \left. \left. + (\mu_i^k - \mu_{i-1}^k) T_{\text{FAU},i}^k \right] \right\}. \quad (12) \end{aligned}$$

When solving the subproblem, only the decision variables belong to room  $i$  are optimized while those such as fresh air temperatures belong to other rooms are kept at their latest available values.

To coordinate the subproblems to satisfy the constraints relaxed, the Lagrangian multipliers are updated in a high-level dual problem as

$$\max_{\lambda, \mu} q, \text{ with } q \equiv \sum_{i=1}^I L_i^* - C_{\text{HVAC}} \sum_{k=1}^K \lambda^k \quad (13)$$

where  $L_i^*$  is the optimal cost for the subproblem of room  $i$  in (12). The updated multipliers are then used to solve subproblems iteratively until certain stopping criteria [3] are satisfied.

#### IV. SOLVING THE SUBPROBLEMS

Here, subproblems are solved by using an event-based optimization approach. In Section IV-A, a novel definition of events is introduced to make the non-stationary optimal policy stationary in the planning horizon through state augmentation. In Section IV-B, a Q-learning algorithm for obtaining the optimal policy is presented.

##### A. Definitions of Events

In an event-based approach, the control of HVAC devices, i.e., the actions, is needed only when certain events occur. For example, an event could be the rising of the indoor temperature from a value between 24 and 25 degrees to a value between 25 and 26 degrees. For an event, actions include the decision of the fresh air flow rate to the room, the FAU outlet air temperature, and the FCU supply air flow rate and temperature. Each feasible value of the decision variable in (10) corresponds to an action.

Traditionally, an event is defined as a set of state transitions. A state transition is the change of the state. To observe a change of the state, two measurements of the state in two successive time instants are needed. To have a faster response to an event, the time step between two time instants should be as low as possible. For our problem, the time step is selected as the sampling period  $\delta$ , i.e., the time interval for measuring the state by sensors. This is the minimal time step which is available.

Since a policy is a mapping from events to actions, all state transitions in an event have the same action. Therefore, state transitions should be aggregated to an event in a way that all state transitions in an event share some common characteristics. For our problem, the state variable in (5) has four elements, including indoor temperature, humidity, and  $\text{CO}_2$  level, and wall temperature. One way of defining events is first to divide the domain of each element into several intervals. For the  $j$ th element of the state  $x_i^t$  in room  $i$ , its domain is divided into  $N_j$  intervals,  $\mathcal{D}_{i,1}(j), \mathcal{D}_{i,2}(j), \dots, \mathcal{D}_{i,N_j}(j)$ , which have no overlaps. It is assumed that, in a time step  $\delta$ , at most one element changes from one of its intervals to another. Since the sampling period  $\delta$  could be several seconds which is relatively smaller as compared to the time constants of a room, the assumption is reasonable. Based on the transitions of each element, a type of events can be defined according to the start interval and the end interval in a time step. For example, an event with the 1<sup>st</sup> element crossing from  $\mathcal{D}_{1,1}(1)$  to  $\mathcal{D}_{1,2}(1)$  at time  $t$  and the other elements remaining within  $\mathcal{D}_{1,1}(2), \mathcal{D}_{1,1}(3)$ , and  $\mathcal{D}_{1,1}(4)$  is defined as

$$\begin{aligned} e_{i,1}^t = & \left\{ \langle x_i^{t-\delta}, x_i^t \rangle \mid x_i^{t-\delta}(1) \in \mathcal{D}_{i,1}(1), x_i^{t-\delta}(1) \in \mathcal{D}_{i,2}(1), \right. \\ & \left. x_i^{t-\delta}(l), x_i^t(l) \in \mathcal{D}_{i,1}(l), l = 1, 2, 4 \right\}. \quad (14) \end{aligned}$$

Events of other elements can be defined similarly. If in one time step, more than one element across their intervals, then we assume that several events occur sequentially. They are stored in a stack and optimal actions are found for them in sequence.

If events are defined in the traditional way as in (14), an event, however, would be time dependent and the optimal policy of a

subproblem would not be stationary. That is because the underlying process and the subproblem objective are non-stationary due to the following time-varying variables.

$c^t$	Energy price.
$C_{L,i}^t$	Cooling load of FAU and FCU for room $i$ at time $t$ and affected by the non-stationary outside temperatures and number of occupants.
$\text{COP}^t$	Coefficient of performance affected by outside temperatures and $C_{L,i}^t$ ; and calculated by using the model of chillers, cooling tower, and pumps [27].
$\lambda^t$	Lagrangian multiplier.

If the optimal policy is nonstationary, we would need to search for the optimal policy within the non-stationary policy space whose size increases exponentially as the number of  $N$  increases [11]. In addition, the Q-factors or performance potentials used to evaluate the event-action pairs would be non-stationary and a lot of simulations would be needed to obtain the Q-factors or performance potentials. This simulations could be too time consuming.

Our idea to overcome the difficulty is to define events in a new way that makes the optimal policy stationary in the planning horizon, i.e., the next  $N$  hours. In the new way, the state variable is first augmented to include the above-mentioned time varying variables in the next  $N$  hours, and events are then defined based on state transitions of the augmented state. Since all the time varying variables that cause the underlying process and the object function non-stationary have already been included in the augmented state, all of the information needed to determine the optimal action for an event is included in the event itself. Therefore for easy computational, it is reasonable to assume that the optimal policy in the planning horizon of the next  $N$  hours is stationary, although the optimal policy may change over time when the problem is solved in a moving window manner. The stationary optimal policy in the planning horizon can be searched within the stationary policy space whose size is significantly smaller than that of non-stationary policy space.

Considering those time-varying variables in the next  $N$  hours are so many and they are mainly used to define events to trigger pre-cooling, their average values in the future are used as approximation to trigger the pre-cooling. The augmented state variable then is

$$y_i^t = [T_{ai}^t, H_i^t, \text{CO2}_i^t, \bar{c}^t, \bar{\text{COP}}^t, \bar{C}_{L,i}^t, \bar{\lambda}^t]^T \quad (15)$$

where  $T_{ai}^t$ ,  $H_i^t$  and  $\text{CO2}_i^t$  are indoor temperature, humidity, and  $\text{CO}_2$  concentration in room  $i$  at time  $t$ , respectively, and  $\bar{c}^t$ ,  $\bar{\text{COP}}^t$ ,  $\bar{C}_{L,i}^t$  and  $\bar{\lambda}^t$  are the average values of  $c^t$ ,  $C_{L,i}^t$ ,  $\text{COP}^t$  and  $\lambda^t$  in the next  $M$  hours, respectively. The value of  $M$  is required be larger than the length of common pre-cooling time period and can be set to six for example.

The wall temperature in the state (5) is not included in the augmented state because the wall temperatures affect the indoor temperature through the heat transfer between the wall and indoor air [3], and the amount of heat transferred has already

been reflected in the indoor air cooling load  $C_{L,i}^t$  [3]. Although taking the wall temperature out of the augmented state is not crucial to make the optimal policy stationary, it does reduce the number of events.

An event is then defined as a set of state transitions of the augmented state in the same way as in (14). The domain of each element of the augmented state is divided into several intervals. Based on the transitions of each element, a type of event can be defined according to the start interval and the end interval in a time step. For an event with the first element crossing from  $\mathcal{D}_{1,1}(1)$  to  $\mathcal{D}_{1,2}(1)$ , it is no longer time-dependent and is defined as

$$e_{i,1} = \left\{ \langle x_i^{t-\delta}, x_i^t \rangle | x_i^{t-\delta}(1) \in \mathcal{D}_{i,1}, (1)x_i^{t-\delta}(1) \in \mathcal{D}_{i,2}(1), \right. \\ \left. x_i^{t-\delta}(l), x_i^t(l) \in \mathcal{D}_{i,1}(l), l = 2, 3, \dots, 7, \forall t \right\}. \quad (16)$$

Events of other elements can be defined similarly.

In the new definition of events, there are seven types of events, each corresponding to one element of the augmented state. The first three types are defined based on the transitions of the current indoor environment, i.e.,  $T_{ai}^t$ ,  $H_i^t$  and  $\text{CO2}_i^t$ . They are used to trigger actions of HVAC devices to maintain the indoor temperature, humidity and  $\text{CO}_2$  level in the comfort range with minimal energy costs. The other four types are defined based on the transition of future information of  $c^t$ ,  $\text{COP}^t$ ,  $C_{L,i}^t$ , and  $\lambda^t$ . They are used to trigger actions of HVAC devices for precooling, so that the cheap electricity can be taken advantage of to save energy costs and the cooling load can be shifted to satisfy the chiller capacity constraints.

Based on the new definition of events, the optimal policy become stationary in the planning horizon of the next  $N$  hours. The stationarity can reduce the computational requirements significantly since we can iteratively update a stationary policy until it converges. As time moves forward, the optimal policy, however, could be different due to the change of time-of-day prices, COP, cooling load, and multipliers.

Note that the event in (16) corresponds to the observable event in Xiren Cao's event-based optimization framework [10], [16]. From the view of event-based method, an event is defined as a set of state transitions. From the view of the state aggregation method, if we define a new state  $y_i^t = (x_i^{t-\delta}, x_i^t)$ , then an event is equivalent to a state aggregation. Therefore, the state aggregation is a special event  $e_t := \{ < y_i^t y_i^{t+\delta}, >: y_i^t \in s_o \}$  where  $s_o$  is a subset of the complete set of  $y_i^t$  [25].

### B. Solving Subproblems Based on Q-Factors

To solve the subproblems, we need to find the event-based optimal policy, i.e., to find the optimal action for each event to minimize the subproblem cost. For an event  $e$ , a common method to find its optimal action  $a^*$  is to evaluate each feasible event-action pair  $(e, a)$  by how it affects the subproblem cost. Similar to dynamic programming, an event-action pair  $(e, a)$  can be evaluated by using the summation of the current cost and the optimal cost-to-go as

$$Q(e, a) = r(e, a) + \min_{a_{\text{next}}} Q(e_{\text{next}}, a_{\text{next}}) \cdot \frac{(N - \tau)}{N} \quad (17)$$

where  $r(e, a)$  is the current cost from the time of the current event  $e$  to the time the next event  $e_{\text{next}}$  occurs, the second term on the right hand is the optimal cost-to-go  $g(e_{\text{next}})$  from the time of  $e_{\text{next}}$  to the end of the next  $N$  hours,  $\tau$  is the time length from the current event  $e$  to the next event  $e_{\text{next}}$ , and  $a_{\text{next}}$  is a feasible action for  $e_{\text{next}}$ .

In the dynamic programming,  $Q(e, a)$  can be estimated backward based on the models of rooms and HVAC. Since in the backward process, Q-factors in all discrete time instants are needed to be estimated, it is time consuming. Our subproblem, however, has the stationary optimal policy, and there is no need to estimate  $Q(e, a)$  based on models. Rather, it can be estimated from historical data by using the Q-learning algorithm [4], which is a model-free reinforcement learning technique. The advantage of Q-learning is that the computational time can be saved since no calculation of state dynamics for all the time instants in the next  $N$  hours is needed. A common disadvantage of Q-learning is that Q-factors are hard to converge if they are non-stationary [4]. However in our problem, Q-factors are stationary.

In the Q-learning algorithm, all Q-factors are first initialized. After an action  $a$  is taken upon an event  $e$  and the next event  $e_{\text{next}}$  occurs, the current cost  $r(e, a)$  is first calculated based on the energy costs measured in the real HVAC. The “old” Q-factor  $Q_{\text{old}}(e, a)$  is then updated by the “new information”  $r(e, a)$  to obtain the new  $Q_{\text{new}}(e, a)$  as [4]

$$Q_{\text{new}}(e, a) = Q_{\text{old}}(e, a)[1 - \alpha(e, a)] + \alpha(e, a) \cdot \left[ r(e, a) + \min_{a_{\text{next}}} Q_{\text{old}}(e_{\text{next}}, a_{\text{next}}) \cdot \frac{(N - \tau)}{N} \right] \quad (18)$$

where  $\alpha$  is the learning rate. In (18), the second term in the second square brackets is the optimal cost-to-go  $g(e_{\text{next}})$  for the next event  $e_{\text{next}}$  and is calculated based on old Q-factors of  $e_{\text{next}}$ . The learning rate  $\alpha$  determines to what extent the  $Q_{\text{old}}(e, a)$  should be updated by the new information  $r(e, a)$ . Q-factors can converge when  $\alpha$  diminishes to zero over time at a proper rate [4]. To ensure the satisfaction of the comfort requirements, any actions dissatisfying the requirements will be assigned an extreme large current cost  $r$ .

Based on Q-factors estimated by Q-learning, the optimal action  $a^*$  for the event  $e$  is then the one that has the minimal Q-factor as [4]

$$a^* = \arg \min_a Q(e, a). \quad (19)$$

The Q-learning method can solve the sub problems optimally as long as the optimal policy is stationary.

To guarantee the convergence of Q-factors, each Q-factor should have a positive probability to be updated. Therefore, when event  $e$  occurs, each feasible action should have a positive probability to be taken to control devices. As in [4], the optimal action  $a^*$  obtained by (19) is taken with the probability of  $1 - \kappa$ , with  $\kappa > 0$  and  $\kappa \ll 1$ . Each of the other feasible actions is taken with the probability of  $\kappa / (|a_e| - 1)$ , where  $|a_e|$  is the number of feasible actions for the event  $e$ .

## V. SOLVING THE DUAL PROBLEM

Here, the dual problem is solved. In Section V-A, the very recent surrogate LR method [23], [24] is briefly introduced. The method as well as other LR-based methods requires the cooling load for all stages in the next  $N$  hours, while only the cooling load for the first event in the next  $N$  hours is provided by the subproblem solutions. In Section V-B, the sample average approximation is combined with the event-based method to approximate the rest of the cooling load. In Section V-C, the iteration of solving the dual problem and subproblems are presented.

### A. Surrogate LR Method

The dual problem is solved by using the surrogate LR method developed in [23], [24]. The method obtains a surrogate subgradient direction for multipliers by solving only one or several subproblems as long as the surrogate optimization condition is satisfied [9]. Unlike the Surrogate Subgradient (SSG) method [25] which may not converge due to the lack of the optimal dual value, the surrogate LR method does not require the optimal dual value. It guarantees the convergence of multipliers by selecting the step sizes for updating multipliers in a way that the distance between multipliers decreases at consecutive iterations.

Based on the chiller capacity constraints (2), the SSG direction related to  $\lambda_n^k$  at time  $k$  and iteration  $n$  is calculated as [3]

$$g^k(\lambda_n^k)E \left( C_{\text{FAU}}^k + \sum_{i=1}^I C_{\text{FCU},i}^k \right) - C_{\text{HVAC}}, \quad k = 1, \dots, K \quad (20)$$

where  $K$  is the number of stages in the next  $N$  hours. Subproblem solutions only provide the FAU cooling load  $C_{\text{FAU}}$  and FCU cooling load  $C_{\text{FCU}}$  for the first event in the next  $N$  hours. The way to obtain the rest of the cooling load is presented in the next subsection.

The multiplier  $\lambda_{n+1}^k$  is then updated in the surrogate subgradient direction as

$$\lambda_{n+1}^k = \max [0, \lambda_n^k + \alpha_n g^k(\lambda_n^k)] \quad (21)$$

where  $\alpha_n$  is the step size at iteration  $n$  and is updated by [24]

$$\alpha_n = \gamma_n \frac{\alpha_{n-1} \|g(\lambda_{n-1})\|}{\|g(\lambda_n)\|} \quad (22)$$

with

$$\gamma_n = 1 - \frac{1}{M \cdot n^p}, \quad 0 < p < 1, \quad M > 1, \quad n = 1, 2, \dots \quad (23)$$

The above way of updating the step size guarantees that the multiplier converges to a unique limit. At convergence, the surrogate dual value provides a lower bound to the primal cost.

### B. Using Sample Average Approximation to Obtain Cooling Load

As shown in (20), to obtain the SSG direction, the FAU and FCU cooling load for all  $K$  stages in the next  $N$  hours are needed. If optimal actions for  $K$  stages are available, then the cooling load in these stages can be calculated by using FAU and FCU models [3]. In the time-based method in [3], optimal actions have to be obtained in each stage to minimize the total cost of all stages. However, in the event-based method in Section IV, to solve the subproblem which has stationary optimal policy,

only the optimal action for the first event in the next  $N$  hours is needed. Based on this optimal action, only the cooling load for the first event can be calculated.

To obtain the rest of the cooling load, our method is to use simulation to estimate all of the possible subsequent events as well as their optimal actions. In the simulation, the optimal action for the first event is applied to room and HVAC models, and the subsequent events and optimal actions are estimated with the consideration of uncertainties in outside temperatures and numbers of occupants. The difficulty is that as the simulation moves forward to the end of the  $N$  hours, the tree of possible events grows and the number of possible events increases exponentially. To overcome this difficulty, the sample average approximation [14], [15] is combined with the event-based method to approximate the cooling load.

The key idea is first to generate scenarios of uncertain outside temperatures and numbers of occupants in the next  $N$  hours. For each scenario, events are triggered based on the sampled random variables, optimal actions are taken based on Q-factors, and cooling load is calculated iteratively until the end of the  $N$  hours. In this process, the optimal action for an event is determined only based on the information known at the time the event occurs. That means no future information of the realization of uncertainties is used in determining the optimal action. The averaged cooling load of all sampled scenarios is regarded as an approximation of its mean value. As for the number of scenarios, it is increased by one at a time until the change in averaged cooling load is below a given threshold (e.g., 1% of the average cooling load).

### C. Solving the Dual Problem and Subproblems Iteratively

After the multipliers are updated in the dual problem, the subproblems are solved iteratively until some stopping criteria are satisfied. When the algorithm stops, the relaxed constraints may still not be satisfied. Two heuristics developed in [3] are then used to obtain feasible solution for the planning horizon of  $N$  hours, and the feasible solution is applied to control the HVAC devices. As time passes, the HVAC energy cost optimization problem in (6) is solved in a moving window manner, i.e., it is solved by looking ahead of  $N$  hours every time an event occurs. The multipliers, cooling load, etc., which are updated in a moving window manner, would be different over time. Although this may cause the optimal policy changing overtime, the optimal policy for a planning horizon of  $N$  hours can still be assumed to be stationary and therefore the event-based approach can be used to solve the optimization for the purpose of easy computation.

## VI. NUMERICAL SIMULATION RESULTS AND DISCUSSION

Our approach is implemented in MATLAB and runs on a PC with 2.67-GHz Intel Core i7 processor and 4 GB of RAM. Three examples taken from reference [3] are considered. In Example 1, a building with two rooms is used to examine the new definition of events. The event-based method with events defined based on augmented states is compared with the one with events defined based on original states. In Example 2, a building with 15 rooms is examined to illustrate how the actions are triggered by events defined based on future information. To see how energy cost is saved and human comfort is improved, the greedy algorithm is used as a comparison with our event-based method.

In Example 3, a building with 144 rooms is considered for the comparison of the event-based approach with the time-based approaches which use stochastic dynamic programming. Three aspects include computational time, response time, and energy cost are compared.

Buildings are occupied from 7:00 am to 10:00 pm, and the time-of-day electricity price is 0.81 RMB/KWh from 7 am to 10 pm and 0.35 RMB/KWh during other hours. The number of hours looking ahead,  $N$ , is selected to 24. The optimal policies of the three examples are obtained from June 1 to August 31 of a typical meteorological year in Beijing. These policies are then applied to detailed building and HVAC models (developed in the building simulation software DeST [5]) for evaluation. The time interval  $\delta$  for the event-based approaches of the three examples is 10 s.

*Example 1: Building With Two Rooms to Examine the Definition of Events Based on the Augmented State:* In this example, a building with two rooms is selected to illustrate that the new definition of events is more effective than the traditional one. Traditionally, events are defined based on state variables. In our approach, events are defined based on the augmented state variable which is a combination of the original state and the time varying variables including energy price, cooling load, COP and multipliers.

Two event-based approaches are considered, with events defined based on the augmented state variable in the first approach and based on the original state variable in the second one. For the second approach, as presented in Section IV-B, the Q-factors and the optimal policy are non-stationary. The method developed in [11], [12] for problems with non-stationary optimal policies is used. In the method, a large number of simulations are needed to estimate the non-stationary Q-factors based on the models of buildings and HVACs. The simulations require large computational efforts and take a long time.

The average energy cost per day is 23.14 RMB for the first approach and 22.99 RMB for the second one. The computational time is 0.17 s for the first one and 38.5 s for the second one. It can be seen that the energy cost of the stationary optimal policy is close to that of the nonstationary optimal policy. Their energy cost are not exactly the same and the small difference is caused by two reasons. The first is the assumption of the stationarity of the optimal policy in the planning time horizon of  $N$  hours. The second is that when an event occurs, the optimal action by Q-learning is only applied by a probability slightly less than one, so that each Q-factor should have a positive probability to be updated for the convergence. However, to obtain the non-stationary optimal policy, the computational time of the method in [11], [12] is so large even for this small problem with two rooms. The large computational time causes the event-based approach no longer having the advantage of fast response to events.

*Example 2: Three-Floor Building With 15 Rooms to Illustrate Energy Savings and Comfort Improvement:* In reference [3], the time-based approach saves energy cost and improves human comfort by pre-cooling. This example use a three-floor building with 15 rooms to illustrate our event-based approach can also achieve that. It also demonstrate how actions are triggered by the events defined based on future information. Similar to [3], a greedy approach is considered for comparison. The greedy



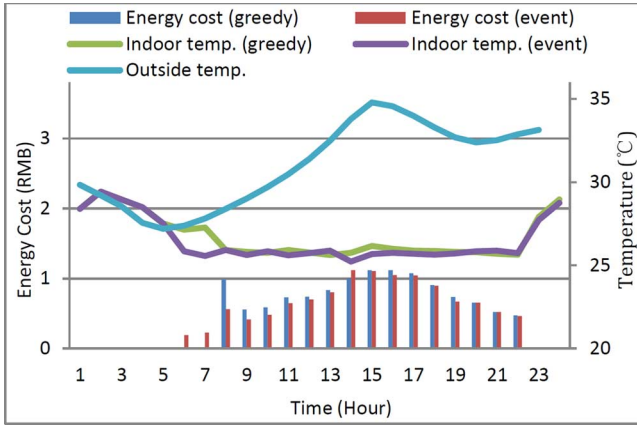


Fig. 2. Energy costs and indoor temperatures under the event-based approach and the greedy approach.

approach minimizes only the current energy cost and does not consider the impact of current control on future energy cost. It is also performed in the LR framework. If the time steps that is looked ahead in the dynamic programming in [3] is reduced to one, then we obtain the greedy algorithm.

The average energy cost per day of the event-based approach is 158.8 RMB with the duality gap of 0.98%. As for the greedy approach, the average energy cost per day is 165.6 RMB. As compared to the greedy approach, the event-based approach saves about 4.3% of energy cost. Considering that the HVAC energy cost is usually huge and accounts for a major part in the building operation cost, the savings is meaningful. For the two approaches, the hourly energy costs and indoor temperatures of a room in a selected day are presented in Fig. 2. In the event-based approach, the event of the average electric price  $\bar{c}^t$  rising above its threshold occurs two hours before the office hour. The optimal action obtained for this event is to precool the room to take advantage of low-price electricity for energy cost reduction. In the greedy approach, no precooling is used.

It can also be seen from the figure that the indoor temperature under the event-based approach drops about 0.5 degree around 2 pm. That is because the event of the average cooling load  $\bar{C}_{L,i}^t$  rising above the threshold occurs due to the high predicted cooling load around 3 pm. The optimal action for this event is precooling. By precooling, the indoor temperature at 3 pm is below 26 degrees, which is the upper limit of comfort range for indoor temperature. On the contrary, under the greedy approach the indoor temperature at 3 pm exceeds 26 degrees, and the comfort requirement for temperature is not satisfied. Therefore, the events defined based on the augmented state variable save energy costs and improve human comfort.

**Example 3: Building With 144 Rooms to Examine Computational Time, Response Time and Energy Cost:** In this example, the event-based approach is compared with the time-based approaches which use stochastic dynamic programming. Following [3], our event-based approach is used to solve the energy cost optimization problem of the same six-floor building with 144 rooms. The results are then used to illustrate that our approach saves computational time significantly and has a much faster response to events, as compared with the time-based

TABLE I  
PERFORMANCE OF EVENT-BASED APPROACH AND TIME-BASED APPROACH

	Time step (minute)	Cost per day (RMB)	Computation time (second)	Response time (second)
Event-Based	-	1501	4.1	4.1
Time-based (SDP)	10	1511	14.6	315
	5	1507	29.7	178
	1	1496	165.5	195

approach [3] which uses the stochastic dynamic programming (SDP) within the LR framework.

Since in this paper, the energy cost optimization problem does not consider the control of lights, blinds, and windows, the time-based approach used for comparison solves the same problem without considering the above devices. Also, in the time-based approach, time needs to be discretized and three time steps including ten minutes, five minutes and one minute are considered. The results under the two approaches are presented in Table I, including energy costs, computational time, and response time (i.e., the time length from events to the execution of decisions in HVAC devices).

As shown in the table, the average computational time of the event-based approach is 4.1 s, saving approximately 70% of the computation time of the time-based approach with the time step of ten minutes. This is because in the event-based approach, the action is optimized only for the first event while in the time-based approach, a day is divided into 144 time stages and the action has to be optimized for each time stage to minimize the total cost of all time stages.

In the event-based approach, an optimization problem is formulated once an event happens. Its response time therefore equals the computational time. On the contrary, the time-based approach formulates the optimization problem only at the beginning of each time stage. Its response time therefore includes not only the computational time but also the time waited to formulate the problem. The waiting time equals the time length from the time an event occurs to the beginning of the next time stage plus. It can be seen from the table that the response time of the event-based approach is 4.1 s. It is much smaller than 315 s of the time-based approach with the time step of ten minutes. Due to the faster response, about 0.7% of energy cost is saved by the event-based approach as compared to the time-based approach with ten minutes as the time step.

In practice, the response time of a few minutes is still practically feasible since it might take more than 10 min to satisfy a given set point of temperature or humidity. The difficulty caused by the delay in response, however, is that occupants would want an HVAC system to take prompt actions (e.g., turn on the fan or increase its speed) once they gave a new set point or when some other events happened. Otherwise, they might think the system is not responsive or reliable and take further actions which may cause energy to be wasted, e.g., keeping on increasing or reducing the set point.

As compared with the time-based approach with the time step of one minute, the event-based approach, however, consumes

0.3% more energy cost. This is because in the Q-learning, the optimal action for an event is applied to control devices with a probability slightly less than one. In this way, every other feasible action has a positive probability to be applied and their Q-factors can get a chance to be updated. Although this time-based approach with 1 min as the time step consumes less energy cost, it is not applicable in practice since its computational time is 178 s, much larger than its time step of one minute. That means when solving the energy cost optimization problem in a moving window manner, the time-based approach cannot solve the problem within a time window.

From the above comparison, we can see that, in general, the event-based approach has a similar level of energy costs to the time-based approaches, but it saves significant computational time and has a much faster response to events.

## VII. CONCLUSION

In this paper, an innovative event-based approach is developed within the LR framework. The HVAC energy cost optimization problem is solved and decisions are executed on an “as needed” basis to reduce computational requirements, have a faster response to events and reduce energy costs. The difficulty is that the optimal policy is non-stationary because the underlying process is nonstationary due to the time-varying uncertainties. Our approach first aggregates the uncertainties into a few key factors such as the COP and cooling load, and then augments the states to include these key factors and other time varying variables in the definition of events. In this way, the optimal policy for the planning horizon is assumed stationary for the purpose of easy computation, although the optimal policy may change over time when the problem is solved in a moving window manner. The problem is then solved by using Q-learning, surrogate LR method, and sample average approximation.

Our approach opens up a new way to solve event-based problems with nonstationary optimal policies. In addition, if these problems are solved within the LR framework, our approach is also applicable by using the event-based method to solve sub-problems and the surrogate LR method as well as sample average approximation to solve a dual problem.

## ACKNOWLEDGMENT

The authors would like to thank M. A. Bragin, University of Connecticut, Storrs, CT, USA, for providing the new surrogate Lagrangian relaxation method [23], [24].

## REFERENCES

- [1] L. Perez-Lombard, J. Ortiz, and C. Pout, “A review on buildings energy consumption information,” *Energy and Buildings*, vol. 40, no. 3, pp. 394–398, 2007.
- [2] J. Xu, P. B. Luh, W. E. Blankson, R. Jerdonek, and K. Shaikh, “An optimization-Based approach for facility energy management with uncertainties,” *HVAC&R Res.*, vol. 11, no. 2, pp. 215–237, 2005.
- [3] B. Sun, P. B. Luh, Q. Jia, Z. Jiang, F. Wang, and C. Song, “Building energy management: Integrated control of active and passive heating, cooling, lighting, shading and ventilation systems,” *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 3, pp. 588–602, Jul. 2013.
- [4] C. Watkins and P. Dayan, “Technical Note: Q learning,” *Mach. Learning*, vol. 8, pp. 279–292, 1992.
- [5] DeST Software. Dept. Building Sci., Tsinghua Univ., 2008 [Online]. Available: <http://www.dest.com.cn>
- [6] A. I. Dounis, P. Tiropanis, A. Argiriou, and A. Diamantis, “Intelligent control system for reconciliation of the energy savings with comfort in buildings using soft computing techniques,” *Energy and Buildings*, vol. 43, pp. 66–74, 2011.
- [7] X. Xu, S. Wang, Z. Sun, and F. Xiao, “A model-based optimal ventilation control strategy of multi-zone VAV air-conditioning systems,” *Appl. Thermal Eng.*, vol. 29, pp. 91–104, 2009.
- [8] S. Soyguder and H. Alli, “An expert system for the humidity and temperature control in HVAC systems using ANFIS and optimization with fuzzy modeling approach,” *Energy and Buildings*, vol. 41, pp. 814–822, 2009.
- [9] X. Zhao, P. B. Luh, and J. Wang, “Surrogate gradient algorithm for Lagrangian relaxation,” *J. Optim. Theory Applicant.*, vol. 100, no. 3, pp. 699–712, 1999.
- [10] X. R. Cao, “Basic ideas for event-based optimization of Markov systems,” *Discrete Event Dynamic Syst.: Theory and Applicant.*, vol. 15, no. 2, pp. 169–197, 2005.
- [11] Y. Zhao, X. Chen, Q. Jia, X. Guan, S. Zhang, and Y. Jiang, “Long-term scheduling for cascaded hydro energy systems with annual water consumption and release constraints,” *IEEE Trans. Autom. Sci. Eng.*, vol. 7, no. 4, pp. 969–976, Oct. 2010.
- [12] Q. Jia, “On solving optimal policies for finite-stage event-based optimization,” *IEEE Trans. Autom. Control*, vol. 56, no. 9, pp. 2195–2200, Sep. 2011.
- [13] B. Sun, P. Luh, Q. Jia, and B. Yan, “Event-based optimization with non-stationary uncertainties to save energy costs of HVAC systems in buildings,” in *Proc. IEEE Conf. Autom. Sci. Eng.*, 2013, pp. 436–441.
- [14] A. J. Kleywegt, A. Shapiro, and T. Homem-De-Mello, “The sample average approximation method for stochastic discrete optimization,” *SIAM J. Optimization*, vol. 12, no. 2, pp. 479–502, 2001.
- [15] B. Verweij, S. Ahmed, A. J. Kleywegt, G. Nemhauser, and A. Shapiro, “The sample average approximation method applied to stochastic routing problems: A computational study,” *Computational Optim. Applic.*, vol. 24, pp. 289–333, 2003.
- [16] X. R. Cao, *Stochastic Learning and Optimization: A Sensitivity-Based Approach*. New York, NY, USA: Springer, 2007.
- [17] Z. Liu, X. Chen, X. Xu, and X. Guan, “A decentralized optimization method for energy saving of HVAC systems,” in *Proc. IEEE Conf. Autom. Sci. Eng.*, 2013, pp. 225–230.
- [18] E. Asadi, M. Silva, C. Antunes, and L. Dias, “A multi-objective optimization model for building retrofit strategies using TRNSYS simulations, GenOpt and MATLAB,” *Building and Environment*, vol. 56, pp. 370–378, 2012.
- [19] H. Huang, L. Chen, M. Mohammadzahari, E. Hu, and M. Chen, “Multi-zone temperature prediction in a commercial building using Artificial Neural Network model,” in *Proc. IEEE Conf. Control Autom.*, 2013, pp. 1948–1949.
- [20] P. Ferreira, A. Ruano, S. Silva, and E. Conceicao, “Neural networks based predictive control for thermal comfort and energy savings in public buildings,” *Energy and Buildings*, vol. 55, pp. 238–251, 2012.
- [21] S. Soyguder and H. Alli, “An expert system for the humidity and temperature control in HVAC systems using ANFIS and optimization with fuzzy modeling approach,” *Energy and Buildings*, vol. 41, pp. 814–822, 2009.
- [22] Y. Ma, F. Borrelli, B. Hencsey, B. Coffey, S. Benguea, and P. Haves, “Model predictive control for the operation of building cooling systems,” *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 3, pp. 796–802, Mar. 2012.
- [23] M. A. Bragin, P. B. Luh, J. H. Yan, N. P. Yu, X. Han, and G. A. Stern, “An efficient surrogate subgradient method within Lagrangian relaxation for the payment cost minimization problem,” in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2012, pp. 1–8.
- [24] M. A. Bragin, P. B. Luh, J. H. Yan, N. P. Yu, X. Han, and G. A. Stern, “Convergence of the surrogate Lagrangian relaxation method,” *J. Optim. Theory Applic.*, vol. 164, no. 1, pp. 173–201, 2015.
- [25] L. Xia, Q. Jia, and X. R. Cao, “A tutorial on event-based optimization—a new optimization framework,” *Discrete Event Dynamic Syst.*, pp. 1–30, 2013.
- [26] X. Zhao, P. B. Luh, and J. Wang, “Surrogate gradient algorithm for Lagrangian relaxation,” *J. of Optimization Theory and Applicant.*, vol. 100, no. 3, pp. 699–712, 1999.
- [27] R. Y. Zhao, C. Y. Fan, and D. H. Xue, *Air Conditioning* (in Chinese). Beijing, China: China Construction Industry, 2009.
- [28] Z. Ren and B. H. Krogh, “State aggregation in Markov decision processes,” in *Proc. 41st IEEE Conf. Decision and Control*, 2002, vol. 4, pp. 3819–3824.
- [29] Q. Jia, “On state aggregation to approximate complex value functions in large-scale Markov decision processes,” *IEEE Trans. Autom. Control*, vol. 56, no. 2, pp. 333–344, Feb. 2011.



**Biao Sun** (S'11) received the B.E. degree in automation from Tsinghua University, Beijing, China, in 2006, and the Ph.D. degree from the Center for Intelligent and Networked Systems (CFINS), Department of Automation, TNLIST, Tsinghua University, in 2014.

His research interests include optimal control for building energy savings, and fault detection and diagnosis of heating ventilation and air conditioning systems (HVAC).



**Peter B. Luh** (S'77–M'80–SM'91–F'95) received the B.S. degree from National Taiwan University, Taipei, Taiwan, the M.S. degree from the Massachusetts Institute of Technology, Cambridge, MA, USA, and the Ph.D. degree from Harvard University, Cambridge, MA, USA.

He has been with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT, USA, since 1980, where he is the SNET Professor of Communications and Information Technologies. He is also a member of the Chair Professors Group at the Center for Intelligent and Networked Systems (CFINS), Department of Automation, Tsinghua University, Beijing, China, and a Chair Professor with the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan. His research interests include smart, green and safe buildings; smart grid, electricity markets, and effective renewable integration to the grid; and intelligent manufacturing systems.

Dr. Luh was the founding Editor-in-Chief of the IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING and an Editor-in-Chief of the IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION.



**Qing-Shan Jia** (S'02–M'06–SM'11) received the B.E. degree in automation and Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2002 and 2006, respectively.

He is currently an Associate Professor with the Center for Intelligent and Networked Systems (CFINS), Department of Automation, TNLIST, Tsinghua University, Beijing, China. He was a Visiting Scholar with Harvard University in 2006 and a Visiting Scholar with the Hong Kong University of Science and Technology in 2010. His research interests include theories and applications of discrete event dynamic systems and simulation-based performance evaluation and optimization of complex systems.



**Bing Yan** (S'11) received the B.S. degree from Renmin University of China, Beijing, China, in 2010. She is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT, USA.

Her research interests include scheduling of manufacturing systems, optimized energy management of smart buildings and eco-communities, design and operation optimization of microgrids, and exergy-based operation optimization of distributed energy system.