Yiyun SUN, Senlin ZHANG, Meiqin LIU, Ronghao ZHENG, Shanling DONG, Xuguang LAN, 2024. Multi-agent evaluation for energy management by practically scaling *α*-rank. *Frontiers of Information Technology & Electronic Engineering*, 25(7):1003-1016. https://doi.org/10.1631/FITEE.2300438

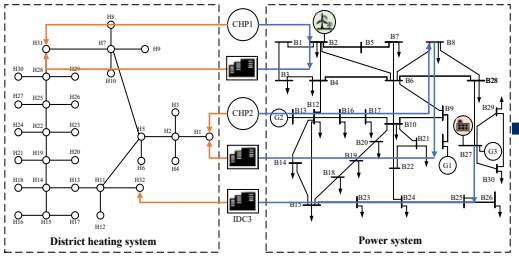
Multi-agent evaluation for energy management by practically scaling *α*-rank

Key words: Energy management; Multi-agent deep reinforcement learning; Strategy evaluation; Power grid system

Yiyun SUN
E-mail: sunyiyun@zju.edu.cn
ORCID: https://orcid.org/0009-0007-5593-8767

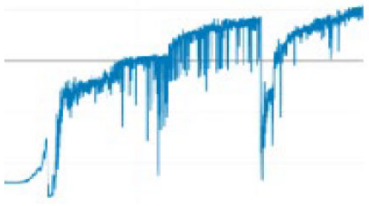
Corresponding author: Meiqin LIU
E-mail: liumeiqin@zju.edu.cn
ORCID: https://orcid.org/0000-0003-0693-6574

Motivation

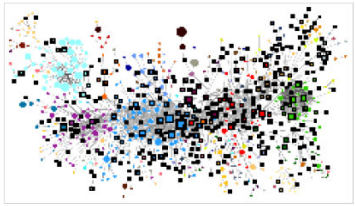


Currently, the increasing number of photovoltaic units being distributed into distribution networks may result in voltage issues. Multi-agent reinforcement learning (MARL) methods have shown great potential in energy management systems to alleviate these issues.

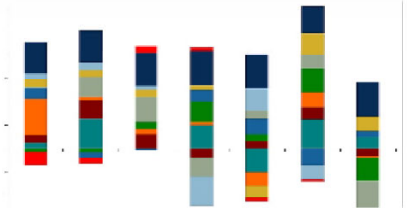
Problem



Most MARL algorithms have difficulty in guaranteeing the efficiency of each agent's strategy, which needs to be evaluated.



The existing strategy evaluation methods applied to MARL have high sampling complexity and are costly when applied to distributed network systems.

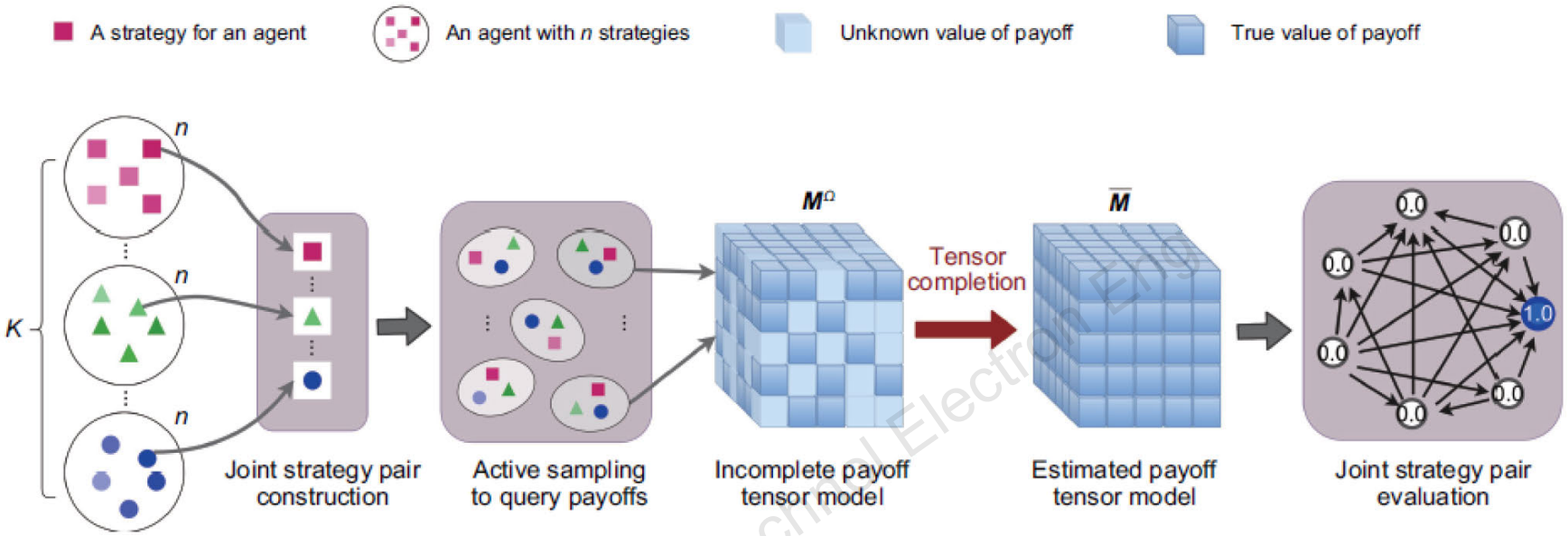


There is a practical problem with using strategy evaluation for the multi-agent system due to noisy payoffs.

Main idea

- The α-rank-based evaluation method to obtain the ranking of all reconstituted joint strategies at both the building and system levels is proposed.
- Unlike existing evaluation methods that require all joint strategy payoffs, we propose an approach to accurately estimate complete payoffs from a subset of key observed payoffs. This approach alleviates the problem of high sampling complexity during evaluation.
- Considering the treatment of observed payoffs with noise in real scenarios, a payoff prediction model is presented in TcEval-AS to accurately predict the real payoff in an energy management task with few interactions.

Method



The overview of the proposed TcEval method

Pipeline:

(1) Use the active sampling method to select key observed payoffs with the maximum information gain.

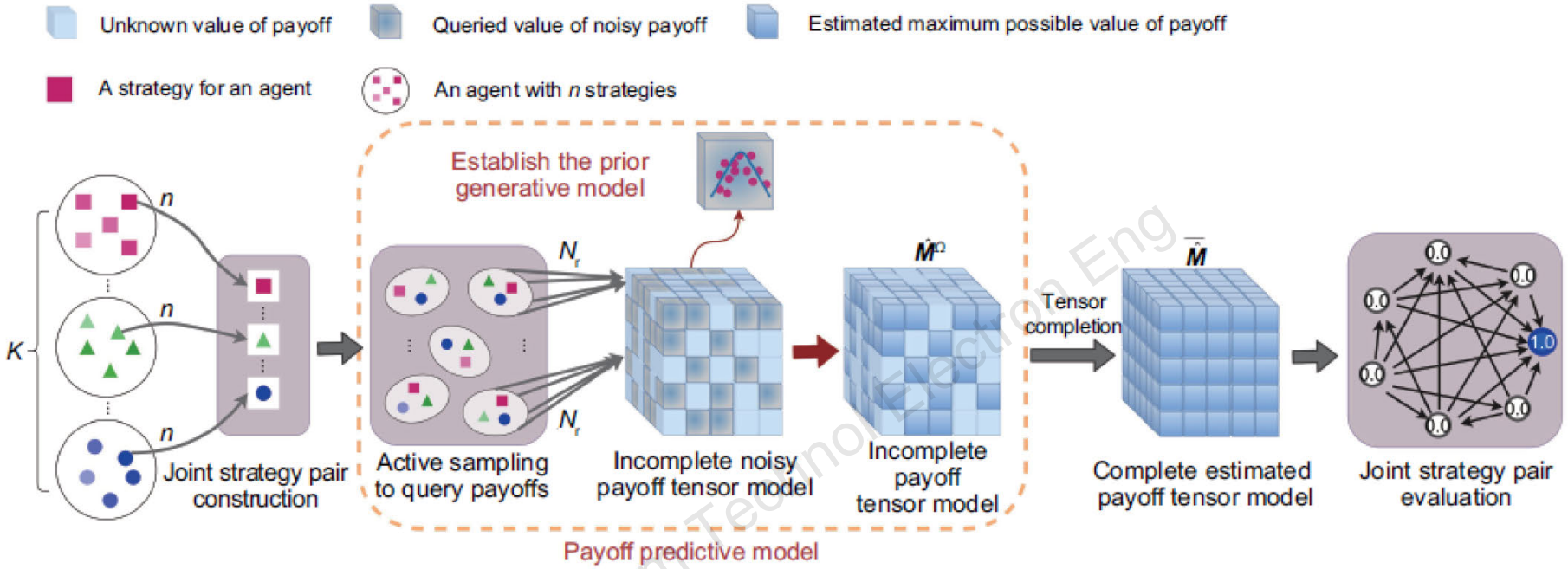
(2) Estimate the complete payoff tensor based on the PS-HOSVD tensor completion algorithm with the selected rank.

(3) Obtain the steady-state strategy distribution for the joint strategy set and calculate the ranks of all joint strategies.

Key point:

The TcEval method is proposed to precisely evaluate all joint strategies through a small portion of joint strategies.

Method (Cont'd)

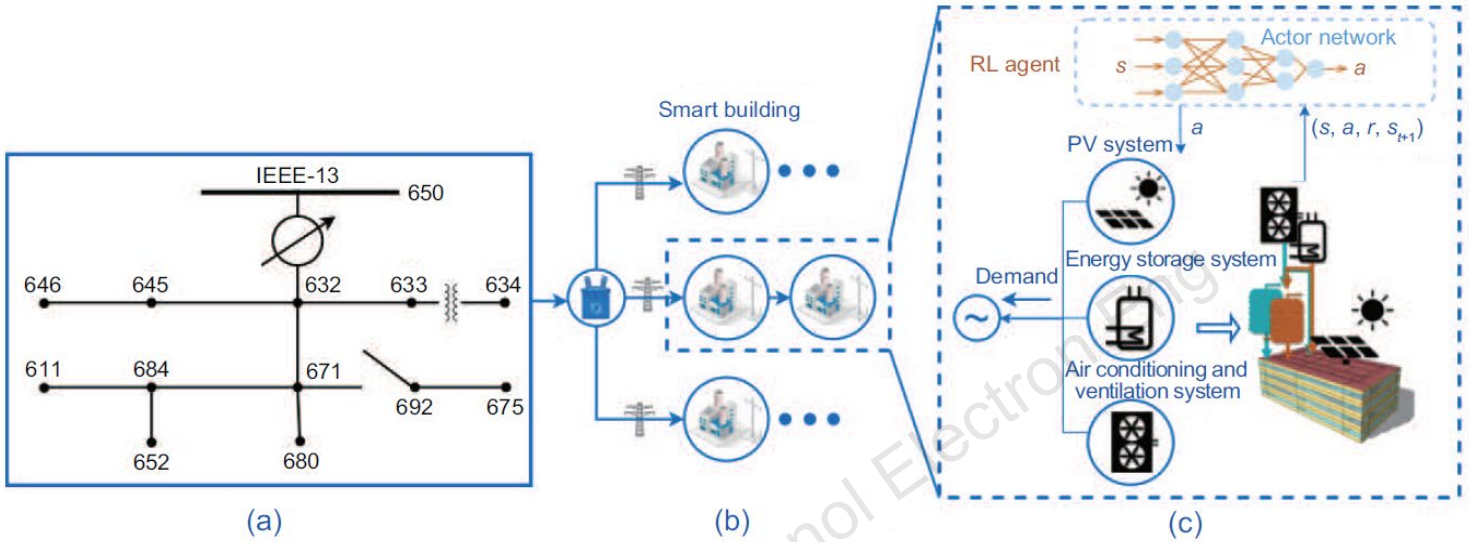


The overview of the proposed TcEval-AS method

Key points:

- To solve the problem of accessing only noisy payoffs, the payoff probabilistic predictor can be represented based on the mathematical prior model.
- The proposed payoff prediction model presented in TcEval-AS aims to accurately predict the real payoff in an energy management task with a few interactions. Additionally, an interaction complexity analysis based on the theory of tensor completion with noisy terms is proposed.

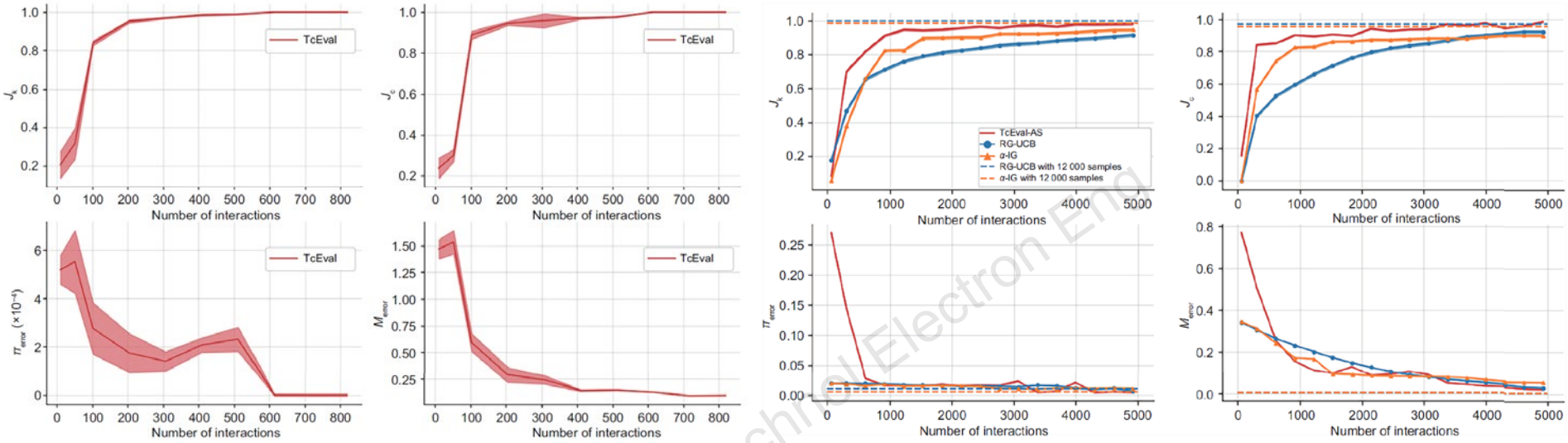
Constructed environment



Structure of the EnergyGrid environment

- EnergyGrid is developed as an adaptation of the GridLearn environment, as shown in the figure.
- The IEEE-13-bus network is selected as the grid model, and building microgrids based on the data from the prototype buildings provided by the U.S. Department of Energy are developed to connect to it.
- The purpose of the energy management task is to achieve the rational strategy, further preserving building occupants' comfort while pursuing system-level objectives.

Major results: Evaluation performance

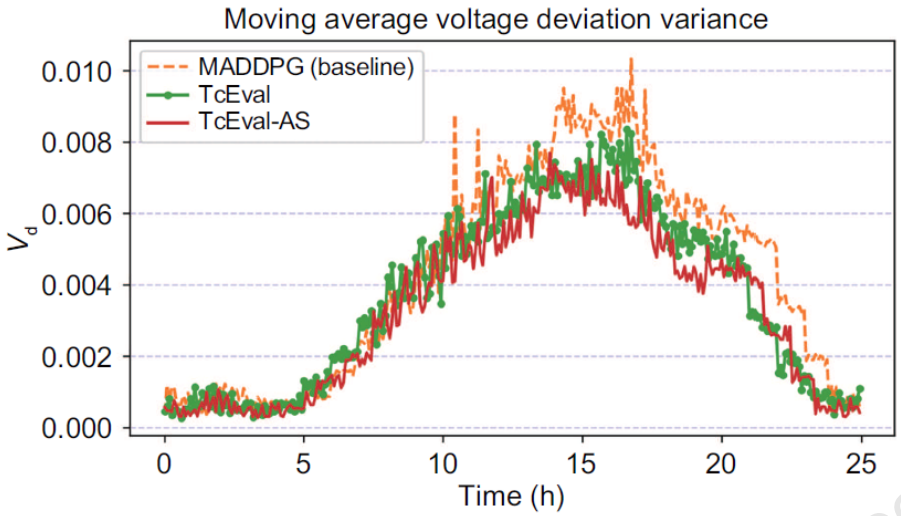


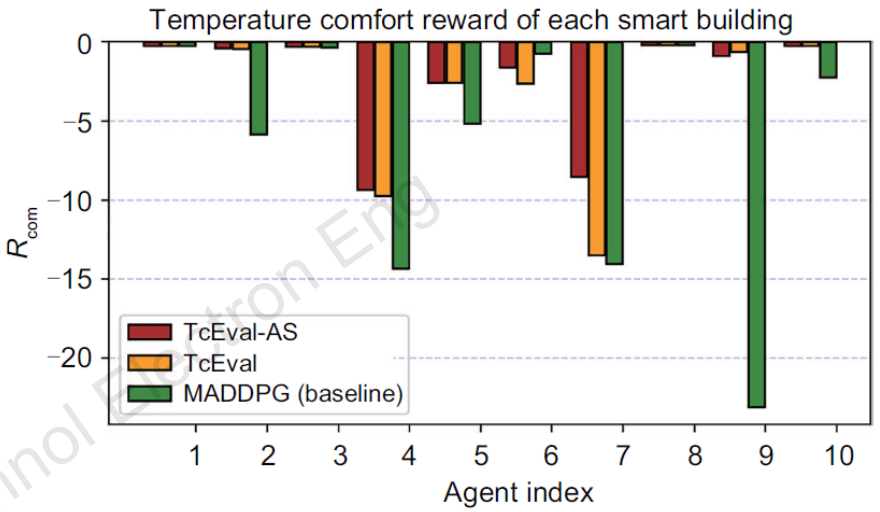
(a) Evaluation performance with precise payoffs **(b) Evaluation performance with noisy payoffs**

Evaluation performance comparison in case study 1 with state-of-the-art evaluation methods

- The proposed TcEval obtains accurate rankings with as few as about 400 interactions, reducing the number of interactions by over 60% compared to the original α-rank.
- TcEval-AS achieves better performance in terms of all metrics compared to RG-UCB and α-IG, with the same number of interactions. These results prove that the proposed method is feasible.

Major results: Control performance of the top-rank joint strategy





Control performance in case study 2 with strategies obtained from different methods

- The top-rank joint strategy evaluated by TcEval-AS usually effectively reduces the overall voltage deviation, especially during high deviations, which ensures the stability of the power system at the system level.
- The baseline joint strategy does not perform well on some agents, possibly because all agents converge on similar control strategies.
- The combination of MARL methods with the strategy evaluation algorithm is an effective roadmap toward trustworthy MARL.

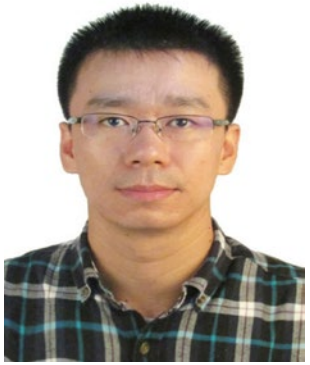
Conclusions

- In this paper, two algorithms, namely, TcEval and TcEval-AS, are proposed for the energy management task in a simulated power grid system.
- The TcEval method is developed to achieve large-scale strategy evaluation by estimating the α-rank with tensor completion.
- Considering the noise prevalent in practice, the proposed TcEval-AS adds a noise processing module based on TcEval, using a mathematical model to fit the noisy payoff distribution and predict the real payoff.
- Joint strategy evaluation is performed in two case studies, and it is shown that the proposed algorithms can achieve better performance than the RG-UCB and α-IG methods while reducing the number of interactions.

Yiyun SUN received the B.E. degree in automation from Northwestern Polytechnical University, Xi'an, China, in 2021. She is currently working towards the Ph.D. degree in control theory and control engineering at Zhejiang University, Hangzhou, China. Her research interests include multi-agent games and multi-agent reinforcement learning.

Senlin ZHANG (Member, IEEE) received the B.E. degree in control theory and control engineering from the Wuhan University of Technology, Wuhan, China, and the M.E. degree in control theory and control engineering from Zhejiang University, Hangzhou, China, in 1984 and 1991, respectively. He is currently a Professor with the College of Electrical Engineering and the National Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China. He has authored more than 200 papers in major journals and international conferences. His current research interests include unmanned undersea systems, intelligent systems, and multi-sensor networks.

Meiqin LIU (Senior Member, IEEE) received the B. E. and Ph.D. degrees in control theory and control engineering from Central South University, Changsha, China, in 1994 and 1999, respectively. She was a Post-Doctoral Research Fellow with the Huazhong University of Science and Technology, Wuhan, China, from 1999 to 2001. She was a Visiting Scholar with the University of New Orleans, New Orleans, LA, USA, from 2008 to 2009. She was a Professor with the College of Electrical Engineering, Zhejiang University, Hangzhou, China, from 2001 to 2021. She is currently a Professor with the National Key Laboratory of Human–Machine Hybrid Augmented Intelligence, Xi'an Jiaotong University, Xi'an, China, and also with the College of Electrical Engineering, Zhejiang University, Hangzhou, China. She has authored more than 200 papers in major journals and international conferences. Her work was supported by the Zhejiang Provincial Natural Science Fund for Distinguished Young Scholars in 2010, and by the National Science Fund for Excellent Young Scholars of China in 2012. She has won a second prize of Science and Technology Award of Zhejiang Province in 2013, and a first prize of Natural Science Award of Chinese Association of Automation in 2019. Her current research interests include theory and application of artificial intelligence, multi-sensor networks, information fusion, and nonlinear systems.



Ronghao ZHENG (Member, IEEE) received the B.S. degree in electrical engineering and the M.S. degree in control theory and control engineering from Zhejiang University, Hangzhou, China, in 2007 and 2010, respectively, and the Ph.D. degree in mechanical and biomedical engineering from the City University of Hong Kong, Hong Kong, China, in 2014. He is currently with the College of Electrical Engineering, Zhejiang University. His research interests include distributed algorithms and control, especially the coordination of networked mobile robot teams, with applications in automated systems and security.

Shanling DONG (Member, IEEE) received the B.E. degree in Automation from Xidian University, Xi'an, China, and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2019. Supported by China Scholarship Council, she was a Visiting Graduate Student at University of California, Riverside, USA, from Sept. 2017 to Sept. 2018. She was a Research Associate and Post Doctoral Research Fellow with the City University of Hong Kong, Hong Kong SAR, China, from Dec. 2019 to Dec. 2020. She is currently with the College of Electrical Engineering, Zhejiang University, Hangzhou, China. Her research interests include Markov jump systems, fuzzy systems, and multi-agent systems.

Xuguang LAN (Senior Member, IEEE) received Ph.D. degree in pattern recognition and intelligent systems from Xi’an Jiaotong University, Xi’an, China, in 2005. He was a visiting scholar in Northwestern University from Sept. 2013 to Oct. 2014, and Ecole Centrale de Lyon from May 2005 to Sept. 2005. Currently, he is a professor of the College of Artificial Intelligence and Robotics in Xi'an Jiaotong University. He has published more than 110 papers in famous journals and conferences in the field of artificial intelligence and robotics, such as IEEE Trans and ICML/CVPR/RSS, and has obtained more than 20 national invention patents. His research interests include computer vision, deep reinforcement learning, robot learning, and multi-agent games.