The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Multi-carrier energy hub management through deep deterministic policy gradient over continuous action space

Tomah Sogabe[*1,2,3] Dinesh Bahadur Malla[*1,3] Tomoyuki Hioki[*2], Kei Takahashi[*2], Masaru Sogabe[*3], Katsuyoshi Sakamoto[*1,2], Koichi Yamaguchi[*1,2]

[*1] Info-Powered Energy System Research Center,
[*2]Department of Engineering Science
The University of Electro-Communications, Chofu, Tokyo, 182-8585, Japan
[*3] Technology Solution Group, Grid Inc., Kita Aoyama, Minato-ku, Tokyo, 107-0061, Japan

Abstract: Multi-carrier energy hub has provided more flexibility for energy management systems. On the other hand, due to the mutual impact of different energy carriers in an energy hub's energy management becomes more challengeable. For energy management purpose Mathematic optimization tools are used, but real-time optimization challenges the optimal management. On the other hand, energy demand and supply are very changeable so optimization objectives may vary or more than one. For real-time management, changing environment and multi-objective options AI is purposed. In this work operation of multi-carrier energy hub optimization has been solved by executing a multiagent AI algorithm, which contain deep deterministic policy gradient(DDPG) algorithm. Research multi-agent simulation results show that AI agent can manage a balance between demand and supply, proper charging and discharging of storage agent to optimize energy hub cost. It also describes the price determination method by using AI, which is good for demand and supply management purpose for a market.

1. Introduction

An energy hub is a conceptual model of multi-carrier energy systems used to represent the interactions of multiple energy conversion and storage technologies [1]. The energy hub concept was emerged because or world's energy crisis and its related problems caused a considerable movement into efficient utilization of energy systems. This comprehensive attitude to energy, which presented in the energy hub idea, persuaded the researchers to design future energy systems based on this idea [2]. In the energy hub concept, the whole of energy systems is investigated instead of individual management of energy carriers such as electricity, natural gas and so on. It also combined renewable energy source with a nonrenewable source to minimize carbon emissions for an environmental objective. The main question in the optimization of energy hub operation problem is that what is an optimum arrangement of energy components in each time for providing demands with minimum cost [1]. The energy system is a very changing environment to manage this feature the energy hub contain energy utility, which is capable of energy conversion, energy storage and direct connection of multi-energy carriers [3]. The objective of energy hub optimization is normally cost minimization, which generally includes both technology investment and operation costs [4].

Multi-carrier energy hub management is a process of energy flow optimization in a hub system, which investigates the flow of energy carriers in an integrated system by considering technical constraints of each system [5]. The mathematic optimization is a popular method in system optimization, for example, MILP, game optimization software, genetic algorithm, dynamic relaxation etc. The mathematics optimization tools are expensive and more difficult with the increase in parameter and objectives. So, we purposed for AI which refers to a Reinforcement learning process where agents to learn optimal behavior under different conditions. Key concepts in reinforcement learning are state, action, reward, and policy. The state refers to the state of the environment calculation at a given time. The action refers to the specific action taken by an agent, e.g. the direction and distance of an agent's movement within a given interval of time. The reward refers to the feedback signal (often a simple scalar value) given to an agent as a result of a specific action taken within a specific state. The policy links the states and actions of an agent and refers to the action(s) with the estimated highest reward value in any given state. The aim of reinforcement learning is normally to facilitate agents in identifying an optimal policy. For the problem-solving purpose, we can use two kinds of actions that can be continuous or discrete. Nature of Environment and problem optimization it's taking action will have determined. In this work, we used continuous actions algorithm DDPG [6]. Because we want to learn how much quantity purchase or sell and what will be the price? So according to the nature of optimization continuous action are preferred. If we use discrete nature actions, it will suggest doing buy or sell or pre-defined quantity of purchase and sell.

Contact: Tomah Sogabe, i-PERC, The University of Electro-communications, sogabe@uec.ac.jp

2. Model and learning algorithm

2.1 Model

Energy hub concept is very board concept, it contains a lot of devices and technologies. Our assumed energy hub input energy carriers are electricity, natural gas and the output side consist of electrical demand, heat demand [7]. The internal devices are the electrical transformer, CHP, boiler, electrical and thermal storage systems, PV and wind form. For hub energy supply and demand management, there are markets, hub's input and output situations are responsible for determining optimal operation based on received database on an agent action. The Energy hub model is figure in Fig.1. For selling and buying price determination purpose we divide the model learning in to two case which is described in section 3.

2.2 Learning algorithm

We used DDPG [6] as our learning algorithm, which is a limiting case of stochastic policy gradient in the actor-critic approach used for solving continuous tasks. To solve complex continuous action tasks, it requires a policy with stable learning and faster convergence, since policy may converge to sub-optimal solutions at the early stage of policy learning. There are mainly two networks in the DDPG they are actor and critic network which are in Fig.1. In the learning process critic network is updated with TD (temporal difference) of state and next state. The actor parameters are updated in the gradient of critic output with respect to policy parameters.

3. Simulation information

For demonstrating how energy demand market and its price affect the hub's operation, the simulation results are dividing into two case studies. These case studies are as follows:

Case1: With constant agent's buy and sell price and variable demand and supply.

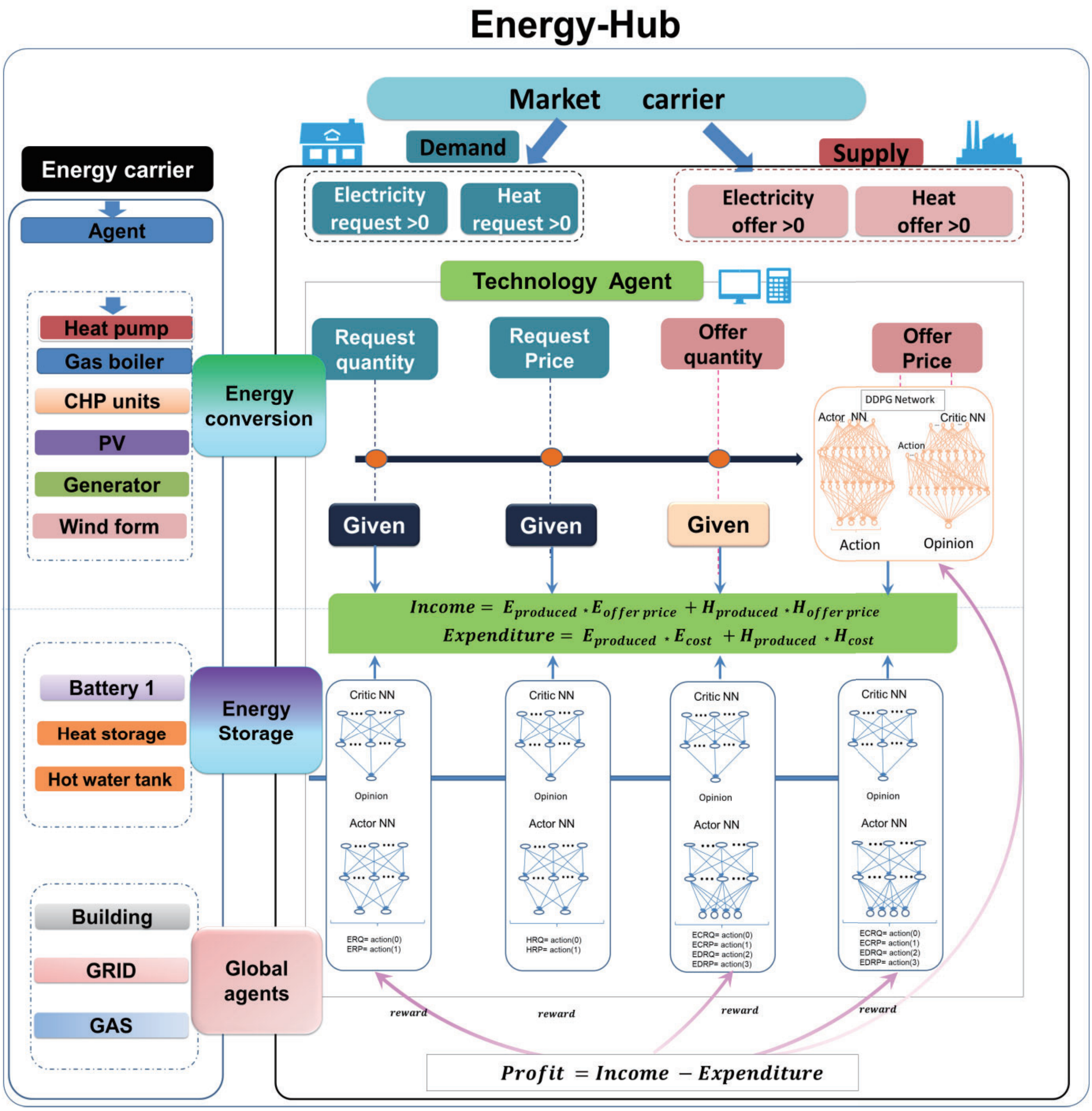

Fig. 1. Multi carrier Learning agent's architecture

Case2: With the variable price, demand, and supply.

In the case1 There is selling and buying price is constant for GRID and house agent. But another agent like storage, Boiler, Power generator's demand is variable. In this case, we assume that selling and buying price is determined by the standard or global agent like GRID in the real world. So, based on the GRID fixed price other agent calculate the quantity to sell and buy, which optimize their profit. In case2 selling and buying price with quantity are variable. So every agent learns what quantity and price make the total profit maximum for this point in time. We assumed that there is a free market where all agent has selling and buying price determination right, and they also fix their buying and selling quantity. After fixing individual price and quantity, we go to the global market where a higher buying price agent with lower selling price agent mechanism fixed the quantity sell and buy.

Based on the agent network state and actions are as follows:

Agent Name	Contribution to states	Actions			
		Buy quantity	Sell quantity	Buy price	Sell price
Boiler	Heat production	-	O	-	O
CHP	Heat production Electricity production	-	O	-	O
Heat storage	Current SOC	O	O	O	O
Photo voltaic	PV generation	-	-	-	O
Power generator	Maximum generation	-	O	-	O
Power storage	Current SOC	O	O	O	O

Fig.2. Agent and action data chart

4. Simulation result and discussion

Here we present simulation results which are based on section 2 model and AI algorithm. And results are divided in two case study which was discussed in section 3. Multi-carrier energy system contains many carriers and According to the content, it's level of difficulty will be ranked difficult or sample. Here in this work, we take a few carrier hub systems to minimize the power use cost per day. Case study 1 and case study 2 contain differ carrier. There are Electrical storage, heat storage, CHP, boiler, windfarm, and GRID in case study 1. A building, solar PV, power generator, power storage, GRID are in case study 2. According to the case study results are discussed below:

Case study 1:

In this case, we have a GRID, wind farm, boiler, CHP, electrical storage, heat storage and demand. Demand contain electricity demand, heat demand and gas demand for houses hold in total for 24 hours 24 discrete form. Here we have a 24-hour total cost for an electricity minimization problem. The Fig.3 shows the

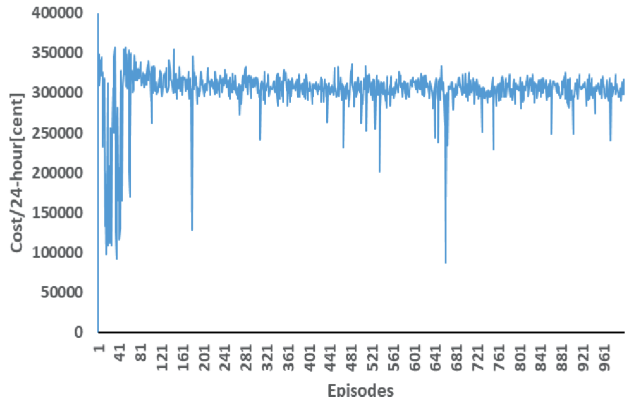

Fig. 3. 24-hour agent contribution total cost

learning time optimization cost for 24 hours. At the beginning of the learning time, the 24-hour cost was 430000[cent]. And after the 1000 episode, its 24-hour total cost is 395316.45[cent]. For optimization purpose storage agent like electricity storage for

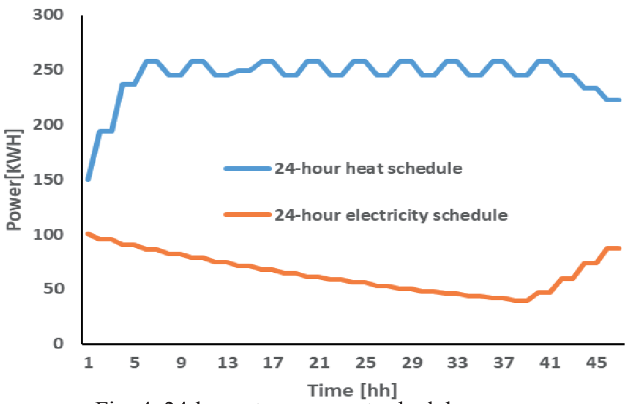

Fig. 4. 24-hour storage agent schedule

electricity and heat storage for heat, storage plays a very significant role. The test time electricity storage and heat storage schedule for 24-hour is in Fig.4. We set electricity storage's maxing storage capacity is 200[KWH] and initial time is half of the maximization. Heat storage max capacity is 300[KWH] and initial is half of the max. We also set minimum level, which is for storage safety, after long time learning agent is able not to discharge below the minimum level. This case study is based on [7] MILP optimization for energy conversion and management. We use a multiagent algorithm (section 2.2) for minimizing the

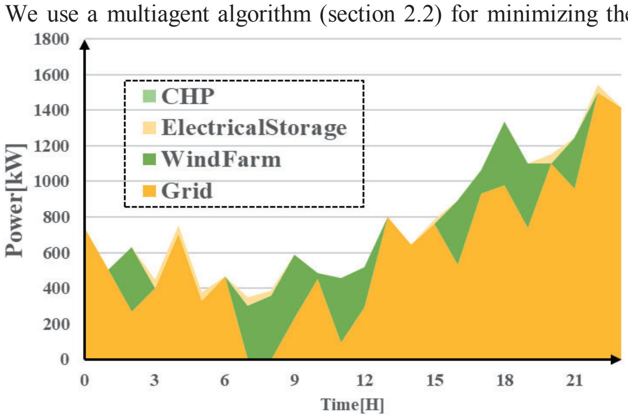

Fig. 5. 24-hour electricity contributor agent

cost. We have storage, CHP and boiler agents, they learn the quantity to supply for demand fulfillment. Here GRID is a global agent and can supply the unbalance quantity of electricity and heat. Fig.5 shows the total contributor for 24-hour electricity almost GRID and wind farm agent supply to the demand but another agent has small capacity and their contribution also small.

Case study 2:

In this case, we have the battery, power generator and PV are student agents, building and GRID are fixed-learning agents Student agent learns to sell/buy price and quantity from their action interact with the environment. In Fig.6 blue line is the

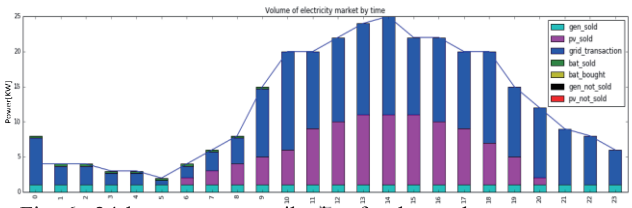

Fig. 6. 24-hour agent contribution for demand

demand of the day and other colored bar are agent contribution for demand. The model is based on the GRID global agent it can buy and can sell too. Start to hour power generator production and GRID are used for the demand of building and after PV starting to produce power it also included. But in increasing in demand PV and generator production does not fulfill the demand so, more power from GRID is use. We use agent wise profit and total profit as a reward in the learning process. In the start of the learning time, all agent does not know how much they can sell

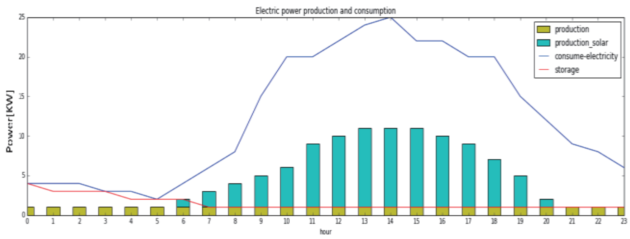

Fig. 7. 24-hour agent contribution for demand other than

and what price will good for sell/ buy. So in the beginning quantity not sold remain but increase in learning they balance the demand and supply. We also use fixed cost and variable cost concept for making profit real. Power generator agent almost fixed for all 24 hours after learning, and battery agent most of the

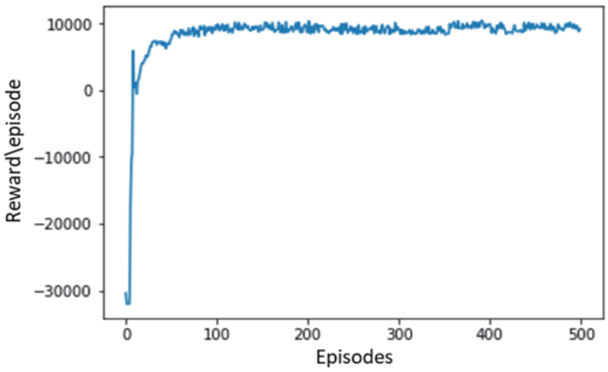

Fig. 8. Reward of multi-agent reinforcement learning

time discharge before to its minimum level but charging actions are very low in Fig.7. The learning time reward graph in Fig.8.

5. Conclusion

Increasing energy scarcity leads to the energy hub concept. It presented a comprehensive attitude towards the energy management problem. Optimal operation of an energy hub with consideration new conditions can help to find out about multi-carrier concept more and move to make it practical. This paper studied presents AI to optimal operation of the energy hub with consideration energy markets and system uncertainties with real-time action. The presented new ways of energy free market operation significantly power up energy market based on forecasting and its sustainability. Artificial intelligence for optimization not only diminish the cost but it also helps to optimize forecast near to very real time. The current work can extend by integrating more power and heat systems in the future. Also, the addition of forecasting agent can make more efficient energy market. Which is not consider in this work and the future works focus on forecasting and real-time optimization.

References

[1] Geidl and G. Andersson. Optimal power flow of multiple energy carriers. IEEE Transactions on Power Systems, 22:145.155, 2007.

[2] Favre-Perrod P. A vision of future energy networks. In: 2005 IEEE power engineering society inaugural conference and exposition in Africa. IEEE; 2005.p. 13–7.

[3] Zhang X, Shahidehpour M, Alabdulwahab A, Abusorrah A. Optimal expansionplanning of energy hub with multiple energy infrastructures. IEEE Trans SmartGrid 2015;6(5):2302–11.

[4] Skarvelis-Kazakos S, Papadopoulos P, Unda IG, Gorman T, Belaidi A, Zigan S.Multiple energy carrier optimisation with intelligent agents. Appl Energy 2016;167:323–35

[5] Shabanpour-Haghighi A, Seifi AR. Simultaneous integrated optimal energy flow of electricity, gas, and heat. Energy Convers Manage 2015;101:579–91

[6] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller. Deterministic policy gradient algorithms. In Int. Conf on Machine Learning, 2014.

[7] M.J. Vahid-Pakdel, Sayyad Nojavan, B. Mohammadi-ivatloo, Kazem Zare. Stochastic optimization of energy hub operation with consideration of thermal energy market and demand response. Energy Conversion and Management 145 (2017) 117–1