Energy Proceedings
Vol 38, 2024
ISSN 2004-2965

Decentralized cooperative power dispatch based on the coupling of MAPPO and digital twin for BIPV cluster with ESS

Yuxin Liu[1], Han Zhu [2,3*], Zhuohui Zhang[2], Zhengrong Li[1], Bin He[2,3]

1 School of Mechanical Engineering, Tongji University, Shanghai, China

2 College of Electronic and Information Engineering, Tongji University, Shanghai, China

3 Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai, China
(*Corresponding Author: zhuhan@tongji.edu.cn)

ABSTRACT

Building Integrated Photovoltaics (BIPV) with Energy Storage Systems (ESS) enables buildings to play a crucial role in on-site PV consumption. However, due to the uncontrollability of PV, buildings often struggle to fully utilize it in real time. This paper proposes a decentralized cooperative power dispatch approach based on multi-agent proximal policy optimization (MAPPO) for cluster consisting of multiple BIPV with ESS. To acquire reliable strategies, a digital twin (DT) is employed as a sample and training environment for MAPPO to minimize cumulative grid power replenishment. An example of a small-scale building cluster is used to demonstrate the coupling of MAPPO and DT. The decentralized dispatch strategy is obtained with a one-hour time step. Verification results indicate a 9.85 MWh boost in PV self-absorption compared to a self-generating self-using strategy. Leveraging DT opens up further possibilities for applying MAPPO to power dispatch challenges.

Keywords: Multi-agent Proximal Policy Optimization, renewable energy, digital twin, decentralized dispatch, building integrated photovoltaics

NONMENCLATURE

Abbreviations	
BIPV	Building Integrated Photovoltaics
ESS	Energy Storage Systems
PV	Photovoltaic
MADRL	Multi-agent deep reinforcement learning
MAPPO	Multi-agent proximal policy optimization
DT	Digital twin
RERs	Renewable energy resources
Symbols	
E_{ci}	Charging power in building i (kW)
E_{fi}	Net outflow of electrical power from building i (kW)
E_g	Net power acquisition from the public grid (kW)
E_i	Power consumption of building i (kW)
E_{ti}	Electrical power converted to heat of building i (kW)
E_w	Power loss in the cluster (kW)
i	A building in the cluster
T	Dispatching cycle duration (h)
S_{ei}	Generated power of PV of i (kW)
$S_{ei,consume}$	Direct invocation of PV power of i (kW)
$S_{ei,grid}$	PV power flow to the grid of i (kW)
SoC_i	Charge status of the battery of i (%)
t	A certain timestep
$V_{si}(t)$	Used capacity of the storage system at the end of the t in building i (kWh)
$V_{max,i}$	Maximum electrical storage capacity in building i (kWh)
η_{Invert}	Inverter efficiency (%)
η_{RTE}	Round-trip efficiency (%)
η_{SDC}	Self-discharge efficiency (%)
$\eta_{Transform}$	Transformer efficiency (%)

This is a paper for 15th International Conference on Applied Energy (ICAE2023), Dec. 3-7, 2023, Doha, Qatar.

1. INTRODUCTION

Building energy accounts for a considerable proportion of total social energy use [1,2]. RERs is a green, clean and efficient source to help the carbon neutrality [3,4]. The introduction of RERs, especially on-site PV, into building side has attracted extensively attention, and BIPV is an extensively used form [5]. BIPV fully applies the electricity generated by distributed PV to building [6].

However, PV generation in BIPV is directly related to solar irradiance, so the supply of electricity would be intermittent and erratic. Districted ESS could be a buffer to help building self-consumption flexibly[7]. ESS in a building realizes a partial transfer of PV generation and transforms the task of balancing power from a short-term issue to a mid-term issue [8]. BIPV with ESS enables every single building to generate, use and store electricity independently and makes the building to be a dealer actively.

For a cluster composed of BIPV with ESS, power dispatch is a key issue. There are two main patterns for cooperative energy dispatch in building clusters: centralized and decentralized. Centralized pattern means that there is a top control center to collect all members information and decide their power transaction[9,10], while decentralized pattern means that each member collects information that affects it on its own and makes its own actions. When dealing with complex and variable scenarios, decentralized dispatch pattern performs better because of its strong flexibility and robustness [11,12].

Another important issue for the BIPV cluster with ESS is how to obtain a decentralized dispatch policy. MADRL is a popular data-driven methods to help multiple individuals to optimize their decisions and to consider the collective interest comprehensively[13]. MADRL could be applied in the building energy decentralized dispatching[14], so it is feasible to introduce MADRL for the decentralized power dispatch in the cluster composed of BIPV with ESS.

Among mature MADRL algorithms, MAPPO as an on-policy method is gradually used in various cooperative scenarios, like human-drone spatial crowdsourcing [15] and Internet of Vehicles Systems [16]. However, every MADRL algorithm would face the challenges of accuracy gap [17,18], time or cost limitation [19,20] and security lack [21] when applied in projects. In order to provide a reliable training environment for MAPPO, DT could be an effective solution by constructing digital virtual environment.

In this paper, we innovatively introduce the coupling of MAPPO and DT to decentralized dispatch in a cluster composed of BIPV with ESS aimed at maximizing the long-term PV consumption within the cluster. To this aim, **Section 2** discusses the optimized and customized dispatch model and MAPPO algorithm, and for the coupling of two emerging technologies, a data flow structure with building energy simulation engine and PV generation simulation engine is proposed. A case study is analyzed to demonstrate the feasibility and effectiveness in **Section 3**. Finally, **Section 4** summarizes the conclusions.

2. METHODOLOGY

In this section, the power relationship of the cluster composed of BIPV with ESS is further elaborated. The energy framework including power generation, consumption and storage would influence the transaction mechanism and decentralized dispatch model. Transaction mechanism will determine the path and constraints of electricity flow, while dispatch model will affect policy acquisition and optimization. Therefore, the paper analyzes the cluster power structure as the premise of the optimization method, so as to realize the decentralized cooperative dispatch.

2.1 Cluster energy framework

In this study, the energy relationship within a cluster includes only electricity. The PV integrated with building is the form of rooftop PV, thus conforming to the maximum utilization of PV without consideration of additional space. Building side demands mainly come from heating and cooling loads, lighting and electrical equipment.

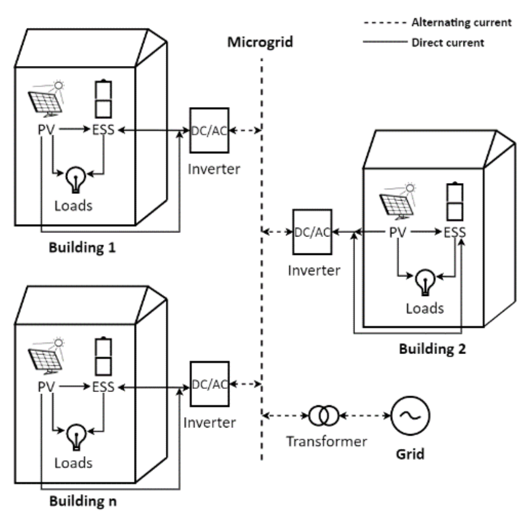

Fig. 1 Energy Framework

The energy framework of a cluster composed of BIPV with ESS is shown in Fig. 1. With on-site PV and ESS, the building in the cluster is no longer just a consumer of electricity, but an energy-using place including PV and ESS is defined as a building.

2.2 Energy transaction mechanism

A cluster composed of BIPV with ESS could be considered as a microgrid. Therefore, the electrical transaction between gird and buildings can be simplified as a transaction between grid and the microgrid, called as an external transaction, and transaction between microgrid and buildings, called as an internal transaction. The former is to keep the local container at a "fixed level", because microgrid has no buffering capacity and ignored transformation loss. The latter is to meet the electricity demand and balance the power inside the cluster. The energy transaction mechanism of the cluster is shown in Fig. 2.

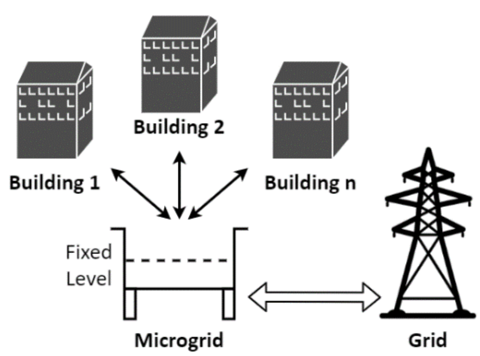

Fig. 2 Energy transaction Mechanism

2.3 Energy dispatch model

2.3.1 Objective function

The optimized dispatch of cluster is aimed to minimize the invoked amount from grid to the whole cluster over a long period of time. The objective function is shown as:

$$\min C = \sum_{t}^{T} E_g(t) \tag{1}$$

Where C is the overall cumulative power from grid in T timesteps. Since this study mainly focus on the power invoked from the grid, the power exported to the grid is not considered.

2.3.2 Constraint conditions

With the battery loss, the state of charge $SOC_i(t)$ at each time step of building i can be expressed as:

$$SOC_i(t) = \frac{E_{c,i}(t)\eta_{RTE}(i)}{V_{max,i}} + SOC_i(t-1)\eta_{SDC} \tag{2}$$

$$SoC_i(t)_{max} \geq SoC_i(t) \geq SoC_i(t)_{min} \tag{3}$$

Where $E_{c,i}(t)$ refers to the charge and discharge amount of ESS in building i at time t, and $E_{c,i}(t) > 0$ indicates charging, and vice versa; η_{RTE} depends on the direction of ESS charging and discharging.

In every building, the power balance relationship is shown as:

$$E_{ci}(t) = S_i(t) - E_{fi}(t) - E_i(t) \tag{4}$$

With the inverter loss, the power balance relationship between buildings and microgrid is expressed as:

$$E'_g(t) + \sum_{i}^{n} \eta_{Invert}(i)E_{fi}(t) = 0 \tag{5}$$

Where $E'_g(t)$ is the external transaction power of microgrid, and $E'_g(t) \geq 0$ indicates the cluster requesting from the grid, and vice versa; n is the number of buildings in cluster; η_{Invert} (i) depends on the direction of $E_{fi}(t)$.

The grid replenishes microgrid request timely, and ignore the overflow:

$$E_g(t) = \begin{cases} \eta_{Transform}E'_g(t) \geq 0, \text{if } E'_g(t) \geq 0 \\ 0, \text{if } E'_g(t) < 0 \end{cases} \tag{6}$$

2.4 Decentralized power dispatch optimization algorithm based on MAPPO

2.4.1 Markovian decision process

In order to solve the power coordination dispatching problem of the above-mentioned clusters, cooperative MAPPO online optimization calculation is adopted. Markov Decision Process (MDP) is the most basic problem model of reinforcement learning, which can describe sequential decision problems. The dispatching problem could be described as a MDP, and the state of the next moment is directly obtained from the action and state of the last step, and the environment parameters of this step, so it is the specified state transfer form.

In MAPPO algorithm, a building combined with PV and ESS is defined as an agent. For every agent, the state of agent i at timestep t is expressed as:

$$s_t^i = (S_i(t), E_i(t), SoC_i(t), holiday(t), t) \tag{7}$$

Where $holiday(t)$ is the judgement value of whether the day is a holiday(weekend) or not; t is the time value in the 24-hour system.

For each agent, the power dispatch is mainly dynamically regulated by controlling the ESS on the basis of meeting the demand. Correspondingly, the action of agent i is defined as:

$$a_t^i = \left(\frac{E_{ci}(t)}{V_{max,i}}\right) \tag{8}$$

Since the dispatch is the result of the cooperation of all buildings in the cluster, all agents have the same goal and all agents share a reward function. The goal is to minimize the accumulated invoked power from grid in power dispatch, while the goal is to maximize the accumulative reward in MAPPO. Thus, the reward function is setting as:

$$R = K\sum_{i}^{n}(E_{fi}(t) - (S_i(t) - E_i(t))) \tag{9}$$

Where K is a positive hyperparameter that needs to be determined by data experiments.

2.4.2 Algorithm solution flow coupled DT

The training process of MAPPO can be summarized as "centralized training, decentralized execution", which belongs to the parallel learning in MADRL training schemes and has high computational efficiency[22].

DT is a rising technology to transform physical world into a mathematical world, so DT could provide a platform for MAPPO to interact, train and validate. In addition, MAPPO needs to optimize policy based on rewards, and DT could feedback instant rewards. The structure for coupling MAPPO and DT is shown as Fig.3. At first timestep t, DT acquires the environment and system parameters to get $S_i(t)$ and $E_i(t)$. $SoC_i(t)$ would be given a staring value. Actor network of MAPPO receives states from DT, and decide an action. Actions from all agents will form a joint action. DT calculates the reward with the joint action and states to update critic network, and updates $SoC_i(t)$ with eq (2).

3. CASE STUDY

3.1 Basic information

The case selects a cluster composed of a hotel, a mall and an office building in Shanghai, China for decentralized dispatch. Since the accuracy of the prediction model is not deeply explored in this study, the meteorological parameters in DT are typical meteorological year parameters, and the occupancy rates are based on ASHRAE90.1. PVsyst and EnergyPlus are used for PV generation simulation and building energy consumption simulation respectively. The dispatching duration is a month, using January (744 h) as an attempt. The interval between two timesteps is 1 h. The Maximum electrical storage capacities of 3 buildings are 500kWh, 200kWh, and 1000kWh. The upper limit of SoC is defined as 100% and lower limit is 0. Since ESS is lithium-ion batteries, $\eta_{SDC} = 99.99\%$, $\eta_{RTE}(i) = \begin{cases} 95\%, \text{ if } E_{c,i}(t) \geq 0 \\ \frac{1}{95\%}, \text{ if } E_{c,i}(t) < 0 \end{cases}$, $\eta_{Invert}(i) = \begin{cases} 97\%, \text{if } E_{fi}(t) \geq 0 \\ \frac{1}{97\%}, \text{if } E_{fi}(t) < 0 \end{cases}$ and $\eta_{Transform} = \frac{1}{98.5\%}$, the hyperparameter K in eq (9) is 0.03.

The actor network and critic network are fully connected with the neural network. The learning rate is 5e-4, and the reward discount is 1. The simulation is based on Python, and the computer is configured with CPU Intel Core i5 and memory 16 GB.

3.2 Results

3.2.1 Training results

The optimized The case se To verify the feasibility of the method, more than 6,000 episodes were trained, each episode consisting of one month (744 steps). The results of cumulative reward and C for each episode obtained by training are shown i*n* Fig. 4. It can be seen

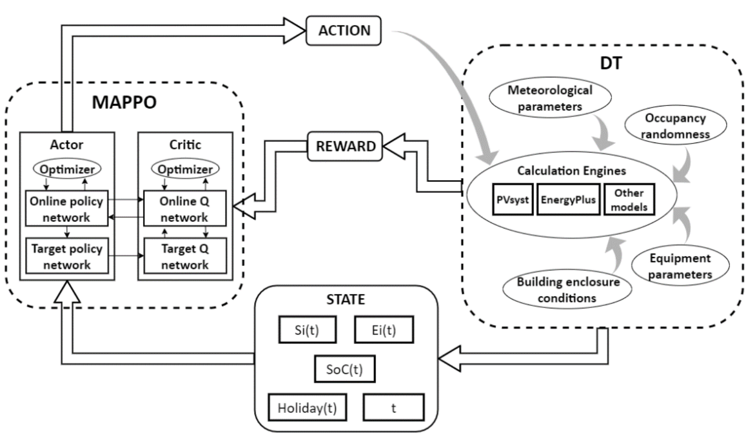

Fig.3. Structure for coupling MAPPO and DT

that the changes of reward and C are corresponding, and both begin to converge around 4000 episodes which indicates that the agents have learned the dispatching policy that minimizes C. The result shows that the setting of reward can effectively help the cluster to reduce C and to achieve the goal of minimize cumulative invoked power from public grid.

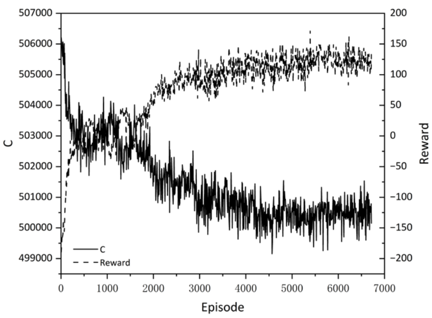

Fig. 4 MAPPO training results

3.2.2 Dispatching results

In order to evaluate the dispatching effect of the converged strategy, 5 sets of data in January were run based on DT. The effect of common priority strategy was compared as control group. The Self-generating & self-using (SGSU) strategy refers to that each building's PV generation flows first to the load, then to the battery, and finally to the microgrid and electricity load is first taken from PV power generation, followed by batteries, and finally from microgrids.

The dispatch results from MAPPO and common priority strategy are shown in *Table 1*. Self-consumption rate of PV means the proportion of PV generation consumed in total consumption. PV consumption rate refers to the proportion of PV generation consumed in the total generation. MAPPO performs better and PV consumption increased by 4875.25 kWh for a month dispatch.

Table 1 Dispatching results

Strategy	MAPPO	SGSU
PV consumption (MWh)	212.80	202.95
Self-consumption rate of PV (%)	37.57	35.83
PV consumption rate (%)	77.62	74.03
Loads (MWh)	557.94	
PV generation (MWh)	274.16	

4. CONCLUSION

The results obtained by case study proves the feasibility of coupling MAPPO and DT in decentralized cooperative power dispatch problem for the cluster composed of BIPV with ESS. The strategy developed by MAPPO training can reduce the cluster's accumulated power from the grid in a month, which means increasing the consumption of local PV power generation to strengthen the carbon neutrality goals. MAPPO can make full use of ESS to make cooperative dispatching more flexible with DT. Furthermore, the construction of DT and the model framework of MAPPO can be optimized according to the actual engineering situation.

ACKNOWLEDGEMENT

This study was supported by “Shanghai 2022 ‘Science and Technology Innovation Action Plan’ science and technology support carbon peak carbon neutral project (Grant No. 22dz1208000)”.

REFERENCE

[1] GlobalABC. 2021 GLOBAL STATUS REPORT FOR BUILDINGS AND CONSTRUCTION. 2021.
[2] Laustsen J. Energy Efficiency Requirements in Building Codes, Energy Efficiency Policies for New Buildings. IEA Information Paper 2008.
[3] Jin C, Lv Z, Li Z, Sun K. Green finance, renewable energy and carbon neutrality in OECD countries. Renew Energy 2023;211:279–84. https://doi.org/10.1016/J.RENENE.2023.04.105.
[4] Paska J, Surma T. Electricity generation from renewable energy sources in Poland. Renew Energy 2014;71:286–94. https://doi.org/10.1016/J.RENENE.2014.05.011.
[5] Li Y, Li L, Deng W, Zhu D, Hong L. Building Integrated Photovoltaic (BIPV) Development Knowledge Map: A Review of Visual Analysis Using CiteSpace. Buildings 2023;13. https://doi.org/10.3390/BUILDINGS13020389.
[6] Wijeratne WMPU, Samarasinghalage TI, Yang RJ, Wakefield R. Multi-objective optimisation for building integrated photovoltaics (BIPV) roof projects in early design phase. Appl Energy 2022;309. https://doi.org/10.1016/J.APENERGY.2021.118476.
[7] Hsieh S, Omu A, Orehounig K. Comparison of solar thermal systems with storage: From building to neighbourhood scale. Energy Build 2017;152:359–72. https://doi.org/10.1016/J.ENBUILD.2017.07.036.

[8] Koohi-Kamali S, Rahim NA, Mokhlis H. Smart power management algorithm in microgrid consisting of photovoltaic, diesel, and battery storage plants considering variations in sunlight, temperature, and load. Energy Convers Manag 2014;84:562–82. https://doi.org/10.1016/J.ENCONMAN.2014.04.072.
[9] Tsikalakis AG, Hatziargyriou ND. Centralized control for optimizing microgrids operation. IEEE Power Energy Soc Gen Meet 2011. https://doi.org/10.1109/PES.2011.6039737.
[10] Olivares DE, Canizares CA, Kazerani M. A centralized energy management system for isolated microgrids. IEEE Trans Smart Grid 2014;5:1864–75. https://doi.org/10.1109/TSG.2013.2294187.
[11] Karavas CS, Kyriakarakos G, Arvanitis KG, Papadakis G. A multi-agent decentralized energy management system based on distributed intelligence for the design and control of autonomous polygeneration microgrids. Energy Convers Manag 2015;v 103:166–79. https://doi.org/10.1016/j.enconman.2015.06.021.
[12] Strenge L, Schultz P, Kurths J, Raisch J, Hellmann F. A multiplex, multi-timescale model approach for economic and frequency control in power grids. Chaos 2020;30. https://doi.org/10.1063/1.5132335.
[13] Nguyen TT, Nguyen ND, Nahavandi S. Deep Reinforcement Learning for Multiagent Systems: A Review of Challenges, Solutions, and Applications. IEEE Trans Cybern 2020;50:3826–39. https://doi.org/10.1109/TCYB.2020.2977374.
[14] Perera ATD, Kamalaruban P. Applications of reinforcement learning in energy systems. Renew Sustain Energy Rev 2021;137. https://doi.org/10.1016/j.rser.2020.110618.
[15] Wang Y, Liu CH, Piao C, Yuan Y, Han R, Wang G, et al. Human-Drone Collaborative Spatial Crowdsourcing by Memory-Augmented and Distributed Multi-Agent Deep Reinforcement Learning. Proc - Int Conf Data Eng 2022;2022-May:459–71. https://doi.org/10.1109/ICDE53745.2022.00039.
[16] Jia Y, Zhang C, Huang Y, Zhang W. Lyapunov Optimization Based Mobile Edge Computing for Internet of Vehicles Systems. IEEE Trans Commun 2022;70:7418–33. https://doi.org/10.1109/TCOMM.2022.3206885.
[17] Mocanu E, Mocanu DC, Nguyen PH, Liotta A, Webber ME, Gibescu M, et al. On-Line Building Energy Optimization Using Deep Reinforcement Learning. IEEE Trans Smart Grid 2019;10:3698–708. https://doi.org/10.1109/TSG.2018.2834219.
[18] Fu Q, Han Z, Chen J, Lu Y, Wu H, Wang Y. Applications of reinforcement learning for building energy efficiency control: A review. J Build Eng 2022;50. https://doi.org/10.1016/j.jobe.2022.104165.
[19] Liu Y, Zhang D, Gooi HB. Optimization strategy based on deep reinforcement learning for home energy management. CSEEJ Power Energy Syst 2020;6:572–82. https://doi.org/10.17775/CSEEJPES.2019.02890.
[20] Cao D, Zhao J, Hu W, Ding F, Huang Q, Chen Z. Attention enabled multi-agent DRL for decentralized volt-VAR control of active distribution system using PV inverters and SVCs. IEEE Trans Sustain Energy 2021;12:1582–92. https://doi.org/10.1109/TSTE.2021.3057090.
[21] Hernandez-Leal P, Kartal B, Taylor ME. A survey and critique of multiagent deep reinforcement learning. Auton Agent Multi Agent Syst 2019;33:750–97. https://doi.org/10.1007/s10458-019-09421-1.
[22] Yu C, Velu A, Vinitsky E, Wang Y, Bayen A, Wu Y. The Surprising Effectiveness of PPO in Cooperative, Multi-Agent Games 2021.