Reinforcement Learning with continuous agent that charges/discharges BESSs

Keywords: Reinforcement Learning (RL), Distributed Energy Resources (DERs), Battery Energy Storage Systems (BESSs)

Supervisor: Xavier Weiss (xavierw@kth.se)

Project description
The ongoing shift towards decentralized energy production, consumption, and storage introduces new operational complexity to operators. However, the presence of these Distributed Energy Resources (DERs) can also be seen as an opportunity to explore novel controls strategies. In this project you will play the role of the System Operator (SO) and commandeer Battery Energy Storage Systems (BESSs) from their owners during extreme events. Your goal is to train a Reinforcement Learning (RL) agent to minimize the Energy Not Supplied (ENS) of the loads as much as possible and prevent a total blackout by charging / discharging these BESSs.

As part of this project, we will provide the scaffolding needed to get started with Grid2Op (the RL library of choice for power system operations), as well as a tested agent implementation that can work with a continuous action space. You will need to figure out how to convert observations of the environment into a form the agent can work with, how to convert the agent's output into a form that Grid2Op can understand, and how to aid the agent so that it can learn an effective policy.

Objective:
Deliver an agent that is able to reduce the ENS compared to a passive (do nothing) baseline on previously unseen data.

References:
Reflection on winners of the L2RPN competition: Marot, Antoine, et al. "Learning to run a power network challenge: a retrospective analysis." *NeurIPS 2020 Competition and Demonstration Track*. PMLR, 2021. Available: https://proceedings.mlr.press/v133/marot21a.html
Article where Discrete Q Network was first introduced: "Playing atari with deep reinforcement learning," V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, 2013. [Online]. Available: https://arxiv.org/abs/1312.5602
Article where Deep Deterministic Policy Gradient (DDPG) was first introduced: "Deterministic policy gradient algorithms," D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, in International conference on machine learning. Pmlr, 2014, pp. 387–395
Previous working using PPO in Grid2Op: "Safe Deep Reinforcement Learning for Power System Operation under Scheduled Unavailability" X. Weiss, S. Mohammadi, P. Khanna, M.R.Hesamzadeh, L. Nordström. IEEE Power & Energy Society General Meeting (PESGM) 2023. doi: 10.1109/PESGM52003.2023.10252619

Useful data-sets:
The time series (load and generation profiles) used by Grid2Op are generated using Chronix2Grid: https://github.com/Grid2op/chronix2grid. You will not need to generate new profiles (unless you want to of course).