International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: 0138

COORDINATED SCHEDULING OF MULTI-ENERGY MICROGRID BASED ON MULTI-AGENT GAME THEORY AND REINFORCEMENT LEARNING

Ge Shaoyun, Li Jifeng*, Liu Hong, Lu Zhiying, Yang Zan, Yan Jun

Key Laboratory of Smart Grid of Ministry of Education, Tianjin University

ABSTRACT

Since the deep integration and close interaction of multiple energy networks have increased the complexity of optimized scheduling of multi-energy microgrids, the present manuscript proposed a coordinated scheduling model in accordance with multi-agent game theory and reinforcement learning. First, a multi-agent division was conducted. Second, a multi-agent economic decision-making model and a game decision model were constructed. Thus, a coordination scheduling method was proposed in accordance with Nash game theory and the Q-learning algorithm. Finally, the efficiency and effectiveness of the proposed method were validated, which were verified by real-world case studies.

Keywords: Multi-energy microgrid, coordinated scheduling, multi-agent, game theory, Nash Q learning.

NONMENCLATURE

Symbols	
$C_{[\cdot]}(t)$	fuel or loss cost of related equipment.
$C_{\mathrm{MEP}}^{\mathrm{OM}}(t)$	operating and maintenance cost of the MEP agent.
$C_{\mathrm{MEP}}^{\mathrm{ST}}(t)$	starting and stopping cost of controllable devices.
N_g	number of agents participating in the game.
$p_B(t)$	electric purchase price from the utility grid.
$p_S(t)$	sale electric price to the utility grid.
$P_{EVd}(t)$	discharge power of the EVO agent.
$P_{PG}(t)$	purchase of electric power from the utility grid.
$p(s'\|s)$	the state transition probability.
$R(s, s', a)$	the reward function value.
s/ s'	current state/ the state of the next period.
α	the learning factor.
β	probability of the EV owner's response scheduling.
γ	the discount factor ($0<\gamma<1$).

1. INTRODUCTION

The construction of the energy Internet in recent years has changed the existing mode of energy supply systems[1]. Centralized control is commonly used to conduct energy management. However, multi-energy systems usually involve multiple energy providers, each with their own goals. Given this, some researchers have already introduced multi-agent and game theories into energy management technologies. Also, several control variables should be considered during the multi-energy regulation management. Accordingly, a more intelligent scheduling method, that uses artificial intelligence (AI) technology, is needed. Currently, the application of AI technology, especially learning algorithms, to power system management includes fault monitoring, fault diagnosis, and load/power prediction[2-3]. The application of AI technology to deeper processes, requires further exploration. Li et al.[4], introduced the hidden Markov decision process and effectively combined the demand response strategy based on the previous predictions. Kofinas et al.[5] focused on multi-agent microgrid systems and effectively regulated the system's continual state space using deep learning methods to improve the efficiency.

The outcome of recent research has laid the foundation for this study, whereas several open problems remain. First, a more effective division of intelligent agents and an interest in game-based relationships requires additional studies. Second, AI technology-based methods that can be applied for planning and operating processes also require further exploration. Thus, this study proposes a coordinated scheduling method of grid-connected multi-energy microgrids in accordance with multi-agent game theory and reinforcement learning that successfully applies AI technology to integrated energy scheduling.

2. STRUCTURE OF MULTI-ENERGY MICROGRID AND DIVISION OF MULTI-AGENT SYSTEMS

A multi-energy microgrid model was established here based on the energy hub mode, consisting of a combined cold heat and power system (CCHP), gas heat pump (GHP), distributed photovoltaics (PV), central air-conditioning (CAC), electricity storage (ES), and heat storage (HS) components providing power, cold, and heat based energies. The composite structure is illustrated in Fig. 1.

Based on different interest pursuits and the actual situations, this study divided the multi-energy microgrid into the following agents: Renewable energy provider (REP), Microgrid energy providers (MEP), Electric vehicle owner (EVO).

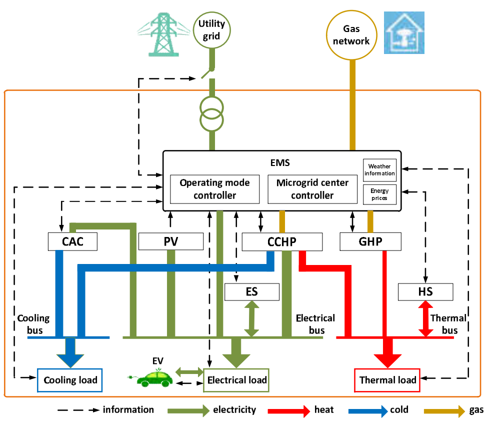

Fig 1 Structure of multi-energy microgrid

3. MULTI-AGENT GAME COORDINATED SCHEDULING MODEL

3.1 Multi-agent Economic Optimized Decision Model

(1) MEP decision model

The MEP agent is capable of optimizing the output of controllable equipment to reduce the operating cost and the cost of purchasing electricity from the utility grid. The specific objective equation is written as:

$$\begin{aligned} I_{\mathrm{MEP}} = \min \sum_{t=1}^{N_T} [& p_{\mathrm{B}}(t) P_{\mathrm{PG}}(t) + C_{\mathrm{CCHP}}(t) + C_{\mathrm{GHP}}(t) \\ & + C_{\mathrm{HS}}(t) + C_{\mathrm{MEP}}^{\mathrm{OM}}(t) + C_{\mathrm{MEP}}^{\mathrm{ST}}(t)] \end{aligned} \quad (1)$$

(2) REP decision model

Because of the different energy prices at different times, the REP agent is capable of arranging the pool energy purchase by adjusting the charge/discharge state of ES devices. The specific objective equation is written as:

$$I_{\mathrm{REO}} = \max \sum_{t=1}^{N_T} [p_{\mathrm{S}}(t) P_{\mathrm{REP}}^{\mathrm{S}}(t) - C_{\mathrm{REP}}^{\mathrm{MO}}(t) - C_{\mathrm{ES}}(t)] \quad (2)$$

(3) EVO decision model

The EVO agent employs an aggregator to control the EVs. EVOs are operated based on profit. Accordingly, the EVO decides whether to discharge to the utility grid according to the electricity price. The objective equation of the EVO is written as:

$$I_{\mathrm{EVO}} = \max \sum_{t=1}^{N_T} \left[p_{\mathrm{S}}(t) P_{\mathrm{EVd}}(t) (\lceil r - \beta(\Delta p(t)) \rceil) \right] \quad (3)$$

3.2 Multi-agent Game Equilibrium Decision Model

A static non-cooperative game problem is raised when multi-agent systems are seeking a maximum benefit. Accordingly, the scheduling plan of the microgrid system is to solve the Nash equilibrium of the game between the multi-agent entities. The objective function is expressed as:

$$G = < N_{\mathrm{g}}, A_{\mathrm{g}}, u_{\mathrm{g}} > \quad (4)$$

The following constraints should be considered when analyzing multi-agent games and formulating coordinated scheduling strategies: 1) Power balance constraints; 2) Energy supply equipment operating constraints; 3) Energy storage equipment operating constraints; 4) Grid connection line capacity restraints.

4. COORDINATED SCHEDULING BASED ON MULTI-AGENT GAME AND NASH Q-LEARNING

4.1 Basic Principle of Nash Q-learning Algorithm

This section combines evolutionary game theory and Q-learning algorithms[6] to formulate multi-agent operating strategies. The game environment for multi-agent systems can be established based on the objective equation (4): the participants are 3 agents; the strategy set is the operation strategy of each agent adopted for the energy supply or storage equipment. The utility formulas respectively correspond to equations (1), (2) and (3). For the Nash Q-learning method, if all the $a_{gi} \in A_{gi}$ can achieve an optimal operating cost/benefit efficiency for each agent, then Q-learning can achieve Nash equilibrium.

Under a game theory environment, the expected total reward and updated Q values of the game agent $N_{g,i}$ can be characterized by the following equation:

$$\begin{aligned} Q^{N_{g,i}} &= (s, a^1, ..., a^{N_g}) = R^{N_{g,i}} \\ &+ \gamma \sum_{s'} p(s' | s, a^1, ..., a^{N_g}) Q^{N_{g,i}}(s', a^1, ..., a^{N_g}) \end{aligned} \quad (5)$$

$$Q_{k+1}^{N_{g,i}}=(s,a^{1},...,a^{N_g})=(1-\alpha_k)Q_k^{N_{g,i}}(s',a^{1},...,a^{N_g}) +\alpha_k\left[R_k^{N_{g,i}}+\gamma\sigma^{1}(s')...\sigma^{N_g}(s')Q_k^{N_{g,i}}(s')\right] \tag{6}$$

where $[\sigma^1(s'),...,\sigma^{Ng}(s')]$ is a mixed-strategy Nash equilibrium solution.

4.2 Coordinated scheduling process based on multi-agent game and Nash Q-learning

Step 1): Initialization of Q values. The initial value of each element (s, a) in the Q value table is set to 0 at the offline pre-learning stage; at the online learning stage, the initial value is set as the Q value table preserved during the pre-learning.

Step 2): Discretize the continual state and action variables to form a <state, action> function. To select state space, the load demand, the output of PV/CCHP/GHP/CAC, the charge/discharge capacity of ES/HS, and the charge/discharge capacity of EVs were selected as the state's input. To facilitate the Q-learning method, the variables are discretized into an interval form. Then, the only state can be determined as: S_k=<k, S_{PV}, S_{Load}, S_{CCHP}, S_{GHP}, S_{CAC}, S_{ES}, S_{HS}, S_{EV}>. According to the equipment output at the specified period and the operation of energy storage devices, the action strategy can be verified as: a_k=<k, a_{CCHP}, a_{GHP}, a_{CAC}, a_{ES}, a_{HS}, a_{EV}>.

Step 3): Calculate the immediate reward of each agent; in the meantime, the future state s' is predicted.

Step 4): Update the Q values and set s←s'.

Step 5): Check whether the learning process converges; if not, set k=k+1 and return to **Step 2)**.

The flow chart of the algorithm is shown in Fig. 2.

5. CASE ANALYSIS AND COMPARISON

5.1 Case Overview

This study took typical industrial parks in China as an example case. The characteristic curve of power/heat/cold demand in the whole year is plotted in Fig. 3. The operating parameters of the equipment are listed in Table 1. The start-stop cost of the controllable unit is 1.94 CNY. There exist 100 EVs in the microgrid system. The EV model is the Nissan Leaf. The electricity prices are differentiated by the following stages: 10:00-15:00 and 18:00-21:00, 0.83 CNY/kW·h; 07:00-10:00, 15:00-18:00 and 21:00-23:00, 0.49 CNY/kW·h; 00:00-07:00 and 23:00-24:00, 0.17 CNY/kW·h; the gas price is 2.28 CNY/m^3. For the Q-learning algorithm: the learning factor α is set to 0.01, the discount factor γ is set to 0.8, and the unit schedule time is Δt=15min. For the state space division: S_{total}=<96, 6, 60, 5, 5, 8, 6, 4, 1>. For the action space division: the PV output contains only 1 state since the output is controlled by the external environment; the controllable unit contains 2 states (operate or stop); the energy storage devices and EVs contains 3 states (charge/leisure/ discharge).

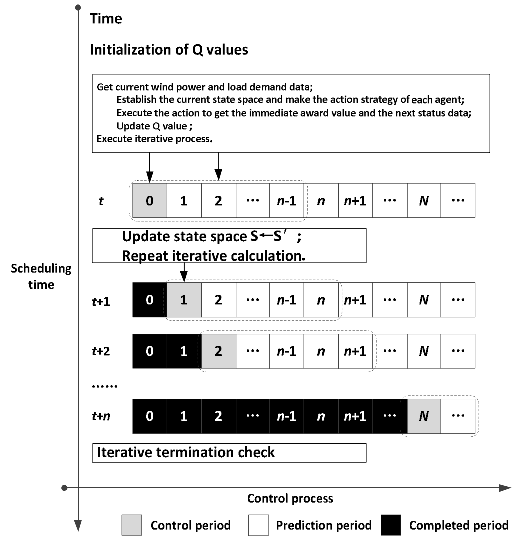

Fig 2 Scheduling method flow chart based on Nash Q learning

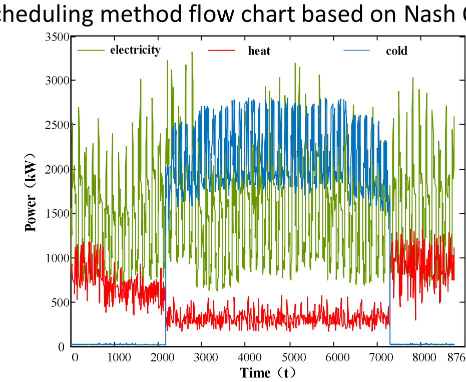

Fig 3 Load demand curve

Table 1 Operation parameters for each type of unit

Devices Type	Rated Capacity /Power	P_{min} (kW)	P_{max} (kW)	Conversion Efficiency	K_m (CNY/kW)
CCHP	2000kW	20	2000	1	1.547
PV	2600kW	0	2600	-	0.02
GHP	1000kW	10	1000	0.8	0.40
CAC	2000kW	20	2000	0.88	0.30
ES	3000kW·h	-	2400	0.9	0.03
HS	1000kW·h	-	800	0.9	0.03
EV	30kW·h	-	6.6	0.9	-

5.2 Existence Proof of Nash Equilibrium and Analysis of Pre-learning

Ref. [6] showed the lemma and proof of Nash equilibrium convergence in the Nash Q-learning algorithm. The offline learning and simulations were performed based on historical data more than 4000 times. The total cost/benefit changes of different multi-agents are shown in Fig. 4. The cost/benefit of each agent tends to be stable after the 2000-th iteration and multi-agent economic game analysis. This demonstrates that the Q-learning algorithm has achieved the trial and gained sufficient experience, making it capable of providing reasonable energy management decisions.

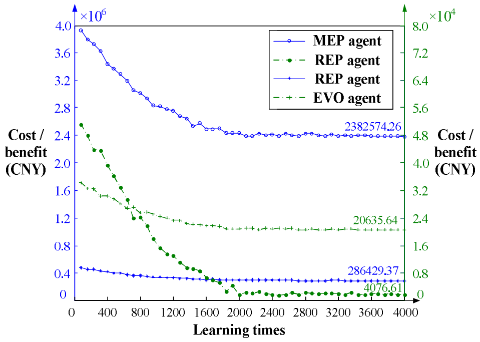

Fig 4 Annual economic change curve of pre-learning process

5.3 Simulation Result Analysis of Online Energy Management

Based on the pre-learning process, this study further analyzed the day-ahead schedule management capability of the method. Taking power energy as an example, the energy management balance and dynamic output changes of the equipment are shown in Fig. 5.

The result shows a vigorous PV output at the period of 11:00-13:00, which is also the peak period of the electricity price. Therefore, the REP agent would deliver the unused PV output in the load supply to the utility grid to gain a benefit. The electricity price valley happens at 0:00-7:00 and 23:00-24:00. Thus, the REP agent would deliver a signal for the ES charging and the purchase electricity from the utility grid. The power load is high and the heat load is low during 18:00-21:00. Accordingly, the CCHP system would deliver an extra heat for HS after satisfying the heat load demand. This can relieve the power supply pressure caused by the insufficient PV output at night.

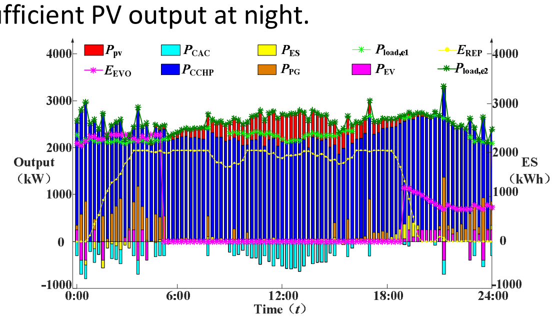

Fig 5 Electric load balance

6. CONCLUSIONS

The major contributions of this study include: 1) The multi-agent theory was introduced, the multi-agents were divided according to their interest pursuits and the game relationship between them were analyzed. 2) The multi-agent decision model was constructed based on different interest pursuits as well as the game decision model considering the interest equilibrium of the multi-agents, and the action strategies of different agents were analyzed via an actual day-ahead scheduling case. 3) An AI method was adopted based on Nash Q-learning to achieve the coordinated scheduling of the multi-energy microgrid. 4) The rationality and feasibility of using AI technology to make multi-energy management decisions were analyzed using a real-world case.

The economic benefits of the multi-agent game case will be further considered, and these benefits will be compared to the traditional centralized whole social economic benefit.

REFERENCE

[1] Lin W, Jin X, Mu Y, et al. A two-stage multi-objective scheduling method for integrated community energy system. Applied Energy 2010; 216:428-441.

[2] Liu R, Yang B, Zio E, et al. Artificial intelligence for fault diagnosis of rotating machinery: A review. Mechanical Systems & Signal Processing 2018; 108:33-47.

[3] Williams K T, Gomez J D. Predicting future monthly residential energy consumption using building characteristics and climate data: A statistical learning approach. Energy & Buildings 2016; 128:1-11.

[4] Li D, Jayaweera S K. Machine-Learning Aided Optimal Customer Decisions for an Interactive Smart Grid. IEEE Systems Journal 2015; 9(4): 1529-1540.

[5] Kofinas P, Dounis A I, Vouros G A. Fuzzy Q-Learning for multi-agent decentralized energy management in microgrids. Applied Energy 2018; 219: 53-67.

[6] Hu, Junling, Wellman, et al. Nash q-learning for general-sum stochastic games. Journal of Machine Learning Research, 2003; 4(4):1039--1069.