Reinforcement Learning-Based Battery Energy Management in a Solar Microgrid

Brida Mbuwir
KU Leuven/EnergyVille
brida.mbuwir@vito.be

Frederik Ruelens
KU Leuven/EnergyVille
frederik.ruelens@esat.kuleuven.be

Fred Spiessens
VITO/EnergyVille
fred.spiessens@vito.be

Geert Deconinck
KU Leuven/EnergyVille
geert.deconinck@esat.kuleuven.be

***Abstract*—The intermittent nature of Renewable Energy Sources (RES) leads to a mismatch between electricity supply and demand, thus, there is a need for energy storage and load management. This paper presents a framework using Reinforcement Learning (RL) to control the operation of a battery storage device in a microgrid. Here, the microgrid considered consists of a photovoltaic (PV) system, inverters, residential consumer and battery storage. The optimal operation of the battery is formulated as a Markov decision process. A deterministic setting is considered with the weather forecast for PV production and electricity consumption known in advance. The agent learns an optimal energy management policy by using its past experiences. The developed solution was tested in a Belgian case with local load and PV production profiles of residential consumers.**

Index terms - microgrid, reinforcement learning, battery, PV, storage.

I. Introduction

The rising concerns about global warming and high penetration of RES have introduced new challenges in the design and operation of power systems [1]. These challenges have led to the smart grid paradigm which has triggered technological advancement towards a green, intelligent and more efficient power grid [2]. It is expected that the future power grid will be a combination of multiple microgrids collaborating with each other [3] [4]. Microgrids are electrical systems consisting of loads, energy storage facilities and RES that can operate in parallel with or disconnected from the main utility grid [5]. Over the years, microgrids have witnessed a great development due to the drop in the cost of PV systems [4]. Due to the increased share of the intermittent RES, microgrids require energy storage facilities and smart load management in order to balance demand and supply.

In the recent years, studies on energy management in microgrids, agent-based methods and RL techniques for developing smart grid applications have gained popularity [6] [7]. Q-learning, a popular RL method has been applied in demand response to minimize the energy consumption of an electric water heater [7] and in the control of energy storage devices in a microgrid [8]. The method allows to find a set of best actions at any state of operation of the system considered. The evaluation of the actions taken in the learning process is based on the rewards or penalties received from the environment [9]. The authors of [8] [10] present the application of Q-learning to model the interactions between different components of a microgrid, with objective to obtain an optimal scheduling plan for energy storage devices. Kuznetsova *et al.* [10] consider a weighted amount of electricity the battery can discharge/charge as the reward to get an optimal battery scheduling plan for a microgrid. However, it remains interesting to explore the depth of RL techniques in a case where the amount of electricity bought or sold from/to the grid is considered as the reward function.

Motivated by the works of [8] and [10], this paper is formulated to build on the existing literature on RL and promote the use of Q-learning in energy management in microgrids. As such, an intelligent decision-maker (agent) using Q-learning to minimize the amount of electricity bought/sold from/to the grid is designed in this framework. The agent develops an optimal battery scheduling strategy which controls the charging and discharging actions of the battery in a continuously changing environment (electricity production and demand).

II. Problem Formulation

This study considers a microgid with residential consumer, PV system, inverters and battery storage facility. The microgrid is connected to the main utility grid through a transformer. The consumer can buy electricity from the grid in case the PV production and energy from the battery are not sufficient. It is assumed that the battery should not be charged using electricity from the main grid.

A value iteration algorithm [9] is considered in the determination of an optimal charging (a^c) and discharging (a^d) action sequence for the battery. The case study does not take into account the technical constraints of the main utility grid and the microgrid. The operational planning of the microgrid is the main focus.

The operational planning of the microgrid is formulated as a sequential decision-making problem using the Markov decision process with state space $\mathcal{S}$, action space $\mathcal{A}$, dynamics and cost/reward function ρ. Inspired by [5] and following the notation of its authors, $\mathcal{S}$, $\mathcal{A}$, the dynamics and ρ are described below. The system evolution is described as a discrete-time process with a time step of 15 minutes over a finite time horizon.

A. State space ($\mathcal{S}$)

Let $s \in \mathcal{S}$ denote a time-dependent vector characterizing the microgrid's state. The state space consists of four components:

- Timing component (S^t): information on the time period relevant for the dynamics of the system - quarter of the day (s^q), day of the week (s^d) and season of the year (s^s),
- PV component (S^{PV}): energy produced by PV panels,
- Battery component (S^B): energy level of the battery,
- Energy demand component (S^L): load.

Thus, the microgrid's state is defined by the vector

$$s = (s^q, s^d, s^s, E^{PV}, E^{bat}, E^l) \in \mathcal{S}$$

B. Action space ($\mathcal{A}$)

The action space consists of charging a^c and discharging a^d actions of the battery, $\mathcal{A} = \{a^c, a^d\}$. The actions are discretized in amounts of energy in [kWh]. The actions are constrained by (1) ensuring the battery is not charged and discharged at the same time.

$$a^c a^d = 0 \tag{1}$$

C. Dynamics

The dynamics of the battery at every time step are presented in (2) constrained by the battery capacity E^{bat}_{max} and efficiency η.

$$E^{bat}_{t+1} = E^{bat}_t + \eta a^c - \frac{a^d}{\eta} \tag{2}$$

$$0 \leq E^{bat}_t \leq E^{bat}_{max} \tag{3}$$

D. Cost/reward function

At time step $t \in (S^t)$, let E^{grid}_t represent the electricity that can be bought or sold from or to the grid, c and p the price of buying and selling electricity to the grid respectively. It is considered that $c > p$ to encourage self consumption. If at time step $t \in (S^t)$ the energy from the PV and battery is insufficient to meet the load, an equivalent amount is bought from the grid as given in (4).

$$E^{grid}_t = E^l_t - E^{PV}_t - E^{bat}_t \tag{4}$$

The cost/reward function ρ at time step t is given by (5).

$$\rho(s_t, a_t, s_{t+1}) = max(0, E^{grid}_t)c + min(0, E^{grid}_t)p \tag{5}$$

An optimization problem with objective to minimize the amount of energy bought or sold from/to the grid as defined in (6).

$$\begin{aligned} \underset{a \in \mathcal{A}, s \in \mathcal{S}}{\text{minimize}} \quad & \sum \rho(s_t, a_t, s_{t+1}) \\ \text{subject to} \quad & (3), (1), (2), \forall t \in (S^t) \end{aligned} \tag{6}$$

III. Expected results

The microgrid in this study consists of a battery of capacity $40kWh$ with $\eta = 90\%$, 5 residential consumers with a combined peak load of $8.47kW$ and 5 PV installations with a combined peak production of $7.79kW$. The expected optimal policy of the agent is as shown in Figure 1. The storage device builds up its reserve during peak production (during the day) in order to meet the load (during less sunny periods and peak demand). Waste of excess energy or curtailment (when battery is fully charged and demand is low) is avoided by selling the excess energy to the grid.

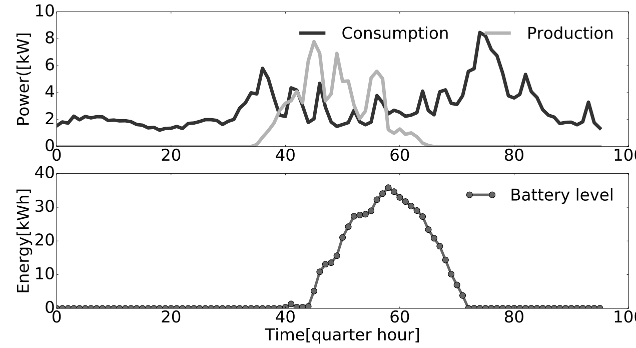

Fig. 1. Optimal policy in a deterministic case on a winter day. Top plot shows the load and PV profiles, bottom plot depicts the battery energy level.

IV. Conclusion

This abstract has discussed the use of RL to efficiently schedule the operation of a battery in a microgrid in a fully deterministic environment. The full paper will present a stochastic scenario where the energy demand and PV production are not known in advance, with the simulations extended over multiple days. Future works will investigate the effect of dynamic electricity pricing and a multi agent scenario where the storage, consumer and utility grid are considered as individual agents interacting with each other. The results of this study will be of interest to companies like ABB, Schneider Electric and Enervalis providing microgrid solutions for: a better integration of RES, an increase in energy reliability, and ensuring grid stability.

References

[1] N. I. Voropai and D. N. Efimov, "Operation and control problems of power systems with distributed generation," in *2009 IEEE Power Energy Society General Meeting*, July 2009, pp. 1–5.

[2] M. Fadaeenejad, A. M. Saberian, M. Fadaee, M. Radzi, H. Hizam, and M. AbKadir, "The present and future of smart power grid in developing countries," *Renewable and Sustainable Energy Reviews*, vol. 29, pp. 828–834, 2014.

[3] B. Zhao, M. Xue, X. Zhang, C. Wang, and J. Zhao, "An MAS based energy management system for a stand-alone microgrid at high altitude," *Applied Energy*, vol. 143, pp. 251–261, 2015.

[4] S. Bacha, D. Picault, B. Burger, I. Etxeberria-Otadui, and J. Martins, "Photovoltaics in microgrids: An overview of grid integration and energy management aspects," *IEEE Industrial Electronics Magazine*, vol. 9, no. 1, pp. 33–46, March 2015.

[5] V. François-Lavet, Q. Gemine, D. Ernst, and R. Fonteneau, "Towards the minimization of the levelized energy costs of microgrids using both long-term and short-term storage devices," *Smart Grid: Networking, Data Management, and Business Models*, pp. 295–319, 2016.

[6] V. N. Coelho, M. W. Cohen, I. M. Coelho, N. Liu, and F. G. Guimarães, "Multi-agent systems applied for energy systems integration: State-of-the-art applications and trends in microgrids," *Applied Energy*, vol. 187, pp. 820–832, 2017.

[7] F. Ruelens, B. Claessens, S. Quaiyum, B. De Schutter, R. Babuska, and R. Belmans, "Reinforcement learning applied to an electric water heater: From theory to practice," *arXiv preprint arXiv:1512.00408*, 2015.

[8] V. François-Lavet, D. Taralla, D. Ernst, and R. Fonteneau, "Deep reinforcement learning solutions for energy microgrids management," in *European Workshop on Reinforcement Learning (EWRL 2016)*, 2016.

[9] L. Busoniu, R. Babuška, B. De Schutter, and D. Ernst, *Reinforcement Learning and Dynamic Programming Using Function Approximators*. Boca Raton, FL: CRC Press, 2010.

[10] E. Kuznetsova, Y.-F. Li, C. Ruiz, E. Zio, G. Ault, and K. Bell, "Reinforcement learning for microgrid energy management," *Energy*, vol. 59, pp. 133–146, 2013.