Coordination of Distributed Energy Resources without Network Models using Reinforcement Learning

Thomas Navidi (tnavidi@stanford.edu)

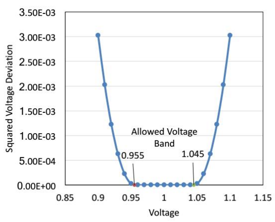
Problem Statement

Goal: Coordinate the operation of distributed storage units in a distribution grid for the purposes of minimizing costs and promoting network reliability

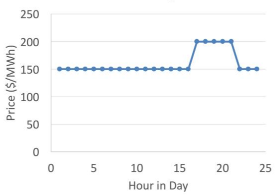
Due to distributed renewable generation such as rooftop PV, distribution grid voltages can deviate from desired operating bounds. However, the coordinated usage of storage units alongside these renewables can prevent voltage issues. Unfortunately, it is not possible to control them without knowledge of the distribution grid models. This research aims to learn a suitable control strategy through reinforcement learning without grid models.

Objective

The first objective is to minimize voltage deviations. The penalty for this is shown on the right.



The second objective is to minimize the cost of electricity. The right shows the price of electricity over time. In order to minimize this we arbitrage by buying at a low price and selling at the high price



$$\frac{1}{2}\sum_{\tau=1}^{24} p_\tau \cdot \Re(s_{0\tau})+150\sum_{i=0}^{N}\sum_{\tau=1}^{24}(\max(v_{i\tau}-V_{\mathrm{tol}+},0)+\max(V_{\mathrm{tol}-}-v_{i\tau},0))^2$$

Daily Reward Equation

Data and Models

Actor Network: Input is hourly power consumption for all nodes in a day. Output is hourly charge power for each storage unit. Dense network with 2 hidden layers of size 100 and 50. ELU activation function and tanh for output layer.

Data Problems: Acquiring training data is hard since 1 point is 1 days worth of data collecting. Labeling the data with storage charging action is hard since the optimization algorithm is too slow to label thousands of points.

Solution: Augment existing data (150 days) with random noise and run a heuristic that achieves near optimal performance quickly for labels. Train the network to mimic the heuristic on augmented data. Use real data for dev and test sets. Later use reinforcement learning to push performance closer to optimal.

Heuristic Training Performance

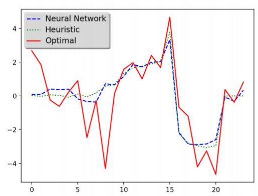
Neural Network Training Loss

Set	Mean Squared Error
Train	5.702E-4
Dev	2.154E-4
Test	1.99E-4

Actor Performance for Storage Control

Method	Arbitrage Profit	Voltage Violations	Total Performance
Neural Net	209.414	0.864	-24.893
Heuristic Control	209.992	0.878	-26.704
Optimal Control	209.577	0.295	60.539

Sample charge profile showing the neural net closely follows the heuristic. X-axis is hours and Y-axis is power.



DDPG Implementation

Critic Network: Input is same as actor input, but also includes actor output. Output is a scalar value for that state and action pair. Dense network with 3 hidden layers of size 100, 50, and 25. First layer is shared with actor network.

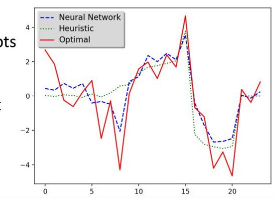
Exploration: Random normal noise with momentum is added to actor output.

Training: Critic trains on several random instances to develop sense of value of state action pairs. Then actor is trained on gradient of critic to push actor output closer to maximizing critic output

Results

Method	Arbitrage Profit	Voltage Violations	Total Performance
DDPG	175.000	0.449	20.035

After many random attempts the network learns a new output to avoid some violations (see discharge at hour 8)



Conclusion and Future Work

In conclusion, DDPG was able to nudge the heuristic controller slightly in the right direction to achieve overall better performance than before. Future work will be to scale this problem to a larger network with multiple storage units. This will require significantly more training data to be collected. Other work includes trying more intelligent exploration policies like modifications based on the time period the violation takes place and whether it is an over-voltage or under-voltage.

- K. Anderson, R. Rajagopal, A. El Gamal, "Coordination of Distributed Storage Under Temporal and Spatial Data Asymmetry," submitted to *IEEE Trans. On Smart Grid*.
- T. P. Lillicrap, J. J. Hunt, A. Pritzel, et. al. "Continuous Control with Deep Reinforcement Learning." Int. Conf. Learning Representations, 2016.