Deep Reinforcement Learning based Smart Building Energy Management

Minh N. H. Nguyen, Tuyen Le Pham, Nguyen H. Tran and Choong Soon Hong
Department of Computer Science and Engineering, Kyung Hee University
Email: [minhnhn, tuyenple, nguyenth, cshong]@khu.ac.kr

Abstract: The increasing electronic devices in modern smart buildings leads to the necessity of an efficient energy management based on smart meters. Furthermore, in U.S. and other countries in Europe has launched the dynamic energy price (Time-of-Used) tariff programs to incentivize consumer to regulate their energy loads according to the time-varying electric prices. Towards this end, electronic devices scheduling becomes an essential feature of the future smart building controllers, which can recommend economical plans for customers. In this paper, we apply a deep reinforcement learning technique to control electric devices more efficient and reduce the energy cost regarding the user dissatisfaction cost.

1. Introduction

Nowadays, the proliferation of smart devices, household renewable energy systems, and smart grids have opened a new research area, i.e., energy management for smart buildings. Different from conventional research, the customer role has transited from a passive role to an active role in the energy management problem [1], [2]. According to the foreknowledge of electricity price from the public utility, customers can schedule their own electrical device loads with respect to daily energy profiles.

In this paper, we apply the state-of-the-art actor-critic approach, deep deterministic policy gradient (DDPG) [3] for energy usage scheduling, which can be considered as an efficient approximated solution approach of the cost minimization problem for nanogrid controllers [2].

2. Smart Building Energy Management

A. System Model

In this section, we propose our nanogrid energy management controller model as in Fig. 1 for smart building. Our model considers the existing of renewable generation system and battery storage. The nanogrid controller can make decisions on device operation to minimize their energy cost and user dissatisfaction cost.

![Fig. 1: Nanogrid controller model.](image)

The energy cost of a smart building can be defined as the buying/selling from/to the grid with an amount of energy (i.e., $E_{grid}^{t} = E_{grid}^{t-1} - E_{grid}^{t}$) as follows:

\[
\Psi(E,d) = \sum_{t=1}^{T} \mu(E_{grid}^{t}, t, d),
\]

where

\[
\mu(E,t,d) = \begin{cases} 
E & \text{if } E \geq 0, \text{ [buying]} \\
E^{-1} & \text{others, [selling]} 
\end{cases}
\]

The electric price function is also a piecewise linear function of energy usage at a certain time. In addition to the energy cost, we include the user dissatisfaction cost based on two assumptions for the controlled electric devices:

A1. The nanogrid controller expects to provide deferrable loads as soon as it can (e.g., electric vehicle charging).

A2. The better for user comfort when using electronic devices (e.g., air conditioner, light.)

Accordingly, we define the following user discomfort cost as

\[
\zeta(x', x^*) = \sum_{t=1}^{T} \omega \times x^d_t (t - t^*)^2 + \kappa Q_{min}^d - Q_{min}^d
\]

B. Problem Formulation

min. $\Psi(E,d) + \zeta(x^d, x^*)$

subject to $E_t^r = x^d_t e^d_t + e^d_t + x^v_t Q_{max}^v + x^b_t e^b_t$, $\forall t$, (3)

$\begin{cases} 
x^d_t, x^v_t \in [0,1], x^b_t \in [-1,1], \forall t, 
\end{cases}$ (4)

$\begin{cases} 
\sum_{t=1}^{T} x^d_t e^d_t = E^d_d (\text{Deferrable load}) 
\end{cases}$ (5)

$\begin{cases} 
x^d_t = 0, \forall t \notin [t, t^*], 
\end{cases}$ (6)

$\begin{cases} 
Q_{min}^d \leq x^d_t Q_{max}^d \leq Q_{max}^d, \forall t, \text{(Variable load)} 
\end{cases}$ (7)

$\begin{cases} 
E_{t+1}^b - E_t^b + x^b_t e^b_t, \forall t = 0, \ldots, T-1, (Battery) 
\end{cases}$ (8)

$\begin{cases} 
E_{min}^b \leq E_t^b \leq E_{max}^b, \forall t, 
\end{cases}$ (9)

The variables of this problem are decisions on deferrable load $x^d$, variable load $x^v$, and battery $x^b$. This optimization problem is energy and user dissatisfaction cost minimization over the control horizon $T$ (e.g., 1 day with 48 half-hour time slots). Since we assume that we can only get the full state information at the beginning of a time slot, then the closed-form solution of this problem is unattainable.
3. Deep Deterministic Policy Gradients

In this section, we investigate the DDPG technique [3] as a solution approach for the cost minimization problem. Using this state-of-the-art deep reinforcement learning approach, we iteratively update the actor network and critic network as in Fig. 2. Three actions of the controller will be the output of the actor network while the critic network will approximate action value function $Q(s, a)$.

For the energy price, we apply the dynamic electric price such as Time-of-Use (TOU) tariff, which includes on-peak, off-peak, mid-peak rates according to the time period in a day in Austin city [4]. The selling rate also follows Value-of-Solar rate in Austin [4].

<table>
<thead>
<tr>
<th>Type</th>
<th>Weekdays (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>On-Peak</td>
<td>[14:00, 20:00]</td>
</tr>
<tr>
<td>Mid-Peak</td>
<td>[6:00, 14:00], [20:00, 22:00]</td>
</tr>
<tr>
<td>Off-Peak</td>
<td>[22:00, 6:00]</td>
</tr>
</tbody>
</table>

Table 1: TOU Tariff 2016 in Austin, U.S. [4]

4. Simulation Results

The energy loads of devices are extracted and generated from the dataset from UK [5]. We train for the first 430 days and test with the last 31 days. Each day is divided into 48 time slots, each time slot duration is 30 minutes. At the beginning of time slot, the controller will collect the state information (i.e., electric price function, required variable load, battery status, renewable generation) and produce the actions $(x^e, x^v, x^b)$. The required period of the deferrable load is set to 17 p.m. ~ 6 a.m. The renewable energy generation is a normal distribution with mean 400Wh in the daytime and will be zero in the nighttime since we consider solar energy source alone. Battery is limited from 100Wh ~ 500Wh. The control unit deferrable load and battery are 200Wh and 100Wh, respectively.

The Fig. 3 shows the costs reduction in the training stage. After the high variance in the first 220 days, the costs have a tendency to become more stable and decrease.
Finally, Fig. 5 illustrates the first 4 days of testing data. With DDPG, deferrable loads are postponed to reduce the energy usage in the on-peak tariff period.

5. Conclusion & Future Work

In this paper, we investigate an efficient energy management approach in a nanogrid controller, which can help to reduce energy cost while sacrificing small level of user satisfaction compared to a simple heuristic strategy. In the future work, we advocate extending the work for the cooperation of nanogrids controller when access the common renewable energy.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (NRF-2016R1D1A1B01015320).

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2017-2013-0-00717)supervised by the IITP(Institute for Information & communications Technology Promotion). *Dr. CS Hong is the corresponding author.

References: