Meta-Reinforcement Learning for Proactive Energy Demand Scheduling in Smart City with Edge Computing

Md. Shirajum Munir, and Choong Seon Hong
Department of Computer Science and Engineering, Kyung Hee University, South Korea.
{ munir, cshong }@khu.ac.kr

Abstract

Renewable energy enabled sustainable energy management ensures a high degree of reliability in order to fulfill the energy demand of a smart city. In such case, renewable energy generation is random over time and also energy consumption of smart city users' is nondeterministic in nature. In this paper, we propose a proactive energy demand scheduling for the smart city using a meta-reinforcement learning (Meta-RL) approach, which not only overcome the challenges of renewable energy aware demand scheduling but also establishes a strong relationship for both energy generation and consumption over time. In order to solve the proposed model distributedly, we consider edge computing for executing local agent to determine individual users' policy with respect to energy consumption and renewable energy generation (users' own sources). On the other hand, microgrid controller determines meta-policy through a meta-agent with Recurrent Neural Network (RNN). Since, a meta-agent accepts local policy as an input with historical observations, which ensures fast and efficient execution of proactive demand scheduling. Finally, we evaluate our proposed model, which demonstrates significant performance gain with high accuracy scheduling.

1. Introduction

In modern development arena, smart city and renewable energy are indispensable toward the ecological growth of urban technology to enable sustainable smart services [1] [2]. A microgrid is capable to fulfill that huge amount of energy demand by enabling the efficient demand scheduling of smart city energy consumptions. However, the challenges come with the unpredictable nature of both energy consumption and renewable generation, which also have a strong relationship over the history of energy consumption and generation [3] [4].

Therefore, to overcome those challenges, a proactive energy demand scheduling is essential such that both energy consumption and renewable energy generation can be considered. In order to do that, we propose a meta-reinforcement learning (Meta-RL)-based [5] energy scheduling model, in which this method is capable to handle both energy consumption and generation with the historical and current observations using Recurrent Neural Network (RNN).

In this paper, we propose a proactive energy demand scheduling model using a Meta-RL learning approach for the smart city energy users. To solve this model distributedly, first, we execute local agent with edge computing facility, which determines a local policy with respect to energy consumption and generation (user's own renewable sources) for each energy users'. Second, we reuse the historical observations and local policy in order to learn the meta-policy for energy scheduling at microgrid controller. Finally, in order to validate model, we have implemented the proposed model and achieves a significant performance gain in order to proactive energy demand scheduling.

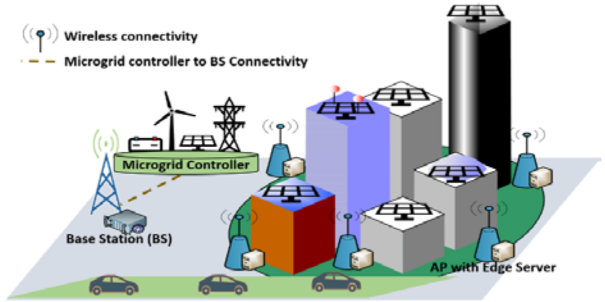

Figure 1: MEC-enabled microgrid-powered smart city scenario

2. System Model and Problem Formulation

We consider a microgrid-powered smart city scenario that includes both renewable and non-renewable energy sources, where each individual building has its own renewable (solar) energy sources. The city connected with edge computing enabled wireless networks to fulfill the smart city services, as seen in Figure 1. We also consider a finite time horizon set $T = \{1, \ldots, T\}$ and each time slot t consists of one-hour duration. Therefore, at time slot t, there is a set $C = \{1, \dots, \mathrm{C}\}$ of energy consumption of C smart city energy users. A set $F = \{1, \dots, \mathrm{F}\}$ determines the renewable energy generation for all users from their individual energy sources. For time slot t the total energy demand for the smart city is denoted by,

$$\boldsymbol{\omega}^{\mathbf{d}}(\mathbf{t}) = \sum_{\forall \mathbf{c} \in \mathbf{C}} \mathbf{d}_{\mathbf{c}}(\mathbf{t}). \quad (1)$$

Total renewable energy generation is as follows:

$$\boldsymbol{\omega}^{\mathbf{r}}(\mathbf{t}) = \sum_{\forall \mathbf{f} \in \mathbf{F}} \mathbf{d}_{\mathbf{f}}(\mathbf{t}). \quad (2)$$

If the energy demand $\omega^{\mathrm{d}}(\mathrm{t})$ is larger than the renewable energy generation $\omega^{\mathrm{r}}(\mathrm{t})$ then microgrid generates non-renewable energy $\omega^{\mathrm{n}}(\mathrm{t})$ to fulfill the energy demand for the city users. The amount of non-renewable energy generation amount is as follows:

$$\omega^{\mathrm{n}}(\mathrm{t}) = \omega^{\mathrm{d}}(\mathrm{t}) - \omega^{\mathrm{r}}(\mathrm{t}). \quad (3)$$

We consider set of actions $A = \{A_1, \dots, A_C\}$ for all energy users, where for a user c, each action a_t consists with two actions tuple $a_t \in <\beta^1, \beta^0>$ at time t. β^1 determines that the user can fulfill energy consumption from it's own renewable sources and β^0 represents a action regarding the energy needs to consume from non-renewable energy. There is a state space set $S = \{S_1, \dots, S_c\}$ for all users, where each state space $s_t \in S_c$ comprises a tuple $\mathrm{s_t} \in <\omega^{\mathrm{d}}(\mathrm{t}), \omega^{\mathrm{r}}(\mathrm{t}), \omega^{\mathrm{n}}(\mathrm{t})>$ of user c.

We assume that every user is capable to communicate with nearby edge server and also the edge server is able to execute the proposed model. In this model, we have formulated for a single energy user c and this model is capable to handle all the users. The transition probability for s_t is defined by [1],

$$P_a(s_{t-1}, s_t) = P(s_t \mid s_{t-1}, a_{t-1}). \quad (4)$$

The policy π with parameter θ is defined as follows:

$$\pi_\theta(\mathrm{a}|\mathrm{s}) = \mathrm{P}_{\mathrm{a_t}}(\mathrm{A_t} = \mathrm{a} \,|\mathrm{S_t} = \mathrm{s}) \quad (5)$$

Now, the reward function depends on the expectation of cumulative rewards and the objective is to maximize the expectation of cumulative reward,

$$R(a,s) = max_r\, E\left[\sum_{l=0}^{\infty} \gamma^l\, r(a_t, s_t)\right], \quad (6)$$

where, γ is the learning rate and γ^l ensures the convergence of the reward calculation. The state-action value function is defined by,

$$Q^{\pi_\theta}(s,a) = E_\pi\left[\sum_{l=0}^{\infty} \gamma^l\, r(a_t, s_t)\, \middle|\, s_{t-1}, a_{t-1}\right]. \quad (7)$$

The state value function for policy π_θ is,

$$V^{\pi_\theta}(s) = E_{\pi_\theta(a_{t-1}|s_{t-1})}[Q^{\pi_\theta}(s,a)]. \quad (8)$$

A set of observations is denoted by $\mathrm{X}_t = \{\mathrm{x}_1, \dots, \mathrm{x}_X\}$, where $\mathrm{X}_t = \{\mathrm{s}_0, \mathrm{a}_0, \mathrm{r}_0, \dots, \mathrm{s}_{t-1}, \mathrm{a}_{t-1}, \mathrm{r}_{t-1}, \mathrm{s}_t\}$ consists with current state s_t and previous history (i.e., state, action, and reward). We use this observation as an input for learning the meta-agent. In this model, we have used the LSTM (long short-term memory) for RNN (recurrent neuron networks) [4]. The objective is to find the optimal policy π_θ and value function $V^{\pi_\theta}(s)$ to determine the action regarding energy demand scheduling decision for each user at time slot t.

Algorithm 1: *Proactive Energy Demand Scheduling*

```
Input : State Space s_t ∈ S_c, ∀ c ∈ C
Output: Meta Policy π_θ, Value V^{π_θ}(s), ∀
c ∈ C , a_t ∈ < β^1, β^0 >
Repeat{
    Repeat{
    Step 1: Local Policy Estimation:
      Calculate:
      X_t = {s_0, a_0, r_0, ..., s_{t-1}, a_{t-1}, r_{t-1}, s_t}
      using eq. (4) to eq. (8)
      X_t Send to microgrid controller
    Step 2: Meta Policy Estimation:
      Calculate :
      Meta Policy π_θ, Value V^{π_θ}(s), ∀
c ∈ C
      Using LSTM model
    } Until ∀ c ∈ C
} Until ∀ t ∈ T
Schedule: a_t ∈ < β^1, β^0 >, t ∈ T, ∀ c ∈ C
```

3. Solution with Meta-Reinforcement Learning

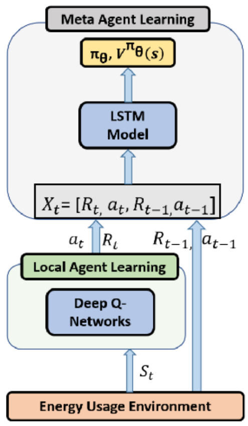

Figure 2: Meta-RL model for a single user

The proposed *proactive energy demand scheduling for smart city* problem is solved distributedly, where first, we obtain the local policy by learning the local agent with respect to energy consumption and renewable energy generation through the nearby edge server. Second, in order to generate meta-policy, we send local policy information to the microgrid controller alone with previous policy observation, so that meta-agent can learn very fast with the optimal decision. This procedure is the same for every energy user and meta-RL [5] model for a single user is shown in Figure 2.

In the proposed proactive energy demand scheduling algorithm (Algorithm 1), step 1 in line 5 to 8 calculates the local policy using the equation (4) to equation (8) and send the calculated observation X_t to microgrid controller (line 9) to determine the meta-policy π_θ and value function $V^{\pi_\theta}(s)$ estimation. The microgrid controller usages observations X_t as an input for LSTM model [4], which includes the historical observation and local policy to determine the proactive energy demand scheduling. Finally, in line 16 of this algorithm generates schedule for every smart city energy users in order to fulfill the energy demand.

4. Performance Evaluation

In this research, we have implemented our proposed model on python platform, along with TensorFlow APIs. To evaluate this Meta-RL model, we have used Solar panel dataset UMassTraceRepository [6] for renewable energy generation and, COMMERCIAL LOAD DATA E PLUS OUTPUT part1(USA CA Los Angeles Intl AP722950TMY3) from openei [7] for smart city energy consumption. The important parameters for model evaluations are as follows: no. of solar units 150, no. of energy consumption institutions 16, no. of LSTM units 48, RL learning rate 0.8, and LSTM look back state 5.

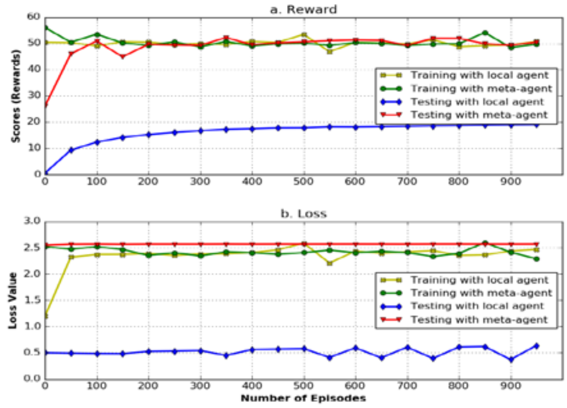

Figure 3: Convergence of Meta-RL

In Figure 3 describes the convergence of the proposed Meta-RL with the local agent-based RL method for both training and testing, where in first subfigure (a. Reward) clearly shows that local agent-based RL takes around 600 episodes (yellow line with x marks) for convergence, on the other hand, proposed model achieves convergences after 200 episodes (green line with o marks) with higher reward. The second subfigure (b. Loss) depicts that for testing, the local agent-based method falls inconsistent (loss is fluctuated) after the 500 episodes, whereas the proposed model gains the stability.

Meta-RL based energy scheduling for the whole smart city is presented in Figure 4, which shows that overall energy scheduling (green line with x marks) almost archives the energy consumption goal. Finally, this figure also shows that when energy consumption is high then scheduling are increased positively with 5.07%, which is one of the trade-off of the proposed model because of high bias.

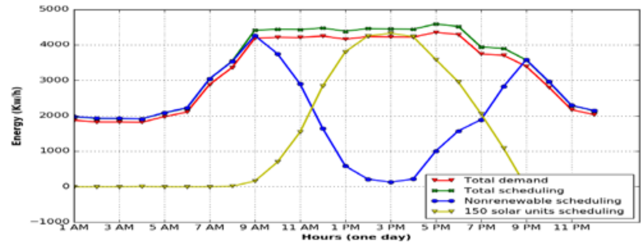

Figure 4: Meta-RL based energy scheduling

5. Conclusion

The proposed proactive energy demand scheduling for the smart city based on meta-reinforcement learning is a novel approach in respect to smart city energy management. We have devised local policy distributedly with the efficient usage of edge computing and microgrid controller determines the meta-policy by considering the local policy and historical observations. In our model evaluation, we have achieved a high accuracy with respect to energy demand scheduling. Additionally, this approach will significantly reduce the risk of smart city energy failure due to nondeterministic nature of both energy consumption and generation.

Acknowledgement:

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2017R1A2A2A05000995). *Dr. CS Hong is the corresponding author.

References

[1] M. S. Munir, S. F. Abedin, M. G. R. Alam, N. H. Tran and C. S. Hong, “Intelligent service fulfillment for software defined networks in smart city,” *2018 International Conference on Information Networking (ICOIN)*, pp. 516-521, 2018. (in Chiang Mai, Thailand)

[2] M. S. Munir, S. F. Abedin, M. G. R. Alam, D. H. Kim and C. S. Hong, " Smart Agent based Dynamic Data Aggregation for Delay Sensitive Smart City Services," *Journal of KIISE*, vol. 45, no. 4, pp. 395-402, April 2018.(in Korea)

[3] Y. Zhang, M. H. Hajiesmaili, S. Cai, M. Chen and Q. Zhu, "Peak-Aware Online Economic Dispatching for Microgrids," *in IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 323-335, Jan. 2018

[4] M. S. Munir, S. F. Abedin, M. G. R. Alam, D. H. Kim, and C. S. Hong, “RNN based Energy Demand Prediction for Smart-Home in Smart-Grid Framework,” *Korea Software Congress 2017*, pp. 437-439, 2017 (in Korea)

[5] J. X. Wang et al., “Learning to reinforcement learn,” *CogSci* , 2017. (In London, UK)

[6] Online: “Solar panel dataset”, UMassTraceRepository: http://traces.cs.umass.edu/index.php/Smart/Smart. (Visited on 3 July, 2018)

[7] Online: “USA_CA_Los.Angeles.Intl.AP.722950_TMY3”, COMMERCIAL_LOAD_DATA_E_PLUS_OUTPUT.part1: https://openei.org/datasets/files/961/pub/. (Visited on 3 July, 2018)