MuZero Based Online Optimization Strategy for a Residential Microgrid

Hang Shuai

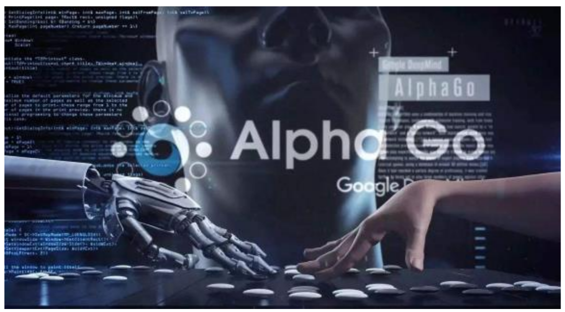
PostDoc Research Associate

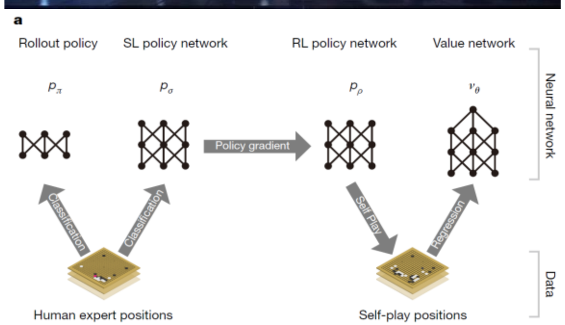
University of Tennessee, Knoxville

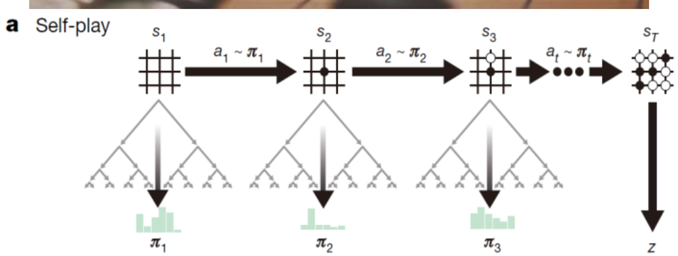
12/27/2020

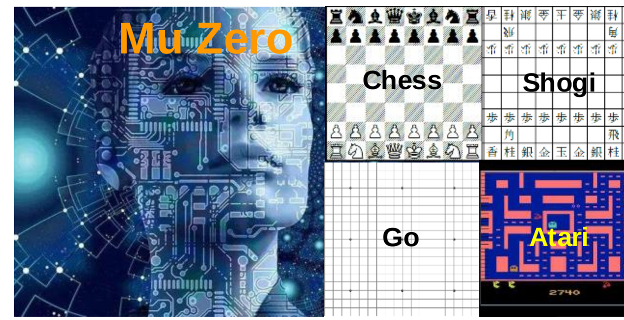
Motivation

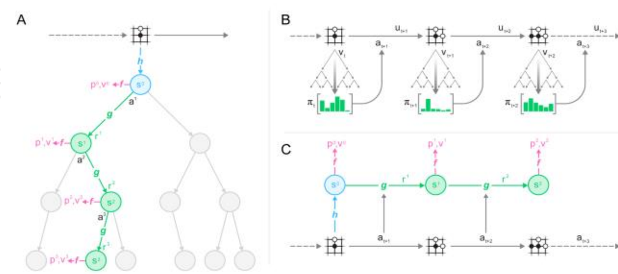
- **Pain points:** Traditional online scheduling strategies are greatly **affected by** the accuracy of renewable energy and load forecasting.
- **Our work:** Develop **the MuZero based optimization algorithm** to learn to operate the system from historical load and renewable energy data without relying on any forecast information.











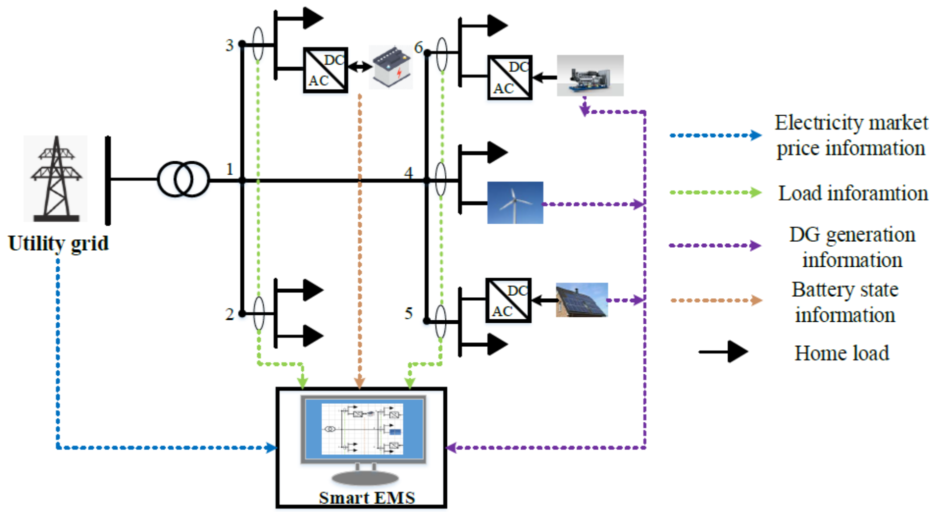
Alpha Go
✓ Supervised learning
✓ Environment model

AlphaGo Zero/Alpha Zero
× Supervised learning
✓ Environment model

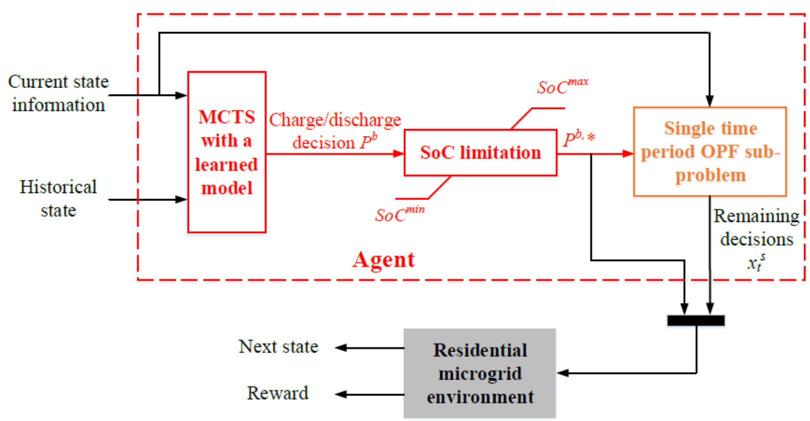
MuZero
× Supervised learning
× Environment model

Microgrid Optimization Model

☐ Online optimization model of microgrid



☐ MuZero based agent outputs the charge/discharge power decision， the remaining decisions are determined by solving the OPF subproblem



The optimization goal is to minimize the operational cost of the microgrid, considering the following constraints:

$$P_g^{DG,min} \le P_g^{DG}(t) \le P_g^{DG,max}, \forall t \in \Gamma, \forall g \in \mathcal{G} \quad (6)$$

$$(P_g^{DG}(t))^2 + (Q_g^{DG}(t))^2 \le (S_g^{DG,max})^2, \forall t \in \Gamma, \forall g \in \mathcal{G} \quad (7)$$

$$0 \le P^{wt}(t) \le \bar{P}^{wt}(t), \forall t \in \Gamma \quad (8)$$

$$(P^{wt}(t))^2 + (Q^{wt}(t))^2 \le (S^{wt,max})^2, \forall t \in \Gamma \quad (9)$$

$$0 \le P^{pv}(t) \le \bar{P}^{pv}(t), \forall t \in \Gamma \quad (10)$$

$$(P^{pv}(t))^2 + (Q^{pv}(t))^2 \le (S^{pv,max})^2, \forall t \in \Gamma \quad (11)$$

$$\begin{cases} 0 \le P_{buy}^{grid}(t) \le P_{buy}^{grid,max} \\ 0 \le P_{sell}^{grid}(t) \le P_{sell}^{grid,max} \end{cases} \forall t \in \Gamma \quad (12)$$

$$0 \le Q^{grid}(t) \le Q^{grid,max}, \forall t \in \Gamma \quad (13)$$

$$\begin{cases} 0 \le P^{ch}(t) \le I^{ch}(t) P^{ch,max} \\ 0 \le P^{dis}(t) \le I^{dis}(t) P^{dis,max} \end{cases} \forall t \in \Gamma \quad (14)$$

$$P^b(t) = I^{dis}(t) P^{dis}(t) - I^{ch}(t) P^{ch}(t), \forall t \in \Gamma \quad (15)$$

$$I^{dis}(t) + I^{ch}(t) \le 1, \forall t \in \Gamma, \{I^{dis}(t), I^{ch}(t)\} \in \{0,1\} \quad (16)$$

$$(P^b(t))^2 + (Q^b(t))^2 \le (S^{b,max})^2, \forall t \in \Gamma \quad (17)$$

$$SoC(t+\Delta t) = SoC(t) + \eta^{ch} \frac{P^{ch}(t)}{E^{max}} \Delta t - \frac{1}{\eta^{dis}} \frac{P^{dis}(t)}{E^{max}} \Delta t \quad (18)$$

$$SoC^{min} \le SoC(t) \le SoC^{max}, \forall t \in \Gamma \quad (19)$$

$$\begin{cases} P_j(t) = P_{ij}(t) - r_{ij} l_{ij}(t) - \sum_{m:(j,m)\in\Upsilon} P_{jm}(t) \\ Q_j(t) = Q_{ij}(t) - x_{ij} l_{ij}(t) - \sum_{m:(j,m)\in\Upsilon} Q_{jm}(t) \end{cases} \forall (i,j) \in \Upsilon, \forall t \in \Gamma \quad (20)$$

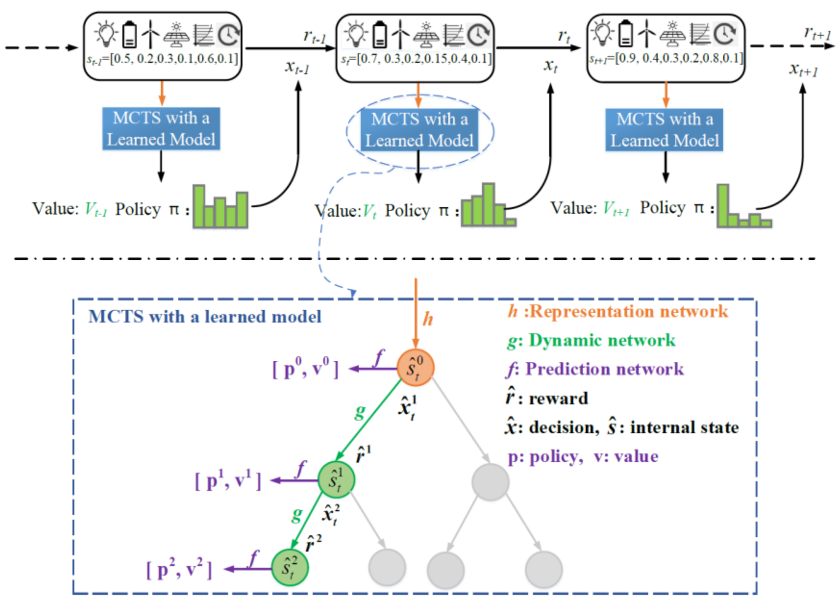
$$v_j(t) = v_i(t) - 2(r_{ij} P_{ij}(t) + x_{ij} Q_{ij}(t)) + (r_{ij}^2 + x_{ij}^2) l_{ij}(t), \forall t \in \Gamma \quad (21)$$

$$V_i^{min} \le |V_i(t)| \le V_i^{max}, \forall t \in \Gamma, \forall i \in \mathcal{N} \quad (22)$$

$$l_{ij}(t) = \frac{P_{ij}(t)^2 + Q_{ij}(t)^2}{v_i(t)}, \forall (i,j) \in \Upsilon, \forall t \in \Gamma \quad (23)$$

Microgrid Optimization Model

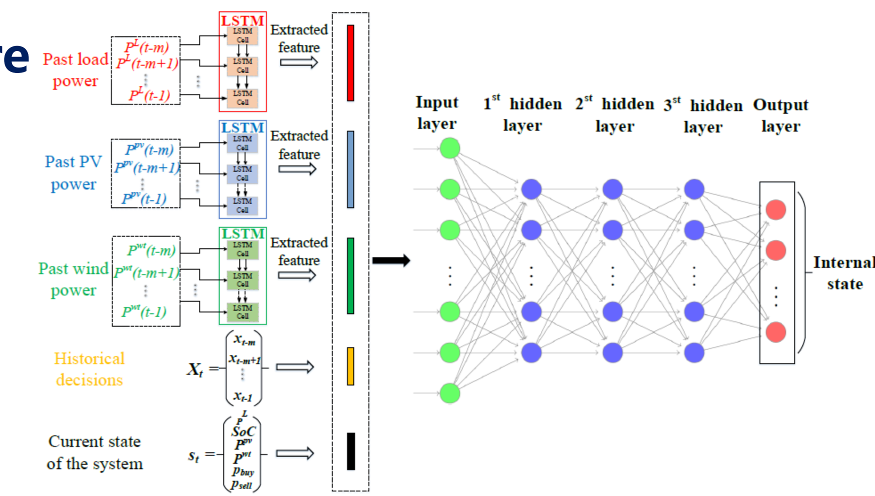
◻ MuZero based optimization algorithm



- **Self-play：** MCTS uses the learned neural network model obtained from the previous iteration to build a tree and generates the training data；
- **Training：** According to the generated data, the neural network parameters are constantly updated.

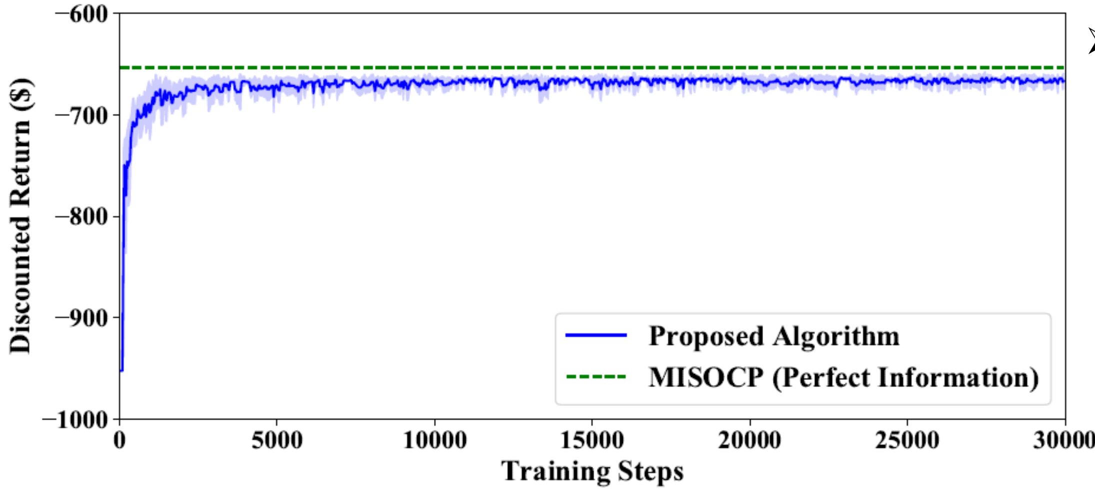
◻ Representation network architecture

- LSTM units extract features from the historical load and renewable power data, and we generate the "internal state" together with the current state information using a fully connected neural network.



Simulation Results

- As the training time increases, the optimal operation cost obtained by the MuZero agent is gradually approaching the optimal value



- The “optimal value" is calculated offline using the MISOCP (Mixed Integer Second-Order Cone Programming) method under the assumption that the prediction information is accurate. The optimal value is only used as a reference.

- The microgrid online scheduling test for 100 consecutive days, and MuZero algorithm performs best

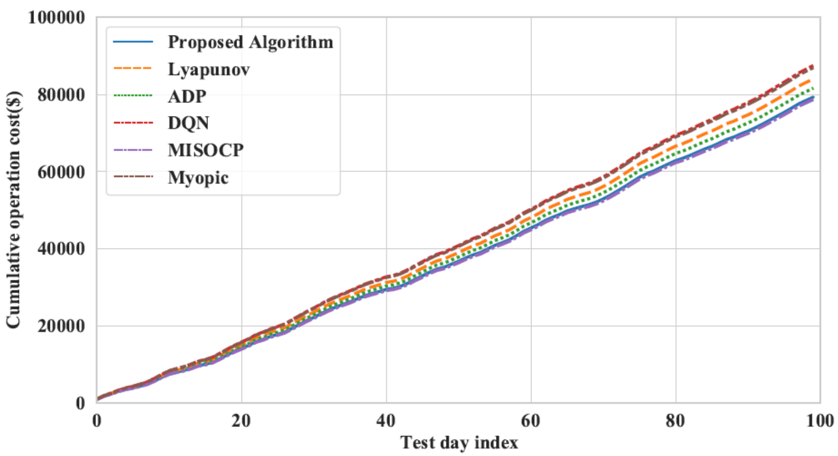


TABLE IV
THE PERFORMANCE IMPROVEMENT OF DIFFERENT OPTIMIZATION METHODS COMPARED TO MYOPIC POLICY ON THE TESTING DATA SET.

Performance improvement		*Mean*	*Maximum*	*Minimum*	*Standard deviation*
Online methods	Proposed algorithm	**9.30%**	16.68%	5.28%	2.12%
	Lyapunov optimization	3.76%	9.89%	1.93%	1.65%
	ADP	6.57%	14.78%	4.16%	1.92%
	DQN	-0.65%	-0.14%	-3.45%	0.77%
Off-line method	MISOCP	10.20%	23.28%	6.45%	3.02%

Reference

H. Shuai, and H. B. He. Online Scheduling of a Residential Microgrid via Monte-Carlo Tree Search and a Learned Model," *IEEE Transactions on Smart Grid*. DOI: 10.1109/TSG.2020.3035127. (Early Access)