Multilateral Decision Strategy of Electric Vehicle Charging Market based on Deep Reinforcement Learning

Ruichang Zhang

December 2021

1 Introduction

The electric vehicle charging market is facing many difficulties. The rapid increase number of electric vehicles will not only bring about higher energy demand, but will also have a huge impact on the stability of the grid during grid-connected charging. Under this situation, one main problem is that electricity demand of consumers do not always match and the supply of producer, which may lead to profit loss and social welfare loss. Solutions need to be found to address the problem. Concretely, consumer need a strategy to seek for utility under a price given by producer and likewise, producer need a similar strategy to maximize its profit. They sense each other by observing the opponent's actions and make their own decisions. Moreover, different from a typical consumer-producer market, there exists another participant called the aggregator who acts as a middleman between consumer and producer in the electric vehicle charging market. The existence of the aggregator makes it harder to explore a perfect solution. Setting objections for each participant are crucial because conflicts always exists in this game. In the proposed work, two circumstances are to be discussed: maximizing each lateral's profit and maximizing social welfare. At last, comparison of each participant's payoff and social welfare under each circumstance will be presented.

This proposed work will design a multilateral decision strategy under plug-in electric vehicle(PEV) charging market, which mainly contains three kind of participants: PEV users, electricity aggregators and the power grid. Firstly, this work will establish a game theoretical framework containing the participants above. Various factors will be assigned to each lateral to represent their legal moves, benefits and costs. Secondly, two assumptions will be brought up to direct the optimization objectives. Assumption one is about to maximizing the total profit of a party and assumption two aims at achieving a maximum social welfare. Finally, this work will be validated under a virtual environment, the benefit and cost of each party under different assumptions will be quantified.

2 Literature Review

With a huge amount of PEVs flooding the market, entities in the electricity market are facing great challenges. The market will fall into chaos if not reasonably regulated. Currently, researchers have proposed solutions from various research fields such as energy allocation, demand response program, operational control, etc.. Especially, demand response program together with controlling of electricity market proves to be an feasible solution for improving the efficiency of market operation [1].

The difficulties with the power system can be transformed into decision and control problems in most cases. Many tradition methods with good performance were proposed, such as convex optimization method [2], mixed integer programming [3, 4], dynamic programming [5, 6], stochastic programming [7, 8], genetic algorithm [8], ant colony optimization [9], etc.. However, it is often the case that conventional methods often meet bottlenecks when attempting to solve decision and control problems due to increasing complexity, uncertainty and high data dimensions. Therefore, data-driven methods toward solving such problems are being extensively studied [10].

Demand response can be defined as the changes in electricity usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity over time, or incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized [11]. DR,by promoting the interaction and responsiveness of the customers, determines short-term impacts on the electricity markets, leading to economic benefits for both the customers and the utility.

Moreover, by improving the reliability of the power system and, in the long term, lowering peak demand, it reduces overall plant and capital cost investments and postpones the need for network upgrades [12, 13].

In a typical power grid model, demand response program acts as a mediator, balancing the load of the grid and the cost of end users in case the market fall into anarchy, damaging total social welfare. But how to efficiently find the Nash equilibrium needs lots of research. Plenty of studies using DRL method has been done to explore such cases. Intuitively, the grid benefits from more profits and less power fluctuation while user benefits from less cost and more utility from each unit of electricity. In [14], a DRL based demand response strategy is proposed by assisting the service provider in purchasing electricity from various consumers to balance power fluctuations and keep grid reliability to solve a real-time incentive-based demand response problem. Further, a DDQN method was used to optimize scheduling space of interruptible load at demand side in the case of realistic daily load of users [15].

Combined with game theory, some work has been down by building a game model and search a Nash equilibrium. A two-stage game model between power companies and customers is proposed in [16] and solved by RL. At the first stage, the customers' optimal power consumption is obtained and at the second stage, the power companies' prices are determined. A dynamic pricing demand response model is proposed in [17], considering the service provider's profit and customers' costs. The retail price is determined by RL according to electricity demand and wholesale electricity prices.

However, when it comes to a real electricity market of PEV charging, the result from a two-sided game seems less practical. In most cases, PEV users do not purchase electricity directly from the grid but from an aggregator, who servers as a middle man between the end user and the grid. It buys power from wholesale electricity market and sells to customers in retail price. This part of the game only benefit from its profits. Some studies used DRL methods to solve such questions under different assumptions. A constrained energy trading game among end-consumers is proposed in [18] by adaptive RL with incomplete information, and finally, bidding strategy converges to Nash equilibrium. Paper [19] propose a microgrid energy trading game model considering renewable energy generation and demand, battery level and trading history, and the Nash equilibrium is obtained by DRL approach. Furthermore, A hierarchical electricity market with bidding and pricing of load serving entity (LSE) is researched in [20], dynamical bid and price response functions are learned by deep neural network and state transition samples are generated by deep deterministic policy gradient (DDPG) algorithm.

Previous works have shown positive performance under their respective assumptions and objectives. However, an integrated comparison of strategies among different entities under different assumptions still waits to be explored. Concretely, the aggregator will probably set different retail electricity price under the assumption of maximum social welfare and maximum profit of itself. Also, factors like the electricity load fluctuation in the grid which do not affect the profit of power grid directly and instantly could also cause higher operation cost by changing states whom are not presented explicitly in the model. As some hidden state may be affected by factors which were neglected before, the model should coordinate more reasonable factors.

3 Proposed Methodology

Method of this proposed work follows three steps: establishing the game model, solving the objection and validating the result under realistic scenarios.

3.1 Modelling

Three entities will be described in detail to simulate the power grid, the aggregator and the PEV user. In general, payoff of each participants can be defined as its own utility minus cost. A tentive framework of the model are illustrated in Fig.1. Market state serves as a common environment which contains all current information of the market. Once an agent make actions, market state will change and others will sense a new state. Based on the new state, the others will make their own decision and further change the state. At last, the initial agent can get its own reward based on the new state. From the perspective of each agent, it makes decision based on the current state and gets a reward. For example, PEV user can provide demand curve as its action, then others responses to the new market state and make their decisions, finally, PEV user obtains retail price and calculate the reward.

Detail of each lateral of the model are showed below:

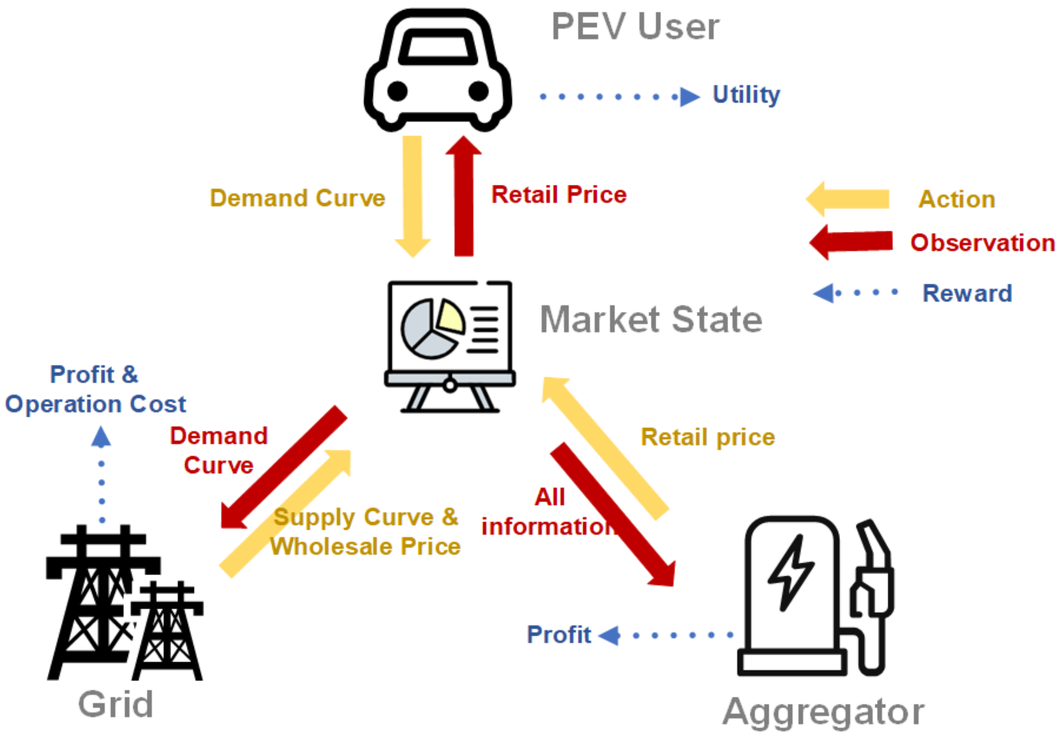

Figure 1: A sequential decision-making framework

3.1.1 PEV user model

A typical PEV user's benefit from the satisfaction brought by each unit of electricity $U(x)$, and suffer from cost of the price $P(x)$. Moreover, the shifting of their demand caused by adjustments of retail price may as well result in satisfaction loss. In the market environment, users can only accept the price and decide how to assign their charging demand over time to maximize their payoff.

In conclusion, user observe retail price of the market and make decision by output their corresponding demand curve.

3.1.2 Power grid model

The objective of the power grid can be concluded as maximizing the profit and minimizing operation cost. The wholesale electricity price are constrained by the market state so the grid has to decide the supply curve of electricity quantity. Also, as the fluctuation range of power output increase, cost of each unit of electricity will rise accordingly.

In conclusion, the grid observes the wholesale market state and decides the supply curve and wholesale price.

3.1.3 Aggregator model

The only purpose of the aggregator is to make money by pocket the difference between wholesale price and retail price. It collect user's behavior distribution and decide the most profitable charging price strategy.

In conclusion, an aggregator observes the wholesale price and distribution of demand curve. Then, it decides the retail price and further affect others' decisions.

3.2 Solving

To solve the above model, DRL methods can be utilized with some modifications. An Markov decision process(MDP) with three agents can be defined as:

$$(S, A_{user}, A_{agg}, A_{grid}, R_{user}, T, \gamma, R_{user}, R_{agg}, R_{grid})$$

Elements of this MDP are illustrated as follows:

- State S : S is consisted of historical user demand distribution, current user demand distribution, wholesale market price, retail market price, electricity supply curve.
- Action $A = [A_{user}, A_{agg}, A_{grid}]$: Users can decide current user demand; Aggregators can decide the retail market price; The grid can decide wholesale market price and supply curve.
- Reward $R = [R_{user}, R_{agg}, R_{grid}]$: Users benefit from the utility from purchasing process; Aggregators benefit from their profit; The grid benefit are decided by electricity profit and operation cost.
- Transition probability T : Given current state and an action from an arbitrary side, next state can be calculated directly based on aforementioned model.
- Discount factor γ : Account for long term reward decline.

Deep reinforcement learning(DRL) is a combination of deep learning and reinforcement learning, which has the ability to make decisions from large amount of data. Due to its robustness and flexibility nature, it show great performance in various kind of decision making tasks. It has been applied to many areas such as video games [21], recommendation system [22], transportation system, etc.. Likewise, deep reinforcement learning technique can also be integrated into power systems to solve problems like demand response program, dynamic pricing [23], PEV charge scheduling [24], etc.. DRL methods like deep deterministic policy gradient (DDPG) can be utilized to solve the problem. Alternatively, MARL method together with game theory may as well play roles in the problem.

There are two feasible deep reinforcement learning solutions to this model: single-agent algorithm and multi-agent algorithm. From the perspective of single-agent, a centralized network controls the whole market and all participants act simultaneously. Reward function can be designed as weighted summation of their own payoff or the social welfare. This joint agent observes all the information at once, and output a joint action set. Potential advantages of the solution is that the decision process is shorter, even Monte Carlo sampling can be utilized. The shortcoming is also apparent for it probably won't converge if the rewards are conflict to some extent. From a multi-agent game model sight, each agent makes decision separately and get the reward after others made their actions. Advantage of multi-agent sight is that the reward is easy to design for each agent. Also, agents make decisions sequentially which is closer to reality. However, an episode will cover a long trajectory and most of the times it may fell into a local optima.

3.3 validating

Finally, numerical results should be analysed to prove the performance of the proposed method.

4 Conclusion

Lots of works has been done to explore the potential of using deep reinforcement learning to solve smart grid scenarios. However, few work has research the strategies among a three-entities electricity charging market. This work will propose a game model and a DRL solution towards this question. The method may have the potential to impact others to explore the question deeper.

References

[1] Linas Gelazanskas and Kelum AA Gamage. Demand side management in smart grid: A review and proposals for future direction. *Sustainable Cities and Society*, 11:22–30, 2014.

[2] Wenbo Shi, Na Li, Chi-Cheng Chu, and Rajit Gadh. Real-time energy management in microgrids. *IEEE Transactions on Smart Grid*, 8(1):228–238, 2015.

[3] Michael Zachar and Prodromos Daoutidis. Microgrid/macrogrid energy exchange: A novel market structure and stochastic scheduling. *IEEE Transactions on Smart Grid*, 8(1):178–189, 2016.

[4] Christos Ordoudis, Pierre Pinson, and Juan M Morales. An integrated market for electricity and natural gas systems with stochastic power producers. *European Journal of Operational Research*, 272(2):642–654, 2019.

[5] Luckny Zéphyr and C Lindsay Anderson. Stochastic dynamic programming approach to managing power system uncertainty with distributed storage. *Computational Management Science*, 15(1):87–110, 2018.

[6] Jean-Laurent Duchaud, Gilles Notton, Christophe Darras, and Cyril Voyant. Power ramp-rate control algorithm with optimal state of charge reference via dynamic programming. *Energy*, 149:709–717, 2018.

[7] Hieu Trung Nguyen, Long Bao Le, and Zhaoyu Wang. A bidding strategy for virtual power plants with the intraday demand response exchange market using the stochastic programming. *IEEE Transactions on Industry Applications*, 54(4):3044–3055, 2018.

[8] Ali İrfan Mahmutoğulları, Shabbir Ahmed, Özlem Çavuş, and M Selim Aktürk. The value of multi-stage stochastic programming in risk-averse unit commitment under uncertainty. *IEEE Transactions on Power Systems*, 34(5):3667–3676, 2019.

[9] Gia Nhu Nguyen, K Jagatheesan, Amira S Ashour, B Anand, and Nilanjan Dey. Ant colony optimization based load frequency control of multi-area interconnected thermal power system with governor dead-band nonlinearity. In *Smart trends in systems, Security and Sustainability*, pages 157–167. Springer, 2018.

[10] Zidong Zhang, Dongxia Zhang, and Robert C Qiu. Deep reinforcement learning for power system applications: An overview. *CSEE Journal of Power and Energy Systems*, 6(1):213–225, 2019.

[11] M.H. Albadi and E.F. El-Saadany. A summary of demand response in electricity markets. *Electric Power Systems Research*, 78(11):1989–1996, 2008.

[12] Charles Goldman. Coordination of energy efficiency and demand response. 2010.

[13] Salman Mohagheghi, James Stoupis, Zhenyuan Wang, Zhao Chang Li, and Hormoz Kazemzadeh. Demand response architecture: Integration into the distribution management system. *2010 First IEEE International Conference on Smart Grid Communications*, pages 501–506, 2010.

[14] Renzhi Lu and Seung Ho Hong. Incentive-based demand response for smart grid with reinforcement learning and deep neural network. *Applied energy*, 236:937–949, 2019.

[15] Biao Wang, Yan Li, Weiyu Ming, and Shaorong Wang. Deep reinforcement learning method for demand response management of interruptible load. *IEEE Transactions on Smart Grid*, 11(4):3146–3155, 2020.

[16] Pavlos Athanasios Apostolopoulos, Eirini Eleni Tsiropoulou, and Symeon Papavassiliou. Demand response management in smart grid networks: A two-stage game-theoretic learning-based approach. *Mobile Networks and Applications*, 26(2):548–561, 2021.

[17] Renzhi Lu, Seung Ho Hong, and Xiongfeng Zhang. A dynamic pricing demand response algorithm for smart grid: reinforcement learning approach. *Applied Energy*, 220:220–230, 2018.

[18] Huiwei Wang, Tingwen Huang, Xiaofeng Liao, Haitham Abu-Rub, and Guo Chen. Reinforcement learning for constrained energy trading games with incomplete information. *IEEE transactions on cybernetics*, 47(10):3404–3416, 2016.

[19] Xingyu Xiao, Canhuang Dai, Yanda Li, Changhua Zhou, and Liang Xiao. Energy trading game for microgrids using reinforcement learning. In *International Conference on Game Theory for Networks*, pages 131–140. Springer, 2017.

[20] Hanchen Xu, Hongbo Sun, Daniel Nikovski, Shoichi Kitamura, Kazuyuki Mori, and Hiroyuki Hashimoto. Deep reinforcement learning for joint bidding and pricing of load serving entity. *IEEE Transactions on Smart Grid*, 10(6):6366–6375, 2019.

[21] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.

[22] Xinshi Chen, Shuang Li, Hui Li, Shaohua Jiang, Yuan Qi, and Le Song. Generative adversarial user model for reinforcement learning based recommendation system. In *International Conference on Machine Learning*, pages 1052–1061. PMLR, 2019.

[23] Feiye Zhang, Qingyu Yang, and Dou An. Cddpg: A deep-reinforcement-learning-based approach for electric vehicle charging control. *IEEE Internet of Things Journal*, 8(5):3075–3087, 2020.

[24] Hepeng Li, Zhiqiang Wan, and Haibo He. Constrained ev charging scheduling based on safe deep reinforcement learning. *IEEE Transactions on Smart Grid*, 11(3):2427–2439, 2019.