Reinforcement Learning in Energy Management: PV & Battery Storage for Consumption Reduction

Sukrit Jaidee
Energy solution
Electricity Generating Authority of Thailand
Nonthaburi, Thailand
sukrit.jai@egat.co.th

Walanchaporn Boon-nontae
Energy solution
Electricity Generating Authority of Thailand
Nonthaburi, Thailand
Walanchaporn.b@egat.co.th

Weerayut Srithiam
Energy solution
Electricity Generating Authority of Thailand
Nonthaburi, Thailand
weerayut.s@egat.co.th

2023 IEEE Conference on Artificial Intelligence (CAI) | 979-8-3503-3984-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/CAI54212.2023.00028

Abstract*—Thailand's steady increase in electricity costs has led to a rapid growth in solar rooftops and batteries installations for energy storage in households. These trends have presented an opportunity for the development of cost-efficient algorithms that optimize energy management. To accomplish this goal, we employed Reinforcement Learning (RL) to optimize energy management by regulating battery charge and discharge, while simultaneously reducing peak demand to mitigate demand charges associated with electricity consumption. We compared various state-of-the-art RL algorithms, including Advantage Actor-Critic (A2C), Augmented Random Search (ARS), Deep Q-Network (DQN), Proximal Policy Optimization (PPO), Quantile Regression DQN (QRDQN), Recurrent PPO (R-PPO), and Trust Region Policy Optimization (TRPO), to a baseline model referred to as Load FIRST. Load FIRST is a rule-based default algorithm commonly used in various solar inverter brands. Our study revealed the promising potential of RL algorithms to optimize battery power management for energy savings, specifically in the rapidly expanding solar rooftop and battery storage market in Thailand. The ARS model, in particular, yielded the most substantial reduction in electricity costs. The cost of electricity generated by the ARS model was 1,068.73 Baht, representing an 18.47% (217.65 Baht) lower than the baseline cost of 1,286.38 Baht. Our results suggested that employing RL algorithms for battery management optimization could reduce both peak demands and electricity costs.

Keywords*—Reinforcement Learning, Energy Management, Photovoltaic, Battery Energy Storage, Energy Optimization

I. INTRODUCTION

Nowadays, the integration of solar photovoltaic (PV) systems and battery energy storage systems has enabled the optimization of energy consumption patterns, especially when used with Time-of-use metering. Additionally, RL has emerged as a promising approach to provide a decision-making framework in complex environments, making it a suitable solution for home energy management systems (HEMS). As a result, our study contributed to the growing knowledge on reinforcement learning-based energy management and promoted sustainable energy consumption.

Our case study for this research was KARIN Smart Home, Srinakarin Dam, Kanchanaburi. The house was equipped with a 10-kW solar capacity and a 10 kW, 20 kWh battery with a 0.5 C-rate. Its air conditioning system comprised three 18,000 BTU/hour units with an EER of 12.24 BTU/hr./watt and two 24,500 BTU/hour units with a SEER of 22.50 BTU/watt. Additionally, the house had three water heaters, eleven 9-watt LED bulbs, and other electricity-consuming equipment. The data used for our study covered the period from November 12, 2022, to March 16, 2023, with a 5-minute resolution. During this time, it is winter in Thailand. However, even in winter, it's still warm by average standards, though it's cooler compared to other periods of the year. The average temperature during this time ranges from 18°C to 32°C, although temperatures can vary somewhat depending on the region. To achieve our goals, we explored various state-of-the-art RL algorithms, including A2C, ARS, DQN, PPO, QRDQN, R-PPO, and TRPO. Our objective was to determine the optimal RL algorithm for controlling battery charging and discharging actions, considering solar PV output, TOU metering, and fluctuating energy demands, to achieve maximum consumption reduction, minimum electricity bills, and lower on-peak and off-peak time costs, as well as demand charge costs.

II. LITERATURE ON RL ENERGY MANAGEMENT

Lu, et al. [1] proposed a novel approach to HEMS using reward shaping based A2C. The method addressed challenges in optimizing energy consumption and cost, considering complex energy markets and demand patterns. The results showed superior performance compared to other RL algorithms in terms of energy cost savings and efficiency. The proposed approach demonstrated promising solution for practical application. N. S. Raman, at el. [2] focused on enhancing electric supply resiliency during natural disasters using rooftop solar PV panels and batteries. The authors proposed a RL-based controller as an alternative to model predictive control (MPC) for managing PV-battery systems. The RL controller demonstrated comparable resiliency performance to MPC by commanding critical loads and batteries, while significantly reducing computational effort, making it a promising approach for improving resiliency in electric supply systems. H. Li, Z. Wan and H. He [3] proposed a deep RL (DRL)-based HEMS for optimal scheduling of home appliances, considering the randomness in utility pricing and residents' activities. The authors develop a DRL approach using PPO to determine the optimal demand response strategy. Simulation results showed the effectiveness of the proposed method in optimizing electricity usage and reducing costs based on time-varying prices. H. Li, Z. Wan and H. He [4] presented a DRL-based HEMS for optimizing electricity consumption and costs, while A. Mathew, et al. [5] proposed a similar DRL-based intelligent system, focusing on adaptive appliance scheduling and cost reduction. Yu, et al. [6] explored a DRL approach for smart HEMS, aiming to optimize energy consumption and load scheduling amid uncertainties. Lastly, Gao J, et al. [7] proposed a HEMS that leveraged personalized federated DRL to optimize energy consumption and costs while preserving user privacy.

III. METHODOLOGY

Various methods can be employed to implement an EMS, including rule-based systems, machine learning, and reinforcement learning. While the rule-based method is computationally fast, it might not always provide optimized results. On the other hand, the machine learning method requires input data and labels, and its performance may vary

979-8-3503-3984-0/23/$31.00 ©2023 IEEE
DOI 10.1109/CAI54212.2023.00028

across different households. The RL method, however, offers the advantage of building a model with or without historical data, and does not require labels; a well-designed reward system is sufficient. Therefore, we consider the RL method to be the most suitable solution. In this study, we used historical data to train the agent within the environment. However, when some historical data was absent, we appropriately simulated the environment's values or imported data from open-source resources, such as Energy Plus [8].

A. Reinforcement Learning

RL is a type of machine learning approach in which an agent learns to make decisions in an environment by receiving feedback to achieve a goal. The components of RL consist of the agent, environment, state, action, reward, policy, value function, and Q-function. The core principle of RL involves an agent issuing actions to explore an environment. After each action, the agent receives an observation consisting of the next state, reward, and done. Through multiple exploration attempts, the agent learns to distinguish between high reward and low reward states, which are stored in the database. The data stored in the database is then used to improve the algorithm by propagating high and low rewards, encouraging the agent to avoid low rewards and focus on high rewards. Ultimately, the algorithm helps the agent make better decisions, resulting in actions that yield higher rewards, as shown in Figure 1.

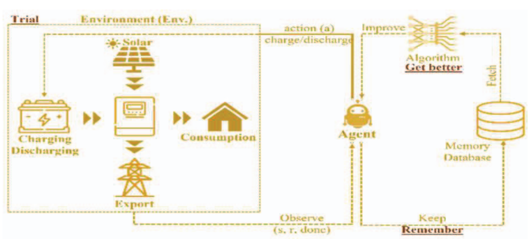

Fig. 1. RL concept

B. Energy Environment

Our states of energy environment consisted of load consumption (kW), solar power (kW), hour (0-24), battery charge (kWh), grid on-peak cost (4.18 THB), grid off-peak cost (2.6 THB), peak demand (kW), net consumption (kWh), energy in the battery charged by the grid (kWh), and self-consumption (kWh). The battery charge represented the amount of energy remaining in the battery, while peak demand denoted the maximum energy demand in each period. Moreover, the two actions involved were charging and discharging the battery. The reward function in this case is a negative electricity cost, which signifies that the agent should strive to take actions that reduce both the number of units of electricity consumed and the peak demand.

IV. RESULTS

We conducted a comparative analysis of several benchmarking algorithms, including Baseline, A2C, ARS, DQN, PPO, QRDQN, R-PPO, and TRPO. Figure 2 presented the electricity cost reduction achieved by each model. The most significant reduction in electricity cost was observed in the ARS model, resulting in a cost of 1,068.73 Baht. In contrast, the baseline model, which employed the LOAD FIRST algorithm (a strategy utilizing solar panel power to meet a building or home's electrical load prior to charging the battery storage), exhibited an electricity cost of 1,286.38 THB. In summary, the electricity cost reduction achieved by the ARS model amounted to 217.65 Baht, corresponding to a decrease of 18.47%.

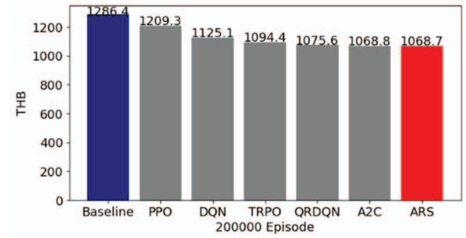

Fig. 2. Electricity cost reduction achievable with each model

V. CONCLUSIONS

The reduction in the electricity bill achieved by the ARS model, which amounted to 18.47%, was within a satisfactory range. However, there is still potential for the model to be more efficient and to further reduce electricity. One possible strategy to achieve this is to employ predictive models, such as LSTM, GRU, or DeepESN [9], to forecast load consumption and solar power generation in advance. This would enable the agent to strategically plan for charging during off-peak hours and storing power in the battery to supply the electrical load during peak times, as well as to reduce peak demand. Other possible approaches include refining the reward function to expedite the model's learning process and fine-tuning the model's hyperparameters. In conclusion, the flexibility and efficiency of RL in energy management are significant advantages. However, it is essential to find a balance between investing costs and the resulting savings.

REFERENCES

[1] R. Lu, Z. Jiang, H. Wu, Y. Ding, D. Wang and H. -T. Zhang, "Reward Shaping-Based Actor–Critic Deep Reinforcement Learning for Residential Energy Management," in IEEE Transactions on Industrial Informatics, vol. 19, no. 3, pp. 2662-2673, March 2023, doi: 10.1109/TII.2022.3183802.

[2] N. S. Raman, N. Gaikwad, P. Barooah and S. P. Meyn, "Reinforcement Learning-Based Home EMS for Resiliency," 2021 American Control Conference (ACC), New Orleans, LA, USA, 2021, pp. 1358-1364, doi: 10.23919/ACC50511.2021.9483162.

[3] H. Li, Z. Wan and H. He, "A Deep Reinforcement Learning Based Approach for Home EMS," 2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 2020, pp. 1-5, doi: 10.1109/ISGT45199.2020.9087647.

[4] Z. Wan, H. Li and H. He, "Residential Energy Management with Deep Reinforcement Learning," 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 2018, pp. 1-7, doi: 10.1109/IJCNN.2018.8489210.

[5] A. Mathew, A. Roy and J. Mathew, "Intelligent Residential EMS Using Deep Reinforcement Learning," in IEEE Systems Journal, vol. 14, no. 4, pp. 5362-5372, Dec. 2020, doi: 10.1109/JSYST.2020.2996547.

[6] Yu et al., "Deep Reinforcement Learning for Smart Home Energy Management," in IEEE Internet of Things Journal, vol. 7, no. 4, pp. 2751-2762, April 2020, doi: 10.1109/JIOT.2019.2957289.

[7] J. Gao, W. Wang and B. Campbell, "Poster Abstract: Residential EMS Using Personalized Federated Deep Reinforcement learning," 2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), Milano, Italy, 2022, pp. 541-542, doi: 10.1109/IPSN54338.2022.00071.

[8] Moriyama, Takao et al. "Reinforcement Learning Testbed for Power-Consumption Optimization." ArXiv abs/1808.10427 (2018): n. pag.

[9] Gallicchio, Claudio and Alessio Micheli. "Deep Echo State Network (DeepESN): A Brief Survey." ArXiv abs/1712.04323 (2017): n. pag.