2020 5th IEEE International Conference on Big Data Analytics

Deep Reinforcement Learning Approach for Autonomous Agents in Consumer-centric Electricity Market

Yuankun Liu
Artificial Intelligence Application Department
China Electric Power Research Institute
Beijing, China
e-mail: yjs-lyk@epri.sgcc.com.cn

Dongxia Zhang
Artificial Intelligence Application Department
China Electric Power Research Institute
Beijing, China
e-mail: zhangdx@epri.sgcc.com.cn

Chunyu Deng
Artificial Intelligence Application Department
China Electric Power Research Institute
Beijing, China
e-mail: dengchunyu@epri.sgcc.com.cn

Xinying Wang
Artificial Intelligence Application Department
China Electric Power Research Institute
Beijing, China
e-mail: wangxinying@epri.sgcc.com.cn

***Abstract*—With the development of distributed energy, many novel electricity market mechanisms are emerging. The consumer-centric electricity market, which includes the peer-to-peer (P2P) model and the community-based market, arouses more attention in academia and industry area. This paper applies the Deep Q-Learning (DQN) for autonomous agents in the consumer-centric electricity market. Both the local energy priority transactions and public shared energy facilities are taking into consideration. We generated a test dataset and compared results in 5 different scenarios. This study verifies that the applied data-driven methods can handle the peer-to-peer (P2P) decision-making problem as well as promote the profitability of the whole community in the electricity market. Furthermore, multi-agent cooperation with public resources is more appropriate than other situations.**

Keywords*—*peer-to-peer energy trading, big data, artificial intelligence, smart grid

I. Introduction

With the increase of sustainable economy and the promoted by environmental policy, nowadays, the electric power system is undergoing changes from traditional fossil-fuel and centralized control system to a decentralized system with more distributed renewable energy. In this process, a large number of distributed solar photovoltaics appear in the end-user side of the electricity system. These changes have become prevalent in many areas, which lead a role transformation for the end-users from passive consumers to prosumers who are actively participating in the energy production, transaction, and demand response [1].

In the electricity market, the choices of energy supply sources for the end-users are limited, especially for prosumers who cannot but sell their surplus power to the utility companies. However, the local electricity market ecosystem is gradually diversifying and deregulating with some new market mechanisms [2]. The peer-to-peer (P2P) trading in the local electricity market has significantly aroused attention recently [3]. The P2P model facilitates the prosumers to share their electrical power locally, which can bring an E-commerce-like business model to activate the 'desire' of the sharing energy [4] shown in fig.1. Meanwhile, the prosumers tend to form an alliance or a collaboration in terms of their common interests in the electricity market, leading another intensively novel pursued business model named as community-based market [5-6], which are shown in fig.2. Both the P2P model and the community-based market keep a similar consumer-centric viewpoint. What distinguishes these new alternative markets are the degree of decentralization and the organizational structures [7].

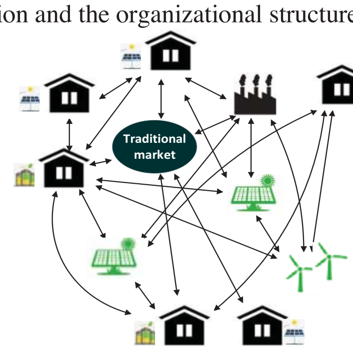

Figure 1. P2P market design

With the increasing trend of decentralized management and cooperative economic principles, the consumer-centric electricity markets, which uphold bottom-up perspective instead of the top-down approach of conventional power system, would empower the sharing and cooperation of consumers, promote the consumption of local renewable energy and reduce financial burden on government and public utilities to invest in distributed clean energy [8-10]. [11] incentivizes the coordination between vast numbers of distributed energy resources, which combine the virtual power plants and peer-to-peer energy trading, leading a potential structure for future prosumer electricity markets.

The key to success in consumer-centric electricity markets is the end user's acceptance of market mechanisms.

978-1-7281-4111-4/20/$31.00 ©2020 IEEE

There are dynamic changes between supply and demand sides and fluctuating prices affected by various factors, causing a decision-making burden for the end-users in electricity markets. In this context, the autonomous agent is one of the approaches to addressing the challenges, which play an intermediate role in balancing the supply and demand between the consumers and producers [12]. On the one hand, it can endow the consumer the opportunity to purchase renewable energy from local prosumers preferentially. On the other hand, the common interests of the community group would be strengthened and consolidated in consumer-centric electricity markets.

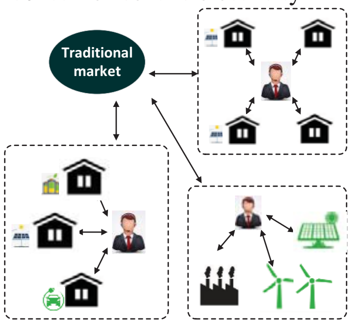

Figure 2. community-based market design

Strategies are the cores of agents in the consumer-centric electricity markets, which are responsible for automatically making the prices publishing about buying and selling electricity, automatically implementing the transaction from producers to consumers in the market. The Power TAC is an open-source platform to simulate the electricity market environment, which has been widely adopted to develop strategies of agents [13]. However, it mainly focuses on the traditional wholesale market, in which the agent automatically purchases electricity from remote fossil fuel power plants [14-15]. In the local retail market, the strategies have been an active research area, and numerous machine learning technologies have been proposed. Supervised learning and unsupervised learning are suitable methods and have been widely applied in terms of new market mechanisms [16-17]. Some scholars designed a tariff market for consumers and used the Markov Decision Processes (MDPs) and reinforcement learning (RL) for learning the strategy of broker agents [18]. The RL techniques have also been applied to learn the strategies in the electricity market in [19-20]. [21] adopts a similar Tariff Market environment proposed in [18] and develops a new strategy based on deep reinforcement learning and multi-agent. [22] applies a SARSA algorithm, another RL technology, to the retail broker strategies.

Three main characteristics are observed when designing the strategy of an autonomous agent in consumer-centric electricity markets:

1. The conventional method relies on the abstract model, in which the accuracy in reflecting the reality cannot be guaranteed. Meanwhile, these conventional ways are based on fixed rules and past information. They lack the adaptability resiliency and flexibility to the non-stationary system.
2. The consumption patterns and geographic location of consumers are diversified and differentiated. The consumer-centric electricity markets should maximize the interests of the community group, which need to take into account the diverse characteristics of consumers and develop distinct pricing strategies. However, the personalization of consumers is not adequately considered in related research work.
3. Some community public resources represented by energy storage are overlooked in some researches. In fact, they can be used as practical tools in the consumer-centric electricity markets to expand overall revenue shared by the members of the community.

In this paper, we prototype a future scenario, which can facilitate both localized energy trading and maximizing the profit of community group. We develop an autonomous agent via deep reinforcement learning play an assistant role in the consumer-centric electricity market serving for the consumers and prosumers.

Our contribution is to:

1. We applied a data-driven method instead of the model-based method, which is suitable for the non-stationary system.
2. A multiagent model system is applied in our work. According to the consumption patterns, we generate the residential household data into four groups, which are separately led by four independent agents. We applied deep reinforcement learning as the strategies. Meanwhile, a cooperative mechanism is applied to coordinate the agents.
3. We apply a sharing economy model to make full use of community public resources and analyze the overall profitability of the community group.

The remainder of the paper is organized as follows: The overview of the background is explained in Section II. In the Section III, the research problem is described in mathematical formulas. The applied cooperative mechanism is explained at last in this Section. Results and discussions are shown in the Section IV. The conclusion and future work are presented In the Section V.

II. Background

A. DL, RL and DRL

Many algorithms of new generation AI technologies are emerging and in the process of development. DL, RL and their combination DRL are representative methods [23]. Deep Learning is a branch of machine learning, which originally resulted from multi-layer artificial neural networks. This algorithm can automatically extract the appropriate features through their multi-level representations and outputs. The 'Deep' here typically refers to the number of layers. There are many different structures of DL. Among them, CNN and RNN are the most popular structure [23]. Reinforcement Learning is also one of the machine learning techniques. Different from supervised learning algorithms which are based on prior knowledge of external, the RL algorithms help the agent become 'smarter' through the interaction with the environment in response to dynamics and uncertainties. There are four basic parts in RL: agent, environment, reward and action. The agent takes some actions to interact with the environment to get some rewards, and the rewards lead the agent to become 'smarter' via a process that expectation goal of obtaining more rewards is gradually achieved. The Fig.3 shows the basic structure of RL.

DRL introduces neural networks to directly express and optimize the value functions, the strategies or the environmental models in an end-to-end manner. DRL can make full use of high-dimensional original data to extract a pattern and build a model, which can also be used as a basis for policy control [24].

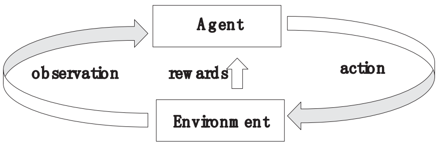

Figure 3. Agent and Environment in RL

The Q-value in the Q-learning, whereby the value of an action taken in a given state, is stored in the Q-table and updated in a similar gradient down manner:

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha(R_{t+1} + \lambda \max_a Q(s_{t+1},a) - Q(s_t,a_t)) \tag{1}$$

In (1), α is the step-size control, where $R_{t+1} + \lambda \max_a Q(S_{t+1},a)$ is the expected rewards that can be obtained by performing action a_t in state s_t. The Q-Learning becomes unrealistic when the state and action space are in high-dimension, because it is too slow to learn the value of the given state individually.

In order to overcome this problem, estimating a value function with approximation is proposed [23]. By adjusting the parameter ω, the function conforms to the value function based on a certain strategy shown in (2).

$$Q(s,a,\omega) \approx Q_\pi(s,a) \tag{2}$$

The task is transformed into solving the parameter ω in the objective function in the formula (3):

$$L(\omega) = \mathrm{E}\left[\left(R_{t+1} + \lambda \max_{a'} Q(s_{t+1},a';\omega') - Q(s_t,a;\omega)\right)^2\right] \tag{3}$$

The gradient descent method is adopted to gradually approximate the parameters that can make the objective function converge to a minimum value.

B. Consumer-centric Electricity Market

Different from the wholesale market, our consumer-centric electricity market is designed for the promotion of localized energy trading under the auxiliary conditions of sharing public resources.

The market proposed here is similar to a combination, and hybrid design of the P2P and community-based market previous mentioned, shown as the fig. 4. The components are introduced by tariff market design in [18] and [21], in which the difference between the tariff market and our consumer-centric electricity market is that the public energy resources owned by community group participate in the decision making and financial income analysis.

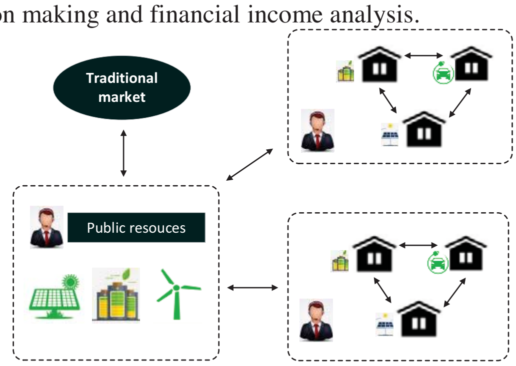

Figure 4. Consumer-centric electricity markets design

The components are outlined as follows:

1. Consumer, $C = \{C_n, n = 1,2,\ldots,N\}$, where each C_n represents a group of households' customers.
2. Producer, $P = \{P_m, \mathrm{m} = 1,2,\ldots,\mathrm{M}\}$, where each P_m represents a group of households' producers.
3. Agents, $A = \{A_k, \mathrm{k} = 1,2,\ldots,\mathrm{K}\}$, where each A_k is the intermediary between the C and P, as well as is responsible for seeking profits in electricity markets. In our design, the local transactions have higher priority. If there is an imbalance between local supply and demand, the agent also buys and sells electricity in the traditional Market.
4. Service Operator, O manages the physical grid and regional infrastructure.
5. Public resources, $PU = \{PU_i, \mathrm{i} = 1,2,\ldots,I\}$, where each PU_i is the public power related equipment, such as energy storage batteries, shared photovoltaics, etc.

It should be highlighted that our market only can run in the daytime, because there is no solar distributed generation at night.

III. Problem Formulation

A. Dataset Characteristics and Analysis

To verify the effect of the algorithm, we generated the dataset randomly. Based on the actual user consumption data and photovoltaic power generation data, we add random (+/−)5%, (+/−) 10%, (+/−) 15%, (+/−) 20% fluctuations to the base data, forming 4 groups of customers. Each group has 8 customers, in which 5 of them own solar power generation. Due to different living habits and consumption patterns, the household consumption load varies. According to their characteristics, it is suitable to manage the consumers and producers separately. Fig. 5 shows the average consumption of 4 groups.

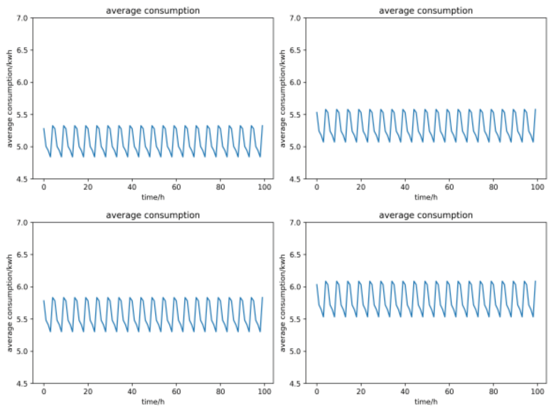

Figure 5. Average consumption loads on four different groups

B. Objective Function and Constraints

This situation can be summarized as a mathematical optimization problem. Our objective function is presented in the formula 4.

$$\max \sum_{k=1}^{K} \sum_{t=1}^{T} (p_{t,C}^{A_k} w_{t,C} - p_{t,\mathrm{P}}^{A_k} w_{t,\mathrm{P}} - \phi_t \left| w_{t,C} - w_{t,\mathrm{P}} - PU_{t,i} \right|) \quad (4)$$

In (4), $w_{t,C}$ and $w_{t,\mathrm{P}}$ represent the consumption and production. $p_{t,C}^{A_k}$ and $p_{t,\mathrm{P}}^{A_k}$ respectively represent the consumer price and the producer price under the autonomous agent A_k. The term $\left| w_{t,C} - w_{t,\mathrm{P}} - PU_{t,\mathrm{i}} \right|$ represents the supply-demand imbalance in the portfolio at timeslot t, where the $PU_{t,\mathrm{i}}$ represents the public energy resources shared by the community. The ϕ_t is the imbalance fee of A_k at time t, which is specified by the Service Operator.

C. Implementation Details

We set up 5 agents in our experiment. As we have 4 different groups of customers, each of them is assigned an agent. The public sharing resources is assigned an agent.

For the 4 different autonomous agents, the state vector is $S_t = [\mathrm{T(t)}, \mathrm{U(t)}]$,where T(t) is the tariff price, U(t) is the average electricity consumption. The action is price publishing, including the purchase price of the consumer and the sale price of the producer. The autonomous agents play the roles to promote the local energy P2P trading model preferentially as well as get more profits for the community group. The rewards of the agents follow the objective function. To promote cooperation among multiple agents, we applied the method in the [21], which is shown in (5).

$$r_t^i = r_t - (\sum_{j \neq i} p_t^j w_{t,C}^j - \sum_{k \neq i} p_t^k w_{t,\mathrm{P}}^k - \varphi_t^i), \mathrm{j} \in \mathrm{C}, \mathrm{k} \in \mathrm{P} \quad (5)$$

In (5), the i represents the customer group. φ_t^i represents the supply-demand imbalance term $\phi_t \left| w_{t,C} - w_{t,\mathrm{P}} - PU_{t,i} \right|$. r_t is reward computed as formula 4.

For the agent of the public sharing resources, we have simplified the process. We use energy storage batteries as a representative. We have specified some passive and strict charging and discharging conditions.

IV. Results And Discussion

In our experiment, we compare the results of 5 different situations. The first situation is the fixed price. The second and third situations are the cooperative and non-cooperative multiagent modes with public resources. The Fourth and fifth cases are the cooperative and non-cooperative multiagent modes without public resources. The accumulation profits of 5 different situations are shown in Fig. 6.

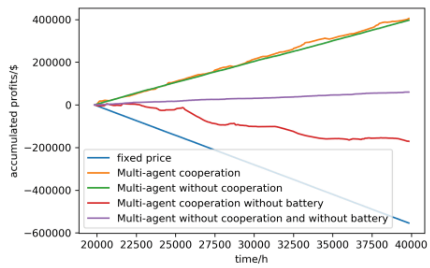

Figure 6. Accumulation profits of 5 different situations

As we can see in Fig. 6, the method of multi-agent cooperation with the assistance of the public resources shows the best results. We compare the training process of multi-agent cooperation with public resources to the other 4 situations, which is shown in the Fig. 7.

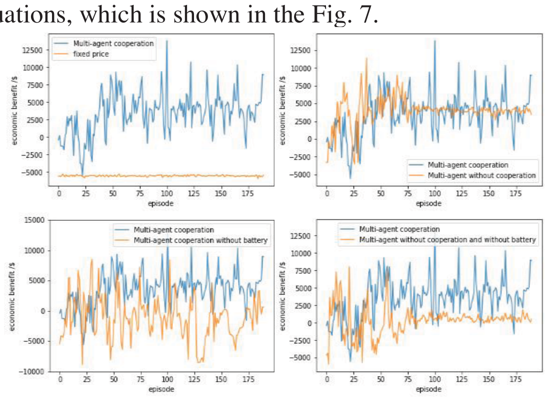

Figure 7. The comparison of training process

The DRL algorithm is an optimizing process continuously interacting with the environment. Due to the characteristics of the applied algorithm, the agent gradually adapts to the environment and obtain more rewards. Meanwhile, economic benefits increase. There are many random choices at the beginning. The agent learns to choose the converging trend and possibilities which are close to the optimization objective after many iterations. As the fig. 8 shows, the convergence of cooperation methods is not as good as non-cooperation methods, but the actual results of which are better.

V. CONCLUSION AND FUTURE WORK

In this paper, a consumer-centric electricity market is designed, which includes the P2P model and the community-based market. We generated the dataset randomly and executed the local peer-to-peer energy transaction by the autonomous agent according to the data characteristics. Deep Q-Learning (DQN) is applied for autonomous agents in this consumer-centric electricity market. Meanwhile, to promote the effective, we verify the algorithm in 5 different situations. It is proved that multi-agent collaboration and public shared facilities help increase the profitability of the community.

Much investigation works need further research to examine the mechanism of the proposed algorithm. It is interesting to make a comparison with traditional model-based algorithms. Besides, partial observable condition should be discussed and analyzed in detail.

ACKNOWLEDGMENT

This research has been funded by the Project Research on Forecasting Method of Smart Grid Big Data Based on Random Projection Neural Networks (61703379) supported by National Natural Science Foundation of China.

REFERENCES

[1] Mengelkamp, Esther, et al. "A blockchain-based smart grid: towards sustainable local energy markets." Computer Science-Research and Development 33.1-2 (2018): 207-214.

[2] Zhang, Chenghua, et al. "Review of existing peer-to-peer energy trading projects." Energy Procedia 105 (2017): 2563-2568.

[3] Zhang, Chenghua, et al. "Peer-to-Peer energy trading in a Microgrid." Applied Energy 220 (2018): 1-12.

[4] Morstyn, Thomas, et al. "Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants." Nature Energy 3.2 (2018): 94.

[5] Liu, Nian, et al. "Energy-sharing model with price-based demand response for microgrids of peer-to-peer prosumers." IEEE Transactions on Power Systems 32.5 (2017): 3569-3583.

[6] Yuan, Chen, Mahesh S. Illindala, and Amrit S. Khalsa. "Co-optimization scheme for distributed energy resource planning in community microgrids." IEEE Transactions on Sustainable Energy 8.4 (2017): 1351-1360.

[7] Sousa, Tiago, et al. "Peer-to-peer and community-based markets: A comprehensive review." Renewable and Sustainable Energy Reviews 104 (2019): 367-378.

[8] Hu, Jing, et al. "Identifying barriers to large-scale integration of variable renewable electricity into the electricity market: A literature review of market design." Renewable and Sustainable Energy Reviews 81 (2018): 2181-2195.

[9] Peng, Donna, and Rahmatallah Poudineh. "Electricity market design for a decarbonised future: An integrated approach." (2017).

[10] Bauwens, Thomas. "Explaining the diversity of motivations behind community renewable energy." Energy Policy 93 (2016): 278-290.

[11] Morstyn, Thomas, et al. "Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants." Nature Energy 3.2 (2018): 94.

[12] Ketter, Wolfgang, John Collins, and Carsten A. Block. "Smart grid economics: Policy guidance through competitive simulation." (2010).

[13] Ketter, Wolfgang, Markus Peters, and John Collins. "Autonomous agents in future energy markets: the 2012 power trading agent competition." Twenty-Seventh AAAI Conference on Artificial Intelligence. 2013.

[14] Urieli, Daniel, and Peter Stone. "An MDP-based winning approach to autonomous power trading: formalization and empirical analysis." Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems. International Foundation for Autonomous Agents and Multiagent Systems, 2016.

[15] Liefers, Bart, Jasper Hoogland, and Han La Poutré. "A successful broker agent for power tac." Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets. Springer, Cham, 2014. 99-113.

[16] Robu, Valentin, et al. "Efficient buyer groups with prediction-of-use electricity tariffs." IEEE Transactions on Smart Grid 9.5 (2017): 4468-4479.

[17] Reddy, Prashant P., and Manuela M. Veloso. "Negotiated learning for smart grid agents: entity selection based on dynamic partially observable features." Twenty-Seventh AAAI Conference on Artificial Intelligence. 2013.

[18] Reddy, Prashant P., and Manuela M. Veloso. "Strategy learning for autonomous agents in smart grid markets." Twenty-second international joint conference on artificial intelligence. 2011.

[19] Angelidakis, Angelos, and Georgios Chalkiadakis. "Factored mdps for optimal prosumer decision-making in continuous state spaces." Multi-Agent Systems and Agreement Technologies. Springer, Cham, 2015. 91-107.

[20] Chowdhury, Moinul Morshed Porag, et al. "Investigation of learning strategies for the spot broker in power tac." Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets. Springer, Cham, 2015. 96-111.

[21] Yang, Yaodong, et al. "Recurrent Deep Multiagent Q-Learning for Autonomous Brokers in Smart Grid." IJCAI. Vol. 18. 2018.

[22] Wang, Xishun, Minjie Zhang, and Fenghui Ren. "A hybrid-learning based broker model for strategic power trading in smart grid markets." Knowledge-Based Systems 119 (2017): 142-151.

[23] Zhang, Dongxia, Xiaoqing Han, and Chunyu Deng. "Review on the research and practice of deep learning and reinforcement learning in smart grids." CSEE Journal of Power and Energy Systems 4.3 (2018): 362-370.

[24] Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." nature 529.7587 (2016): 484.