Unlocking the Flexibility of District Heating Pipeline Energy Storage with Reinforcement Learning

Stepanovic, K.; Wu, Jichen; Everhardt, Rob; de Weerdt, M.M.

Publication date
2022

Document Version
Final published version

Citation (APA)
Stepanovic, K., Wu, J., Everhardt, R., & de Weerdt, M. M. (2022). *Unlocking the Flexibility of District Heating Pipeline Energy Storage with Reinforcement Learning*. Abstract from 34th Benelux Conference on Artificial Intelligence (BNAIC) and the 30th Belgian Dutch Conference on Machine Learning (Benelearn), Mechelen, Belgium.

Unlocking the Flexibility of District Heating Pipeline Energy Storage with Reinforcement Learning *

Ksenija Stepanovic[1 [0000−0003−4273−4785]], Jichen Wu[1,2], Rob Everhardt[2], and Mathijs de Weerdt[1 [0000−0002−0470−6241]]

[1] Delft University of Technology, The Netherlands
[2] Flex Technologies, The Netherlands

1 Motivation and Approach

Energy storage systems are crucial for providing flexibility to the power systems [2]. An important type of energy storage can be found in pipeline energy storage, which originates from the thermal inertia of the district heating network (DHN) [11]. Utilization of the pipeline energy storage decouples heat and electricity production of a combined heat and power (CHP) production plant [4]. Decoupling of heat and electricity enables a CHP operator to achieve profit gain by trading electricity with an external grid at favourable moments.

To mitigate stability, feasibility, and computational complexity challenges of previous approaches in the control of the combination of a CHP and a DHN [6, 7, 12, 3, 4], we develop a control strategy that relies on reinforcement learning. The primary contributions of our work are: (1) we model the CHP economic dispatch with DHN pipeline storage as a Markov decision process, (2) we empirically evaluate choices for state space and reward signal construction, and (3) we demonstrate the effectiveness of our approach when compared with an optimal linear model without grid dynamics (LP) [1], and MINLP, an exhaustive mixed-integer nonlinear program [6]. In this paper we introduce a Q-learning algorithm for the DHN control by defining the action space, state space and reward signal [10, 5, 8]. We present our RL framework in Figure 1.

2 Results and Discussion

We aim to understand the strengths and weaknesses of Q-learning compared to previous approaches such as LP and MINLP, in particular in dealing with the environment and model uncertainty under limited sensor information. We evaluate these on four criteria: performance, feasibility, stability, and time-flexibility. In Figure 2, we observe that both Q^{full} and MINLP successfully exploit fluctuations in electricity price for profit gain by producing more heat when electricity price is low, and more electricity when electricity price is high. Compared to an optimal LP without grid dynamics, Q^{full} achieves a profit gain in the order of magnitude €10^2, while MINLP achieves €10^3. However, as Q^{full} successfully approximates the complex dynamics of a DHN, it results in a fewer constraint violations than MINLP.

* This extended abstract is based on a paper published at journal Energies [9].

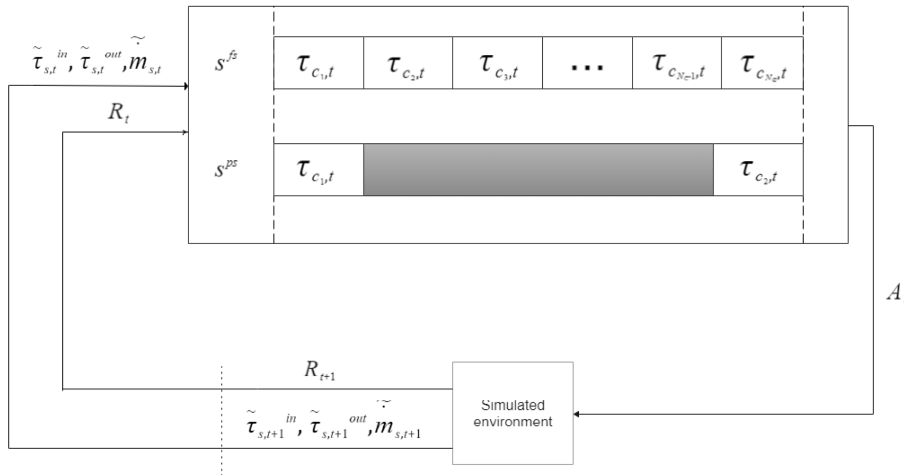
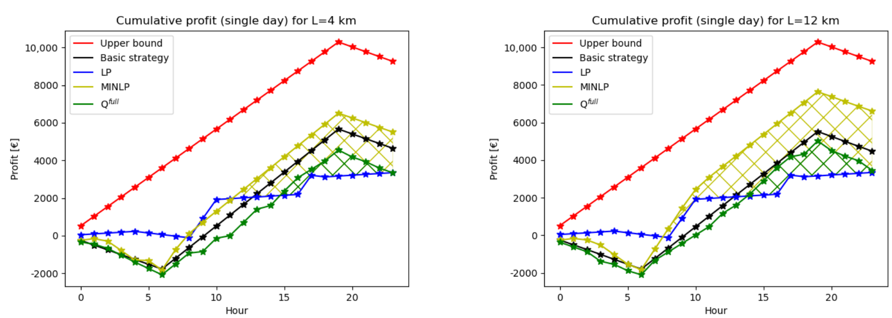


Fig. 1. The reinforcement learning agent–simulation environment interaction. Based on the RL agent's action at the time-step t, A_t, the simulation environment at the time-step $t+1$ outputs observations of the supply network inlet temperature $\widetilde{\tau}_{s,t+1}^{in}$, supply network outlet temperature $\widetilde{\tau}_{s,t+1}^{out}$ and mass flow $\widetilde{m}_{s,t+1}$, and reward R_{t+1}. These observations form a full s^{fs} and partial state s^{ps} Q-learning, Q^{full} and Q^{par}.



(**a**) (**b**)

Fig. 2. (**a**) The single-day cumulative profit for $L = 4$ km. (**b**) The single-day cumulative profit for $L = 12$ km.

The MINLP does not find a primal bound of the model for 120 days for pipe length 4 km and 113 days for pipe length 12 km. In contrast to the MINLP, the Q^{full} provides a stable output for all evaluation days. The obtained results show that Q-learning requires more computational power during algorithm training but has better time scale flexibility compared to MINLP (provides the response on unseen scenarios in a few seconds). Therefore, it is more suitable for real-time energy markets.

References

1. Abdollahi, E., Wang, H., Rinne, S., Lahdelma, R.: Optimization of energy production of a CHP plant with heat storage. In: 2014 IEEE Green Energy and Systems Conference (IGESC). pp. 30–34. IEEE (2014)
2. Averfalk, H., Werner, S.: Economic benefits of fourth generation district heating. Energy **193**, 116727 (2020)
3. Grosswindhager, S., Voigt, A., Kozek, M.: Predictive control of district heating network using fuzzy DMC. In: 2012 Proceedings of International Conference on Modelling, Identification and Control. pp. 241–246. IEEE (2012)
4. Guelpa, E., Verda, V.: Thermal energy storage in district heating and cooling systems: A review. Applied Energy **252**, 113474 (2019)
5. Hausknecht, M., Stone, P.: Deep recurrent Q-learning for partially observable MDPs. In: 2015 AAAI fall symposium series (2015)
6. Li, Z., Wu, W., Shahidehpour, M., Wang, J., Zhang, B.: Combined heat and power dispatch considering pipeline energy storage of district heating network. IEEE Transactions on Sustainable Energy **7**(1), 12–22 (2015)
7. Runvik, H., Larsson, P.O., Velut, S., Funkquist, J., Bohlin, M., Nilsson, A., Razavi, S.M.: Production planning for distributed district heating networks with jmodelica. org. In: Proceedings of the 11th International Modelica Conference, Versailles, France, September 21-23, 2015. pp. 217–223. Linköping University Electronic Press (2015)
8. Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., et al.: A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. Science **362**(6419), 1140–1144 (2018)
9. Stepanovic, K., Wu, J., Everhardt, R., de Weerdt, M.: Unlocking the flexibility of district heating pipeline energy storage with reinforcement learning. Energies **15**(9), 3290 (2022)
10. Sutton, R.S., Barto, A.G.: Reinforcement learning: An introduction. MIT press (2018)
11. Vandermeulen, A., van der Heijde, B., Helsen, L.: Controlling district heating and cooling networks to unlock flexibility: A review. Energy **151**, 103–115 (2018)
12. Verrilli, F., Srinivasan, S., Gambino, G., Canelli, M., Himanka, M., Del Vecchio, C., Sasso, M., Glielmo, L.: Model predictive control-based optimal operations of district heating system with thermal energy storage and flexible loads. IEEE Transactions on Automation Science and Engineering **14**(2), 547–557 (2016)