Large-Scale Home Energy Management Using Entropy-Based Collective Multiagent Reinforcement Learning Framework

Extended Abstract

Yaodong Yang, Jianye Hao*, Yan Zheng, Xiaotian Hao, Bofeng Fu
College of Intelligence and Computing, Tianjin University
Tianjin, China
yydapple@gmail.com,{jianye.hao,yanzheng,xiaotianhao}@tju.edu.cn,jxmg686000@163.com

ABSTRACT

Smart grids are contributing to the demand-side management by integrating electronic equipment, distributed energy generation and storage, and advanced meters and controllers. With the increasing adoption of distributed energy generation and storage systems, residential energy management is drawing more and more attention, which is regarded as being critical to demand-supply balancing and peak load reduction. In this paper, we focus on a microgrid in which a large-scale modern homes interact together to optimize their electricity cost. We present an Entropy-Based Collective Multiagent Deep Reinforcement Learning (EB-C-MADRL) framework to address it. Experiments demonstrate that EB-C-MADRL can reduce both the long-term group power consumption cost and daily peak demand effectively compared with existing approaches.

KEYWORDS

Energy and emissions; Agent solutions of significant social and economic impact; Other innovative application areas

ACM Reference Format:

Yaodong Yang, Jianye Hao*, Yan Zheng, Xiaotian Hao, Bofeng Fu . 2019. Large-Scale Home Energy Management Using Entropy-Based Collective Multiagent Reinforcement Learning Framework. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 3 pages.

1 INTRODUCTION

Meeting the growing energy demand due to the presence of more volatile types of loads raises a major challenge for the power grid [8, 12]. To satisfy demand that varies sharply, companies usually have to install additional generation capacity to meet the peak demand and charge end-users higher costs. At the same time, the increasing renewable generation is naturally intermittent, which makes the power grid hard to maintain the demand-supply balance. The peak load and supply-demand imbalance have received more and more attention by energy generation and dsitribution companies [15].

The home energy demand-side management (DSM) [9] has been proposed to handle the above problems, such as dynamic programming [14], game theory [5] and reinforcement learning (RL) [6]. However, these works only consider incomplete subsets of the home power systems and require rigid schedules for end users' appliances usage. Recently, smart homes combined with the distributed energy generation (DG) and distributed energy storage (DS) show the great possibility for the revolution of the power grid [3, 7]. It provides us with opportunities of unfreezing the rigid schedule for users. RL based DSM techniques for the smart home was first investigated in [1] and then extended in [11] with electric vehicles (EV).

However, these smart home DSM works focus on optimizing the energy activities for a single household without considering the group aggregate effect which would result in overloads on the transformer [2]. To this end, we research on the user-friendly DSM techniques for a smart home community. We propose an entropy-based collective multiagent reinforcement learning (MARL) framework to address the large-scale energy cost optimization problem.

2 MICROGRID ELECTRICITY MARKET

At the beginning of each time slot t, the home EMS needs to decide two actions based on its own state: $P_{c,t}$ for power trading amount and $C_{e,t}$ for the EV charging rate. Our microgrid market mechanism has two trading processes: the internal trading process and the external trading process. Households trade inside the group first to satisfy the demand of each other. If the internal trading cannot fully meet the group, then the external smart grid will deal with the unsatisfied demand. To encourage users to actively participate in such a microgrid, we set the internal power price $p_{in,t}$ the average of external power selling price $p_{os,t}$ and external power buying price $p_{ob,t}$ for customers. Extra aggregate demand or supply would be processed by external trading after internal trading. The final cleaning electricity price for the trading power $P_{c,t}$ is:

$$
\begin{aligned}
p_{s,t} &= \begin{cases} \frac{p_{in,t}\psi_{b,t}+p_{os,t}(\psi_{s,t}-\psi_{b,t})}{\psi_{s,t}}, & if\,\psi_{s,t} \geq \psi_{b,t} \\ p_{in,t}, & if\,\psi_{s,t} < \psi_{b,t} \end{cases} \\
p_{b,t} &= \begin{cases} p_{in,t}, & if\,\psi_{s,t} \geq \psi_{b,t} \\ \frac{p_{in,t}\psi_{s,t}+p_{ob,t}(\psi_{b,t}-\psi_{s,t})}{\psi_{b,t}}, & if\,\psi_{s,t} < \psi_{b,t} \end{cases},
\end{aligned} \tag{1}
$$

where $p_{s,t}$ and $p_{b,t}$ are the power selling price and buying price at time t. $\psi_{s,t}$ and $\psi_{b,t}$ are the total power selling and buying amount. Through the incentive mechanism, we turn the smart home community a multiagent system, where each agent's reward is determined by trading prices affected by the total group. Promoting the group coordination can be solved by MARL approaches.

3 EB-C-MARL FRAMEWORK

3.1 Collective Group Behavior

The massive market dynamic property raises huge challenges. One primary problem is that each agent's policy is changing as training

* Corresponding author: Jianye Hao.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

progresses, and the environment becomes non-stationary from the perspective of any individual agent [4]. Even if we could obtain actions from other agents, in the large-scale multiagent systems, the joint action space of the agents grows exponentially with the number of agents and makes the value function learning extremely hard [13]. However, in market settings where agents are influenced from their collective action effect, we could represent such collective influence by the market dynamics abstraction to avoid above issues.

Collective DQN. Each agent is coordinating with the mircogird market instead of directly interacting with any individual. Thus, we abstract market macro-actions to replace other agents' joint action to simplify the multiagent Q-function significantly.

$$Q^i(s, a^1, a^2, \ldots, a^N \equiv Q^i(s, a^i, a^{market}), \tag{2}$$

where the abstraction of market dynamics a^{market} includes the seller group collective action a_s, the buyer group collective action a_b and group EV charging distribution $\vec{C}_e$. One additional privacy benefit is that each household only need to access to its own states without knowing any other. Then we obtain Equation 3:

$$Q^i(s, a^i, a^{market}) \approx Q^i(o^i, a^i, a_s, a_b, \vec{C}_e). \tag{3}$$

The abstractions of current market dynamics cannot be exactly obtained as all households make decisions at the same time. Instead we propose using yesterday's group collective actions to approximate current market dynamics by human life's daily periodicity:

$$Q^i(o^i, a^i, a_s, a_b, \vec{C}_e) \approx Q^i(o^i, a^i, a_s^{'}, a_b^{'}, \vec{C}_e^{'}), \tag{4}$$

where $a_s^{'}$, $a_b^{'}$ and $\vec{C}_e^{'}$ are group action statistics at one day ago.

Collective A2C. Similarly, collective actions enhance A2C.

$$\begin{aligned} \pi^i(s, a^1, \ldots, a^{i-1}, a^{i+1}, \ldots, a^N) &\equiv \pi^i(s, a^{market}) \\ \approx \pi^i(o^i, a_s, a_b, \vec{C}_{e,t}) &\approx \pi^i(o^i, a_s^{'}, a_b^{'}, \vec{C}_{e,t}^{'}). \end{aligned} \tag{5}$$

3.2 Reward Shaping with Individual Entropy

For reducing the daily peak load, we use **individual entropy** to diversify household EV charging to different time slots. The uncoordinated RL learning will result in high peak load as EV would charge in the low electricity price period coincidentally. Inspired by [10], we utilize a more accurate individual entropy in the reward function to diversify the EV charging behavior. Intuitively, if one household chooses a low-frequency action, a higher bonus would be assigned to the household as it contributes more to the system's action entropy H_t. The h_t^i for user i is calculated as follows:

$$h_t^i = \frac{-\log p_{a_t^i}}{N}, \tag{6}$$

where p_{a_t} is the frequency of action a_t performed at t. h_t^i gives the incentive to choose a different action from current high-frequency actions. Therefore, it helps reduce the peak load by mitigating the phenomenon that households charge EV concurrently. h_t^i is accurate credit assignment of the system's entropy which represents the distribution degree of EV charging behavior:

$$\sum_i h_t^i = \sum_i \frac{-\log p_{a_t^i}}{N} = \sum_{a_t^i} \frac{-n_{a_t^i} \log p_{a_t^i}}{N} = \sum_{a_t^i} -p_{a_t^i} \log p_{a_t^i} = H_t. \tag{7}$$

4 EXPERIMENTS AND ANALYSIS

4.1 Validating the Collective Group Behavior

We first validate the collective group behavior abstraction and compare the proposed control algorithms with a rule method and DQN. The rule-based control algorithm is called *Naive-greedy* policy described in [1], which charges the EV when arriving home and sell the energy when there is a power surplus. Then we augment both DQN and A2C with market dynamics approximations to validate the collective group behavior abstraction. Table 1 shows the results and collective A2C has the least electricity operating cost.

Table 1: Group Power Operating Results

Algorithm	Operating Cost ($)	Peak Load (kwh)
Naive Greedy	-263195.44	453302.63
DQN	-111133.42	421048.18
A2C	-92173.61	478321.76
Collective DQN	-93087.09	429021.03
Collective A2C	-88878.34	465816.24

4.2 Validating the Individual Entropy

Despite achieving the least cost, collective A2C still has high peak loads by the uncoordinated EV charging. To mitigate the new peaks, we enhance collective DQN and collective A2C with individual entropy to encourage agents to diversify EV charging. Table 2 gives the results of related methods and EB-C-MADRL. Compared with DQN, entropy-based collective A2C (EB-C-A2C) achieves 24.69% cost reduction and 5.15% peak load reduction.

Table 2: Group Power Operating Results

Algorithm	Operating Cost ($)	Peak Load (kwh)
Naive Greedy	-263195.44	453302.63
DQN	-111133.42	421048.18
Collective A2C	-88878.34	465816.24
EB-C-A2C	-83689.13	399381.48

5 CONCLUSION

In this paper, we focus on a large-scale smart home EMS optimization problem. We propose EB-C-MADRL to learn home EMS control policies in a community microgird market. Simulation experiments exhibit superior performance of our method in terms of the electricity operating cost saving and the daily peak load reduction.

ACKNOWLEDGMENTS

The work is supported by the National Natural Science Foundation of China (Grant Nos.: 61702362, U1836214), Special Program of Artificial Intelligence, Tianjin Research Program of Application Foundation and Advanced Technology (No.: 16JCQNJC00100), and Special Program of Artificial Intelligence of Tianjin Municipal Science and Technology Commission (No.: 569 17ZXRGGX00150).

REFERENCES

[1] Heider Berlink and Anna Helena Reali Costa. 2015. Batch Reinforcement Learning for Smart Home Energy Management. In *Proceedings of the 24th International Conference on Artificial Intelligence*. AAAI Press, 2561–2567.

[2] Ivana Dusparic, Colin Harris, Andrei Marinescu, Vinny Cahill, and Siobhan Clarke. 2013. Multi-agent residential demand response based on load forecasting. In *Proceedings of the 1st IEEE Conference on Technologies for Sustainability (SusTech)*. 90–96.

[3] Thillainathan Logenthiran, W. Li, and W. L. Woo. 2016. Intelligent multi-agent system for smart home energy management. In *Proceedings of the IEEE Innovative Smart Grid Technologies - Asia*. 1–6.

[4] Ryan Lowe, YI WU, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel, and Igor Mordatch. 2017. Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments. In *Proceedings of the 31th Advances in Neural Information Processing Systems*. 6379–6390.

[5] Amir-Hamed Mohsenian-Rad, Vincent W. S. Wong, Juri Jatskevich, Robert Schober, and Alberto Leon-Garcia. 2010. Autonomous Demand-Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid. *IEEE Transactions on Smart Grid* 1, 3 (2010), 320–331.

[6] Daniel O'Neill, Marco Levorato, Andrea Goldsmith, and Urbashi Mitra. 2010. Residential demand response using reinforcement learning. In *Proceedings of the 1st IEEE International Conference on Smart Grid Communications*. 409–414.

[7] Peter Palensky and Dietmar Dietrich. 2011. Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads. *IEEE Transactions on Industrial Informatics* 7, 3 (2011), 381–388.

[8] Valentin Robu, Enrico Gerding, Sebastian Stein, David C. Parkes, Alex Rogers, and Nicholas R. Jennings. 2013. An Online Mechanism for Multi-Unit Demand and its Application to Plug-in Hybrid Electric Vehicle Charging. *Journal of Artificial Intelligence Research* 48 (2013), 175–230.

[9] Goran Strbac. 2008. Demand side management: Benefits and challenges. *Energy Policy* 36, 12 (2008), 4419–4426.

[10] Tanvi Verma, Pradeep Varakantham, and Hoong Chuin Lau. 2018. Entropy Controlled Non-Stationarity for Improving Performance of Independent Learners in Anonymous MARL Settings. *arXiv preprint* abs/1803.09928 (2018).

[11] Di Wu, Rabusseau Guillaume, FranÃğois lavet Vincent, Precup Doina, and Boulet Benoit. 2018. Optimizing Home Energy Management and Electric Vehicle Charging with Reinforcement Learning. In *Proceedings of the 16th Adaptive Learning Agents*.

[12] Yaodong Yang, Jianye Hao, Mingyang Sun, Zan Wang, Changjie Fan, and Goran Strbac. 2018. Recurrent Deep Multiagent Q-Learning for Autonomous Brokers in Smart Grid. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. International Joint Conferences on Artificial Intelligence Organization, 569–575.

[13] Yaodong Yang, Rui Luo, Minne Li, Ming Zhou, Weinan Zhang, and Jun Wang. 2018. Mean Field Multi-Agent Reinforcement Learning. In *Proceedings of the 35th International Conference on Machine Learning*, Vol. 80. PMLR, 5567–5576.

[14] Yuan-Yih Hsu and Chung-Ching Su. 1991. Dispatch of direct load control using dynamic programming. *IEEE Transactions on Power Systems* 6, 3 (1991), 1056–1061.

[15] Baosen Zhang, Ramesh Johari, and Ram Rajagopal. 2015. Competition and Coalition Formation of Renewable Power Producers. *IEEE Transactions on Power Systems* 30, 3 (2015), 1624–1632.