

# 基于MuZero的微电网在线调度策略

帅航

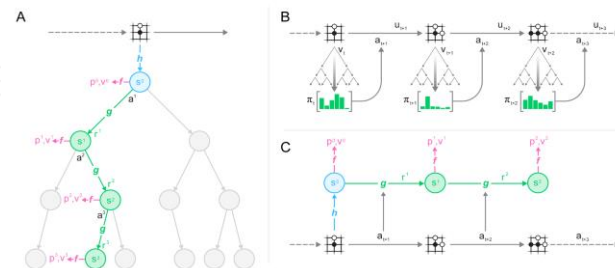
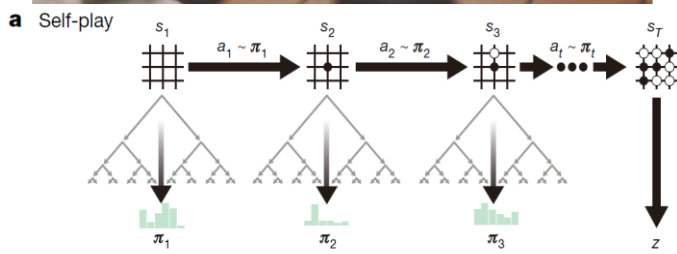
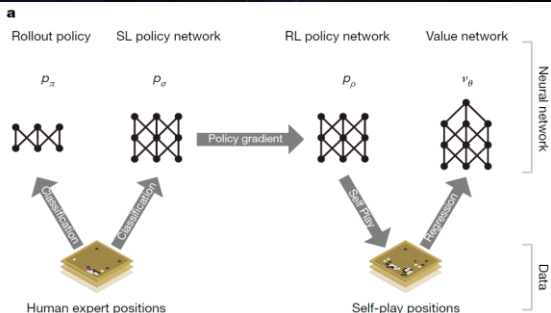
田纳西大学博士后研究员

hangshuai523@gmail.com

2020年12月

# 研究背景

- 针对问题：传统在线调度策略受可再生能源与负荷短期**预测精度**影响大
- 解决方案：应用**深度强化学习**，从历史负荷及可再生能源功率数据中**自主**  
**学习**不依赖于预测的调度策略，应对调度中可再生能源等带来的不确定性



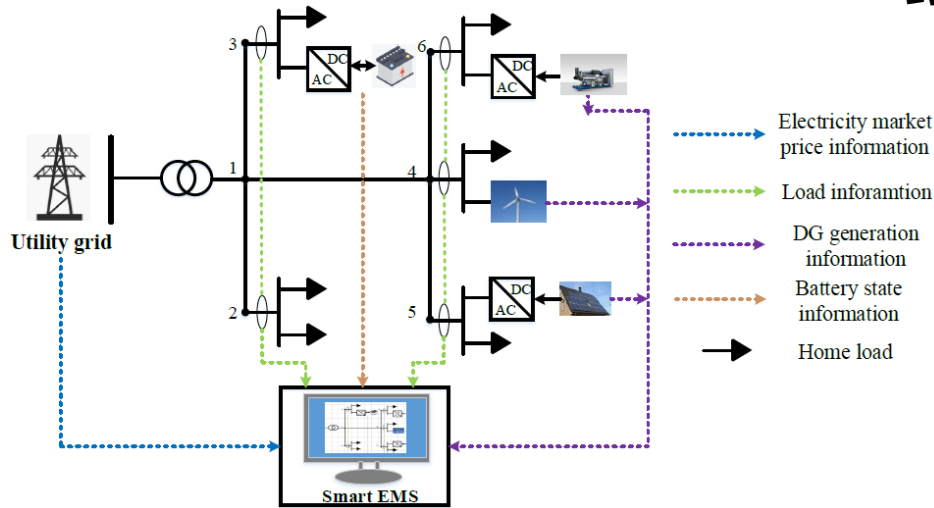
Alpha Go  
需要监督学习  
需要环境动态模型

AlphaGo Zero/Alpha Zero  
无需监督学习  
需要环境动态模型

MuZero  
无需监督学习  
无需环境动态模型

# 微电网在线调度建模

## □ 微网在线调度模型



优化目标为**最小化微网运行费用**，考虑以下约束条件：

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

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

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

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

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

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

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

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

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

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

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

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

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

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

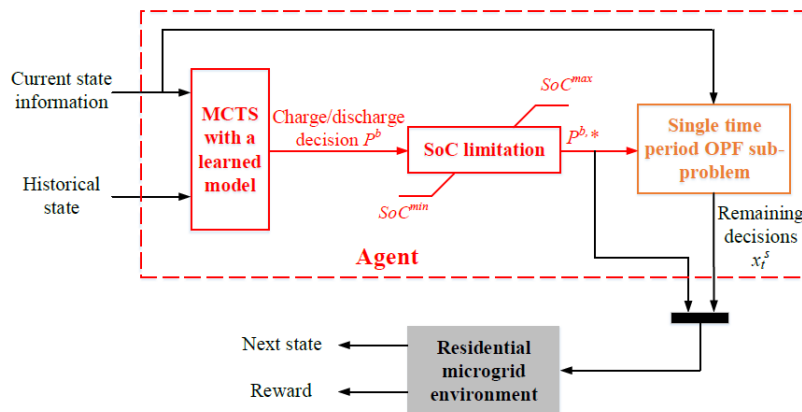
$$\begin{cases} P_{ij}(t) = P_{ij}(t) - r_{ij} l_{ij}(t) - \sum_{m:(j,m) \in \Upsilon} P_{jm}(t) \\ Q_{ij}(t) = Q_{ij}(t) - x_{ij} l_{ij}(t) - \sum_{m:(j,m) \in \Upsilon} Q_{jm}(t) \end{cases} \forall (i, j) \in \Upsilon, \forall t \in \Gamma \quad (20)$$

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

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

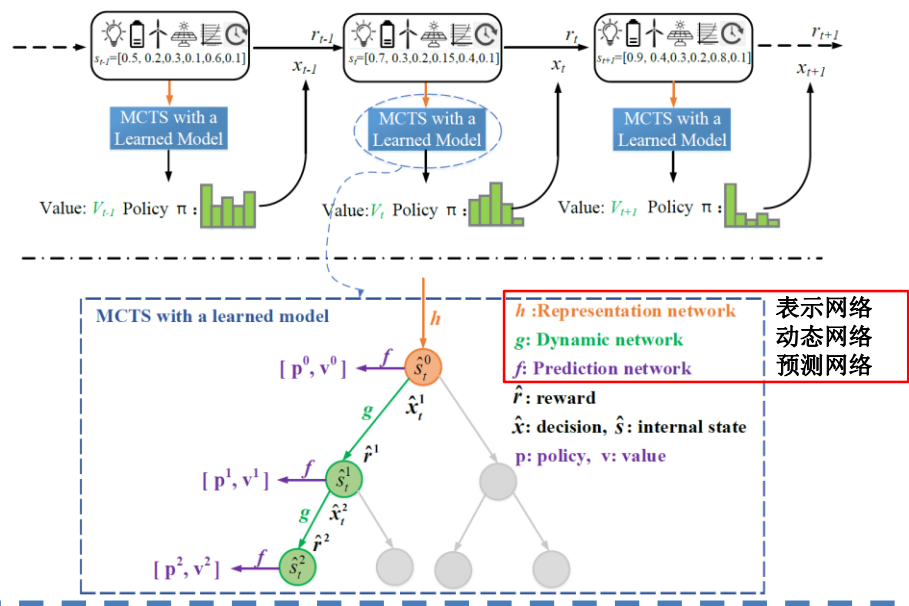
$$l_{ij}(t) = \frac{P_{ij}(t)^2 + Q_{ij}(t)^2}{v_i(t)}, \forall (i, j) \in \Upsilon, \forall t \in \Gamma \quad (23)$$

## □ MuZero智能体决策电池充放电功率，剩余的决策变量通过求解最优潮流子问题获得



# 基于MuZero的微网在线调度算法设计

## 基于MuZero的微网在线调度算法设计

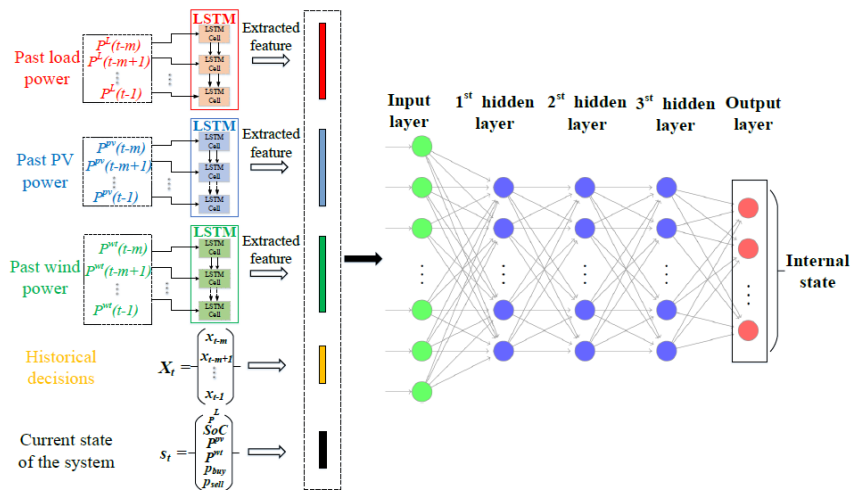


➤ Self-play过程：MCTS利用上一次迭代得到的神经网络模型(learned neural network model)建立树，产生大量训练样本；

➤ Training过程：根据样本数据，神经网络参数不断更新。经过训练，能很好地模拟系统的动态特性,并计算出当前最优策略和值函数。

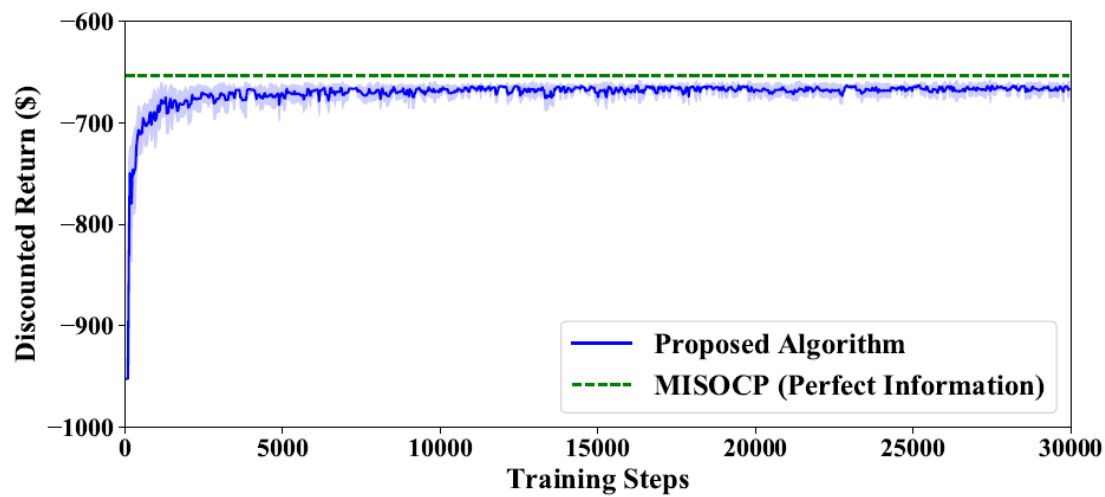
## LSTM网络提取历史数据的特征量（表示网络）

➤ 充分挖掘负荷及可再生能源历史数据中的信息，与当前状态一起生成“内部状态”，作为根节点，供MCTS生成树。



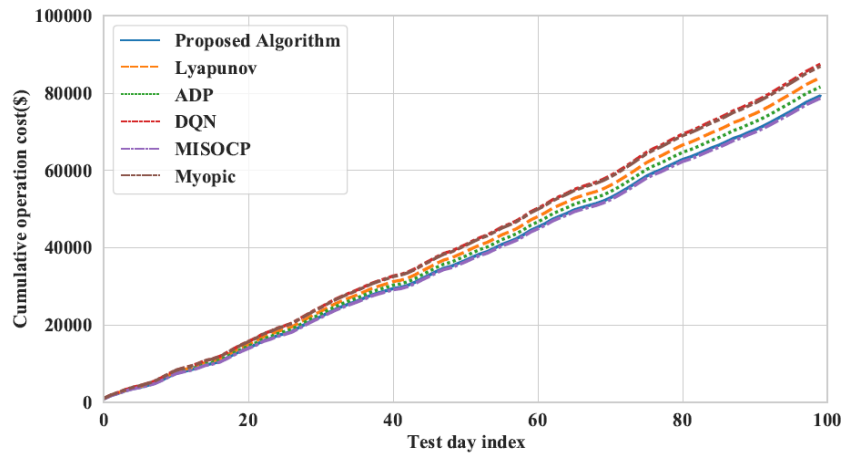
# 基于MuZero的微网在线调度仿真研究

□ 随着训练时间增长，MuZero智能体决策得到的最小运行费用逐渐逼近最优值



➤ “最优值”是在预测信息准确无误前提下，利用MISOCP（混合整数二阶锥规划）方法离线计算得到，仅作为衡量在线调度方法优劣的参考值。

□ 连续100天微网在线调度运行测试，采用MuZero智能体决策的运行费用最小



不同方法相对myopic策略的性能提升对比

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%

# 参考文献

---

**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](https://doi.org/10.1109/TSG.2020.3035127). ([Early Access](#))