Model design for the analysis of market-based redispatch – applying deep reinforcement learning algorithms

Viktor Zobernig[1, 2], Ksenia Tolstrup[1], Jochen Cremer[2], Laurens de Vries[2]
[1] AIT Austrian Institute of Technology, Giefinggasse 6, Vienna, Austria, viktor.zobernig@ait.ac.at, ksenia.tolstrup@.ait.ac.at [2] Delft University of Technology, Postbus 5, Delft, Netherlands, V.Z.Zobernig@tudelft.nl, J.L.Cremer@tudelft.nl, L.J.deVries@tudelft.nl

Overview

The growing share of renewable energy sources, progressive energy market integration, and a slow pace of grid expansion significantly increased grid congestion volumes and, consequently, the need for redispatch measures in the coupled European electricity grids. Implementing proper incentive mechanisms for redispatch could encourage the entry of more flexible resources to reduce the cost of solving congestion. In addition to other short-term and ancillary service markets, redispatch markets could provide an extra avenue to incentivize participation of existing and new flexible resources to support the system.

On the flipside, studies have shown that a two-staged market design is prone to gaming that exploits the opportunities between the day-ahead electricity market and a potential redispatch market. This behavior is often referred to as “inc-dec gaming”. The conditions for the occurrence of inc-dec gaming and its impact on socio-economic costs are inconclusive in literature: some raise concerns and others interpret the effects as rather moderate [1], [2].

The applied approaches are limited by assuming certain characteristics about bidding behavior or are focused on specific national conditions (e.g., in Germany). State-of-the-art optimization methods have to discretize the set of bids to choose from, as well as the availability of adjusting the offered quantities. This limits the possibility of identifying the maximum impact of strategic behavior on the market [3]. These studies also assume that market participants have critical information about parameters of their competitors.

This paper addresses these shortcomings by proposing an agent-based model of the two-staged market, representing potentially strategic market actors using a deep reinforcement learning (DRL) algorithm. Generally, DRL has proven its capability to simulate self-optimizing participants with partial information exploring energy market domains [3]. Using DRL makes it possible to implement agents without any assumption about their bidding behavior that maximize their own reward (e.g., profits). Additionally, this design ensures observing changes in market inefficiency arising from any gaming strategy.

We propose this model design to improve the integration of existing uncertainties. Further, to introduce an approach that is not limited to identify strategic behavior that leads to an equilibrium but identifies any strategic patterns. With that we want to support the analysis of sufficient redispatch market design and the implementation of potential countermeasures.

Methods

We apply an agent-based model (ABM) approach including DRL algorithm to allow for a possibility for the agents, i.e., market actors, to act strategically. The model environment consists of a day-ahead market (DAM), which gets cleared without any regard to network constraints, a physical grid model and a redispatch market. The whole model process is illustrated in Fig. 1. Similarly, to the logic of a remedial action optimization used by transmission system operators, the grid model is implemented to check whether the dispatch resulting from a DAM clearing leads to any congestion by using a power flow calculation. The redispatch market gets cleared by calculating an optimal power flow including the dispatched quantities from the day-ahead market together with the possible range of deviation of those and potential additional offers, based on the bids of redispatch market participants. It is assumed that activated redispatch gets remunerated pay-as-bid and the day-ahead market price is through uniform pricing.

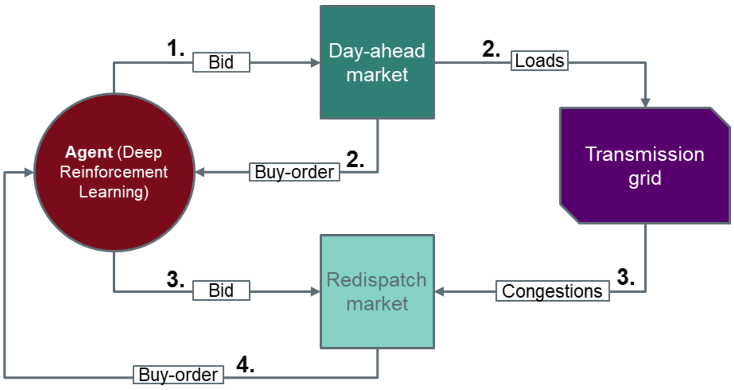

Figure 1: Simplified illustration of the agent-based model together with deep reinforcement learning agents.

Agent behavior is trained by using the Deep Deterministic Policy Gradient (DDPG) algorithm, which is also applied to analyze inc-dec gaming in [3]. The DRL-based agents learn to act optimally through the Markov Decision Process. We further adapt this algorithm by splitting up the action space from one action for both markets, to two separate actions for each of the markets. Thus, the agent decides first the quantity and at which price it wants to offer it on the DAM, regarding to a continuous action space, and then does the same for the redispatch market, considering the previous market outcome. The implemented agents represent either gas or wind plants. In general, the model is constructed in a way that it easily can get extended by introducing further electricity markets (e.g., balancing) to study more complex strategies.

Results

We structured the results in three parts: (1) verifying the learned agent behavior by comparing the results to naïve rule-based agents that always offering their maximum available capacity at their short-term marginal costs; (2) comparing our adapted approach to the DDPG algorithm and (3) demonstrating the advantage of using DRL by first, increasing the uncertainty of the occurrence of congestion through increasing the fluctuation of wind availability and secondly by comparing the outcome of this scenario with fluctuation, once by having full foresight of the predicted wind availability and once without.

Generally, it is assumed that DRL is capable of outperforming rule-based agents. However, investigations have shown that DDPG does not always converge to an optimal behavior and sometimes performed even worse than the rule-based agents. For instance, an agent that is situated in a region where it is very likely to get regulated downwards, a well-performing gaming strategy would be to increase its offered quantity on the day-ahead market while decreasing its price bid to ensure its activation. in this way, the agent could increase the congestion and with that its ability of selling more on the redispatch market for a much higher price compared to the day-ahead. Nevertheless, as long as the agent does not influence the occurring congestion with this type of behavior, they would earn less than without behaving strategically, which might hinder the agent of finding the optimal solution. Hence, finding the optimum can become difficult, because the gradients might indicate to improve the action in the wrong direction. Thus, we decided to train the agents behavior on the two markets individually but keeping maximizing their cumulative profits. In this way, we manage to simplify the space of optimal solutions for the agent without introducing any additional assumptions.

In a next step we use our modified agents to analyse their changes in behavior when relaxing the probability of the occurrence of congestion through increasing the fluctuation of wind availability. In optimization methods commonly applied for solving such problems, it is assumed that agents are aware of the available capacity of others in order to identify their optimal bidding strategy. Thus, we intend to examine how agent behavior changes 1) if they know about the wind availability, indicated through their observation space, and 2) if they do not assess this information. The second scenario demonstrates a situation where the agents have to decide their action under uncertainty. With that we want to show the impact of including such information. This approach improves commonly applied methods in two ways 1) it allows to identify any strategic behavior pattern 2) it is more flexible regarding the model environment design, i.e., neglecting important assumption of having information that it is not available in reality.

Conclusions

We introduced a new DRL algorithm architecture to improve agent behavior in sequential day-ahead and redispatch markets We further proposed different scenarios, distinguishable by the information that is available for the agents, to demonstrate some advantages of using DRL instead of optimization methods. It allows to investigate more complex bidding strategies in multi-staged electricity market. Further, this model enables a more accurate analysis of different market designs as well as the investigation of sufficient countermeasures. Based on that currently proposed monitoring methods to detect gaming might get further improved. These findings support the investigation of remuneration methods for redispatch reducing socio-economic costs by incentivizing a higher number of flexibility providers without facing the risk of gaming.

For further studies it is suggested to also introduce additional source of uncertainties to the model in order to analyze the impact of inc-dec gaming more accurately. A drawback of this method is that DRL is facing the problem of non-stationarity when more DRL agents are introduced to the model, which has a significant impact on their performances. Hence, we see that further improvements of this approach are needed in the future.

References

[1] L. Hirth und I. Schlecht, „Market-Based Redispatch in Zonal Electricity Markets“, *SSRN Journal*, 2018, doi: 10.2139/ssrn.3286798.

[2] M. Sarfati und P. Holmberg, „Simulation and Evaluation of Zonal Electricity Market Designs“, *Electric Power Systems Research*, Bd. 185, S. 106372, Aug. 2020, doi: 10.1016/j.epsr.2020.106372.

[3] Y. Ye, D. Qiu, M. Sun, D. Papadaskalopoulos, und G. Strbac, „Deep Reinforcement Learning for Strategic Bidding in Electricity Markets“, *IEEE Transactions on Smart Grid*, Bd. 11, Nr. 2, S. 1343–1355, März 2020, doi: 10.1109/TSG.2019.2936142.