ENERGY TRADING STRATEGIES FOR STORAGE DEVICES USING APPROXIMATE DYNAMIC PROGRAMMING

Ioannis Boukas, University of Liege, +32 483 72 86 53, ioannis.boukas@uliege.be

Damien Ernst, University of Liege, dernst@uliege.be

Bertrand Cornelusse , University of Liege, bertrand.cornelusse@uliege.be

Overview

The efficient integration of renewable energy resources (RES) in future power systems as directed by the recent worldwide energy policy drive has given rise to challenges related to the security, sustainability and affordability of the power system ("The Energy Trilemma"). In particular, the effect of high RES penetration to the modern electricity markets has been the subject of extensive research.

In that context, energy storage is expected to play a key role in the future power system and energy markets. There are many different services that can be provided from storage devices based on the connection point with the grid, ranging from energy arbitrage and reserves to distribution deferral and increased self-consumption. In this study we will focus on the energy arbitrage in the wholesale market and in particular in the intra-day market. Wholesale transactions are expected to move closer to real time for two reasons, namely in order to reward flexibility sources and to exploit more accurate forecasts of RES production and load [1].

In this paper, the problem faced by a storage device operator participating in the Continuous Intraday (CID) market is considered. A modeling framework is implemented in order to simulate the operation of the storage device and its interaction with the market. The goal of the trading agent is the maximization of the cumulative revenues received over the entire trading horizon. Approximate Dynamic Programming (ADP) is used to develop a trading strategy that is compared to the benchmark called "rolling intrinsic". The results indicate that the trading agent is able to outperform the benchmark after a period of training. Real data from the CID market are used for training and validation purposes.

Methods

Participation in the CID market is a continuous double auction process similar to stock exchange, where each traded product corresponds to a time-slot of physical delivery. Trading starts when the gate opens and finishes when each products' gate closes 30 mins before the physical delivery. The trading process takes place through a centralized order book for each product. Participants submitting orders must specify the type of each order (buy or sell) the volume traded and the price they are willing to pay/receive per unit, as well as many other specifications related to the order execution and validity. Then the orders are sorted based on their price and a transaction takes place when the price someone is willing to pay in order to buy energy is at least equal to the price that someone requests in order to produce this amount of energy. Orders are treated according to the first come first serve (FCFS) principle.

We model this process as a Markov Decision Process (MDP), where the state of the system at the following time-step depends on the state at the previous time step, the action taken and some stochastic exogenous parameter. The global state is composed of the state of the storage device and the state of the CID market. The storage state contains the estimated state of charge of the storage device for each product/time slot at each time-step in the trading process. As CID market state we define the orders available for each product at time step. Finally, the actions available are to either accept or not each order in the order book. The rewards coming from a transition are considered as the revenues collected from the transactions made.

Due to the continuously changing size and the high dimensionality of the problem an equivalent but reduced state representation is used similar to [2]. We motivate the use of high-level actions that map into the original action space and are able to extend the actions proposed in [2]. Thus, additionally to either "Idle" (i.e. accept no order) or "Optimize based on current knowledge" (i.e. run an order acceptance optimization model) the agent can also open several positions in the market. We show that introducing an auxiliary variable it is possible to abstract these three actions. An imbalance penalty is used in order to penalize deviations of the actual production/consumption from the contracted schedule.

An approximate policy iteration method is used to solve the problem specified. The goal is the derivation of a policy that according to the state of the market and the storage device can take positions in the market in order to maximize the accumulated profits and the efficient use of RES. Moreover, the analysis of the policy illustrates the adoption of different strategies according to different parameters such as the time distance to physical delivery etc.

Results

A case study is used to evaluate the results of the proposed methodology. The participation of a pumped-hydro storage unit in the CID market is considered. The technical characteristics of the storage device such as, technical minimum and maximum output, capacity and roundtrip efficiency are explicitly taken into account. The CID market is simulated using historical order data from the CID market. The performance of the policy trained on a part of the data, is evaluated on a different dataset under two metrics. The total expected rewards and the total risk of the policy with respect to the "rolling intrinsic" are compared. The effect of the policy on the operation of the storage device is illustrated.

Conclusions

The strategic participation of a storage device operator in the Continuous Intraday market is investigated in this paper. The problem is formulated as a Markov Decision Process and due its complex nature a reduced state representation is used. A new action space is proposed that succeeds to generalize the possible actions and to capture the different behaviours of a trading agent. An optimal policy is derived using Approximate Dynamic Programming. The optimal policy consists in identifying the optimal action that should be taken based on the state that is observed. The proposed methodology is evaluated on a test case where the trained policy is compared to the benchmark. A comparison based on the profitability and the risk of the trading strategies is taking place. Results are produced regarding the evolution of the state of charge depending on the market state. Future work should be directed in the enlargement of the trading horizon and the products that are considered.

References

[1] I. Pérez Arriaga and C. Knittel et al, *Utility of the Future. An MIT Energy Initiative response*. 2016.

[2] I. Boukas, D. Ernst, A. Papavasiliou, and B. Cornelusse, "Intra-day Bidding Strategies for Storage Devices Using Deep Reinforcement Learning," vol. 14, no. October, 2018.