ISSN 2004-2965 Energy Proceedings, Vol. 19, 2021

International Conference on Applied Energy 2021
Nov. 29 - Dec. 5, 2021, Thailand/Virtual
Paper ID: 161

A Home Energy Management System Optimization Model Based on DNN and RL Adapting to Users' Uncertain Behaviors

Kezheng Ren [1], Jun Liu [1], Zeyang Wu [1*]

1 Shaanxi Key Laboratory of Smart Grid, Xi'an Jiaotong University, Xi'an, China.

ABSTRACT

About 35% of the total energy consumption comes from house appliances. Home energy management system (HEMS) is a household optimal system that can improve the efficiency of electricity consumption, increase the consumption of new energy and reduce carbon emissions. At present, the research of HEMS is simulated under the fixed load setting, and the adaptability of the model to the users' uncertain behaviors is not considered. In this paper, an optimization model of HEMS based on deep neural network (DNN) and reinforcement learning (RL) algorithm is presented. The model aims to minimize the electricity cost and the comfort cost. For the residential case in this paper, the model can reduce 34.2% total cost. The simulation results show that the proposed model can better adapt to the uncertain behaviors of users than the optimization model based on genetic algorithm (GA).

Keywords: home energy management system, deep neural network, reinforcement learning, uncertain behavior

1. INTRODUCTION

Household appliances are the largest source of energy consumption in addition to heating, accounting for 30% of residents' carbon emissions [1]. Building energy conservation is an important content of energy conservation and emission reduction, including heater, refrigeration, air conditioner, lights, etc. And now households use a large number of new interactive devices, such as distributed energy, energy storage, electric vehicles, which will probably bring challenges to the power system due to their stochastic nature [2].

With the gradual increase in the power rating of household appliances, more and more energy is wasted, which results in lower efficiency. At the same time, it brings more challenges to the stability of the power system. So, demand response (DR) becomes an important concept. DR makes full use of the adjustable capacity of the demand-side elastic load to achieve the balance of power supply and demand and saves more money for users [3]. It is shown that 15% energy can be saved with DR [4].

HEMS must meet the power demand of users instead of save electricity bills only. Studies have shown that electricity is a resource whose value is much higher than its price. HEMS should comprehensively consider electricity bills and comfort to obtain satisfactory results for users and encourage users to change their electricity consumption habits. And HEMS should be able to adapt to different home environments and different users' habits [5]. Therefore, it is very important to study whether a HEMS model can adapt to the uncertain behavior of users. Generally, the simulation of HEMS is based on deterministic behaviors. All information of loads is known and users' habits will not change, or the randomness of new energy equipment is added for optimization, but it does not judge whether a model is adaptive to the uncertain behavior of users.

In this study, a data-driven HEMS optimization model based on DNN and RL is proposed to improve household electricity efficiency. And the behavior violation probability (BVP) is defined to measure the adaptability of the model to users' uncertain behavior.

2. THE HEMS MODELING CONSIDERING LOAD DEMAND RESPONSE

As shown in Fig.1, this paper establishes a HEMS optimization solution process that combines DNN and

RL. Electricity price and PV generation are forecasted by prediction model based on DNN, and then the trends and load data are sent to the decision-making model base on multi-agent Q learning to make optimal arrangement for different types of loads in the next hour.

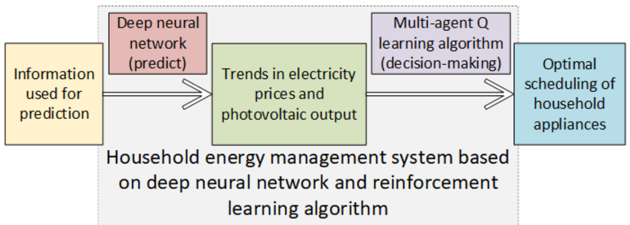


Fig. 1. Home energy management system optimization model process

2.1 *Load demand response model*

Different household appliances have different operating characteristics, such as adjustable operation time, adjustable operation power and other physical quantities like temperature and charge. As shown in the Fig.2, the household appliances can be divided into the following categories [6].

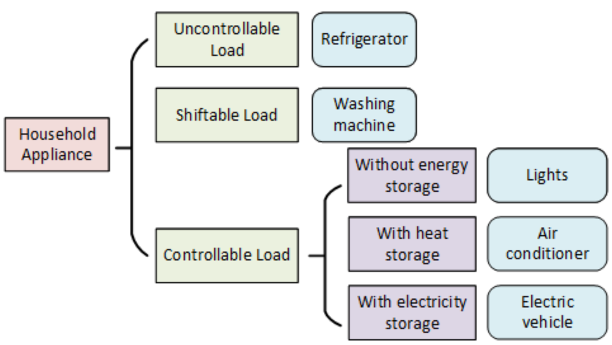


Fig. 2. Household appliances classification

Uncontrollable load runs at rated power and cannot be adjusted. The shiftable load can be started at the time when the electricity price is low within the operation time. Its power is uncontrollable and its comfort is determined by its start-up time. The power of controllable load can be adjusted during the running time. The comfort of controllable load without energy storage is related to its power, the comfort of controllable load with heat storage is related to the indoor temperature and the comfort of controllable load with electricity storage is related to the state of charge (SOC).

2.2 *Deep neural network prediction model*

Electricity price and PV generation are both uncertain, which bring difficulty to schedule. Different kinds of prediction algorithms have been proposed [7]. This article uses the forecasting method based on DNN to reduce the impact of uncertainty to load control. The prediction models for electricity price and PV generation include one input layer, three hidden layers and one output layer. The accuracy of the prediction model is closely related to the selection of features [8]. 31 parameters such as year, month, day, time period, week, whether it is weekend, whether it is a holiday, and the electricity price in the past 24 hours, are used to predict the electricity price in the current time period. 28 parameters including year, month, day, time period and PV output in the past 24 hours are used to predict the PV generation trend.

2.3 *HEMS optimal scheduling model based on multi-agent Q learning algorithm*

In QL, the agent observes the state and selects an action every hour, and then updates the Q value with the obtained reward. The Q-value update equation is as follows [6]:

$$\begin{aligned} &Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \\ &\theta\left[r(s_t,a_t) + \Upsilon * \max(s_{t+1},a_{t+1}) - Q(s_t,a_t)\right] \end{aligned} \tag{1}$$

Where $r(s_t,a_t)$ is the reward; γ is the discount factor indicating the relationship between the importance of future awards and current awards; θ is the learning rate controlling the degree to which the new Q value updates the old Q value.

This paper adopts the multi-agents Q learning to avoid dimensional explosion [9]. Each load has its own agent, and the optimal state is selected independently among multiple agents. At the same time, the energy from PV and EV is dispatched to other loads. To avoid the complexity of the distribution process reducing the practicality of HEMS, energy from PV and EV first supplies the uncontrollable load, and then supplies the controllable load without energy storage and load with heat storage. The shiftable load does not receive this part of energy due to the uncertain operating time.

The Markov elements for different types of loads, including state, action and reward are set as follows:

$$\begin{aligned} &s_t^{\mathrm{un}} = \{t\} \ a_t^{\mathrm{un}} = \{1\} \ P_N^{\mathrm{un}} \\ &r_t^{\mathrm{un}} = -\sigma_t^{\mathrm{un}} p_t \left(P_t^{\mathrm{un}} - E_t^{\mathrm{d,un}}\right) \end{aligned} \tag{2}$$

$$\begin{aligned} &s_t^{\mathrm{s}} = \{t,c\} \ a_t^{\mathrm{s}} = \{0,1\} \ P_N^{\mathrm{s}} \\ &r_t^{\mathrm{s}} = -\sigma_t^{\mathrm{s}} p_t P_t^{\mathrm{s}} - k_{\mathrm{s}} (t - t_{start})^2 \end{aligned} \tag{3}$$

$$\begin{aligned} &s_t^{\mathrm{c}} = \{t\} \ a_t^{\mathrm{c}} = \{0.4,0.5,0.6,0.7,0.8,0.9,1\} \ P_N^{\mathrm{c}} \\ &r_t^{\mathrm{c}} = -\sigma_t^{\mathrm{c}} \left[p_t \left(P_t^{\mathrm{c}} - E_t^{\mathrm{d,c}}\right) + k_c \left(1 - P_t^{\mathrm{c}} / P_{\max}^{\mathrm{c}}\right)^2 \right] \end{aligned} \tag{4}$$

$$s_t^{\mathrm{ac}}=\{t, temp\}\ a_t^{\mathrm{ac}}=\left\{\frac{m}{15}\right\}P_N^{\mathrm{ac}}, m=-15,-13,\cdots,13,15$$
$$r_t^{\mathrm{ac}}=-\sigma_t^{\mathrm{ac}}\left[p_t\left(\left|P_t^{\mathrm{ac}}\right|-E_t^{\mathrm{d,ac}}\right)+k_{\mathrm{ac}}\left(T_t^{\mathrm{in}}-T_{\mathrm{set}}^{\mathrm{in}}\right)^2\right] \tag{5}$$

$$s_t^{\mathrm{ev}}=\{t, soe\}\ a_t^{\mathrm{ev}}=\left\{\frac{m}{15}\right\}P_N^{\mathrm{ev}}, m=-15,-13,\cdots,13,15$$
$$r_t^{\mathrm{ev}}=-\sigma_t^{\mathrm{ev}}\begin{bmatrix}p_tP_t^{\mathrm{ev}}+\kappa_{\mathrm{ev}}k_{\mathrm{ev}}\left(SOE_t^{\mathrm{ev}}-SOE_{\exp}^{\mathrm{ev}}\right)^2\\+\beta\left|P_t^{\mathrm{ev}}\right|\end{bmatrix} \tag{6}$$

Where s_t represents state at time t; a_t represents action at time t; r_t represents reward at time t; σ_t represents use state at time t, which is 0 or 1; p_t represents electricity price at time t; P_t represents power of uncontrollable load at time t; E_t^{d} represents dispatch power from new energy equipment at time t; k is the comfort factor; t_{start} is the time that shiftable load can be operated; κ_{ev} represents the charging anxiety coefficient; β represents battery degradation cost coefficient.

The transition probability of QL is the ε-greedy strategy, which weighs the exploitation and exploration.

The objective function of this paper is the opposite of the sum of rewards for each load. The objective is to minimize the sum of electricity cost and comfort cost:

$$F=\min(-(r_t^{\mathrm{un}}+r_t^{\mathrm{s}}+r_t^{\mathrm{c}}+r_t^{\mathrm{ac}}+r_t^{\mathrm{ev}})) \tag{7}$$

2.4 Adaptability to uncertain behaviors of EV

"Behavior violation parameter" is defined to measure the adaptability of the model to user behavior. The smaller the BVP, the more the model adapts to uncertain behaviors. EV's uncertain behaviors are defined as follows: from 19:00 to 23:00, EV may choose to stay at home, go to mall or travel every hour. Their probabilities are 0.7, 0.2 and 0.1 respectively, and only one behavior is considered one night. At 6:00, users will choose to stay at home or go to work earlier, with probabilities of 0.8 and 0.2. The three behaviors of go to mall, travel and go to work earlier need to satisfy the energy storage conditions, which are 55%, 60% and 80% respectively. If the energy storage is insufficient, it is considered that the scheduling result violates the user's behavior.

$$bvp=\left(\sum_{i=1}^{k}\left(\sum_{j=1}^{n_i}\Delta soc\%\right)/N_i\right)\times 100\% \tag{8}$$

Where i represents the number of the uncertain behavior; j represents the times of users' behavior being violated; $\Delta soc\%$ represents the degree of violation of users' behavior, it can be different according to the characteristics of loads and for EV is the difference between the actual soc and the expected soc; N_i represents the total times of the uncertain behavior occurs.

3. RESULTS AND DISCUSSION

3.1 Study Case

The simulation case of this paper is a detached house. The household appliances include a solar panel, an uncontrollable load, refrigerator (RFEG), a shftable load, washing machine (WM), four controllable loads, two controllable loads without energy storage, lights (L1, L2), one load with heat storage, air conditioning (AC), and one with electricity storage, electric vehicle (EV). The parameters of the loads are illustrated in Table.1. For the EV, the capacity is 8 kWh, the minimum and maximum of energy storage percentage is 30% and 90%, the expected energy storage percentage is 80% and the battery degradation cost coefficient is $1.25\times 10^{-4}\$\cdot\mathrm{kWh}^{-2}$.

Table. 1. Load parameter

Loads	Power rating /kW	Start time /h	End time /h	operation hours /h	Comfort factor
REFG	0.5	0	23	24	/
WM	0.7	19	23	2	0.0005
L1	1.2	6	23	18	0.02
L1	1.2	6	23	18	0.03
AC	1.2	0	233	24	0.01
EV	0.8	18	7(+1D)	13	0.005

For QL, the discount factor γ is set to 0.9, and the learning rate θ is set to 0.1. In order to speed up the convergence speed, ε is set to 0.7.

3.2 Results

3.2.1 DNN prediction results

Fig.3 shows the prediction results for electricity price and PV generation in 72 hours. Although there are certain errors, the trend of predicted data and real data is consistent. In order to show that the prediction model of DNN is conducive to the further decision-making of HEMS, this paper compares the prediction results of the autoregressive moving average model (ARMA) and DNN, and bring the two forecasting results into later decision-making progress to compare their impact on HEMS. As shown in Table.2, DNN has higher prediction accuracy and performs better in latter optimization progress, reducing more cost.

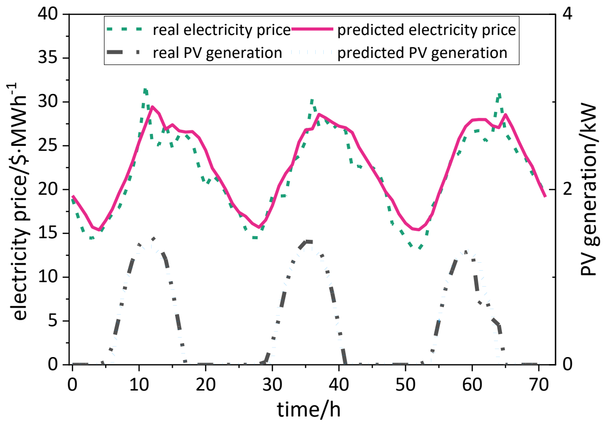


Fig. 3. Prediction model results

Table. 2. Comparison of HEMS scheduling results based on DNN and ARMA prediction data

	DNN	ARMA
Electricity cost/$	0.2302	0.2506
Comfort cost/$	0.2284	0.2138
Total cost/$	0.4586	0.4644

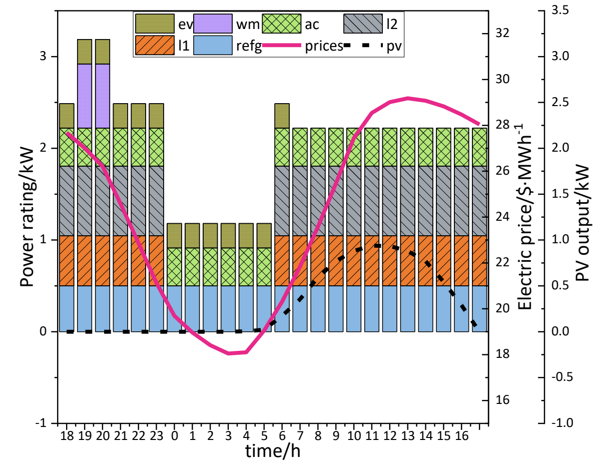
3.2.2 Optimal schedule results of multi-agents QL

Fig.4 is the schedule result of HEMS without DR (a) and with DR(b). In both cases, the total power consumed by each load is the same. The purple curve is the electricity price. It can be seen that the peak power price appears in 10:00-20:00. In Fig.4(a), it can be seen that the peak power consumption appears in the peak period of electricity price in 19:00-20:00, which results in high electricity bills. In Fig.4(b), the peak load of the scheme considering DR is about 6% lower than that without DR. And the power peak and price peak are staggered to achieve the purpose of shifting the load and saving electricity bill.

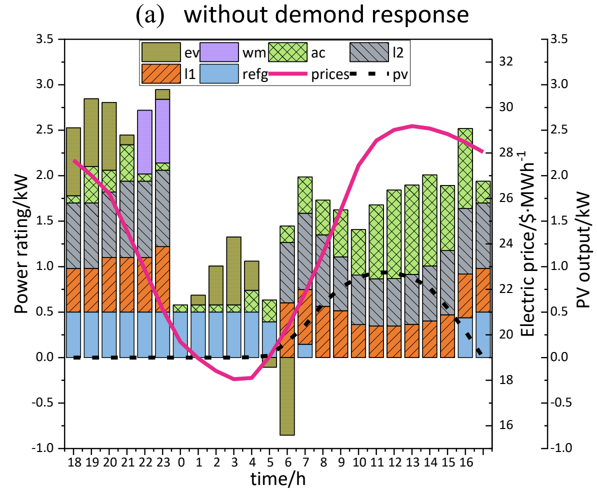
For EV, it only charges without DR. But with DR, it charges a lot at the lowest point of electricity price in 2:00-4:00, and then the energy is released to other loads in 5:00-6:00, thus approximately achieving the load transfer.

The black dotted line is the PV generation and it shows that solar power is mainly concentrated between 8:00 and 15:00. During that time, uncontrollable load does not need energy from the power grid and the total household power consumption reduces a lot.

The electricity cost without and with DR is 1.26$ and



(a) without demond response



(b) with demond response

Fig. 4. Load arrangement

1.09$, and the comfort cost is 1.20$ and 0.53$. We can see that compared with operating at average power, the optimization solution process proposed in this paper can both reduce the household electricity cost and comfort cost.

As shown in Fig.5, the total cost for four schemes is summarized. The total cost of household appliances without DR is 2.46$. Considering DR, the cost fell to 1.83$, when the EV only charges. If we use EV as a storage battery which can charge and discharge, the total cost can be reduced to 1.80$. Then the PV panel is considered to the simulation case, and the total cost is 1.62$. Therefore, considering DR, household with PV and EV can save 34% total cost.

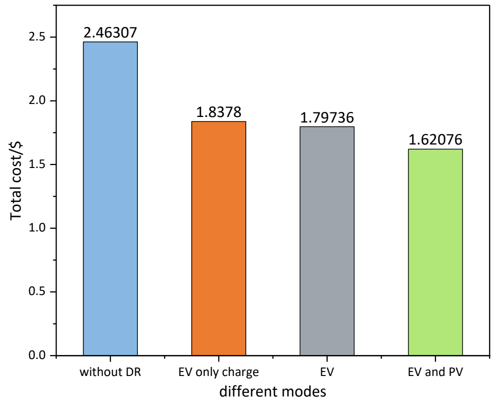


Fig. 5. Results of the total cost of different schemes

3.2.3 Uncertain behavior simulation results

In this paper, 1000 uncertain behavior simulations are carried out and compared with HEMS optimization method based on GA. As shown in Table.3, the total cost of the EV based on RL is lower. And the bvp of optimization method based on GA is more than twice that of the method proposed in this paper. Fig.6 shows the soc curves of two optimization methods. The mutation process of GA is uncertain, while RL algorithm considers the impact of actions on the present and the future. So, the soc curve of RL is more stable and the method based on RL is more adaptable for uncertain behaviors.

Table. 3. Adaptability of hems based on QL and GA to users' uncertain behavior

method	Average total cost/$	bvp
RL	0.120734	2.10%
GA	0.127184	4.62%

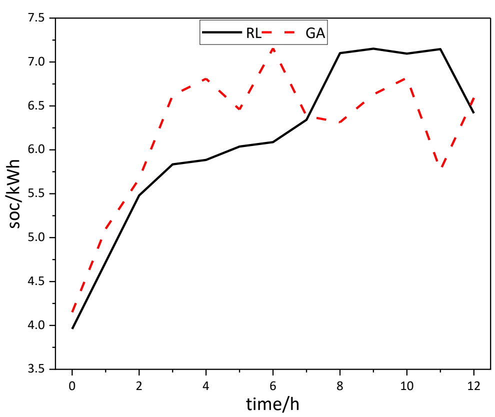


Fig. 6. State of charge curves of two optimization methods

4. CONCLUSIONS

This paper establishes a data-driven optimization model for HEMS system. The model uses DNN to predict uncertain electricity price and PV generation, then applies the multi-agents QL algorithm to schedule each household appliances independently. This model makes the full use of new energy equipment, reduces electricity costs, and improve users' comfort. The results show that HEMS can save 34.2% total cost for the household. And the *bvp* shows that the model proposed in this paper is more suitable for the uncertain behaviors of users than optimization model based on GA.

ACKNOWLEDGEMENT

This work was supported by the National Key Research and Development Program of China under Grant 2018YFB0905000.

REFERENCE

[1] Limin J, Hobngwei L, Zongfeng Q. Research and Practice on the Carbon Neutrality Standards and Certification of Smart Home Appliances. J China Standardization, 2021(10):44-50+54.

[2] Jun L, Yefu C, Chao D. "Distributionally Robust Chance-Constraint Optimal Power Flow with Wasserstein-Moment Metric in Active Distribution Networks," ICAE2018 , 2018: 1-5.

[3] Iqbal S. A Comprehensive Review on Residential Demand Side Management Strategies in Smart Grid Environment. J Sustainability, 2021;13(13) : 7170-7170.

[4] Sharda S, Sharma K, Singh M. A real-time automated scheduling algorithm with PV integration for smart home prosumers. J Building Engineering 2021;44

[5] José R. V and Zoltán N. Reinforcement learning for demand response: A review of algorithms and modeling techniques. J Applied Energy, 2019;235 : 1072-1089.

[6] Lu R,Hong S H, Yu M. Demand Response for Home Energy Management Using Reinforcement Learning and Artificial Neural Network.J IEEE Transactions on Smart Grid. 2019;1-1.

[7] Jun L, Wanliang F, Xudong Z, Chunxiang Y. "An Improved Photovoltaic Power Forecasting Model With the Assistance of Aerosol Index Data," IEEE Transactions on Sustainable Energy. 2015, 6(2): 434-442.

[8] Jiacheng L, Jun L, Hongyan Z, Zeyang W, Xiaoming L, Liu W. "Short-Term Photovoltaic Output Prediction Based on Diffusion Kernel Density Estimation and Improved mRMR Feature Selection," Automation of Electric Power Systems. 2021.

[9] Xu X, Jia Y, Xu Y et al. A Multi-agent Reinforcement Learning based Data-driven Method for Home Energy Management.J IEEE Transactions on Smart Grid. 2020;PP(99):1-1.

