ZTE COMMUNICATIONS

中兴通讯技术(英文版)

Special Topic:
Evolution of AI Enabled Wireless Networks

RIS element

z

y

0

DR

DR

DR

DT

DT

...

DT

User

x

D2D pairs

The 9th Editorial Board of ZTE Communications

CONTENTS

Special Topic ▶ | **Evolution of AI Enabled Wireless Networks**

ZTE COMMUNICATIONS

ZTE Communications is a peer-reviewed international ICT journal (CN 34-1294/TN and ISSN 1673-5188) featuring industry-university-institute cooperation. It is published quarterly as a printed publication and can also be freely accessed at https://zte.magtechjournal.com.

ZTE Communications was founded in 2003 and has a readership of more than 100 000. The English version is distributed to universities, colleges, and research institutes in more than 100 countries. The journal covers a wide range of topics in the field of ICT. The Editorial Board members are distinguished domestic and international experts. The journal has become an integrated forum for university academics and industry researchers from around the world.

The journal uses ScholarOne Manuscripts, a professional web-based manuscript submission and peer-review tracking system. Authors must submit manuscripts electronically to https://mc03.manuscriptcentral.com/ztecom.

2023 Special Topics in *ZTE Communications*

Issue 1: Federated Learning over Wireless Networks
CUI Shuguang, The Chinese University of Hong Kong, Shenzhen
YIN Changchuan, Beijing University of Posts and Telecommunications
ZHU Guangxu, Shenzhen Research Institute of Big Data

Issue 2: Evolution of AI Enabled Wireless Network
WANG Ling, Northwestern Polytechnical University
GAO Yin, ZTE Corporation

Issue 3: Reinforcement Learning and Intelligent Decision
GAO Yang, Nanjing University
FENG Yanghe, National University of Defense Technology

Issue 4: 3D Point Cloud Processing and Applications
SUN Huifang, Pengcheng Laboratory
LI Ge, Peking University
CHEN Siheng, Shanghai Jiao Tong University
LI Li, University of Science and Technology of China
GAO Wei, Peking University

Special Topic on
Evolution of AI Enabled Wireless Networks

Guest Editors

WANG Ling

GAO Yin

Evolving to 6G wireless communications, "Internet connecting things" is becoming "Internet connecting intelligence". In order to adapt to this trend, wireless devices should be capable of conducting efficient and robust communications and also complete complex services of smart applications. However, the current wireless networks are trapped in delivering a massive number of signals while assuring their accuracy, and the limited resources block this evolution. The big data technology is a persistently-developing paradigm, which facilitates the change of network functions from "transmit every bit" to "what and how to transmit". In particular, artificial intelligence (AI) technologies allow intelligent devices to pre-process information according to environment and service's demands. Additionally, forwarding critical information in a more efficient and accurate way may support complicated AI tasks such as virtual reality, augmented reality, and autonomous driving. However, there still exist many fundamental research challenges, such as high quality wireless connectivity as well as high accuracy and robust sensing capability, which need to be solved for accommodating AI techniques towards 6G wireless communications.

In this special issue, a series of articles are presented to propose innovative solutions to enabling AI techniques over wireless networks. These papers cover a wide range of topics, including 6G wireless communication protocols, semantic communications, green energy efficiency concerns, network architecture designs, and the application of AI techniques in future wireless networks. The call-for-papers of this special issue have brought excellent submissions in both quality and quantity. After two-round reviews, ten excellent papers have been selected for publi-

cation in this special issue which is organized as follows.

The first paper titled "Intelligent 6G Wireless Network with Multi-Dimensional Information Perception" focuses on the critical issues and proposes three application scenarios in 6G wireless systems. The intelligent wireless network and information perception require a deep fusion of AI and wireless communications in 6G systems. Therefore, the fusion of AI and 6G networks is discussed for the enhancement of 5G-advanced technologies and future wireless communication systems. The wireless AI technology architecture with 6G multi-dimensional information perception is then introduced.

The second paper titled "Deep Learning-Based Semantic Feature Extraction: A Literature Review and Future Directions" provides an overview of the applications of semantic feature extraction in various fields, aiming to provide insights into the potential of this technology to advance the development of artificial intelligence. The applications of semantic feature extraction in natural language processing, hyperspectral image analysis, disease diagnosis and medical image analysis, and autonomous driving are focused. The development trends and challenges are also explored.

The third paper titled "Content Popularity Prediction via Federated Learning in Cache-Enabled Wireless Networks" proposes a privacy-preserving algorithm based on federated learning (FL) and long short-term memory (LSTM), which is referred to as FL-LSTM, to predict content popularity and reduce the risk of privacy leakage. The performance of the proposed FL-LSTM is close to the centralized LSTM and better than other benchmark algorithms in terms of privacy protection. Meanwhile, the caching policy in this paper raises about 14.3% of the content bit rate.

The fourth paper titled "Federated Learning for 6G: A Survey from Perspective of Integrated Sensing, Communication and Computation" contributes to the understanding of FL in the context of wireless networks and provides insights into addressing the challenges and optimizing the design for the integration of FL into future 6G networks. This paper provides a

comprehensive survey of FL, with special emphasis on the design and optimization of integrated sensing, communication and computation (ISCC). Subsequently, the challenges are highlighted and the state of the art in potential solutions is reviewed. Design guidelines are also provided for the incorporation of FL and ISCC.

The fifth paper titled "Future Vision on Artificial Intelligence Assisted Green Energy Efficiency Network" proposes AI/machine learning (ML)-assisted energy-saving strategies to achieve optimal performance in terms of cell shutdown duration and energy efficiency. 5G new radio is designed to enable denser network deployments, which raises significant concerns about network energy consumption. The AI/ML based energy saving schemes achieve great performance on power consumption and energy efficiency. Moreover, further consideration on future wireless communication networks is put forward.

The sixth paper titled "Machine Learning Driven Latency Optimization for Internet of Things Applications in Edge Computing" introduces a machine learning-enabled orchestration framework, which utilizes the states of edge resources and application resource requirements to facilitate a resource-aware offloading scheme for minimizing the average latency of emerging Internet-of-Things applications. Moreover, a variant bin-packing optimization model is further proposed, which co-locates applications firmly on edge resources to fully utilize available resources.

The seventh paper titled "Multi-User MmWave Beam Tracking via Multi-Agent Deep Q-Learning" proposes a multi-user beam tracking algorithm by using a distributed deep Q-learning method to reduce overhead cost. By online learning of users' moving trajectories, the proposed algorithm learns to scan a beam subspace, aiming to maximize the average effective sum-rate. Considering practical implementation, the continuous beam tracking problem is considered as a non-Markov decision process and a simplified training scheme of deep Q-learning with low complexity is developed. Furthermore, a scalable state-action-reward is designed for scenarios with different users and antenna numbers.

The eighth paper titled "RIS-Assisted UAV-D2D Communications Exploiting Deep Reinforcement Learning" proposes a reconfigurable intelligent surface (RIS) model to rebuild the wireless channels for mitigating the strong interference caused by line-of-sight (LoS) air-to-ground channels in a device-to-device (D2D) communication system underlying cellular networks enabled by unmanned aerial vehicle (UAV). A RIS softmax deep double deterministic (RIS-SD3) policy gradient method is also proposed, which could smooth the optimization space as well as reduce the number of local optimizations.

The ninth paper titled "SST-V: A Scalable Semantic Transmission Framework for Video" provides a highly-efficient solution to video transmission by proposing a scalable semantic transmission algorithm, named scalable semantic transmission

framework for video (SST-V), which jointly considers the semantic importance and channel conditions. Specifically, a semantic importance evaluation module is designed to extract more informative semantic features according to the estimated importance level, facilitating high-efficiency semantic coding. By further considering the channel condition, a cascaded learning based scalable joint semantic-channel coding algorithm is proposed, which autonomously adapts the semantic coding and channel coding strategies to the specific signal-to-noise ratio (SNR).

The last paper titled "UAV Autonomous Navigation for Wireless Powered Data Collection with Onboard Deep Q-Network" proposes to jointly optimize the UAV's flight trajectory, the sensor selection and operation modes in order to maximize the average data traffic of all sensors within a wireless sensor network (WSN) during finite UAV's flight time, while ensuring the energy required for each sensor by wireless power transfer (WPT). For the sake of considering a practical scenario, the UAV has no prior knowledge of sensor locations. Therefore, the deep Q-network (DQN) is employed to execute the navigation based on the UAV position, the battery level state, channel conditions and current data traffic of sensors within the UAV's coverage area.

To conclude, it is hoped that this special issue will serve as a valuable resource for researchers, practitioners, and students who are interested in AI techniques over wireless networks. We also hope that it will inspire further research in this field, leading to new and innovative solutions that will drive the evolution of AI techniques. Finally, we would like to express our sincere gratitude to all the authors, reviewers, and editorial staff who have contributed to the success of this special issue. Hopefully, the articles in this special issue are both insightful and informative for prospective readers in the field.

Biographies

WANG Ling received his BS, MS, and PhD degrees in electronic engineering from Xidian University, China in 1999, 2002 and 2004, respectively. From 2004 to 2007, he worked at Siemens and Nokia Siemens Networks. Since 2007, he has been with the School of Electronic and Information, Northwestern Polytechnical University, China, and was promoted to Professor in 2012. His current research interests include array processing and smart antennas, wideband communications, cognitive radio, adaptive anti-jamming for satellite communications, satellite navigation, and date link systems.

GAO Yin is a senior expert in wireless communication standardization and industrial relationship of ZTE Corporation and the 3GPP RAN3 Chair. She possesses 16 plus years of experience in the telecommunication industry and 14 plus years in the 3GPP RAN3 working group. She received her bachelor's and master's degrees from Xidian University, China. She joined ZTE and worked on software development and system design of the wireless communication system as early as 2005. Starting in 2007, she has been focusing on advanced radio access technology research and standardization with over 500 3GPP contributions and over 50 granted patents globally. She has been served as RAN3 Vice Chair of 3GPP RAN3 since 2017 and was elected as Chair of 3GPP RAN3 in May 2021.

Intelligent 6G Wireless Network with Multi-Dimensional Information Perception

YANG Bei, LIANG Xin, LIU Shengnan, JIANG Zheng, ZHU Jianchi, SHE Xiaoming

(China Telecom Research Institute, Beijing 102209, China)

Abstract: Intelligence and perception are two operative technologies in 6G scenarios. The intelligent wireless network and information perception require a deep fusion of artificial intelligence (AI) and wireless communications in 6G systems. Therefore, fusion is becoming a typical feature and key challenge of 6G wireless communication systems. In this paper, we focus on the critical issues and propose three application scenarios in 6G wireless systems. Specifically, we first discuss the fusion of AI and 6G networks for the enhancement of 5G-advanced technology and future wireless communication systems. Then, we introduce the wireless AI technology architecture with 6G multidimensional information perception, which includes the physical layer technology of multi-dimensional feature information perception, full spectrum fusion technology, and intelligent wireless resource management. The discussion of key technologies for intelligent 6G wireless network networks is expected to provide a guideline for future research.

Keywords: 6G wireless network; artificial intelligence; multi-dimensional information perception; full spectrum fusion

1 Introduction

From the perspective of 6G vision and requirements, there are several typical application scenarios, including smart cities, intelligent transportation, industrial intelligence, immersive Extended Reality (XR), holographic communication, sensory interconnection, and integrated perception and communications[1 – 7]. Intelligence and perception have become two keywords in 6G scenario applications[8 – 15]. Therefore, fusion is becoming a typical feature and a key challenge that distinguishes 6G from previous generations of communication systems.

Studies on the spectrum are essential for mobile communication systems. In the future, the 6G spectrum will expand to frequency bands with more abundant spectrum resources, e.g., terahertz and visible light. Larger bandwidth, higher frequency and full spectrum technology become the trend of evolution. Millimeter wave technology has been supported in the 5G wireless communication system standards. In the 6G era, millimeter wave technology will be mature and widely used. At the same time, the exploration of the terahertz band becomes a popular topic of research on 6G[1 – 3, 13 – 15].

For the full spectrum of 6G, we think the coverage of medium and low-frequency bands should be mainly guaranteed. Millimeter waves and terahertz provide a large bandwidth to form a multi-frequency collaborative ubiquitous networking pattern. Operators will face more fragmented spectrum resources. In order to improve the efficiency of resource utilization and network operation management, these fragmented spectrum resources must be fused. For example, 6G full spectrum fusion can be achieved by designing architecture and a series of key technologies.

2 Key Issues and Scenarios in 6G

Based on the aforementioned vision of 6G and the trend of spectrum fusion, we introduce some perspectives on the design of native intelligence systems for 6G. Intelligence, as a key capability of wireless communication networks, is becoming an enabling technology for emerging businesses. From 5G to 6G, artificial intelligence (AI) will complete the role transformation from assistance to endogeny. The fundamental design principle of wireless AI is to create a future high-efficiency and sustainable network. We need to efficiently utilize the fragmented spectrum and improve energy efficiency.

Therefore, we introduce three key scenarios in 6G for the fusion of AI and networks:

1) Scenario 1: assisting the enhancement of 5G-advanced

technology to improve network efficiency

This scenario mainly includes the existing network architecture based on 5G evolution, defining the wireless AI technology framework, mechanisms and signaling processes to achieve the integration of traditional technologies and AI, boost the technology enhancement of 5G evolution, and improve the efficiency of networks and system performance. Currently, the initial research work on intelligent air interface is underway in 3GPP Release 18. It lays the groundwork for the evolution of 5G systems and the design of 6G systems.

2) Scenario 2: AI native enabling typical high-value applications

Emerging 6G businesses, such as metaverse, immersive XR, digital twins and fully automated driving, will pose high requirements for future wireless communication systems. An AI native wireless network in 6G will become a key enabling technology for high-value application scenarios to meet the demands of high-value application scenario users for the ultimate experience and ultra-high network operational efficiency.

3) Scenario 3: AI native building efficient and sustainable networks for operators

With the goal of full spectrum fusion in 6G, AI native wireless networks will address the super-high complexity and challenges of network management and operation brought about by multi-dimensional network deployment in future mobile communication systems. AI native will be used to realize resource awareness and dynamic control of 6G wireless networks with multiple operators, multiple frequency bands and multiple modes, improving network energy efficiency and resource utilization efficiency and building a green, efficient and sustainable 6G network.

3 Wireless AI Technology Architecture

Therefore, we design and propose a 6G wireless AI architecture for multi-dimensional information perception, which consists of three aspects: physical layer technology for multi-dimensional feature information perception, intelligent full-spectrum fusion, and intelligent wireless resource management.

Achieving the deep integration between AI and wireless communications requires the support of big data of wireless communications, which enables AI to perceive, extract and integrate multi-dimensional information to solve problems with high complexity. Data is the essential requirement of the design of wireless communication systems with internal AI. Facing the multi-objective optimization problems, future wireless AI will be based on multi-dimensional information perception, e.g., scene information, service information, network state information, terminal measurement information, and so on.

In addition, the current design of the wireless protocol stack and the demand for multi-dimensional information perception from the bottom to the top are contradictory. Therefore, when designing the 6G protocol stack, we need to consider matching the requirements of multi-dimensional information perception for wireless AI. In this section, we introduce the design from three aspects in detail.

3.1 Multi-Dimensional Feature Information Perception

We consider the physical layer technology of multi-dimensional feature information perception to be achieved through AI-driven feature extraction, perception, and fusion of multi-dimensional wireless data (e. g., channel impulse response, received signal power, and so on) in the frequency, time, space, or angle domain. This physical layer technology of multi-dimensional feature information perception includes AI-driven extra large-scale massive multiple-input multiple-output (XL-MIMO) technology and AI-enabled localization technology.

3.1.1 AI-Driven XL-MIMO Technology

AI-based XL-MIMO technology, such as intelligent channel state information (CSI) enhancement and intelligent beam management, can integrate with some traditional functions through AI, achieving the replacement of certain functional modules and helping to improve the utilization efficiency of network resources and system performance.

Different from the codebook-based CSI acquisition method in traditional wireless communication systems, intelligent CSI enhancement technology includes intelligent CSI compression feedback and intelligent CSI prediction. Intelligent CSI feedback enhancement can increase the number of reference signals used for obtaining channel state information. As shown in Fig. 1, AI technology is used to compress and reconstruct the multi-dimensional channel feature information of CSI feedback, reducing the overhead of air interface CSI feedback or improving CSI reconstruction accuracy.

Intelligent CSI feedback technology usually deploys an AI model at both the base station (BS) side and the user equipment (UE) side, including CSI compression and CSI reconstruction. During the stage of training an AI model, the UE estimates multi-dimensional channel feature information based on reference signals and feeds the multi-dimensional channel feature information back to the BS as input and label data for
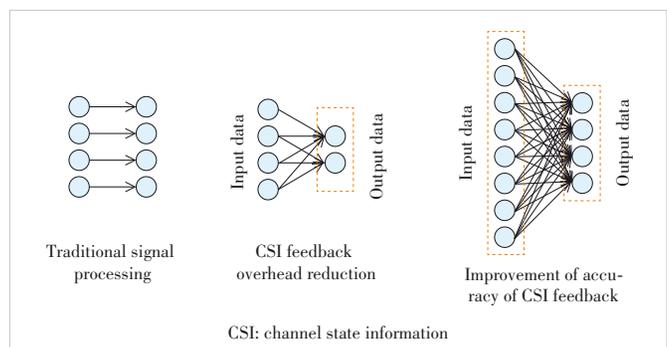
▲Figure 1. Enhancement of artificial intelligence (AI)-based CSI feedback

the training of the AI model. During the inference stage of the AI model, the UE sends the estimated channel feature vector to the encoder for CSI compression. Then, the compressed CSI codeword is reported to the BS. The BS uses the decoder to reconstruct the channel feature vector from the codeword that is fed back. The channel feature vector is used for precoding. We propose to utilize AI techniques to extract CSI features in a fully efficient manner, with lower inference complexity but comparable performance to traditional codebook methods. With the application of intelligent CSI feedback technology in dense urban areas and other scenarios, we estimate that more than 60% of uplink feedback overhead can be saved[16]. In addition, further improvement in compression performance can be achieved through intelligent retraining and optimization of quantization modules of AI models.

In traditional methods, time delay exists among CSI measurement, feedback and precoding, which results in a system performance loss since the current wireless channel uses an outdated CSI. This problem is more pronounced in high-speed mobility scenarios. As shown in Fig. 2, intelligent CSI prediction technology can utilize channel correlation features in the time domain to predict future CSI based on AI algorithms. In addition, CSI prediction technology can be enhanced by exploiting multi-dimensional channel feature information such as multi-user and bidirectional channel reciprocity. The traditional codebook information and uplink reference signal measurement results are used as inputs to extract and fuse features through AI algorithms for CSI prediction, thereby improving the accuracy of CSI prediction.

The intelligent CSI enhancement technology that perceives multi-dimensional feature information has broad application prospects and is of great significance for improving network performance and resource utilization efficiency.

The widespread use of massive antenna arrays poses great challenges to future beam management. As shown in Fig. 3, intelligent beam management can use AI algorithms to predict beams in the spatial and temporal domains, reduce beam measurement pilot overhead and measurement latency, and significantly improve beam management accuracy and system performance. Fig. 3(a) illustrates our proposition that implements AI algorithms to solve the conventional beam management problem through a classification model at high frequency, which can significantly reduce the required number of beam directions for measurement. Specifically, instead of measuring 32 BS beam directions and eight UE beam directions as in traditional methods, measuring only eight and four beam directions for the BS and UE respectively is sufficient. This reduction in measurement time by 87.5% results in an improved beam prediction accuracy of more than 45%[17]. Therefore, AI-based beam management can provide new ideas for defining mechanisms for future ultra-high frequency XL-MIMO beam management, simplifying relevant signaling processes while providing excellent prediction performance.

3.1.2 AI-Enabled Precision Positioning Technology

Traditional positioning technology mainly solves the problem of indoor positioning, and its main idea is to calculate the position of a user based on the measurement of multiple station points, especially the line-of-sight (LOS) path. Common methods include the time of arrival (TOA), time difference of arrival (TDOA), multi-round trip time (Multi-RTT), and so on.

The introduction of AI technology is mainly used to solve the impact of non-line-of-sight (NLOS) paths on positioning calculation and the problem of inaccurate positioning in heavy NLOS environments. By utilizing AI to extract multi-dimensional feature information of measurement signals and conduct binary classification for LOS and NLOS paths, the interference of NLOS path information on positioning calculation can be mitigated. The CSI with higher LOS probability is used to calculate the UE position, which can improve positioning performance, especially when there are few indoor antennas to improve positioning accuracy.

In scenarios with heavy NLOS, traditional positioning algorithms are no longer applicable because the measurement sig-
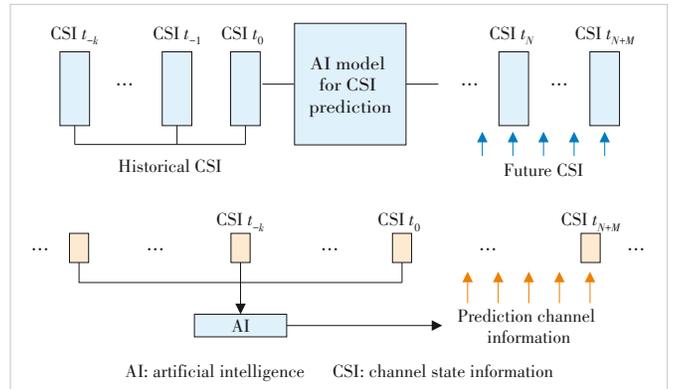
▲ Figure 2. AI-based CSI prediction
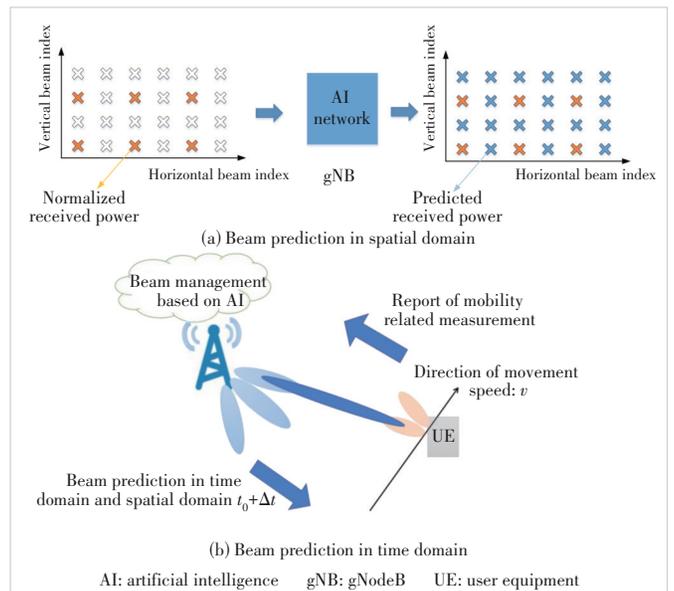
▲ Figure 3. Intelligent beam management

nals between the UE and BS have few LOS paths. AI algorithms can learn the mapping between multi-dimensional feature information and location information. We propose that in heavy NLOS indoor factory scenarios, using AI algorithms to extract and fuse multi-dimensional measurement signal features such as channel impulse response (CIR), TOA and reference signal received power (RSRP) can improve the positioning accuracy from 10 meters to decimeter level.[18]

As shown in Fig. 4, there are two ways to integrate AI with positioning technology. One is to input multi-dimensional information such as CSI, reference signal received power, and delay power spectral estimation value into the AI model, and extract and fuse features through the model to directly output the final positioning coordinates. The other method is to combine AI with traditional positioning algorithms, which is used to identify and optimize the input information of traditional positioning algorithms, thereby improving positioning accuracy.

AI models can practically be deployed in the location management function (LMF), the BS side, or the UE side for assisted positioning. When an AI model is deployed in the LMF as an input node, the BS and UE need to provide feedback on the required measurement positioning information, which introduces additional delay and overhead due to the large feedback overhead. When the AI model is deployed at the BS or UE side, the deployed node can, based on AI, select location-related measurement information from some paths with high LOS probability or optimized positioning measurement information and feedback it to the LMF. The LMF generates the accurate position coordinates of the UE, which reduces the delay and feedback overhead. As the BS and UE need to collect data for model training, the intelligent positioning enhancement technology at the BS side has a more important value of research, considering that UE has relatively weaker computing and storage resources and the mobility problem of UE will lead to the storage of a large amount of model information.

3.2 Intelligent Full Spectrum Fusion Architecture

6G full spectrum fusion will expand and improve the dimension and complexity of wireless network management. AI native can achieve intelligent full spectrum fusion for multi-tasking, solve the spectrum fragmentation problem faced by operators, and create efficient and sustainable networks.

In terms of architecture design, the basic principles are to achieve layered and distributed decomposition of complex problems through multi-level architecture. One is to virtualize physical resources such as spectrum resources, computing and storage resources through centralized and distributed computing and storage resources, and to face 6G different scenes and business needs; unified resource packaging is achieved through intelligent control methods to form virtual resource units and achieve full spectrum virtual resource fusion. The second point is to enable intelligent and flexible loading of the virtual protocol stack through AI native and shield the differences in multi-tasking in the protocol stack, frame structure and other aspects; in this way full spectrum protocol stack fusion will be achieved and channels between various protocol stack layers will be opened to achieve cross-layer optimization. The third point is to enable multi-task collaboration through AI native, to effectively integrate multi-level computing, communication, and storage resources through cloud-network-edge-terminal deep fusion, using a unified resource model and service-oriented distributed architecture, and to achieve intelligent collaborative networking of multiple frequency bands, multiple systems and multiple operators.

Intelligent full spectrum fusion can specifically include multi-access fusion, communication and sensing fusion, and fusion of multiple duplex modes. In actual network operation and management, the optimization goals of multi-spectrum fusion are multi-dimensional. The network needs to carry out multi-task collaboration to avoid strategy conflicts caused by individual decisions, thus realizing the intelligent full spectrum fusion mechanism of future 6G networks. Multi-task driven intelligent full spectrum fusion ultimately virtualizes physical resources into cloud-based resource pools, as shown in Fig. 5. Targeting various scenarios and business requirements of 6G, resources are packaged and formed into virtual space, time, frequency, and computing resource units for user allocation through intelligent regulation. Through the intelligent and flexible loading of AI native enabled virtual protocol stack, the fusion of the full spectrum protocol stack is achieved.

3.2.1 Intelligent Multi-Access Fusion

The current status of mobile communications is characterized by the coexistence of multiple generations of communication standards. It can be predicted that in the future, networks will be in a long-term situation of coexistence of 4G, 5G and 6G multi-frequency bands and multi-access standards. Emerging services pose greater challenges to the network in terms of transmission rate, delay, reliability, security, and so on. Based on the machine learning (ML) framework and utilizing various AI learning algorithms such as federated learning and rein-

(a) Direct positioning based on AI

CIR, RSRP, PDP, etc. → Positioning algorithm for LOS/NLOS recognition based on AI → UE position

(b) AI-assisted positioning

CIR, RSRP, PDP, etc. → LOS/NLOS recognition based on AI → Remove NLOS path → Traditional/AI-based positioning algorithm → UE position

AI: artificial intelligence
CIR: channel impulse response
LOS: line-of-sight
NLOS: non-line-of-sight
PDP: power delay profile
RSRP: reference signal received power
UE: user equipment

▲Figure 4. AI-based precision positioning technology
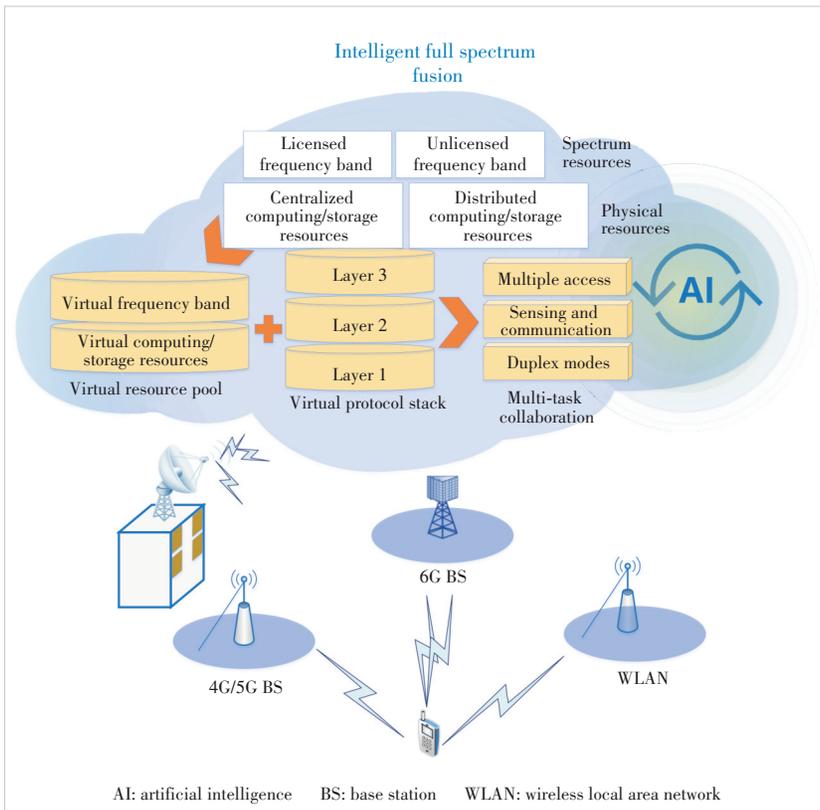
▲Figure 5. Multi-task driven intelligent full spectrum fusion

forcement learning, intelligent multi-access technology maximizes the network capacity in terms of coverage, capacity, energy efficiency and other aspects by integrating the advantages of multi-frequency bands and multi-standard networks through AI native, which improves the operational efficiency of the network.

For future business scenarios, multi-access convergence is mainly reflected in the need for AI-native solutions to the following fusion requirements:

1) With the scarcity of low-frequency wireless resources, the 6G frequency band will extend to significantly higher frequencies than 4G and 5G networks, forming a multi-frequency collaborative networking scenario. High frequencies have the advantage of large bandwidth and high transmission rates, but suffer from poor coverage due to path loss and penetration loss. Intelligent multi-frequency collaborative networking can be achieved by fusing the 6G high-frequency band with the 4G/5G network frequency band, providing both wide-area coverage and high transmission rates.

2) 6G networks and wireless local area networks (WLAN) can be intelligently integrated for indoor environments and vertical industries. When 6G and WLAN are fused at the access network level, WLAN will access the 6G base station at the wireless access network side to obtain business flows and forward them to terminals. AI-driven intelligent multi-access collaborative technology can achieve load and environment

awareness for multi-access networks, enabling intelligent flow distribution for 6G and WLAN business flows.

Finally, multi-connection of 6G+5G/4G+WLAN can achieve the fusion of advantages of all network standards. Intelligent network architecture design and AI algorithm solutions that combine centralized and distributed control are urgently needed for wireless resource management and dynamic adjustment between multiple standards to meet the requirements of future 6G businesses.

3.2.2 Intelligent Fusion of Communication and Sensing

With the development of the fusion of communication and sensing towards 6G full spectrum, intelligent wireless air interfaces will achieve multi-dimensional feature information extraction, sensing, and fusion of multiple communication signals, communication and sensing signals and multiple sensing signals. The full spectrum fusion technology for AI native and sensing will realize the mutual promotion and fusion development of intelligence and sensing, and further enhance the physical layer of 6G wireless networks and wireless resource management technology. In practical applications, AI-based positioning technology will be limited by the wireless datasets obtained by traditional communication signals. Towards 6G, emerging services such as immersive XR, metaverse and holographic communication will require higher precision positioning. Extending to the 6G full spectrum, the collaborative and accurate multi-frequency perception positioning with fusion sensing signals will greatly improve positioning accuracy and significantly promote the development of future 6G intelligent factories, autonomous driving, smart homes and other applications with wireless AI big data.

3.2.3 Intelligent Fusion of Duplex Modes

With the fusion of various intelligent wireless duplex modes, future networks may support the coexistence of multiple duplexing modes, such as traditional duplexing, flexible duplexing and full duplexing. In multi-frequency cooperative scenarios, CSI, beam direction information and terminal trajectory information working in multi-frequency, same-frequency and adjacent-frequency bands have explicit or implicit correlations. AI-driven XL-MIMO technology may use multi-frequency cooperative networking and flexible duplexing fusion mechanisms to obtain higher-dimensional feature information, further enabling AI-driven improvements in network resource utilization efficiency and system performance. In terms of intelligent CSI acquisition and intelligent beam man-

agement, intelligent pooling of multi-frequency and large bandwidth resources can be fully utilized for resource allocation and variable-rate AI model architecture. This fusion of multi-dimensional information such as geographic location and user orientation can meet intelligent high precision, high efficiency, and low-power consumption requirements of future XL-MIMO technology.

3.3 Intelligent Radio Resource Management

In 6G systems, AI native will become an engine for wireless networks to achieve self-learning, self-operation, self-maintenance and self-evolution. By leveraging multi-task collaborative optimization and intelligent ubiquitous multi-level network resource coordination, intelligent wireless resource management can be enabled, and deep perception of user scenarios and collaborative optimization of network services can be achieved, effectively enhancing wireless system performance and fully ensuring user service experience. To achieve intelligent wireless resource management, a multi-level architecture design is required to build a self-intelligent wireless network, which mainly includes the following three aspects:

1) Intelligent multi-level network resource collaborative management

AI technology enables 6G wireless networks, which can timely perceive user needs through real-time data collection and analysis, activate services on demand and carry out multi-objective joint optimization to promote the network, to achieve the best balance between system energy consumption and performance. The deep fusion of AI technology and 6G wireless networks includes the following several key points:

• Establishing an efficient data knowledge graph, designing a universal database storage model, preprocessing the collected data by classification, establishing a dataset sharing mechanism, fully mining the features in the data, and achieving the accurate perception of user needs;

• Implementing multi-scenario and multi-task collaborative optimization, which integrates and fully utilizes the data and computing resources in the network and effectively avoids duplicate data training and maximizes system performance;

• Constructing an AI model repository to improve the robustness and generalization of AI models, providing technical support for diversified network optimization services, and assisting the network in achieving self-optimization and self-evolution.

2) Network optimization with multi-task collaboration

By deeply integrating 6G wireless networks with AI technology and based on big data in wireless networks, the optimization problems in such areas as mobility management, network energy conservation and service offloading can be effectively solved. With the collection and management of global network data, AI algorithms can process and train user measurement information, network configuration parameters, service requirements and traffic, as well as external environmental in-

put information to accurately predict user movement trajectories and traffic of the BS. This can recommend more reasonable network configurations and strategies to effectively ensure the continuity of user services in the network, while reducing the energy consumption of the BS and lowering energy costs of the operator. Adaptive and predictive AI and ML algorithms can help networks establish autonomy, break through traditional technical barriers, and create high-quality intelligent 6G applications[19].

As the mobile network is composed of various overlapping network services, all service characteristics and goals need to be considered. Taking wireless network optimization as an example, the optimization goals in actual networks are multi-dimensional, so network optimization requires multi-task collaboration to avoid configuration conflicts caused by separate decisions and achieve the global optimization of the wireless communication system, thereby satisfying user needs for extreme differentiation and high-quality service experience.

3) Deep-aware network service collaborative optimization

Intelligent wireless resource management will be based on multi-dimensional information of AI native scenarios and businesses, network capability openness and native perception capability of the network to enable the collaborative optimization of the network and businesses[20]. This will realize the intelligent wireless network slicing resource management, mobility management and network energy saving.

Network slicing is a key technology for operators to support on-demand networking and resource scheduling for both business-to-business and business-to-customer services. In order to provide extremely differentiated end-to-end service quality assurance for 6G, AI-driven intelligent wireless resource management will break down the barriers between services and wireless network perception. We suggest that intelligent resource scheduling can be used to optimize wireless slice resource allocation based on the reinforcement learning framework and through learning, perception and prediction of dynamic wireless environments, application scenarios, and multi-dimensional user business information. This may significantly improve network resource utilization efficiency and ensure service experience for different priority businesses[21–23]. In addition, lightweight and distributed intelligent slice resource management utilizing algorithms such as small sample learning, transfer learning, federated learning and collective learning is an important research direction for the future, for the sake of user behavior and business privacy, information security, and distributed deployment of computing and storage resources in the network.

6G networks have higher dynamism, multi-layering and dimensionality, which leads to more frequent inter-cell handovers, posing huge challenges to 6G network mobility management. On the one hand, a cell-free network architecture design can be used to reduce the frequent interaction of inter-cell handover signaling. On the other hand, with the AI native ca-

pabilities of the network, intelligent mobile trajectory prediction can be achieved for terminals, deducing the optimal inter-cell handover scheme and thereby ensuring communication connectivity and dynamic load balancing of the network. AI can use deep reinforcement learning to solve complex decision-making problems and optimize inter-cell handover strategies in real time, minimizing transmission latency and ensuring reliable wireless connections. In the future, in large-scale vehicle-to-vehicle communication scenarios of fully autonomous driving, 6G networks need to meet the high-speed mobility and latency-sensitive requirements of vehicles. AI technology can learn user behavior, such as scenes and vehicles, through deep learning and long short-term memory (LSTM) networks, predict the future motion state and trajectory of a period of time, and effectively avoid frequent handovers and handover failures[24].

The introduction of 6G millimeter-wave and terahertz multi-frequency and high-bandwidth will increase network operation energy consumption and carbon emissions. With the increase in ultra-high-frequency BS coverage density and the diversity of business requirements of the UE, the tidal phenomenon of 6G communication demands will become more prominent, leading to large energy consumption for the network when BSs are in low-traffic periods in the ultra-dense networking scenarios. Therefore, multi-dimensional scene and business information perception at the BS side is necessary to achieve adaptive adjustment of wireless network scene businesses. Through deep learning and LSTM based on the spatiotemporal correlation of historical demands of the BS, the future business and traffic can be predicted[25], and a reinforcement learning network can be used to design semi-static or dynamic BS switch-off energy-saving states. Through a dynamic balance between the user business guarantee and network energy-saving multi-BS collaborative intelligent networking strategy, overall network energy consumption can be reduced.

4 Conclusions

In this paper, we propose a wireless AI technology architecture with multi-dimensional information perception to solve the key challenge to developing AI native 6G systems, which includes multi-dimensional feature information perception technology in the physical layer, intelligent full spectrum fusion technology, and intelligent wireless resource management.

At present, ongoing studies in 3GPP focus on integrating AI/ML with traditional physical layer techniques and radio access network level signaling enhancement. This paper provides our thoughts and a guideline on the roadmap of the evolution of 5G-Advanced to the intelligent 6G wireless network. The time left to study and clarify the integration of AI/ML with B5G or design an AI-native 6G RAN is limited. Researching the proposed intelligent radio resource management and intelligent full spectrum fusion technology such as distributed learning-based technologies and AI/ML-based mobility en-

hancement is crucial for better preparation for 6G. Relevant research results show that AI and ML are powerful tools for improving the performance of existing or future wireless communication systems. It will be exciting to see the development of 6G wireless systems in the coming years, which will be built upon the foundation of the proposed application scenarios and wireless AI technology architecture.

References

[1] IMT-2030 (6G) Promotion Group. 6G overall vision and potential key technologies white paper [R]. 2021

[2] IMT-2030 (6G) Promotion Group. 6G network architecture vision and key technology outlook white paper [R]. 2021

[3] YLIANTTILA M, KANTOLA R, GURTOV A, et al. 6G white paper: research challenges for trust, security and privacy. University of Oulu: 6G Research Visions, No.9 [R]. 2020

[4] China Center for Information Industry Development Think Tank. 6G concept and vision white paper [R]. 2020

[5] Datang Mobile Communication Equipment Co., Ltd. Global coverage and intelligent connection of scenarios: 6G vision and technology trends white paper [R]. 2020

[6] China Mobile Communications Corporation Research Institute. White Paper: 2030+ vision and requirements [R]. 2020

[7] VIVO Communication Research Institute. 6G vision, needs, and challenges [R]. 2020

[8] DE LIMA C, BELOT D, BERKVENS R, et al. 6G white paper on localization and sensing. University of Oulu: 6G Research Visions, No.12 [R]. 2020

[9] YOU X H, WANG C X, HUANG J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts [J]. Science China information sciences, 2021, 64(1): 110301. DOI: 10.1007/s11432-020-2955-6

[10] Samsung. 6G: the next hyper-connected experience for all [R]. 2020

[11] VISWANATHAN H, MOGENSEN P E. Communications in the 6G era [J]. IEEE access, 2020, 8: 57063 – 57074. DOI: 10.1109/access.2020.2981745

[12] UNISOC. 6G: boundless, but with artificial intelligence [R]. 2020

[13] IOWN. Innovative optical and wireless network global forum vision 2030 and technical directions [R]. 2020

[14] Next G Alliance. Next G alliance report: roadmap to 6G [R]. 2022

[15] Hexa-X. Initial 6G architectural components and enablers [R]. 2021

[16] 3GPP TSG RAN1#110bis-e. Evaluation on AI/ML for CSI feedback enhancement: R1-2208769 [R]. China Telecom, 2022

[17] 3GPP TSG RAN1#110bis-e. Evaluation on AI/ML for beam management: R1-2208771 [R]. China Telecom, 2022

[18] 3GPP TSG RAN1#110 bis-e. Evaluation on AI/ML for positioning accuracy enhancement: R1-2206689 [R]. China Telecom, 2022

[19] YOU X H, WANG C X, HUANG J, et al. White paper: 6G wireless network: vision, enabling technology and new application paradigm [R]. 2020

[20] YANG F Y, LIU Y, YANG B. Some thoughts on 6G networks [J]. ZTE technology journal, 2021, 27(2): 2 – 5. DOI: 10.12142/ZTETJ.202102002

[21] YANG B, XU Y Q, SHE X M, et al. Reinforcement learning based dynamic resource allocation for massive MTC in sliced mobile networks [C]//14th International Conference on Advanced Infocomm Technology (ICAIT). IEEE, 2022: 298 – 303. DOI: 10.1109/ICAIT56197.2022.9862802

[22] YANG B, LI Q Y, SHE X M, et al. Intelligent dynamic resource allocation for network slicing based on actor-critic algorithm [C]//14th International Conference on Advanced Infocomm Technology (ICAIT). IEEE, 2022: 334 – 340. DOI: 10.1109/ICAIT56197.2022.9862788

[23] YANG B, WEI F S, SHE X M, et al. Intelligent random access for massive-machine type communications in sliced mobile networks [J]. Electronics, 2023, 12(2): 329. DOI: 10.3390/electronics12020329

[24] YANG H L, ALPHONES A, XIONG Z H, et al. Artificial-intelligence-enabled intelligent 6G networks [J]. IEEE network, 2020, 34(6): 272 – 280. DOI: 10.1109/MNET.011.2000195

[25] WU Q, CHEN X, ZHOU Z, et al. Deep reinforcement learning with spatio-temporal traffic forecasting for data-driven base station sleep control [J]. IEEE/ACM transactions on networking, 2021, 29(2): 935 – 948. DOI: 10.1109/TNET.2021.3053771

Biographies

YANG Bei (yangbei1@chinatelecom.cn) received her MS degree in electronics and communication engineering from Beijing University of Post and Telecommunications, China in 2014. Her current research interests include artificial intelligence for next-generation radio access networks, network slicing, and industrial IoT technology.

LIANG Xin received his BS degree in information engineering from Beijing University of Posts and Telecommunications (BUPT), China in 2019. He is currently pursuing his PhD degree in information and communication engineering at BUPT. He has been working for China Telecom since July, 2022. His current research interests include wireless channel feedback, channel modeling, and deep learning.

LIU Shengnan received her master's degree in communication and information system from Beijing University of Posts and Telecommunications, China. She is an intermediate engineer with China Telecom Research Institute where she engages in 5G standard and 6G technology related research. Her research interests include mobility enhancement, SON/MDT enhancement, and AI for NG-RAN.

JIANG Zheng received his PhD degree in communication and information system from Beijing University of Posts and Telecommunications, China in 2008. He is currently a professor-level senior engineer with China Telecom Research Institute. His research interests include RAN architecture, MIMO, and conjunction of sensing and communication for 6G.

ZHU Jianchi received his MS degree in communication and information system from Beijing University of Posts and Telecommunications, China in 2007. He is currently a senior engineer with China Telecom Research Institute. His research interests include 5G standardization in 3GPP RAN1 and cutting-edge technology for 6G.

SHE Xiaoming received his PhD degree in electronic engineering from Tsinghua University, China in 2004. He is currently a professor-level senior engineer with China Telecom Research Institute. His research interests include 5G standardization in 3GPP RAN and cutting-edge technology for 6G.

Deep Learning-Based Semantic Feature Extraction: A Literature Review and Future Directions

DENG Letian, ZHAO Yanru

(Northwest Agriculture and Forestry University, Xianyang 712100, China)

Abstract: Semantic communication, as a critical component of artificial intelligence (AI), has gained increasing attention in recent years due to its significant impact on various fields. In this paper, we focus on the applications of semantic feature extraction, a key step in the semantic communication, in several areas of artificial intelligence, including natural language processing, medical imaging, remote sensing, autonomous driving, and other image-related applications. Specifically, we discuss how semantic feature extraction can enhance the accuracy and efficiency of natural language processing tasks, such as text classification, sentiment analysis, and topic modeling. In the medical imaging field, we explore how semantic feature extraction can be used for disease diagnosis, drug development, and treatment planning. In addition, we investigate the applications of semantic feature extraction in remote sensing and autonomous driving, where it can facilitate object detection, scene understanding, and other tasks. By providing an overview of the applications of semantic feature extraction in various fields, this paper aims to provide insights into the potential of this technology to advance the development of artificial intelligence.

Keywords: semantic feature extraction; semantic communication; deep learning

1 Introduction

Artificial intelligence (AI) has become a rapidly growing field in recent years, with semantic communications being a critical component of AI[1 – 3]. Semantic communication has gained increasing attention in recent years due to its significant impact on various fields[4 – 7]. It involves the recognition of semantics and understanding of human language, thereby enabling faster and more accurate transmission of information. The significance of the semantic communication lies in its ability to improve communication efficiency, provide more accurate information, and convey intentions more effectively. Deep learning-based semantic feature extraction plays a critical role in enabling effective communications by extracting meaningful features from data and encoding them in a way that can be easily transmitted and interpreted by agents[7 – 8]. Its importance has been recognized in a wide range of AI domains, including natural language processing (NLP), medical imaging, remote sensing, and autonomous driving.

This survey paper focuses on semantic feature extraction, which is a key step in semantic communications, providing a comprehensive literature review of its applications in various AI domains. This paper aims to review the current research status and development trend of deep learning-based semantic feature extraction.

Firstly, we introduce the concept and research background of semantic feature extraction, including the technical basis of speech interaction, NLP, sentiment computing, knowledge graph and machine translation[9 – 12].

Secondly, we introduce in detail the current research status of deep learning-based semantic feature extraction applications, including speech interaction applications based on speech recognition, intelligent dialogue systems based on NPL, sentiment computing applications based on sentiment analysis, intelligent question answering systems based on knowledge graph and cross-language interaction applications based on machine translation[13 – 18].

Finally, we discuss the limitations of AI-driven semantic communication applications in actual applications and future development directions, including technical difficulties, data problems, security and privacy issues, as well as the application scenarios and commercial value of AI-driven semantic communication applications in the future.

2 Deep Learning-Based Semantic Feature Extraction Methods in Multiple Fields

AI technology, such as deep learning methods, plays a significant role in semantic feature extraction. The process of semantic feature extraction can be divided into two categories: the NLP-based and image-based.

NLP-based semantic feature extraction involves computer processing of natural language, including speech recognition, text classification, named entity recognition, part-of-speech tagging, dependency analysis, and semantic role labeling[9 – 12]. It also includes knowledge graphs that provide a graphical representation of knowledge consisting of entities, attributes, and relationships[13 – 15]. Additionally, sentiment analysis techniques are used to analyze sentiment information contained in text, including sentiment recognition and sentiment analysis[16 – 18]. Semantic representation techniques focus on converting natural language into a form that can be processed by computers, such as word vector representation, sentence vector representation, semantic role labeling, and semantic dependency analysis[19 – 22]. Together, these technologies form the technical system of semantic communication and enable computers to better understand the meaning and semantic information of natural language, leading to better semantic communications.

In the computer vision, semantic communication technologies include image recognition, target detection, visual semantic segmentation, image classification, image retrieval, and image clustering[23 – 24]. Image recognition utilizes machine learning and deep learning to recognize and classify objects in images[25 – 26]. Target detection involves using machine learning to detect specific targets in images, such as faces or text, and identifying their characteristics[27]. Visual semantic segmentation involves segmenting objects in images into different categories, shapes, and colors[28 – 30]. Image classification allocates images to different categories based on the objects present in them. Image retrieval retrieves images related to user input from an image database[31 – 32]. Image clustering involves grouping similar images according to object shape, color, and other characteristics to improve image management[33 – 34]. These technologies leverage deep learning and machine learning to improve the accuracy and efficiency of image processing.

Next, we will describe in detail the application of semantic feature extraction in five explicit scenarios and a summary is shown in Table 1.

2.1 Natural Language Processing

Semantic feature extraction plays a crucial role in natural language processing and has various applications across numerous downstream tasks[35 – 36]. These tasks include information extraction[37 – 38], sentiment analysis[39 – 40], and knowledge graph construction[41]. Gaining a better understanding of users' semantic expressions allows for more efficient processing of their queries and accurate information provision. In named entity recognition (NER)[42], for instance, it is essential to identify entities within sentences. Generally, these entities exhibit distinct semantic features, such as being nouns. By accurately recognizing these features, more effective NER methods can be developed. Similarly, sentiment analysis relies heavily on semantic feature extraction. This process involves identifying mentions (typically nouns) and classifying their sentiment polarities (usually adjectives). Part-of-speech tagging serves as an explicit semantic feature, while implicit features can also be calculated using deep learning techniques based on word embeddings. In their study, BAO et al.[43] proposed a deep learning-based sentiment analysis method, employing a meta-based self-training approach with a meta-weighter. They trained a teacher model to generate in-domain knowledge (semantic features) for supervised learning, using the generated pseudo-labels in a student model. Knowledge graph construction often encompasses multiple tasks, such as named entity recognition[44], relation extraction[45], and anaphora resolution[46]. In their research, HE et al.[41] proposed a multi-task framework for constructing knowledge graphs. They employed a shared encoder to extract common semantic features for all mentioned tasks. More specifically, they developed an end-to-end information extraction system using a multitask-based artificial neural network model for constructing genealogical knowledge graphs from online obituaries. In conclusion, the application of semantic feature extraction in the field of natural language processing enhances its intelligence, facilitating a deeper understanding of users' semantic expressions. This, in turn, enables more efficient processing of user queries and the provision of more accurate information.

2.2 Hyperspectral Image Analysis

Semantic communication in the application of remote sensing images is a technique used to extract information from data acquired by hyperspectral imaging sensors[47]. Hyperspectral imaging sensors collect data in many narrow, contiguous spectral bands, essentially producing a 3D data cube with two spatial dimensions and one spectral dimension[48]. Hyperspectral image analysis involves processing this data cube to extract information about the materials or objects in the imaged scene. The main techniques of hyperspectral image analysis include spectral unmixing, classification, and anomaly detection. Firstly, spectral unmixing separates the contributions of different materials in each pixel of the image. The work in Ref. [49] proposed a dynamical model for unmixing a time series of hyperspectral images, with a simplified version of the model used to derive an efficient spectral unmixing algorithm that is demonstrated on synthetic and real multi-temporal hyperspectral images. The work in Ref. [50] proposed a new method for solving the sparse hyperspectral unmixing problem without relaxation, using a multi-objective optimization approach and a binary coding technique, which is demonstrated to be effective on both synthetic and real hyperspectral datasets. Secondly, the hyperspectral image classification assigns

▼Table 1. Semantic communication application scenarios

Scenario	Task	Method	Model Name	Model Structure	Year
Hyperspectral image	Classification	Markov random field	NE-MFAS[51]	SVM + MRF	2017
		Group sparse coding	MSKGSC[53]	Kernel sparse representation	2016
		Semantic manifold learning	MFAS[52]	SVM + MRF	2016
	Object detection	Multi-view noisy learning	MOL[71]	VGG16 + MIL	2023
		Few-shot learning	AUD-Net[54]	ResNet + multi-head attention	2022
Medical field	Classification	Vision transformer	i-ViT[66]	Vision transformer	2021
		Multi-task learning	MTL-CRD[68]	A semi-supervised multi-task learning framework	2023
		Uncertainty-based model	UMA[69]	Uncertainty-based model acceleration	2022
		Multi-instance learning	HIB[67]	Information bottleneck + hierarchical multi-instance learning	2022
	Segmentation	Segmentation network	W-Net[70]	Composite high-resolution network	2021
	Clinical prediction	Representation learning	CSEDrug[61]	Pretrain + RNN	2022
		Information theory	DAPSNet[60]	RNN + attention + information bottleneck	2023
Natural language processing	Sentiment analysis	Prompt tuning	Survey[40]	Transformer	2022
		Meta-learning	MSM[39]	Teacher-student/BERT	2022
		Meta-learning	MLB[43]	BERT	2021
	Knowledge graph	Multi-task learning	MTL-2[41]	Bi-LSTM	2021
		Multi-task learning	MTL-1[37]	Bi-LSTM	2019
	Named entity	Few-shot learning	Copner[42]	BERT	2022
	Relation extraction	Prompt tuning	VPP[36]	BERT	2023
		Semi-supervised learning	UG-MCT[45]	BERT	2022
		Continual learning	JCBIE[44]	BERT	2022

AUD-Net: a unified detector anomaly detection network
BERT: Bidirectional Encoder Representations from Transformers
Bi-LSTM: Bidirectional LSTM
CRD: cancer region detection
CSEDrug: a comprehensive DDI controllable model
DAPSNet: Dual Attention and Patient Similarity Network
HIB: a multi-instance learning model
I-ViT: an integer-only quantization scheme for vision transformers
JCBIE: joint continual learning biomedical information extraction
LSTM: long short-term memory
MFAS: Multimodal Fusion Architecture Search
MIL: multiple instance learning
MLB: a meta-learning model

MOL: multi-view noisy learning
MRF: modified random forest
MSKGSC: a group sparse coding model
MSM: meta-based self-training method with a meta-weighter
MTL: multi-task learning
NE: network element
RNN: recurrent neural network
SVM: support vector machine
UG-MCT: Uncertainty-Guided Mutual Consistency Training framework
UMA: uncertainty-based model acceleration
VGG16: a convolution neural network architecture
VPP: virtual prompt pre-training model
W-Net: a segmentation network

each pixel to a particular material or object class. The work in Refs. [51] and [52] proposed a method for improving hyperspectral image classification performance by combining multiple features in the same semantic space with local and nonlocal spatial constraints using an extended Markov random field model. The work in Ref. [53] proposed a new method for hyperspectral image classification using adaptive spatial partition of pixels into clusters via group sparse coding that integrates spectral and spatial information to improve classification accuracy and provide distinctive classification maps. Lastly, anomaly detection identifies pixels with unusual spectral characteristics that may indicate the presence of a target material. The work in Ref. [54] proposed a unified detector anomaly detection network (AUD-Net) inspired by few-shot learning to perform anomaly detection across multiple hyperspectral images (HSIs) without repeated training, which addresses the challenges of generalization to different HSIs with different spectral sizes and achieves strong generalization. Hyperspectral image analysis has applications in a wide range of fields, including remote sensing, environmental monitoring, mineral exploration, agriculture, and biomedical imaging. In conclusion, the application of semantic communications in remote sensing images can greatly enhance the user's understanding and provide more accurate information, making the remote sensing image analysis more intelligent and thus better managing the local resources.

2.3 Clinical Application

In recent years, the application of semantic communication technology in the medical field has become increasingly popu-

lar[55 – 57]. In Refs. [55] and [56], an attention-based graph convolutional network (GCN) was proposed to convert unstructured pathological reports into structured data for computer analysis, improving pathologists' workflow and providing more accurate assistance for diagnosis and treatment, with promising results demonstrated on a dataset from TCGA. Semantic communication technology can also be used for intelligent diagnosis, which can use a large amount of case data to carry out intelligent diagnosis according to the patient's medical history and examination results, thus providing doctors with more accurate diagnosis results[58 – 59]. The work in Ref. [58] proposed a general healthcare representation model that uses simple convolution operations and up/down sampling to adaptively extract distinct individual key features, achieving superior performance and model complexity compared to other baseline models on the MIMIC-III dataset. In addition, semantic communication technology can be used for intelligent treatment, which can provide more effective treatment plans according to the patient's medical history and examination results, thus more effectively treating patients[60 – 61]. Two drug recommendation studies[60 – 61] proposed a drug combination prediction model that leverages multifaceted drug knowledge and loss functions to improve drug encoding and drug-drug interaction (DDI) control, achieving superior accuracy, effectiveness and safety compared to state-of-the-art methods. Moreover, doctors can better understand the condition of patients and provide more effective treatment plans for them by using semantic communication technology to query relevant cases according to the patient's medical history and examination results. Semantic communication technology can also be used for intelligent drug development, which can use a large amount of drug data to find new drugs according to the patient's medical history and examination results, thus more effectively treating patients[62 – 63]. The work in Ref. [64] proposed a framework that combines pathological images and medical reports to generate a personalized diagnosis result for an individual patient, using nuclei-level image feature similarity and content-based deep learning methods to extract structured prognostic information and assign importance to different factors, with promising results demonstrated on TCGA data for renal cell carcinoma. In conclusion, the application of semantic communication technology in the medical field can improve medical efficiency and help treat patients better, providing more effective treatment plans for patients and making medical care more intelligent.

2.4 Medical Image Analysis

Semantic communication is a critical component in medical image analysis, where accurate interpretation and communication of medical image findings are crucial for clinical diagnosis and decision-making. In medical image analysis, semantic communication involves the ability to convey the meaning of medical image features and findings, such as the presence of

tumors, lesions, or abnormalities, to clinical experts and other stakeholders. Semantic communication is especially important in cases where medical images need to be interpreted by multiple experts or in cross-institutional settings, where the interpretation of medical images may vary due to differences in expertise or experience[62]. Therefore, the development of semantic communication methods and tools for medical image analysis has become an active research area in recent years. Deep learning-based semantic segmentation and classification algorithms have shown great potential in enabling accurate and efficient semantic communication in medical image analysis. By leveraging the power of deep learning, these algorithms can automatically extract and classify medical image features, and provide intuitive visual representations of the underlying medical conditions. These approaches have the potential to greatly improve the efficiency and accuracy of medical image analysis and interpretation, and ultimately enhance patient care and outcomes. Based on the different granularity of semantic feature extraction, medical image analysis can be divided into classification, detection, and segmentation. Classification involves assigning a label or category to an image or region of interest based on its characteristics, which is used in tasks such as tumor diagnosis and tissue classification[65]. The work in Ref. [66] proposed an instance-based vision transformer to learn robust representations of histopathological images for the pRCC subtyping task by extracting finer features from instance patches and capturing both cellular and cell-layer level patterns by position-embedding, grade-embedding, and self-attention. The work in Ref. [67] proposed a hierarchical Multi-Instance Learning (MIL) framework with an Information Bottleneck (IB) to handle patient-level labels and exploit the correlation among leukemia subtypes for better accuracy and generalization in childhood acute leukemia classification without the need for cell-level annotations. Detection involves identifying the presence or location of specific features or anomalies in an image, which is used in tasks such as lesion detection and localization. The work in Ref. [68] proposed a semi-supervised multi-task learning framework for whole slide image classification to improve the performance on both cancer region detection and subtyping tasks by capturing the interaction of the two tasks, and to preserve the sequential relationship of the tasks using a weight control mechanism, demonstrated to be effective in accuracy and generalization on four large datasets with different cancer types. Segmentation is the process of dividing an image into multiple regions or segments based on their characteristics, which is used to identify and isolate specific structures or features in medical images. The work in Ref. [69] proposed a contrastive learning framework with multi-granularity views for tissue segmentation by designing three contrastive learning tasks from global to local, which can capture fine-grained patterns in the learned representations for transfer learning to various tissue segmentation tasks in histopathological images, demonstrated to be superior to exist-

ing contrastive learning methods. The work in Ref. [70] proposed a Composite High-Resolution Network for nuclei grading, i.e., a special task of nuclei segmentation and classification, which includes a W-Net for segmentation, two high-resolution feature extractors for nuclei classification, and a head-fusion block for label generation.

2.5 Autonomous Driving

The field of autonomous driving has gained significant attention due to the continuous development of artificial intelligence technology. Semantic communication technology has emerged as a crucial factor in this field.

To enable natural dialogue between passengers and autonomous or driverless cars, these vehicles must possess the ability to comprehend natural language. This necessitates the use of semantic analysis technology, which can identify the actual intention and emotional tendencies of language, leading to better passenger satisfaction. Furthermore, autonomous and driverless cars must be capable of perceiving their surrounding environment, identifying various traffic signs, and adhering to driving rules on the road. This requires significant support from semantic modeling and computer vision technology to ensure that the vehicles respond appropriately to traffic signals, pedestrians, and other vehicles. In case of an emergency, voice alarms can assist drivers and passengers in reacting quickly. Voice alarms can warn passengers of problems and allow for further operations based on their responses. All in all, semantic communication technology plays an indispensable role in the development of autonomous driving and driverless cars. It enhances the ability of these vehicles to understand passenger needs, perceives the surrounding environment accurately, and provides timely alarms to passengers in emergencies. These factors provide the foundation for the future development of intelligent transportation.

3 Development Trends

The progress and extensive application of artificial intelligence technology have significantly supported the research and implementation of semantic communication applications. Intelligent voice interaction is projected to become the standard for AI-driven semantic communication applications, with voice technology continuously advancing and proliferating. Additionally, multi-modal interaction will become more prevalent as multi-modal technology is applied to different media, offering more diverse interactions. Personalized interaction will be prioritized by analyzing user preferences and needs. Further, semantic understanding and generation technologies, along with knowledge graphs, will continue to advance, making human-computer interaction more natural and intelligent. While semantic communication applications will bring more convenience and efficiency to daily life, work, and learning, they are also anticipated to encounter new challenges, such as security and semantic ambiguity. In conclusion, the continuous development

and expansion of semantic communication applications will provide both opportunities and challenges for society.

4 Challenges

Semantic communication applications powered by artificial intelligence encounter several challenges that require resolution. Firstly, there is a lack of profound semantic comprehension, making it difficult for the application to fully understand the user's expression, which may result in potential misinterpretation. Secondly, the application's self-learning ability is limited, which restricts its capacity to increase knowledge from user input. Thirdly, the application's computational complexity and memory capacity limitations may negatively affect its processing capabilities. Lastly, the application's weak anti-interference ability makes it vulnerable to external interference and noise.

To address these challenges, the application must have a more accurate and profound understanding of the user's semantic expression and the ability to learn from user input. It must also enhance its computational complexity and memory capacity to process more information in less time and have a stronger anti-interference ability.

Several approaches have been developed to tackle these challenges, including NLP techniques to enhance semantic comprehension, deep learning models such as convolutional neural networks (CNNs) for self-learning, distributed computing and cloud computing to increase computational complexity and memory capacity, and noise cancellation and signal processing to improve anti-interference ability.

In conclusion, addressing these challenges will enhance the application's ability to provide a better user experience, bringing convenience and value to various aspects of life. The difficulties encountered by semantic communication applications driven by artificial intelligence can be addressed through various approaches, including NLP, deep learning, distributed computing, cloud computing, noise cancellation, and signal processing

5 Conclusions

This paper provides a review of semantic communication applications powered by AI, examining its background, technology, applications, and development trends. Semantic communication applications have become an essential direction for the development of AI technology, with widespread applications and significant impact. They enhance human-computer interaction efficiency and quality, improve user experience, and address issues such as language barriers and information overload, bringing convenience and value to all aspects of life. As technology advances, semantic communication applications will become increasingly intelligent, natural, personalized, and multimodal, promoting AI technology development in broader fields. However, challenges such as user privacy protection and semantic ambiguity require continuous

innovation and improvement to ensure healthy application development. In conclusion, semantic communication applications will be extensively adopted in various fields, bringing convenience and transforming people's lives and work in the near future.

References

[1] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521 (7553): 436 – 444. DOI: 10.1038/nature14539

[2] HAMET P, TREMBLAY J. Artificial intelligence in medicine [J]. Metabolism, 2017, 69: S36 – S40. DOI: 10.1016/j.metabol.2017.01.011

[3] DICK S. Artificial intelligence [J]. Harvard data science review, 2019, 1(1):1 – 8. DOI: 10.1162/99608f92.92fe150c

[4] BAO J, BASU P, DEAN M K, et al. Towards a theory of semantic communication [C]//IEEE Network Science Workshop. IEEE, 2011: 110 – 117. DOI: 10.1109/nsw.2011.6004632

[5] LUO X W, CHEN H H, GUO Q. Semantic communications: overview, open issues, and future research directions [J]. IEEE wireless communications, 2022, 29(1): 210 – 219. DOI: 10.1109/MWC.101.2100269

[6] QIN Z J, TAO X M, LU J H, et al. Semantic communications: principles and challenges [EB/OL]. (2021-12-30) [2023-03-01]. https://arxiv. org/abs/2201.01389

[7] XIE H Q, QIN Z J, LI G Y, et al. Deep learning enabled semantic communication systems [J]. IEEE transactions on signal processing, 2021, 69: 2663 – 2675. DOI: 10.1109/TSP.2021.3071210

[8] YANG W T, DU H Y, LIEW Z, et al. Semantic communications for future Internet: fundamentals, applications, and challenges [J]. IEEE communications surveys & tutorials, 2022, 25: 213 – 250. DOI: 10.1109/COMST.2022.3223224

[9] CHOWDHARY K R. Natural language processing [M]//Fundamentals of artificial intelligence. New Delhi: Springer India, 2020: 603 – 649. DOI: 10.1007/978-81-322-3972-7_19

[10] NADKARNI P M, OHNO-MACHADO L, CHAPMAN W W. Natural language processing: an introduction [J]. Journal of the American medical informatics association, 2011, 18(5): 544 – 551. DOI: 10.1136/amiajnl-2011-000464

[11] HIRSCHBERG J, MANNING C D. Advances in natural language processing [J]. Science, 2015, 349(6245): 261 – 266. DOI: 10.1126/science.aaa8685

[12] MANNING C D, SCHUTZE H. Foundations of statistical natural language processing [M]. Cambridge: MIT Press, 1999

[13] CHEN X J, JIA S B, XIANG Y. A review: knowledge reasoning over knowledge graph [J]. Expert systems with applications, 2020, 141: 112948. DOI: 10.1016/j.eswa.2019.112948

[14] CHEN Z, WANG Y H, ZHAO B, et al. Knowledge graph completion: a review [J]. IEEE access, 2020, 8: 192435 – 192456. DOI: 10.1109/ACCESS.2020.3030076

[15] ZOU X H. A survey on application of knowledge graph [J]. Journal of physics: conference series, 2020, 1487(1): 012016. DOI: 10.1088/1742-6596/1487/1/012016

[16] MEDHAT W, HASSAN A, KORASHY H. Sentiment analysis algorithms and applications: a survey [J]. Ain shams engineering journal, 2014, 5(4): 1093 – 1113. DOI: 10.1016/j.asej.2014.04.011

[17] ZHANG L, WANG S, LIU B. Deep learning for sentiment analysis: a survey [J]. Wiley interdisciplinary reviews: data mining and knowledge discovery, 2018, 8(4): e1253. DOI: 10.1002/widm.1253

[18] HUSSEIN D M E D M. A survey on sentiment analysis challenges [J]. Journal of king Saud university: engineering sciences, 2018, 30(4): 330 – 338. DOI: 10.1016/j.jksues.2016.04.002

[19] GRIFFITHS T L, STEYVERS M, TENENBAUM J B. Topics in semantic representation [J]. Psychological review, 2007, 114(2): 211 – 244. DOI: 10.1037/0033-295x.114.2.211

[20] VIGLIOCCO G, VINSON D P. Semantic representation [M]//The oxford handbook of psycholinguistics, GASKELL M G ed. Oxford: Oxford University Press, 2012: 195 – 216. DOI: 10.1093/oxfordhb/9780198568971.013.0012

[21] ABEND O, RAPPOPORT A. The state of the art in semantic representation [C]//

55th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2017: 77 – 89. DOI: 10.18653/v1/p17-1008

[22] VIGLIOCCO G, METEYARD L, ANDREWS M, et al. Toward a theory of semantic representation [J]. Language and cognition, 2009, 1(2): 219 – 247. DOI: 10.1515/langcog.2009.011

[23] VOULODIMOS A, DOULAMIS N, DOULAMIS A, et al. Deep learning for computer vision: a brief review [J]. Computational intelligence and neuroscience, 2018: 1 – 13. DOI: 10.1155/2018/7068349

[24] WILEY V, LUCAS T. Computer vision and image processing: a paper review [J]. International journal of artificial intelligence research, 2018, 2(1): 22. DOI: 10.29099/ijair.v2i1.42

[25] PAK M, KIM S. A review of deep learning in image recognition [C]//4th International Conference on Computer Applications and Information Processing Technology (CAIPT). IEEE, 2018: 1 – 3. DOI: 10.1109/CAIPT.2017.8320684

[26] UCHIDA S. Image processing and recognition for biological images [J]. Development, growth & differentiation, 2013, 55(4): 523 – 549. DOI: 10.1111/dgd.12054

[27] LONG T, LIANG Z N, LIU Q H. Advanced technology of high-resolution radar: target detection, tracking, imaging, and recognition [J]. Science China information sciences, 2019, 62(4): 1 – 26. DOI: 10.1007/s11432-018-9811-0

[28] THOMA M. A survey of semantic segmentation [EB/OL]. (2016-02-21)[2023-03-01]. https://arxiv.org/abs/1602.06541

[29] HAO S J, ZHOU Y, GUO Y R. A brief survey on semantic segmentation with deep learning [J]. Neurocomputing, 2020, 406: 302 – 321. DOI: 10.1016/j.neucom.2019.11.118

[30] LATEEF F, RUICHEK Y. Survey on semantic segmentation using deep learning techniques [J]. Neurocomputing, 2019, 338: 321 – 348. DOI: 10.1016/j.neucom.2019.02.003

[31] LU D, WENG Q. A survey of image classification methods and techniques for improving classification performance [J]. International journal of remote sensing, 2007, 28(5): 823 – 870. DOI: 10.1080/01431160600746456

[32] NATH S S, MISHRA G, KAR J, et al. A survey of image classification methods and techniques [C]//International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT). IEEE, 2014: 554 – 557. DOI: 10.1109/ICCICCT.2014.6993023

[33] NAZ S, MAJEED H, IRSHAD H. Image segmentation using fuzzy clustering: a survey [C]//6th International Conference on Emerging Technologies (ICET). IEEE, 2010: 181 – 186. DOI: 10.1109/ICET.2010.5638492

[34] DHANACHANDRA N, CHANU Y J. A survey on image segmentation methods using clustering techniques [J]. European journal of engineering and technology research, 2017, 2(1): 15 – 20. DOI: 10.24018/ejeng.2017.2.1.237

[35] HE K, MAO B, ZHOU X Y, et al. Knowledge enhanced coreference resolution via gated attention [C]//IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2023: 2287 – 2293. DOI: 10.1109/BIBM55620.2022.9995551

[36] HE K, HUANG Y C, MAO R, et al. Virtual prompt pre-training for prototype-based few-shot relation extraction [J]. Expert systems with applications, 2023, 213: 118927. DOI: 10.1016/j.eswa.2022.118927

[37] HE K, WU J L, MA X Y, et al. Extracting kinship from obituary to enhance electronic health records for genetic research [C]//4th Social Media Mining for Health Applications (#SMM4H) Workshop & Shared Task. Association for Computational Linguistics, 2019: 1 – 10. DOI: 10.18653/v1/w19-3201

[38] WANG H N. Development of natural language processing technology [J]. ZTE technology journal, 2022, 28(2): 59 – 64. DOI: 10.12142/ZTETJ.202202009

[39] HE K, MAO R, GONG T L, et al. Meta-based self-training and re-weighting for aspect-based sentiment analysis [J]. IEEE transactions on affective computing, 2022, PP(99): 1 – 13. DOI: 10.1109/TAFFC.2022.3202831

[40] MAO R, LIU Q, HE K, et al. The biases of pre-trained language models: an empirical study on prompt-based sentiment analysis and emotion detection [J]. IEEE transactions on affective computing, 2022, early access: 1 – 11. DOI: 10.1109/TAFFC.2022.3204972

[41] HE K, YAO L X, ZHANG J W, et al. Construction of genealogical knowledge graphs from obituaries: multitask neural network extraction system [J]. Journal of medical Internet research, 2021, 23(8): e25670. DOI: 10.2196/25670

[42] HUANG Y C, HE K, WANG Y G, et al. COPNER: contrastive learning with prompt guiding for few-shot named entity recognition [C]//29th International Conference on Computational Linguistics. International Committee on Computational Linguistics, 2022: 2515 – 2527

[43] BAO H, HE K, YIN X M, et al. BERT-based meta-learning approach with looking back for sentiment analysis of literary book reviews [M]//Natural language processing and chinese computing. Cham: Springer International Publishing, 2021: 235 – 247. DOI: 10.1007/978-3-030-88483-3_18

[44] HE K, MAO R, GONG T L, et al. JCBIE: a joint continual learning neural network for biomedical information extraction [J]. BMC bioinformatics, 2022, 23 (1): 549. DOI: 10.1186/s12859-022-05096-w

[45] MAO B, JIA C, HUANG Y C, et al. Uncertainty-guided mutual consistency training for semi-supervised biomedical relation extraction [C]//IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022: 2318 – 2325. DOI: 10.1109/bibm55620.2022.9995416

[46] LI Y F, MA X Y, ZHOU X Y, et al. Knowledge enhanced LSTM for coreference resolution on biomedical texts [J]. Bioinformatics, 2021, 7(17): 2699 – 2705. DOI: 10.1093/bioinformatics/btab15

[47] AMIGO J M, BABAMORADI H, ELCOROARISTIZABAL S. Hyperspectral image analysis: a tutorial [J]. Analytica chimica acta, 2015, 896: 34 – 51. DOI: 10.1016/j.aca.2015.09.030

[48] KHAN M J, KHAN H S, YOUSAF A, et al. Modern trends in hyperspectral image analysis: a review [J]. IEEE access, 2018, 6: 14118 – 14129. DOI: 10.1109/ACCESS.2018.2812999

[49] HENROT S, CHANUSSOT J, JUTTEN C. Dynamical spectral unmixing of multitemporal hyperspectral images [J]. IEEE transactions on image processing, 2016, 25(7): 3219 – 3232. DOI: 10.1109/TIP.2016.2562562

[50] XU X, SHI Z W. Multi-objective based spectral unmixing for hyperspectral images [J]. ISPRS journal of photogrammetry and remote sensing, 2017, 124: 54 – 69. DOI: 10.1016/j.isprsjprs.2016.12.010

[51] ZHANG X R, GAO Z Y, JIAO L C, et al. Multifeature hyperspectral image classification with local and nonlocal spatial information via Markov random field in semantic space [J]. IEEE transactions on geoscience and remote sensing, 2018, 56(3): 1409 – 1424. DOI: 10.1109/TGRS.2017.2762593

[52] ZHANG X R, GAO Z Y, AN J L, et al. Joint multi-feature hyperspectral image classification with spatial constraint in semantic manifold [C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2016: 481 – 484. DOI: 10.1109/IGARSS.2016.7729119

[53] ZHANG X R, SONG Q, GAO Z Y, et al. Spectral-spatial feature learning using cluster-based group sparse coding for hyperspectral image classification [J]. IEEE journal of selected topics in applied earth observations and remote sensing, 2016, 9(9): 4142 – 4159. DOI: 10.1109/JSTARS.2016.2593907

[54] NING H Y, ZHANG X R, QUAN D, et al. AUD-net: a unified deep detector for multiple hyperspectral image anomaly detection via relation and few-shot learning [J]. IEEE transactions on neural networks and learning systems, 2022, 99: 1 – 15. DOI: 10.1109/TNNLS.2022.3213023

[55] WU J L, TANG K W, ZHANG H C, et al. Structured information extraction of pathology reports with attention-based graph convolutional network [C]//IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2021: 2395 – 2402. DOI: 10.1109/BIBM49941.2020.9313347

[56] WU J L, ZHANG R N, GONG T L, et al. BioIE: biomedical information extraction with multi-head attention enhanced graph convolutional network [C]//IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022: 2080 – 2087. DOI: 10.1109/BIBM52615.2021.9669650

[57] HE K, HONG N, LAPALME-REMIS S, et al. Understanding the patient perspective of epilepsy treatment through text mining of online patient support groups [J]. Epilepsy & behavior, 2019, 94: 65 – 71. DOI: 10.1016/j. yebeh.2019.02.002

[58] LIU Y, WU J L, WEI Y H, et al. AEFNet: adaptive scale feature based on elastic-and-funnel neural network for healthcare representation [C]//IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022: 2043 – 2050. DOI: 10.1109/BIBM52615.2021.9669339

[59] LI C, XU X D, ZHOU G H, et al. Implementation of national health informatization in China: survey about the status quo [J]. JMIR medical informatics, 2019, 7(1): e12238. DOI: 10.2196/12238

[60] WU J L, DONG Y X, GAO Z Y, et al. Dual attention and patient similarity network for drug recommendation [J]. Bioinformatics, 2023, 39(1): btad003. DOI: 10.1093/bioinformatics/btad003

[61] WU J L, QIAN B Y, LI Y, et al. Leveraging multiple types of domain knowledge for safe and effective drug recommendation [C]//31st ACM International Conference on Information & Knowledge Management. ACM, 2022: 2169 – 2178. DOI: 10.1145/3511808.3557380

[62] WU J L, MAO A Y, BAO X R, et al. PIMIP: an open source platform for pathology information management and integration [C]//Proceedings of 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022: 2088 – 2095. DOI: 10.1109/BIBM52615.2021.9669424

[63] WU J L, ZHANG R N, GONG T L, et al. A precision diagnostic framework of renal cell carcinoma on whole-slide images using deep learning [C]//International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2021: 2104 – 2111. DOI: 10.1109/BIBM52615.2021.9669870

[64] WU J L, ZHANG R N, GONG T L, et al. A personalized diagnostic generation framework based on multi-source heterogeneous data [C]//IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022: 2096 – 2103. DOI: 10.1109/BIBM52615.2021.9669427

[65] MAO A Y , WU J L, BAO X , et al. A two-stage convolutional network for nucleus detection in histopathology image [C]//IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2021: 2051 – 2058, DOI: 10.1109/BIBM52615.2021.9669344

[66] GAO Z Y, HONG B Y, ZHANG X L, et al. Instance-based vision transformer for subtyping of papillary renal cell carcinoma in histopathological image [M]// Medical image computing and computer assisted intervention. Cham: Springer International Publishing, 2021: 299 – 308. DOI: 10.1007/978-3-030-87237-3_29

[67] GAO Z Y, MAO A Y, WU K F, et al. Childhood leukemia classification via information bottleneck enhanced hierarchical multi-instance learning [J]. IEEE transactions on medical imaging, 2023: early access. DOI: 10.1109/tmi.2023.3248559

[68] GAO Z Y, HONG B Y, LI Y, et al. A semi-supervised multi-task learning framework for cancer classification with weak annotation in whole-slide images [J]. Medical image analysis, 2023, 83: 102652. DOI: 10.1016/j. media.2022.102652

[69] GAO Z Y, JIA C, LI Y, et al. Unsupervised representation learning for tissue segmentation in histopathological images: from global to local contrast [J]. IEEE transactions on medical imaging, 2022, 41(12): 3611 – 3623. DOI: 10.1109/tmi.2022.3191398

[70] GAO Z Y, SHI J B, ZHANG X L, et al. Nuclei grading of clear cell renal cell carcinoma in histopathological image by composite high-resolution network [M]//Medical image computing and computer assisted intervention. Cham: Springer International Publishing, 2021: 132 – 142. DOI: 10.1007/978-3-030-87237-3_13

[71] WANG G C, ZHANG X R, PENG Z L, et al. MOL: towards accurate weakly supervised remote sensing object detection via multi-view noisy learning [J]. ISPRS journal of photogrammetry and remote sensing, 2023, 196: 457 – 470. DOI: 10.1016/j.isprsjprs.2023.01.011

Biographies

DENG Letian (2536059342@qq.com) is with the College of Mechanical and Electronic Engineering, Northwest A&F University, China, majoring in electronic information engineering. His research interests include artificial intelligence technology, semantic communication, and computer vision.

ZHAO Yanru received her PhD in agricultural mechanization engineering from Zhejiang University, China in 2018. She conducted joint doctoral training at Washington State University, USA from 2016 to 2017. Now she is an associate professor at the School of Mechanical and Electronic Engineering of Northwest A&F University, China. Her research interests include agricultural information intelligent sensing technology and equipment, high-throughput plant phenotype analysis, accurate orchard management, and intelligent research and development of field equipment. She serves as a reviewer with Computers and Electronics in Agriculture, Biosystem Engineering, and Analytical Methods and Fuel. She is a member of China Artificial Intelligence Society.

Content Popularity Prediction via Federated Learning in Cache-Enabled Wireless Networks

YAN Yuna[1], LIU Ying[2], NI Tao[2], LIN Wensheng[1],
LI Lixin[1]

(1. Northwestern Polytechnical University, Xi'an 710072, China；
2. Shanghai Satellite Engineering Research Institute, Shanghai 200240, China)

Abstract: With the rapid development of networks, users are increasingly seeking richer and high-quality content experience, and there is an urgent need to develop efficient content caching strategies to improve the content distribution efficiency of caching. Therefore, it will be an effective solution to combine content popularity prediction based on machine learning (ML) and content caching to enable the network to predict and analyze popular content. However, the data sets which contain users' private data cause the risk of privacy leakage. In this paper, to address this challenge, we propose a privacy-preserving algorithm based on federated learning (FL) and long short-term memory (LSTM), which is referred to as FL-LSTM, to predict content popularity. Simulation results demonstrate that the performance of the proposed algorithm is close to the centralized LSTM and better than other benchmark algorithms in terms of privacy protection. Meanwhile, the caching policy in this paper raises about 14.3% of the content hit rate.

Keywords: content popularity prediction; privacy protection; federated learning; long short-term memory

1 Introduction

Due to the explosive development of smart devices in networks, data traffic has increased unprecedentedly in recent years. With the limited communication resources, backhaul link congestion will occur in the peak period at times, which leads to poor quality of experience (QoE)[1]. Content caching is considered to be a promising solution to improving the QoE of users. For a traditional approach, almost all contents are placed on the cloud server. However, since a large number of popular files are easy to be repeatedly requested by users, the popular files can be cached in advance at local base stations (BSs), which not only guarantees the hit rate of content, but also is helpful to reduce the users' waiting time, alleviates the pressure on the core network and relieves traffic congestion[2 – 4].

In the past, traditional content-caching strategies, such as Least Recently Used (LRU)[5] and Least Frequently Used (LFU)[6], were used in the deployment phase. However, different users have different content preferences and these preferences are often time-varying, so the fixed content deployment cannot take full advantage of the network caching. Therefore, in order to further improve the performance of network caching, using machine learning (ML) to accurately predict popular files in the future attracts the interest of researchers during the deployment of caching content files[7]. WON and KIM[8] proposed a preference prediction neural network model based on DeepFM to predict the user's preference for movies, which improved the prediction accuracy by considering the interaction of low-order and high-order features of the input data. LI et al.[9] proposed proactive edge caching for device-to-device (D2D) assisted wireless networks. In this paper, the authors adopt bidirectional long short-term memory (LSTM) networks, graph convolutional networks and attention mechanisms to learn user preferences. JIANG et al.[10] proposed LSTM to predict the users' content request distribution, thereby achieving higher accuracy and better

versatility.

The above methods can be classified as centralized ML, where the original training data sets need to be uploaded to the central server. However, the prediction of content popularity often involves personal information (e.g., home addresses, shopping, etc.) as training samples, which results in the risk of privacy leakage. As a privacy-preserving distributed learning framework, federated learning (FL) was proposed to tackle this challenge by training a global statistical model without accessing users' private raw data[11 – 12]. Recently, FL has been put forward to solve the challenging problems in wireless networks[13 – 14]. In particular, an FL-based approach was provided by FAROOQ et al.[15] to build a flood forecasting model. Specifically, the local training parameters are aggregated to build the global model. By transferring the training parameters instead of sending huge data sets, the leakage of the data privacy will be greatly decreased. WANG et al.[16] proposed an efficient content popularity prediction of privacy-preserving (CPPPP) scheme based on federated learning and Wasserstein generative adversarial network (WGAN), which achieves a high cache hit ratio. In this system, the server aggregated the users' updates using federated averaging, and each user performed training on its local data using WGAN, which could achieve high cache efficiency and protect the privacy of users. Therefore, considering privacy protection, FL can also be applied to the content popularity prediction in cache-enabled wireless networks.

Motivated by the above discussions, we propose a privacy-preserving algorithm for content popularity prediction named FL-LSTM, which combines LSTM with FL. Due to the unique design structure, LSTM is suitable for processing and predicting time series, such as content popularity prediction[10]. According to the aggregation mechanism of FL[11], the global content popularity prediction model will be built based on local training parameters. Thus, the FL-LSTM algorithm can inherently improve security performance and obtain reliable prediction performance.

The main contributions of this work are summarized as follows:

• We investigate a content popularity prediction problem in cache-enabled wireless networks, and aim at minimizing the mean-square error (MSE) and maximizing the cache hit rate. Considering the significance of privacy-preserving, a novel content popularity prediction algorithm FL-LSTM based on LSTM and FL is proposed. The algorithm avoids the direct transmissions of raw user data, which preserves user privacy.

• By utilizing a real-world dataset, the simulation results demonstrate that the proposed algorithm achieves similar performance to the centralized LSTM and better prediction ability than other state-of-the-art schemes.

The remainder of this paper is organized as follows. In Section 2, the communication system model and problem formulation are introduced. The privacy-preserving FL-LSTM algorithm is proposed in Section 3. In Section 4, the simulation re-

sults and experiment result analysis are shown. Finally, conclusions are drawn in Section 5.

2 System Model and Problem Formulation

2.1 System Model

In this paper, we consider a cache-enabled wireless network as illustrated in Fig. 1, which consists of a cloud server, multiple BSs and the users served by each associated BS. There are B BSs in this specific region and the set of BSs is denoted by $\mathcal{B} = \{1, 2\cdots, b, \cdots, B\}$. The content library is denoted by $\mathcal{F} = \{1, 2\cdots, f, \cdots, F\}$, where we assume the requested files have the same size. The cloud server $\mathcal{O}$ contains the whole content library and each BS can only store a limited number of files. To simplify the model, we assume that each BS is equipped with a cache C_m of an equal size, where $C_m = n$ represents BS b can only cache n files from the cloud server.

It is assumed that the contents are requested and fetched during the discrete time periods and the set of time periods is expressed as $\mathcal{T} = \{1, 2\cdots, t, \cdots, T\}$. In each time slot t: 1) According to users' previous request information, the local BSs and the cloud server jointly build the content popularity prediction model; 2) based on the prediction results, the related BS will cache the relevant contents from $\mathcal{O}$ in advance; 3) when the requested file is stored in the local BS, the associated users will directly obtain the requested contents; 4) otherwise, the requested file is fetched from the cloud server.

2.2 Problem Formulation

Based on the network discussed above, the basic framework and operation process of the communication system are introduced. It is obviously known that content popularity prediction is perceived as the key to the success of the system. For a specific file f, the popularity will change over time, and its popularity sequence is expressed as $\mathcal{P}_{b,f} = \{p_{b,f}^1, p_{b,f}^2, \cdots, p_{b,f}^t\}$, $p_{b,f}^t \in [0,1]$. Therefore, the popularity pre-
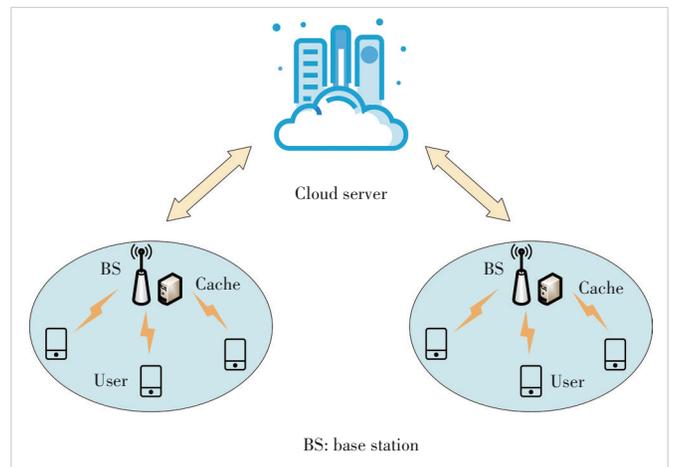
▲Figure 1. Scenario of a cache-enabled wireless network

diction of a file is transformed into a time series prediction problem, and the real and predicted values are expressed as $p^t_{b,f}$ and $\tilde{p}^t_{b,f}$, respectively. Moreover, the MSE is adopted to evaluate the accuracy of the prediction as follows:

$$\text{MSE} = \frac{1}{T} \sum_{t=1}^{T} \left| \tilde{p}^t_{b,f} - p^t_{b,f} \right|^2 . \tag{1}$$

In this paper, content popularity is defined as the ratio of the number of requests for a file to the number of requests for all files within a time slot. If a file is frequently requested by users, the higher popularity of the content is, the more likely it is to be accessed again in the next time slot. In time slot t, the content popularity of a file f can be expressed as $q^f_{b,t} = \text{req}_{t,f} \Big/ \sum_{i=1}^{f} \text{req}_{t,f}$, where $\text{req}_{t,f}$ represents the number of user requests for the file f in a time slot t. Therefore, the popularity of the content library is denoted as $\mathcal{P}_{b,t} = \left\{ p^1_{b,t}, p^2_{b,t}, \cdots, p^f_{b,t} \right\}$, $p^f_{b,t} \in [0,1]$ and $\sum_{f=1}^{F} p^f_{b,t} = 1$, where the order of files is in a descending order according to the popularity.

Due to the limited storage capacity of each BS b, after the prediction task is completed, each BS b needs to sort the predicted popular files, select the contents that are more popular with users to cache in advance and replace the files with low popularity. If the contents are cached locally, this operation does not need to be repeated. For a certain discrete time slot t, the selected pre-cached files in local BS b are formulated as $\mathcal{A}_{b,c} = \left\{ a^1_{b,c}, a^2_{b,c}, \cdots, a^f_{b,c} \right\}$, where $a^f_{b,c} = 1$ if the file f is cached in BS b, otherwise $a^f_{b,c} = 0$ and $\sum_{f=1}^{F} a^f_{b,c} \leq n$.

Furthermore, when we measure the cache-enabled wireless network, the cache hit rate is considered as an important metric of caching performance. The hit rate of each BS during each time slot is defined as $h_b = \frac{1}{n} \sum_{f=1}^{F} a^f_{b,c} \times p^f_{b,t}$, referring to the probability that the precached file is popular content, which is used to represent the effectiveness of the content cache. Therefore, the hit rate of the network is averaged to be the total hit rate $\bar{h}$ during each training episode.

Due to the limited information collected by a single BS, combining several local BSs by the cloud server is necessary to obtain the whole popularity of the content library. However, it will also cause a privacy leakage issue to some extent. Last but not least, the objective of this paper is to predict the content popularity accurately and maximize the cache hit rate during each time slot while preserving users' privacy.

3 FL-LSTM for Content Popularity Prediction

3.1 Literature Overview

1) LSTM: LSTM was first proposed by HOCHREITER et al.[17].

Although the recurrent neural network (RNN) can be used to process and predict the sequence data, the processing and prediction effect of LSTM is better than that of the RNN as the time scale of the processing sequence increases, and the phenomena of "gradient vanishing" and "gradient explosion" of the RNN can be avoided through LSTM. Therefore, LSTM is selected as the benchmark algorithm to predict popular files. The illustration of the LSTM algorithm is shown in Fig. 2. A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate, where the "gate" structure is an approach to data control. The formulas of the three gate structures are defined as follows. The forget gate f_t decides what kind of previous information will be forgotten, i.e.,

$$f_t = \sigma \left(\boldsymbol{W}_{fx} \cdot \boldsymbol{x}_t + \boldsymbol{W}_{fh} \cdot \boldsymbol{h}_{t-1} + \boldsymbol{b}_f \right), \tag{2}$$

where $\boldsymbol{x}_t$ is the input vector of the current time and $\boldsymbol{h}_{t-1}$ is the hidden status of the previous time. $\boldsymbol{W}_{fx}$, $\boldsymbol{W}_{fh}$ and $\boldsymbol{b}_f$ are the input weight, recurrent weight and corresponding bias of the forget gate f_t, respectively. The input gate i_t is used to select which information will be recorded, i.e.,

$$i_t = \sigma \left(\boldsymbol{W}_{ix} \cdot \boldsymbol{x}_t + \boldsymbol{W}_{ih} \cdot \boldsymbol{h}_{t-1} + \boldsymbol{b}_i \right), \tag{3}$$

where $\boldsymbol{W}_{ix}$, $\boldsymbol{W}_{ih}$ and $\boldsymbol{b}_i$ are the input weight, recurrent weight and corresponding bias of the input gate i_t, respectively. $\boldsymbol{x}_t$ and $\boldsymbol{h}_{t-1}$ are weighted to update the value of the input gate through the sigmoid function $\sigma(\cdot)$. The candidate memory cell state $\tilde{\boldsymbol{C}}_t$ is updated by Eq. (4), i.e.,

$$\tilde{\boldsymbol{C}}_t = \tanh \left(\boldsymbol{W}_{cx} \cdot \boldsymbol{x}_t + \boldsymbol{W}_{ch} \cdot \boldsymbol{h}_{t-1} + \boldsymbol{b}_c \right), \tag{4}$$

where $\boldsymbol{W}_{cx}$, $\boldsymbol{W}_{ch}$ and $\boldsymbol{b}_c$ are the input weight, recurrent weight and corresponding bias of the candidate memory cell, respectively. And the tanh function $\tanh(\cdot)$ can control the range of its value to $[-1, 1]$. The new memory cell $\boldsymbol{C}_t$ controls the input and forget mechanism, which updates the state of the unit at
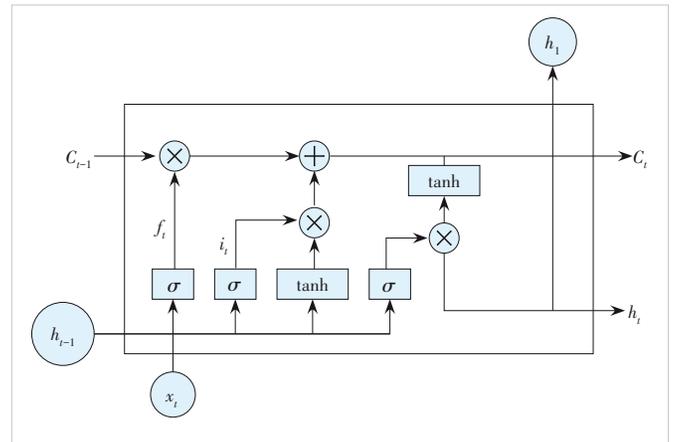
▲Figure 2. Illustration of long short-term memory (LSTM) algorithm

the previous time based on the output of the forget gate and the input gate, i.e.,

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \tag{5}$$

The output gate o_t decides what value should be output by Eqs. (4) – (7), and then the output of the hidden layer h_t is obtained by Eq. (7), i.e.,

$$o_t = \sigma\left(W_{ox} \cdot x_t + W_{oh} \cdot h_{t-1} + b_o\right), \tag{6}$$

$$h_t = o_t \odot \tanh\left(C_t\right), \tag{7}$$

where W_{ox}, W_{oh} and b_o are the input weight, recurrent weight and corresponding bias of the output gate o_t, respectively.

2) Fedavg: The Fedavg algorithm was proposed by KHAI et al.[18] and has several advantages in privacy protection and distributed training. In addition, users do not need to upload all the data, but only upload the local parameters needed by the model, which greatly reduces the overall communication overhead. In this algorithm, the cloud server starts FL training by sharing global model parameters with the base station. Then, each base station selects samples from the local data subset to perform the stochastic gradient descent (SGD) in order to update the local model and share the updated model weight with the cloud server. After that, the cloud server aggregates all the updated local model weight parameters and averages them to generate the global model. Compared with the centralized machine learning, the algorithm has some differences. The algorithm flow is presented as follows:

(a) At the beginning of training, the global model parameters W^o in the cloud server are initialized and then sent to the local BSs as W_t^b.

(b) The BS b trains the local dataset and updates W_t^b to W_{t+1}^b after the training epochs, i.e., $W_{t+1}^b \leftarrow \mathrm{LocalUpdate}\left(b, w_t^b\right)$.

(c) The cloud server aggregates each BS's W_{t+1}^b and then generates a new global model W_{t+1}^o. This formula of aggregation can be described as:

$$W_{t+1}^o = \frac{1}{B} \sum_{b=1}^{B} W_{t+1}^b. \tag{8}$$

(d) Afterwards, the W_{t+1}^o will be broadcasted to all the BSs and the next round of training is started.

3.2 Proposed FL-LSTM Algorithm

Based on the aforementioned LSTM and Fedavg algorithms, we develop the FL-LSTM algorithm. The algorithm can predict the content popularity accurately while preserving the users' privacy. The illustration of the FL-LSTM algorithm is shown in Fig. 3. Then we will introduce the details of this algorithm.

Firstly, the initial LSTM network is adopted on the cloud server as the global model, and each BS will build the local LSTM net-

work with the initial parameters W^o from the global model.

After that, to avoid sharing the raw data directly, the predicted model based on LSTM will be trained locally. The raw data $\mathcal{D}$ are divided uniformly and posted on each BS, with d_b representing the local dataset of BS b. The previous k moments are selected for prediction and the network can be characterized as follows: the input time series $\mathcal{X} = \{x_1, x_2, \cdots, x_t\}$ are defined as the historical popularity, through the vector $\mathcal{X}$ to predict the output vector $\mathcal{Y}$. The output time series $\mathcal{Y} = \{y_1, y_2, \cdots, y_t\}$ are defined as the predicted popularity, and $\mathcal{H} = \{h_1, h_2, \cdots, h_t\}$ means the information of the hidden layer.

The local weight parameters are collectively referred to as W_t^b. The small BS adopts the SGD optimizer, and the weight parameters will be updated to W_{t+1}^b according to Eqs. (2) – (7).

According to the aggregation mechanism in Eq. (8), the cloud server aggregates the updated parameters W_{t+1}^b which are uploaded by all local BSs to generate a new global model. Subsequently, the parameters of the global model W_{t+1}^o are downloaded to each BS and then the next round of training will be started. Model training and parameter updating are repeated until the algorithm is terminated when the maximum number of iterations T is reached. The objective of this algorithm is to minimize the MSE denoted in Eq. (1). Specifically, the proposed FL-LSTM algorithm is summarized in Algorithm 1.

Algorithm 1. Content popularity prediction based on FL-LSTM

1: Initialize the system: cloud server $\mathcal{O}$, local BSs $\mathcal{B}$.

2: Initialize the LSTM network by Eqs. (2) – (7): $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{H}$; the global weight W^o; the local weight W^b and the local sampled batch size k.

3: Initialize the round index t and the local training epoch index n.
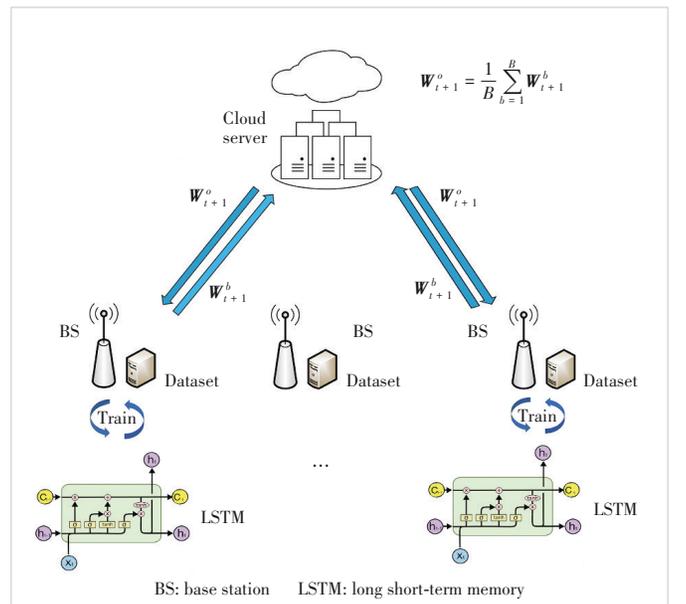
▲Figure 3. Illustration of FL-LSTM algorithm

BS: base station   LSTM: long short-term memory

4: **for** round=$1,\cdots, t ,\cdots, T$ **do**

5: At time-step t, the global model $\boldsymbol{W}_t^o$ is broadcasted to each BS as $\boldsymbol{W}_t^b$.

6: (For each BS b, start the local training.)

7: **for** epoch=$1,\cdots, n,\cdots, n_{max}$ **do**

8: Samples a batch $\{x_i, \tilde{y}_i\}_{i=1}^k$ from d_b;

9: Update the local LSTM network by the loss functions in Eq. (1):

10: $\boldsymbol{W}_{n+1}^b \leftarrow \text{SGD}\left(\nabla_\omega \frac{1}{k} \sum_{i=1}^\kappa (y_i - \tilde{y}_i)^2\right)$

11: **end for**

12: The cloud server aggregates each BS's $\boldsymbol{W}_{t+1}^b$ and updates $\boldsymbol{W}_t^o$ to $\boldsymbol{W}_{t+1}^o$ by Eq. (8).

13: $t \leftarrow t + 1$.

14: **end for**

4 Simulations and Discussions

4.1 Datasets

The MovieLens 1M Dataset[19] is used to evaluate the performance of the proposed FL-LSTM algorithm in this paper. The dataset contains 1 000 209 ratings of approximately 3 900 movies made by 6 040 users. Each sample in the data set includes a user ID, movie ID, user rating, and time stamp when commenting. We assume that the number of ratings by users can reflect the popularity of relevant movies. The process of dataset construction is as follows: Based on this assumption, we choose the ten movies with the most ratings as the prediction objects and divide one hundred discrete time slots based on the provided time stamp. Next, we count the number of ratings by users according to each time slot and consider it as the number of requests. Then, the request times are normalized to calculate the content popularity of the corresponding movie file. Finally, according to the content popularity of each film in one hundred discrete time slots, the LSTM model is used to predict the content popularity of each film in the next moment. By ranking the results, we can predict which movies will be popular at the next moment.

4.2 Simulation Setup

The real MovieLens 1M Dataset was used to construct the content popularity prediction datasets. The top 10 rated movies are selected as the forecast object with their trends for 100 days. In the time series prediction problem, we set $k = 10$, which means that we use the previous 10 time-slots data to predict the popularity of the next moment. In addition, the number of BSs is set as $B = 5$ and the dataset is divided uniformly and handed out to each BS.

4.3 Performance Comparison

In Fig. 4, the performance of different algorithms is compared under the task of predicting the content popularity for 100 days. The LSTM algorithm and the Autoregressive Inte-

grated Moving Average Model (ARIMA) algorithm[20] are selected as the benchmark algorithms. As shown in Fig. 4, the LSTM and FL-LSTM algorithms have similar prediction results, and their accuracy is significantly better than the ARIMA algorithm. The ARIMA algorithm depends on the statistical characteristics of the data and the performance is limited by parameter estimation, so it is difficult to achieve high accuracy. In contrast, the LSTM and the FL-LSTM algorithms have the same core prediction network, which reflects obviously the superiority of performance for the time series prediction problem.

Fig. 5 indicates the convergence of the MSE loss with the

ARIMA: Autoregressive Integrated Moving Average Model   FL: federated learning   LSTM: long short-term memory
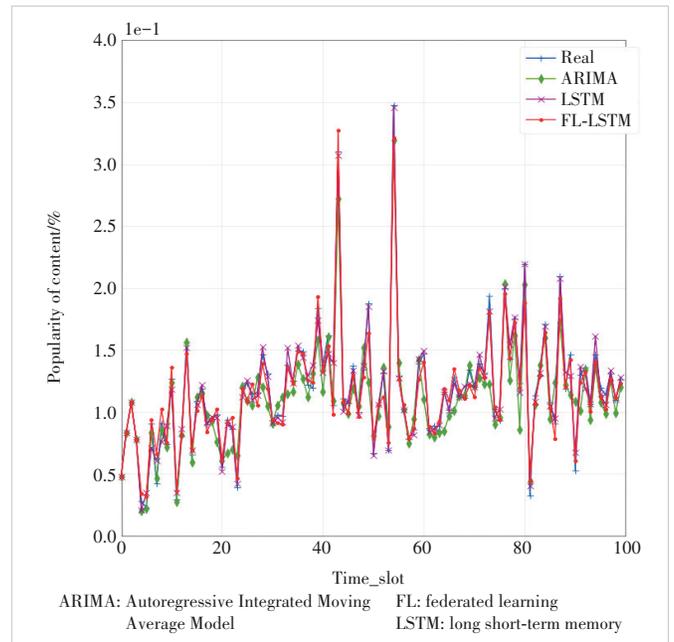
▲Figure 4. Content popularity prediction by different algorithms

FL: federated learning   MSE: mean-square error
LSTM: long short-term memory
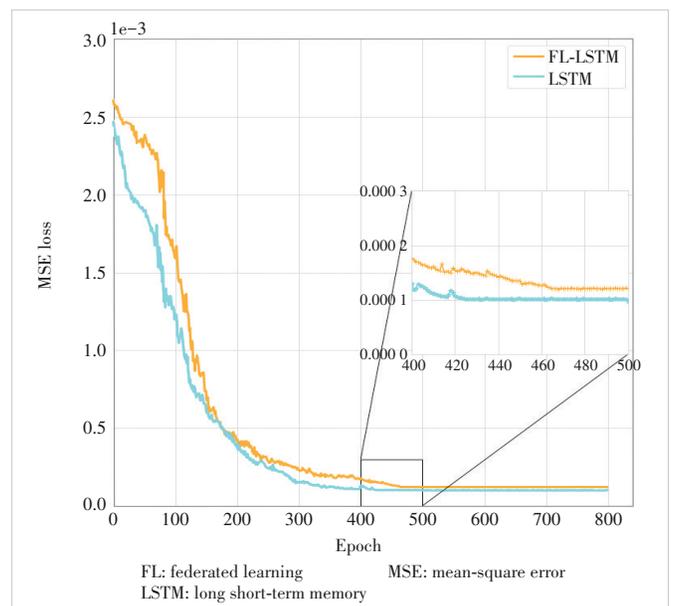
▲Figure 5. MSE loss of LSTM and FL-LSTM algorithms

LSTM algorithm and the FL-LSTM algorithm, where the MSE loss is related to the accuracy of the prediction. In the case of setting the same simulation parameters, both of these algorithms can complete the convergence within the range of 400 – 500 training rounds. At the same time, the value of MSE is very small, and both can be below 0.000 2, which proves the superiority of the prediction performance. In addition, in the same simulation environment, the convergence of the FL-LSTM is slightly slower than the LSTM algorithm and the loss is similar, i.e., only 4.2% between the two algorithms. The main reasons why the federated learning scheme is slightly inferior to the centralized one are as follows: 1) The user data distribution of each BS is different, which also leads to a certain delay in updating the global model during the distributed FL-LSTM training; 2) when weight aggregation is carried out in federated learning, fractional parts of weight parameters are generally truncated to improve uploading efficiency. Therefore, there will be a certain degree of numerical precision loss in the process of average weighted sum. For the sake of privacy protection, it is acceptable to sacrifice a little bit of performance.

Fig. 6 shows how the total hit rate changes as the memory capacity of the BSs increases. Cache capacity is defined as the ratio of the number of files cached to the total number of files in a file set. There is a significant upward trend when increasing the memory capacity. The reason is that the users' requested contents are more likely to be accurately predicted and cached in the base station. However, for the LRU and the random caching (RC), although it takes into account historical popularity information, the content popularity will not be pre-

dicted in advance, which will suffer more inaccurate caching and cause a lower hit rate by at least a 14.3% difference. The optimal performance is obtained under the real content popularity, which is an ideal situation. In addition, compared with traditional cache algorithms, although the FL-LSTM algorithm proposed in this paper requires additional resources for model prediction, the time required to execute model prediction is relatively small. From Fig. 6, it can also be observed that when the memory capacity increases, the hit rate of the FL-LSTM algorithm will gradually approach the optimal value, which is only a 2.3% performance loss at 70% capacity. Therefore, this shows the superiority of the algorithm proposed in this paper.

5 Conclusions

In this paper, we investigate the content popularity prediction problem in cache-enabled wireless networks. Meanwhile, a novel prediction algorithm FL-LSTM based on LSTM and FL is proposed for privacy preservation. The proposed algorithm can not only predict the content popularity accurately but also protect the users' privacy information. Moreover, the performance of the FL-LSTM is validated on the real-world dataset compared to other algorithms. Simulation results demonstrate that the performance of the proposed algorithm just declined slightly, only 4.2% compared with the centralized LSTM and is better than other state-of-the-art schemes while the privacy can be well preserved.

▲Figure 6. Total hit rates versus memory capacity
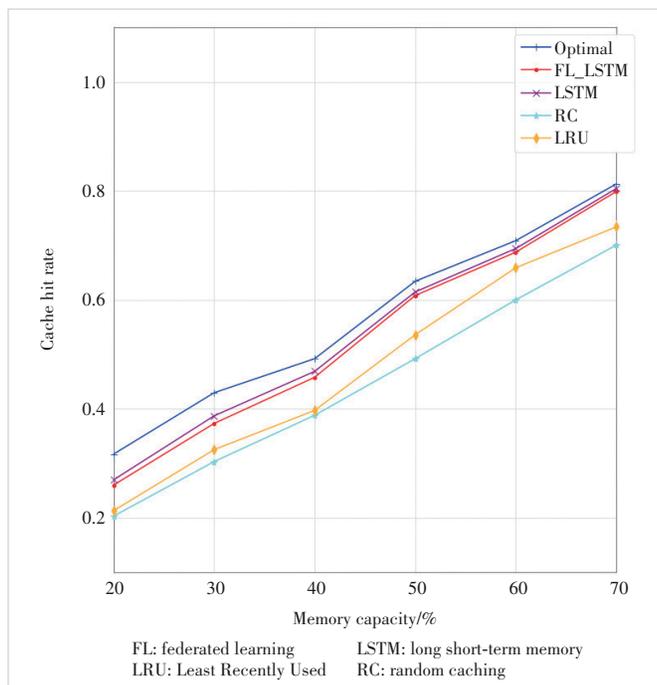
References

[1] WANG C M, HE Y, YU F R, et al. Integration of networking, caching, and computing in wireless systems: a survey, some research issues, and challenges [J]. IEEE communications surveys & tutorials, 2018, 20(1): 7 – 38. DOI: 10.1109/COMST.2017.2758763

[2] NDIKUMANA A, TRAN N H, HO T M, et al. Joint communication, computation, caching, and control in big data multi-access edge computing [J]. IEEE transactions on mobile computing, 2020, 19(6): 1359 – 1374. DOI: 10.1109/TMC.2019.2908403

[3] PASCHOS G S, IOSIFIDIS G, TAO M X, et al. The role of caching in future communication systems and networks [J]. IEEE journal on selected areas in communications, 2018, 36(6): 1111 – 1125. DOI: 10.1109/JSAC.2018.2844939

[4] WEI Y F, YU F R, SONG M, et al. Joint optimization of caching, computing, and radio resources for fog-enabled IoT using natural actor-critic deep reinforcement learning [J]. IEEE Internet of Things journal, 2019, 6(2): 2061 – 2073. DOI: 10.1109/JIOT.2018.2878435

[5] AHMED M, TRAVERSO S, GIACCONE P, et al. Analyzing the performance of LRU caches under non-stationary traffic patterns [EB/OL]. (2013-01-21)[2023-03-01]. https://arxiv.org/abs/1301.4909

[6] JALEEL A, THEOBALD K B, STEELY S C, et al. High performance cache replacement using re-reference interval prediction (RRIP) [C]//37th annual international symposium on computer architecture. ACM, 2010: 60 – 71. DOI: 10.1145/1815961.1815971

[7] LI L X, XU Y, YIN J Y, et al. Deep reinforcement learning approaches for content caching in cache-enabled D2D networks [J]. IEEE Internet of Things journal, 2020, 7(1): 544 – 557. DOI: 10.1109/JIOT.2019.2951509

[8] WON D U, KIM H S. A prediction scheme for movie preference rating based on DeepFM model [C]//International Conference on Information Networking (ICOIN). IEEE, 2022: 385 – 390. DOI: 10.1109/ICOIN53446.2022.9687136

[9] LI D Y, ZHANG H X, DING H, et al. User preference learning-based proactive edge caching for D2D-assisted wireless networks [J]. IEEE Internet of Things journal, 2023, early access. DOI: 10.1109/JIOT.2023.3244621

[10] JIANG Y X, FENG H J, ZHENG F C, et al. Deep learning-based edge caching in fog radio access networks [J]. IEEE transactions on wireless communications, 2020, 19(12): 8442 – 8454. DOI: 10.1109/TWC.2020.3022907

[11] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data [EB/OL]. (2016-02-17) [2023-03-01]. https://arxiv.org/abs/1602.05629

[12] SHI Y M, YANG K, JIANG T, et al. Communication-efficient edge AI: algorithms and systems [J]. IEEE communications surveys & tutorials, 2020, 22(4): 2167 – 2191. DOI: 10.1109/COMST.2020.3007787

[13] KHAN L U, PANDEY S R, TRAN N H, et al. Federated learning for edge networks: resource optimization and incentive mechanism [J]. IEEE communications magazine, 2020, 58(10): 88 – 93. DOI: 10.1109/MCOM.001.1900649

[14] LIM W Y B, LUONG N C, HOANG D T, et al. Federated learning in mobile edge networks: a comprehensive survey [J]. IEEE communications surveys & tutorials, 2020, 22(3): 2031 – 2063. DOI: 10.1109/COMST.2020.2986024

[15] FAROOQ M S, TEHSEEN R, QURESHI J N, et al. FFM: flood forecasting model using federated learning [J]. IEEE access, 2023, 11: 24472 – 24483. DOI: 10.1109/ACCESS.2023.3252896

[16] WANG K L, DENG N, LI X H. An efficient content popularity prediction of privacy preserving based on federated learning and Wasserstein GAN [J]. IEEE Internet of Things journal, 2023, 10(5): 3786 – 3798. DOI: 10.1109/JIOT.2022.3176360

[17] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural computation, 1997, 9(8): 1735 – 1780. DOI: 10.1162/neco.1997.9.8.1735

[18] DOAN K N, VAN NGUYEN T, QUEK T Q S, et al. Content-aware proactive caching for backhaul offloading in cellular network [J]. IEEE transactions on wireless communications, 2018, 17(5): 3128 – 3140. DOI: 10.1109/TWC.2018.2806971

[19] HARPER F M, KONSTAN J A. The MovieLens datasets [J]. ACM transactions on interactive intelligent systems, 2016, 5(4): 1 – 19. DOI: 10.1145/2827872

[20] GUO J Q, HE H W, SUN C. ARIMA-based road gradient and vehicle velocity prediction for hybrid electric vehicle energy management [J]. IEEE transactions on vehicular technology, 2019, 68(6): 5309 – 5320. DOI: 10.1109/TVT.2019.2912893

Biographies

YAN Yuna is currently working toward her master's degree under the supervision of Prof. LI Lixin with the School of Electronics and Information, Northwestern Polytechnical University, China. Her research interests include federated learning, deep learning and semantic communications.

LIU Ying is an engineer of Shanghai Satellite Engineering Research Institute, China, mainly engaged in satellite system design and satellite communications.

NI Tao is a senior engineer of Shanghai Satellite Engineering Research Institute, China, mainly engaged in satellite system design and satellite communications.

LIN Wensheng received his BE degree in communication engineering and ME degree in electronic and communication engineering from Northwestern Polytechnical University, China in 2013 and 2016. He received his PhD degree in information science from the Japan Advanced Institute of Science and Technology in 2019. He is currently an associate professor with the School of Electronics and Information, Northwestern Polytechnical University. His research interests include network information theory, distributed source coding, and age of information.

LI Lixin (lilixin@nwpu.edu.cn) received his BS degree (Hons.), MS degree (Hons.), and PhD degree from Northwestern Polytechnical University (NPU), China in 2001, 2004, and 2008, respectively. From 2009 to 2011, he was a postdoctoral research fellow with NPU. In 2011, he joined the School of Electronics and Information, NPU, where he is currently a full professor and chair of Department of Communication Engineering. In 2017, he held the visiting scholar position with the University of Houston, USA. He holds 26 granted patents and has authored or coauthored five books and more than 200 peer-reviewed papers in many prestigious journals and conferences. His research interests include 5G/6G wireless networks, federated learning, game theory, and machine learning for wireless communications. He was the recipient of the 2016 NPU Outstanding Young Teacher Award, which is the highest research and education honors for young faculties in NPU. He was an exemplary reviewer for *IEEE Transactions on Communications* in 2020.

Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation   *Special Topic*

ZHAO Moke, HUANG Yansong, LI Xuan

Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation

ZHAO Moke, HUANG Yansong, LI Xuan

(Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: With the rapid advancements in edge computing and artificial intelligence, federated learning (FL) has gained momentum as a promising approach to collaborative data utilization across organizations and devices, while ensuring data privacy and information security. In order to further harness the energy efficiency of wireless networks, an integrated sensing, communication and computation (ISCC) framework has been proposed, which is anticipated to be a key enabler in the era of 6G networks. Although the advantages of pushing intelligence to edge devices are multi-fold, some challenges arise when incorporating FL into wireless networks under the umbrella of ISCC. This paper provides a comprehensive survey of FL, with special emphasis on the design and optimization of ISCC. We commence by introducing the background and fundamentals of FL and the ISCC framework. Subsequently, the aforementioned challenges are highlighted and the state of the art in potential solutions is reviewed. Finally, design guidelines are provided for the incorporation of FL and ISCC. Overall, this paper aims to contribute to the understanding of FL in the context of wireless networks, with a focus on the ISCC framework, and provide insights into addressing the challenges and optimizing the design for the integration of FL into future 6G networks.

Keywords: integrated sensing; communication and computation; federated learning; data heterogeneity; limited resources

1 Introduction

With the continuous integration and advancement of communications and the popularization and application of artificial intelligence (AI), the next-generation communication system will not only facilitate huge data rates but also enable the intelligent industry of the Internet of Things (IoT)[1]. The number of connected devices worldwide is estimated to reach 29.3 billion[2] by 2023. The entire IoT network will provide low-latency, high-precision, scalable and flexible services powered by AI and non-contact sensing techniques[3]. In conventional wireless networks, high-quality environmental data are gathered by sensing and then conveyed via data transmission links, which is finally computed in downstream tasks. These separate processes pose challenges to meeting the stringent requirements of ultra-low latency, high reliability, and high capacity in 6G networks.

An integrated network can realize closed-loop information flow and wide-area intelligent cooperation (Fig. 1). It profoundly integrates wireless sensing func-
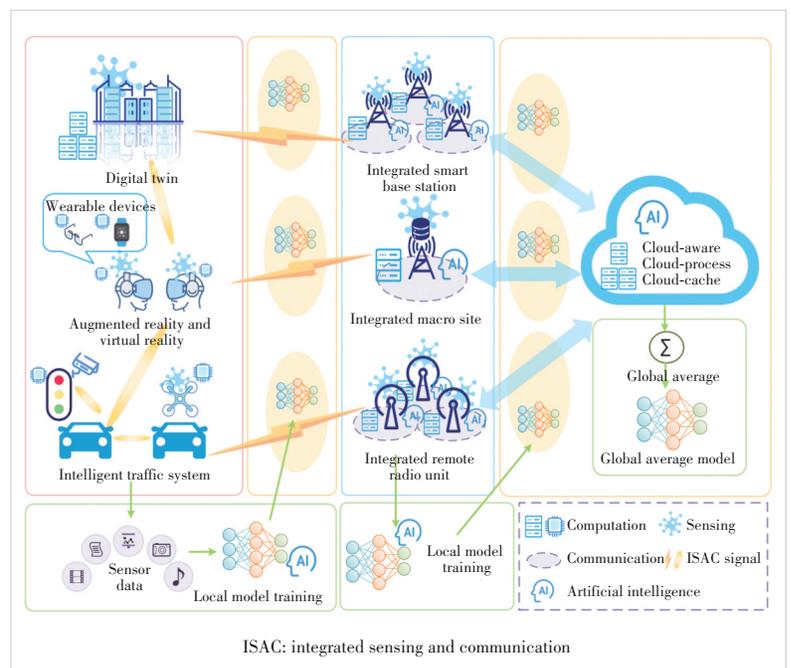
▲Figure 1. Application scenarios of integrated communication, sensing and computation

Special Topic | Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation

ZHAO Moke, HUANG Yansong, LI Xuan

tions, including but not limited to traditional positioning, detection, imaging, and wireless transmission. It also leverages widely distributed computing power to process additional computations, thus realizing the cross-fusion of perceptual communication computing and supporting typical application scenarios such as the smart city, intelligent healthcare, and smart home in 6G systems. To enhance the ability of 6G networks to perform endogenous intellectual sensing and adapt computational power, both academia and industry have preliminarily explored integrated sensing, communication and computation (ISCC) frameworks. On the other hand, the cloud computing platform is utilized for centralized data processing and training by machine learning (ML). Nonetheless, the vast volume of data created by intelligent terminals at the network edge may require substantial communication resources. When computation workloads are distributed for multiple tasks and all data are uploaded to a cloud platform, secure data privacy becomes difficult. Therefore, sending all data to the cloud for computing and learning may be impractical.

In recent years, China and the European Union have respectively introduced relevant regulations such as the Data Security Law and the General Data Protection Regulations[4], which states the regulatory requirements to ensure privacy and security while sharing data. For AI technology, federated learning (FL) was proposed for the sake of low latency and high accuracy[5]. Edge computing pushed cloud services from the network core to the edge closer to IoT devices[6]. Communication transmission delays can be efficiently decreased by intelligently combining terminal equipment, edge server and cloud center to participate in model training and data processing at the edge. Specifically, FL refers to edge intelligent sensing devices that use their computing capabilities to learn local data and obtain models based on different tasks. FL is a widely used distributed learning model, which uses wireless networks to bring a global ML model that improves computing ability and keeps data confidential to a certain extent[7].

In a typical FL training process, the central server broadcasts the global model to each edge device available. The edge device learns from the local data and obtains a local model. The regional model parameters are uploaded to the central server for aggregation to generate a new global model. This process is repeated iteratively to obtain the final global model. The federation has four leading performance indicators: latency, energy, reliability, and large-scale connectivity[8]. Because there is no need to share and transmit raw data and a cluster-like communication structure is adopted, FL is more suitable for large-scale intelligent devices and widely distributed deployment environments. In this paper, we will examine the issues faced by FL and the latest advancements in FL to investigate the future 6G network of universal computing. We will present the challenges in three categories: addressing terminal/data heterogeneity and model variances, executing FL within the constraints of universal

computing resources, and bolstering privacy protection. By introducing the fundamental concepts of FL, summarizing the advantages and disadvantages of existing research, and investigating application schemes for different task scenarios, this paper discusses the research trend of FL in the future edge intelligence system. Section 2 demonstrates the basic framework of FL. In Section 3, we present techniques used for ISCC. In Section 4, we highlight several challenges when implementing FL in the ISCC framework, including participant selection, adaptive aggregation, incentive mechanism, model compression, and privacy protection. Furthermore, we review the solutions to these problems along with their advantages and disadvantages. Finally, we design guidelines for the incorporation of FL and ISCC as well as a range of typical FL applications in Section 5.

2 Framework of Federated Learning

In traditional ML systems, users are required to upload local data to cloud servers with solid computing power for centralized model training, which includes central servers and several edge devices for data collection, as shown in Fig. 2. This scenario generates energy consumption and communication delays during the data upload process. Additionally, there is also a risk of privacy disclosure for privacy-sensitive participating nodes. In response to this problem, researchers have conducted the research on distributed ML. In 2017, Google first proposed FL technology[7]. Since then, FL has received significant attention from the academic community.

In a typical training cycle of FL, the dedicated edge server initially broadcasts a global machine-learning model to participating edge devices. Subsequently, the edge devices utilize their local data to calculate their respective model updates and transfer them to the edge server for further aggregation and global model updates. The FL training process is carried out iteratively in multiple communication rounds. The dataset of N edge devices is $\{D_1, D_2, D_3, \cdots, D_N\}$. The traditional method is to upload the dataset and train the model in the central server, whereas FL coordinates the local training of many data users
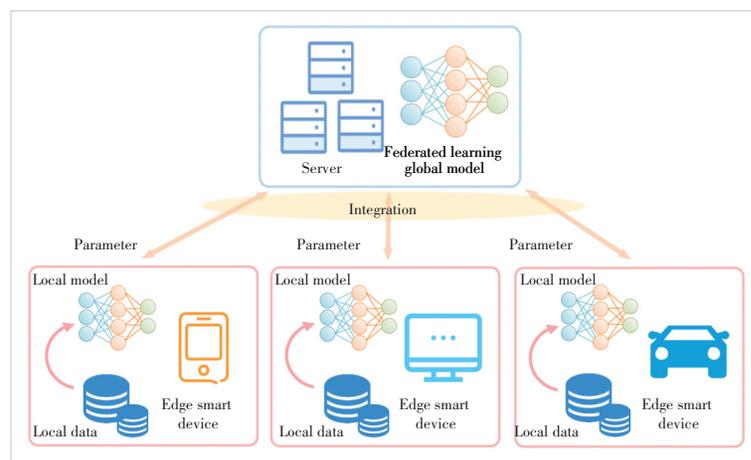
▲Figure 2. Basic schematic diagram of federated learning (FL)

Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation | *Special Topic*

ZHAO Moke, HUANG Yansong, LI Xuan

through the network to update the parameters interactively with the global model on the server side. It cooperatively optimizes the common objective function to obtain the final ML model. For client k with dataset D_k, the loss function can be expressed as $F_k(\omega) = \frac{1}{D_k} \sum_{j \in D_k} f_j(\omega)$. $f_j(x_j, y_j, \omega)$ is the loss function of the j-th data sample related to a specific ML model, where (x_j, y_j) represents the sensing data sample, while ω is the parameter of the learning model. In each learning round t, the steps detailed below are performed.

• Sensing data collection: each edge device k is equipped with sensing capabilities and collects data samples for local model training.

• Global model broadcast: each edge device k downloads the global model parameter $\omega^{(t)}$ through wireless communication links from the central server.

• Local model training: each edge device k uses the global model $\omega_k^{(t)}$ to update its local model, where $\omega_k^{(t)}$ is the local model parameter set of client k in the round of t. Therefore, $\omega_k^{(t)} = \omega^{(t-1)} - \eta \nabla g_k(\omega^{(t-1)})$, where η indicates the learning rate. As long as the local gradient $g_k(\omega^{(t-1)})$ is obtained, $\omega_k^{(t)}$ can be calculated.

• Local model uploading: $\omega_k^{(t)}$ is sent to the server via the uplink wireless channel by using the communication mode of the client device.

• Model parameter aggregating: the local models received by the server from all devices are aggregated to obtain a new global ML model, that is, $\omega^{(t+1)} = \frac{1}{K} \sum_{k \in K} \omega_k^{(t)}$.

By leveraging the FL framework, the initial dataset is stored locally and trained on edge devices. This eliminates the need for sharing data with other devices or servers and further ensures that only global models are achieved via the transmission of model parameters. This advantage may address the limitations of traditional ML methods. As shown in Fig. 3[9], FL achieves the construction of unified data among multiple nodes, providing higher quality services for big data applications by increasing data sample size, data types, data features, and data dimensions, and creating value for the future development of society. In comparison to traditional methods that gather data and train models based on cloud platforms, FL may be better equipped to handle dispersed computing tasks, while simultaneously preserving the privacy of user data. Additionally, FL may help alleviate the exponential increase in cost that arises from an increased data volume. It is also believed to be more user-friendly for larger mobile terminal scales and to provide advantages for a wider distributed deployment environment. By enabling the sharing and fusion of heterogeneous device data, FL may provide powerful support for future 6G environments.

3 Integrated Sensing, Communication and Computation

In traditional wireless networks, sensing, communication and computation are designed separately for various purposes. The isolated design principle is difficult to adapt to the strict requirements of emerging 6G applications, such as autopilot and virtual reality, which demand ultra-low latency, ultra-high reliability, and high capacity. Therefore, a new paradigm has emerged, which integrates communication and computation and comprehensively considers the application of data in downstream in a task-oriented way. As shown in Fig. 4, sensing, computation, and communication are highly coupled with FL in this new paradigm. Because of the fact that radio signals can be utilized for wireless communications and environmental sensing simultaneously, intelligent devices can analyze information about the detected target via wireless sensors in terms of range, positioning, imaging, etc. LIU et al. proposed a resource allocation approach toward ambient intelligence[10]. LI et al. introduced ISCC into over-the-air computation (AirComp) to improve spectrum efficiency and sensing performance, where function calculation from different user data is implemented by utilizing the overlay feature in wireless signal transmission in the air[11]. However, there is still a paucity of studies on the FL under the umbrella of ISCC in wireless networks.

3.1 Integrated Sensing and Communication

Integrated sensing and communication (ISAC) refers to the integration of sensing and communication into the unified de-
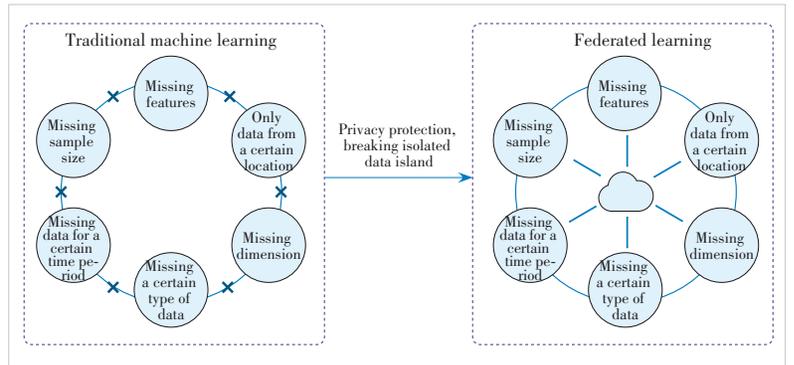
▲Figure 3. Comparison of traditional machine learning (ML) and federated learning (FL)[9]
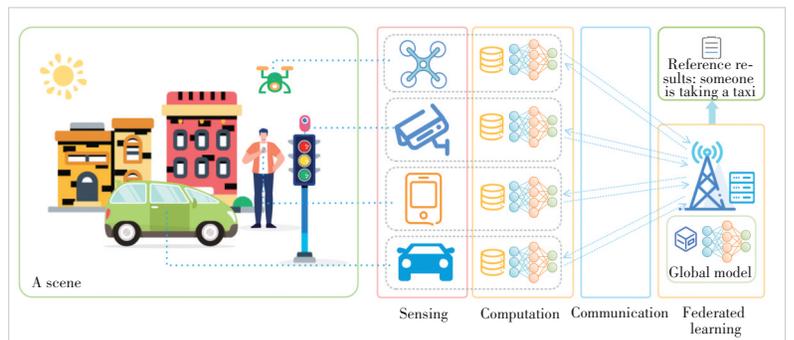
▲Figure 4. Sensing, communication and computation coupled with federated learning (FL)

Special Topic | Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation

ZHAO Moke, HUANG Yansong, LI Xuan

sign of wireless networks to improve spectral efficiency and achieve mutual benefit through sensor-assisted communications and communication-assisted sensing. Compared to traditional wireless networks, ISAC can use wireless infrastructure, spectrum and power resources for simultaneous communication and sensing, which is believed to improve system performance at a lower cost. Meanwhile, the primary challenge in ISAC is the tradeoffs between performance caused by the sharing wireless resources and the contradiction between sensing and communication. These tradeoffs include information theory limitations, physical performance, propagation channel, and cross-layer indicators[12]. There are three perception tasks: detection, estimation and recognition, which are all performed based on the collected signal or data information related to the sensing object.

The integration gain can be obtained through the development of a dual-function waveform that can sense and communicate simultaneously based on shared resources. The leading methods to attain this include scheduling orthogonal or non-overlapping wireless resources (time division/frequency division/space division/code division), using separate signal waveforms, and balancing communication and sensing performance for signal waveform sharing[12]. For example, LIU et al.[13] proposed a privacy protection vertical FL scheme based on distributed ISAC for cooperative object/human motion recognition. The method uses a dedicated frequency-modulated continuous wave signal for each edge device's target sensing and data exchange. It then converts the sensing data into a low-dimensional intermediate vector and transmits it to the edge device. LI et al presented two new FL algorithms that use compression sensing to reduce the communication burden in an IoT system[14]. JEON et al. proposed a compression sensing method for FL on large-scale multi-input multi-output communication systems, which is superior to the traditional linear beamforming method[15]. It can also reduce the performance gap between FL and centralized learning through reconstruction. Based on the above, orthogonal or non-overlapping wireless resources may help to reduce functional interference, but there will be resource competition between sensing and communication. And effective management of waveform interference is necessary to separate waveforms at the same frequency. These methods can improve the efficiency of the system's spectral, hardware, and information processing efficiency, but they come with higher computational complexity.

3.2 Integrated Communication and Computation

As data volumes from edge devices rapidly increase, it becomes challenging for edge servers to receive large amounts of data from edge devices quickly through wireless links due to limited wireless communication resources. This issue may be addressed through asynchronous communication and computation resource management and AirComp. AirComp is used to integrate computation into communication, which improves communication and computation efficiency, protects user privacy,

enhances user experience, and reduces delays caused by multiple access according to Ref. [16]. Compared with AirComp, communication and computing are in order in asynchronous communication and computation resource management. Optimizing resource management for asynchronous communication and computation can minimize energy consumption and delay to a certain extent, but scheduling complexity must be reduced. Compared with asynchronous communication and computation resource management, AirComp needs to consider the interference between AirComp and conventional communication applications, expand the scale of equipment, and conduct extensive performance evaluation under actual settings.

3.2.1 Asynchronous Computation and Communication

The ability of edge devices to update and upload the model status information to the edge server largely depends on their wireless channel qualities. When the edge devices operate under poor wireless channel conditions, it leads to longer model update time, which may delay the follow-up training. During model training, it is necessary to allocate wireless resources properly to improve learning performance, so that limited wireless resources can be fully utilized by collaborative asynchronous computation and communication resource management.

Various optimization algorithms can be employed to solve the problem of energy-delay allocation, or communication and computation performance can be adjusted to reduce energy consumption. On the one hand, it can be achieved by increasing bandwidth utilization. To further reduce the communication delay, ZHU et al.[17] roposed a broadband analog aggregation access scheme, which exploits the waveform superposition characteristics of the multi-access channel to achieve integrated communication and computation. The communication delay is independent of the number of edge devices connected to the access channel. On the other hand, it can also speed up the learning process to solve the problem of communication and resource constraints. For example, it can maximize the FL algorithm convergence rate under the conditions of power and energy[18 - 19]. However, most scenarios are characterized by co-existing sensing, communication and computation. Resource competition between communication and sensing also affects the performance of FL.

3.2.2 AirComp

Existing research mainly applies AirComp to different scenarios, or uses various optimization methods to minimize the signal mean square error[20 - 22] and also optimizes the model in FL by different means based on AirComp.

AMIRI et al. introduced error accumulation and how to sparsify gradient estimates based on AirComp. They proposed saving untransmitted gradient vectors in an error accumulation vector, updating the local model, and computing a new gradient vector based on this error accumulation vector in the next iteration[23]. CAO et al. proposed a "one-time" aggregation to im-

Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation | *Special Topic*

ZHAO Moke, HUANG Yansong, LI Xuan

prove communication efficiency while considering a new power control design to maximize the convergence rate[24]. The results show that the proposed power control strategy achieves a significantly faster convergence rate in FL than the fixed power control benchmark strategy. Since the error accumulation vector and gradient thinning correct the gradient computation process and make more efficient use of the bandwidth, this scheme improves the accuracy of the model based on AirComp. YANG et al. proposed a general integrated communication and computation scheme based on AirComp[25]. However, the experiment revealed that the model's accuracy would gradually decline due to the parameter aggregation mistake in AirComp caused by signal distortion.

4 Challenges in Federated Learning and Their Solutions

FL may partially solve the problem of limited computing and communication resources while preserving client privacy to a certain extent. Despite this, such technology still faces issues such as data heterogeneity, insufficient training accuracy, and low training efficiency. Various schemes are proposed to resolve these issues, as summarized in Table 1.

4.1 Participant Selection

Intelligent edge devices have limited computing capacity, making centralized data processing challenging. Thus, it is crucial to choose edge devices based on data heterogeneity and local models. However, present methods, like the time fairness scheme and throughput fairness scheme[26], may overlook the differences between learning tasks, leading to poor learning performance. Furthermore, due to the unique characteristics of edge intelligence, model updating requires wireless channel resources. Therefore, selecting the correct participants for each FL training round is vital[27].

4.1.1 Selection Based on Participants and Training Quality

Terminal equipment exhibits heterogeneity and non-independent identically distributed characteristics, resulting in substantial differences between terminal equipment models[28]. It is crucial to carefully select high-quality participants, efficiently train models, and ensure their robustness. To achieve this, different strategies are proposed. One strategy involves selecting high-quality participants. For example, LAI et al. proposed a client selection framework, which enables the identification and selection of valuable clients for training[29]. In another study, MONDAL et al. presented a distributed participant selection algorithm that minimizes the costs of energy consumption and data transmission while selecting the least number of participants with the same coverage[30]. ZHANG et al. utilized the FL framework to train lightweight neural networks that establish the relationship between context and sensor data quality[31]. Their approach leverages participants' context information to predict sensing data quality.

The other strategy is that participants' selection can be based on the quality of their local models. KATHAROPOULOS et al. proposed a power-of-choice strategy commonly used in the queue system[32]. According to their analysis, selecting the loss value as an important metric for the client can improve the convergence rate of the entire model. SATTLER et al. devised the clustered FL algorithm which divides the client into two partitions using their cosine similarity and checks partition consistency by testing the gradient norm of the client[33].

4.1.2 Selection Based on Improving Resource Management

Due to the large number of participants, the upload link may become congested, and differing participants could result in unproductive training rounds[34]. Selecting the best user cluster that aligns with limited communication and computational resources will ultimately help improve training efficiency, reduce training time, and enhance model accuracy. RIBERO et al. sug-

▼Table 1. Challenges in federated learning (FL) and their state-of-the-art solutions

Challenge	Specific Method	Advantages and Disadvantages
Participant selection	Participating clients are selected based on the heterogeneous nature of the data, quality of participants and training, and resource constraints.	Selecting participants can make full use of resources and is conducive to continuous training. However, when the data scale is too large, the overall performance cannot be guaranteed in the scenario of edge intelligence applications, and the training process needs to be optimized.
Adaptive aggregation	The best tradeoff is found between local updates and global parameter aggregation under a given resource budget to speed up the local training process.	By adapting the frequency of global aggregation, the performance of the model can be improved, and the utilization of available resources can be improved. However, the convergence of adaptive aggregation schemes currently only considers convex loss functions.
Incentive mechanism	FL requires an effective incentive mechanism for participation and balances rewards and limited communication and computing resources to improve data quality.	By quantifying data quality, the overall benefit of FL is generally improved, but due to the heterogeneity of the environment, the excitation obtained by different edge devices in FL does not match, making it difficult to balance game rewards and resource consumption.
Model compression	The transmission model is compressed to improve the communication efficiency between the server and client. Knowledge distillation exchanges model outputs, allowing edge devices to adopt larger local models.	Client-to-server parameter compression may cause convergence problems, increase computational complexity, and reduce training accuracy. Knowledge distillation alleviates the problem of independent and identical distribution of data to a certain extent, but the quality of wireless channel will affect the accuracy of model training.
Privacy protection	Privacy protection may be achieved through the inference of attacks, the encryption of data and models, and the improvement of privacy protection performance by blockchain technology.	FL may solve the privacy leakage problems caused by the model parameter sharing and multi-party communication and cooperation mechanism of FL. However, further research is needed when it comes to the security problems caused by data poisoning and the removal of traces left by participants' data in the local model, etc.

Special Topic | Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation

ZHAO Moke, HUANG Yansong, LI Xuan

gested that only transmitting client updates with a significant amount of information during each training round reduces the transmission pressure of FL[35]. This approach decreases communication costs during the training process while ensuring model accuracy by selecting the clients participating in each update. ABDULRAHMAN et al. proposed a multi-criteria participant selection algorithm that considers participants' central processing unit, memory, energy and task completion time for FL in IoT's resource-constrained environment[36]. This algorithm maximizes the number of participants while minimizing the number of communication rounds. To ensure long-term performance, XU et al. explored FL in typical wireless networks, identifying issues related to participant selection and bandwidth allocation in long-term client energy constraints. They proposed an online optimization algorithm based on Lyapunov to address these issues[37].

4.2 Adaptive Aggregation

In FL, the model updating procedure is primarily split into two steps: local model updating at clients and global aggregation, which involves uploading model parameters to the server and aggregating them into a global model. The adaptive aggregation problem of FL specifically aims at bandwidth aggregation and model parameter aggregation[38]. With limited resources, local model updating and global aggregation are modified to accelerate convergence and improve accuracy.

GUHA et al. suggested a single-round communication federation learning system to reduce the communications between clients and servers[37]. The entire training is carried out on the edge device, and only the local model parameters are uploaded and aggregated after the movement. Based on the greedy algorithm, HADDADPOUR et al. proposed a hypercluster algorithm that trained each local model several times using the client's local data and selected the model with the minimum training loss[39]. WANG et al. proposed a control algorithm to achieve the ideal tradeoff between the local update and the global aggregation[40]. Analyzing the convergence boundary of distributed gradient descent of FL, it minimizes the training loss under a given resource budget. ZHANG et al. proposed an FL framework with adaptive local aggregation, which captured the personalized data required by the client in the global model, downloaded the global model and local model for the adaptive aggregation, and initialized the local model on each client before trained in each iteration[41].

In the same training iterations, adaptive aggregation FL reaches better performance than the synchronous aggregation of all clients. With effective utilization of computation and communication resources, it obtains lower training loss and higher model accuracy and reduces the load of edge servers.

4.3 Incentive Mechanism

The incentive mechanism quantifies the quality of data that edge devices provide to reduce energy consumption and im-

prove model accuracy with the guarantee of data privacy and the lowest possible computation and communication costs[42]. The incentive mechanism for FL participants usually regards the edge device as the seller and provides them with training services while the server is regarded as the buyer of the data.

To improve the energy efficiency of model transmission, FENG et al. proposed a cooperative relay network-assisted parameter transmission scheme and corresponding service pricing mechanism[43], modeling the relationship between edge devices and FL servers as a Stackelberg game model[44]. SUN et al. investigated the air-ground dynamic digital twinning and joint learning and, on this basis, studied the FL incentive mechanism based on the Stackelberg game and proposed an adaptive adjustment incentive mechanism for the best user and customer selection in dynamic networks[45]. To ensure that the incentive budget is proportional to the value of the FL model and prevent the server from being forced to pay redundant rewards, RICHARDSON et al. proposed an incentive scheme based on influence to prevent the participants from receiving rewards due to redundant data[46]. Optimizing the FL incentive mechanism can effectively limit the number of the participants who falsely contribute to the work, reduce their motivation to phonily report expenses, and thus improve the overall performance of FL.

4.4 Model Compression

While the number of mobile devices rises sharply, it is challenging for mobile virtual network operators to provide low-cost and reliable access services for users due to deficient network infrastructure. The amount of uploaded data is also gradually growing in tandem with the widespread use of powerful ML on edge devices, resulting in significant bandwidth consumption and a decline in communication efficiency. Therefore, reducing the communication overhead in FL becomes an impending issue, which can be addressed by data compression, knowledge distillation, asynchronous parameter update, etc.[47].

4.4.1 Compressed Data Transmission

Compressing the transmission data is an effective measure to improve transmission efficiency. In the FL framework, model parameter compression technologies, such as network pruning, quantization and weight sharing, can be applied to reduce communication costs. Based on the uploaded information on gradient changes in the FL process, the model output value or intermediate value can also be compressed by gradient compression in different levels[48].

However, during data compression, noise is inevitable and will cause a discrepancy between the convergence result and the ideal solution, negatively impacting the effectiveness of FL. ROTHCHILD et al. proposed to reduce the number of communication rounds in FL by directly retrieving the latest gradient value without updating its position in the vector[49].

Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation | *Special Topic*
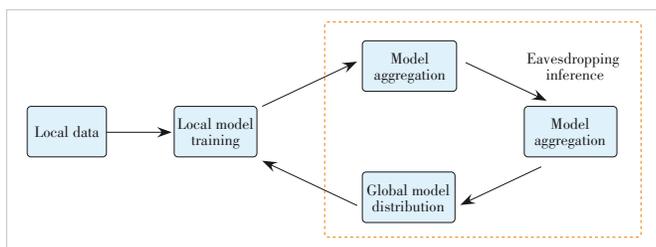
ZHAO Moke, HUANG Yansong, LI Xuan

4.4.2 Knowledge Distillation

Knowledge distillation (KD) can transfer knowledge from one neural network to another by exchanging soft predictions rather than the whole model[47]. KD loss includes mild loss and local training loss. KD is employed to mine the global model knowledge[50]. In addition, LI et al. proposed the federated model distillation algorithm to train heterogeneous models in a way that protects privacy[51]. LIN et al. adopted integrated distillation to migrate the knowledge of all heterogeneous client models to the global model in each global iteration[52]. However, these algorithms ignore the further personalized needs of clients participating in training. ZHANG et al. proposed the knowledge transfer personalized FL algorithm, which parameterizes the similarity of paired clients and uses KD to transfer the personalized soft prediction knowledge to the local[53]. CHO et al. used cluster co-distillation to migrate the understanding of clients with similar data distribution to the local model[54]. DIVI et al. used KD to solve the problem of FL data heterogeneity and proposed the personalized FL algorithm, which was carried out in two stages[55]. The first stage simplies FL. In the second stage, each user selects the best teacher model from the global model of each iteration and distills it to achieve personalization.

4.5 Privacy Protection

FL is a method that is believed to effectively solve the problem of user privacy disclosure, since it does not require edge devices to upload their data. The user equipment builds a network model, generates local model parameters based on local data, and uploads them to the central server, which aggregates local model parameters to the global model, effectively protecting user privacy and easing the burden of communication bandwidth. The gradient should be protected because even if the training data stay inside the local area, attackers can still exploit these shared gradients to reverse the content of the original training data[56], exposing the training data to the public. At the same time, malicious participants or collaborators can use the intermediate information transmitted during the FL training process to launch member inference attacks[57] or data reconstruction attacks[58], exposing participants in FL to privacy disclosure threats. Fig. 5 shows potential privacy attacks in the processes of FL.

By encrypting gradient parameters, it is possible to solve the privacy disclosure issue that might arise during the process of uploading and downloading model parameters. At the same time, some researchers have focused on improving encryption efficiency[59]. FL can consistently use the blockchain consensus mechanism to establish authentic interaction in an untrusted environment. The benefit generated from the blockchain reward mechanism can also encourage knowledge sharing in FL. The combination of blockchain technology and FL can improve data privacy and achieve performance isolation[60].

5 Applications and Prospects

FL is currently in use globally. For example, HART et al. used the federated averaging algorithm to predict the next word of the mobile phone keyboard input method[61]. MUHAMMAD et al. applied FL to the recommendation system[62]. Applying FL to predict the flow of urban global cellular networks can augment the data sets and improve the prediction accuracy of the model, without the problems of complexity and no real-time. Based on the FL framework, the central node collects the model updates transmitted by the edge base station for aggregation, so as to obtain a global model with good performance. The algorithm collects the data of vehicles and their tasks as input and allocates the multi-dimensional resources according to the output results of the model to meet the time-varied resource requirements and efficiently accomplish the computation tasks in the Internet of Vehicles system. Network function virtualization technology can transform traditional network hardware resources into virtual network resources. The two-way gated loop unit based on the distributed FL framework can predict virtual network function resource requirements.

Depending on its low latency and large data processing capacity, FL can also be used in the 6G era, which blends communication, sensing and computation together. For instance, due to various influences in the air[63], large-scale unmanned aerial vehicles (UAV) swarm to avoid collisions and quickly reach the destination. FL based on the wireless network can design the flight route of a UAV swarm well and better solve UAV persistent online decision-making problems by collecting sensing data from the surrounding environment[64], supporting UAV-assisted mobile edge computing with ultra-reliable and low-delay communication[65].

Despite the advancements of FL, it can be improved in several aspects, which are shown as follows.

• Various intelligent devices, including mobile phones, cameras, UAVs, depth sensors and radio sensors, produce data samples in various modalities and have a wide range of computational capacities. Researchers should put more resources to deal with the multimodal adaptive problem in FL based on ISCC.

• The applications of FL combine sensing, communication and computation together. In other words, ISCC can better conduct FL in the future. However, the resource management problems among these three aspects still need to be solved, and the hardware devices that integrate such three functions still need to be developed.

▲ Figure 5. Possible privacy attacks on federated learning (FL)

Special Topic | Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation

ZHAO Moke, HUANG Yansong, LI Xuan

• With the development of FL, privacy attacks against FL are also expanding. It is still necessary to improve encryption efficiency based on existing encryption algorithms and continue to explore the combination of blockchain and FL to form a new edge computing paradigm with higher security.

• The number of IoT devices is increasing, while the data generated by the equipment are also expanding rapidly. In the face of scenarios that include various IoT devices[66], the FL frameworks need to couple with the access of intelligent edge devices with different task attributes. Therefore, it still requires constant exploration of practical coupling design for them and efficiency improvement of FL while ensuring the accuracy of the model.

6 Conclusions

This paper summarizes the development of FL and classifies related technologies according to the challenges that FL faces. Among these technologies, ISCC is the most significant one for its high coupling with FL. Besides, this paper introduces the research on device, data and model heterogeneity in FL and demonstrates different challenges and the existing work about FL, including participant selection, adaptive aggregation, incentive mechanism and game model, model compression, and privacy protection. In the end, the applications and prospects of FL in reality are presented.

References

[1] LETAIEF K B, SHI Y M, LU J M, et al. Edge artificial intelligence for 6G: vision, enabling technologies, and applications [J]. IEEE journal on selected areas in communications, 2022, 40(1): 5 – 36. DOI: 10.1109/JSAC.2021.3126076

[2] Cisco. Cisco annual internet report (2018 – 2023) white paper [R]. 2020

[3] DUNNE R, MORRIS T, HARPER S. A survey of ambient intelligence [J]. ACM computing surveys, 2021, 54(4): 73. DOI: 10.1145/3447242

[4] CUSTERS B, SEARS A M, DECHESNE F, et al. EU personal data protection in policy and practice [M]. Switzerland: Springer Nature, 2019

[5] ZHU G X, LIU D Z, DU Y Q, et al. Toward an intelligent edge: wireless communication meets machine learning [J]. IEEE communications magazine, 2020, 58(1): 19 – 25. DOI: 10.1109/MCOM.001.1900103

[6] YU S, CHEN X, YANG L, et al. Intelligent edge: leveraging deep imitation learning for mobile edge computation offloading [J]. IEEE wireless communications, 2020, 27(1): 92 – 99. DOI: 10.1109/MWC.001.1900232

[7] MCMAHAN H B, MOORE E, RAMAGE D, et al. Federated learning of deep networks using model averaging [EB/OL]. (2016-02-17)[2023-03-05]. https://arxiv.org/abs/1602.05629

[8] YANG Z H, CHEN M Z, WONG K K, et al. Federated learning for 6G: applications, challenges, and opportunities [J]. Engineering, 2022, 8: 33 – 41. DOI: 10.1016/j.eng.2021.12.002

[9] China Academy of Information and Communications Technology. Federal learning scenario application research report (2022) [R]. 2022

[10] LIU P X, ZHU G X, WANG S, et al. Toward ambient intelligence: federated edge learning with task-oriented sensing, computation, and communication integration [J]. IEEE journal of selected topics in signal processing, 2023, 17(1): 158 – 172. DOI: 10.1109/JSTSP.2022.3226836

[11] LI X Y, LIU F, ZHOU Z Q, et al. Integrated sensing and over-the-air computation: dual-functional MIMO beamforming design [C]//1st International Conference on 6G Networking (6GNet). IEEE, 2022. DOI: 10.1109/6GNet54646.2022.9830500

[12] LIU F, CUI Y H, MASOUROS C, et al. Integrated sensing and communications: towards dual-functional wireless networks for 6G and beyond [J]. IEEE journal

on selected areas in communications, 2022, 40: (6): 1728 – 1767. DOI: 10.1109/JSAC.2022.3156632

[13] LIU P X, ZHU G X, JIANG W, et al. Vertical federated edge learning with distributed integrated sensing and communication [J]. IEEE communications letters, 2022, 26(9): 2091 – 2095. DOI: 10.1109/LCOMM.2022.3181612

[14] LI C X, LI G, VARSHNEY P K. Communication-efficient federated learning based on compressed sensing [J]. IEEE Internet of Things journal, 2021, 8(20): 15531 – 15541. DOI: 10.1109/JIOT.2021.3073112

[15] JEON Y S, AMIRI M M, LI J, et al. A compressive sensing approach for federated learning over massive MIMO communication systems [J]. IEEE transactions on wireless communications, 2021, 20(3): 1990 – 2004. DOI: 10.1109/TWC.2020.3038407

[16] JIANG T, SHI Y M. Over-the-air computation via intelligent reflecting surfaces [C]//IEEE Global Communications Conference (GLOBECOM). IEEE, 2020: 1 – 6. DOI: 10.1109/GLOBECOM38437.2019.9013643

[17] ZHU G X, WANG Y, HUANG K B. Broadband analog aggregation for low-latency federated edge learning [J]. IEEE transactions on wireless communications, 19 (1): 491 – 506. DOI: 10.1109/TWC.2019.2946245

[18] LIN F P-C, HOSSEINALIPOUR S, AZAM S S, et al. Semi-decentralized federated learning with cooperative D2D local model aggregations [J]. IEEE journal on selected areas in communications, 2021, 39(12): 3851 – 3869. DOI: 10.1109/JSAC.2021.3118344

[19] HOSSEINALIPOUR S, BRINTON C G, AGGARWAL V, et al. From federated to fog learning: distributed machine learning over heterogeneous wireless networks [J]. IEEE communications magazine, 2020, 58(12): 41 – 47. DOI: 10.1109/MCOM.001.2000410

[20] FU M, ZHOU Y, SHI Y M, et al. UAV-assisted multi-cluster over-the-air computation [J]. IEEE transactions on wireless communications, 2022, early access. DOI: 10.1109/TWC.2022.3227768

[21] LIU W C, ZANG X, LI Y H, et al. Over-the-air computation systems: optimization, analysis and scaling laws [J]. IEEE transactions on wireless communications, 2020, 19(8): 5488 – 5502. DOI: 10.1109/TWC.2020.2993703

[22] ZHU G X, XU J, HUANG K B, et al. Over-the-air computing for wireless data aggregation in massive IoT [J]. IEEE wireless communications, 2021, 28(4): 57 – 65. DOI: 10.1109/MWC.011.2000467

[23] AMIRI M M, GUNDUZ D. Federated learning over wireless fading channels [J]. IEEE transactions on wireless communications, 2020, 19 (5): 3546 – 3557. DOI: 10.1109/TWC.2020.2974748

[24] CAO X W, ZHU G X, XU J, et al. Optimized power control design for over-the-air federated edge learning [J]. IEEE journal on selected areas in communications, 2022, 40(1): 342 – 358. DOI: 10.1109/JSAC.2021.3126060

[25] YANG K, JIANG T, SHI Y M, et al. Federated learning via over-the-air computation [J]. IEEE transactions on wireless communications, 2020, 19(3): 2022 – 2035. DOI: 10.1109/TWC.2019.2961673

[26] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data [C]//20th International Conference on Artificial Intelligence and Statistics. AISTATS, 2017: 1178 – 1187

[27] SINGH P, SINGH M K, SINGH R, et al. Federated learning: challenges, methods, and future directions [M]//Federated learning for IoT applications. Cham: Springer International Publishing, 2022: 199 – 214. DOI: 10.1007/978-3-030-85559-8_13

[28] WANG L P, WANG W, LI B. CMFL: mitigating communication overhead for federated learning [C]//39th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2019: 954 – 964. DOI: 10.1109/ICDCS.2019.00099

[29] LAI F, ZHU X F, MADHYASTHA H, et al. Oort: informed participant selection for scalable federated learning [EB/OL]. (2020-10-12)[2023-03-06]. https://arxiv.org/pdf/2010.06081v1.pdf

[30] MONDAL S, MITRA S, MUKHERJEE A, et al. Participant selection algorithms for large-scale mobile crowd sensing environment [J]. Microsystem technologies, 2022, 28(12): 2641 – 2657. DOI: 10.1007/s00542-022-05271-2

[31] ZHANG W, LI Z, CHEN X. Quality-aware user recruitment based on federated learning in mobile crowd sensing [J]. Tsinghua science and technology, 2021, 26 (6): 869 – 877. DOI: 10.26599/TST.2020.9010046

[32] KATHAROPOULOS A, FLEURET F. Biased importance sampling for deep neural network training [EB/OL]. (2017-05-31)[2023-03-05]. https://arxiv.org/abs/1706.00043

[33] SATTLER F, MÜLLER K, SAMEK W. Clustered federated learning: model-

Federated Learning for 6G: A Survey From Perspective of Integrated Sensing, Communication and Computation | *Special Topic*

ZHAO Moke, HUANG Yansong, LI Xuan

agnostic distributed multitask optimization under privacy constraints [J]. IEEE transactions on neural networks and learning systems, 2021, 32 (8): 3710 – 3722. DOI: 10.1109/TNNLS.2020.3015958

[34] WANG T, WANG P, CAI S B, et al. Mobile edge-enabled trust evaluation for the Internet of Things [J]. Information fusion, 2021, 75: 90 – 100. DOI: 10.1016/j.inffus.2021.04.007

[35] RIBERO M, VIKALO H. Communication-efficient federated learning via optimal client sampling [EB/OL]. (2020-07-30) [2023-03-07]. https://arxiv.org/abs/2007.15197

[36] ABDULRAHMAN S, TOUT H, MOURAD A, et al. FedMCCS: multicriteria client selection model for optimal IoT federated learning [J]. IEEE Internet of Things journal, 2021, 8(6): 4723 – 4735. DOI: 10.1109/JIOT.2020.3028742

[37] XU J, WANG H Q. Client selection and bandwidth allocation in wireless federated learning networks: a long-term perspective [J]. IEEE transactions on wireless communications, 2021, 20 (2): 1188 – 1200. DOI: 10.1109/TWC.2020.3031503

[38] SHI W Q, SUN Y X, HUANG X F, et al. Scheduling policies for federated learning in wireless networks: an overview [J]. ZTE communications, 2020, 18 (2): 11 – 19. DOI: 10.12142/ZTECOM.202002003

[39] HADDADPOUR F, MAHDAVI M. On the convergence of local descent methods in federated learning [EB/OL]. (2019-10-31)[2023-03-06]. https://arxiv.org/abs/1910.14425

[40] WANG S Q, TUOR T, SALONIDIS T, et al. Adaptive federated learning in resource constrained edge computing systems [J]. IEEE journal on selected areas in communications, 2019, 37(6): 1205 – 1221. DOI: 10.1109/JSAC.2019.2904348

[41] ZHANG J Q, HUA Y, WANG H, et al. FedALA: adaptive local aggregation for personalized federated learning [EB/OL]. (2022-12-02) [2023-03-06]. https://arxiv.org/abs/2212.01197

[42] WANG Z L, HU Q, LI R N, et al. Incentive mechanism design for joint resource allocation in blockchain-based federated learning [EB/OL]. (2022-02-18)[2023-03-08]. https://arxiv.org/abs/2202.10938

[43] FENG S H, NIYATO D, WANG P, et al. Joint service pricing and cooperative relay communication for federated learning [C]//International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, 2019. DOI: 10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00148

[44] TADELIS S. Game theory: an introduction [M]. USA: Princeton University Press, 2013

[45] SUN W, XU N, WANG L, et al. Dynamic digital twin and federated learning with incentives for air-ground networks [J]. IEEE transactions on network science and engineering, 2022, 9 (1): 321 – 333. DOI: 10.1109/TNSE.2020.3048137

[46] HU P, GU H L, QI J, et al. Design of two-stage federal learning incentive mechanism under specific indicators [C]//2nd International Conference on Big Data Economy and Information Management (BDEIM). IEEE, 2021. DOI: 10.1109/BDEIM55082.2021.00103

[47] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network [EB/OL]. (2015-03-09)[2023-03-05]. https://arxiv.org/abs/1503.02531

[48] AHN J-H, SIMEONE O, KANG J. Wireless federated distillation for distributed edge learning with heterogeneous data [C]//IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE, 2019. DOI: 10.1109/PIMRC.2019.8904164

[49] ROTHCHILD D, PANDA A, ULLAH E, et al. Fetchsgd: cmmunication-efficient federated learning with sketching [C]//37th International Conference on Machine Learning. PMLR, 2020:8253 – 8265

[50] LEE G, JEONG M, SHIN Y, et al. Preservation of the global knowledge by not-true distillation in federated learning [EB/OL]. (2021-06-06)[2023-03-07]. https://arxiv.org/abs/2106.03097

[51] LI D L, WANG J P. FedMD: heterogenous federated learning via model distillation [EB/OL]. (2019-10-08)[2023-03-01]. https://arxiv.org/abs/1910.03581

[52] LIN T, KONG L J, STICH S U, et al. Ensemble distillation for robust model fusion in federated learning [C]//34th Conference on Neural Information Processing Systems. NeurIPS, 2020: 2351 – 2363

[53] ZHANG J, GUO S, MA X S, et al. Parameterized knowledge transfer for personalized federated learning [C]//35th Conference on Neural Information Processing Systems. NeurIPS, 2021. DOI: 10.48550/arXiv.2111.02862

[54] CHO Y J, WANG J Y, CHIRUVOLU T, et al. Personalized federated learning for heterogeneous clients with clustered knowledge transfer [EB/OL]. (2021-09-16) [2023-03-05]. https://arxiv.org/abs/2109.08119

[55] DIVI S, FARRUKH H, CELIK B. Unifying distillation with personalization in federated learning [EB/OL]. (2021-05-31) [2023-03-06]. https://arxiv.org/abs/2105.15191

[56] MU Y Q. Deep leakage from gradients [EB/OL]. (2022-12-15) [2023-03-05]. https://arxiv.org/abs/2301.02621

[57] NASR M, SHOKRI R, HOUMANSADR A. Comprehensive privacy analysis of deep learning: passive and active white-box inference attacks against centralized and federated learning [C]//IEEE Symposium on Security and Privacy (SP). IEEE, 2019. DOI: 10.1109/SP.2019.00065

[58] ZHU L G, LIU Z J, HAN S. Deep leakage from gradients [C]//33rd Conference on Neural Information Processing Systems. NeurIPS, 2019: 14774 – 14784

[59] ZHOU C Y, FU A M, YU S, et al. Privacy-preserving federated learning in fog computing [J]. IEEE Internet of Things journal, 2020, 7 (11): 10782 – 10793. DOI: 10.1109/JIOT.2020.2987958

[60] HOU D K, ZHANG J, MAN K L, et al. A systematic literature review of blockchain-based federated learning: architectures, applications and issues [C]//2nd Information Communication Technologies Conference (ICTC). IEEE, 2021: 302 – 307. DOI: 10.1109/ICTC51749.2021.9441499

[61] XU J, GLICKSBERG B S, SU C, et al. Federated learning for healthcare informatics [J]. Journal of healthcare informatics research, 2021, 5(1): 1 – 19. DOI: 10.1007/s41666-020-00082-4

[62] MUHAMMAD K, WANG Q Q, O'REILLY-MORGAN, et al. Fedfast: going beyond average for faster training of federated recommender systems [C]//26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 2020: 1234 – 1242. DOI: 10.1145/3394486.3403176

[63] LI X X, LI Z T, OUYANG Y, et al. Using UAV to detect truth for clean data collection in sensor cloud systems [J]. ZTE communications, 2021, 19(3): 30-45. DOI: 10.12142/ZTECOM.202103005

[64] SHIRI H, PARK J, BENNIS M. Communication-efficient massive UAV online path control: federated learning meets mean-field game theory [J]. IEEE transactions on communications, 2020, 68 (11): 6840 – 6857. DOI: 10.1109/TCOMM.2020.3017281

[65] ZHANG P Y, XIE L F, XU J. Joint placement and resource allocation for UAV allocation for UAV assisted mobile edge computing networks with URLLC [J]. ZTE communications, 2020, 18(2): 49 – 56+82. DOI: 10.12142/ZTECOM.202002007

[66] ZHANG G X, DING X J, QU Z C. Space-terrestrial integrated architecture for Internet of Things [J]. ZTE communications, 2020, 18(4): 3 – 9. DOI: 10.12142/ZTECOM.202004002

Biographies

ZHAO Moke received her BE degree in electronic engineering from Beijing University of Posts and Telecommunications (BUPT), China in 2022. She is pursuing her master's degree in electronic engineering at BUPT. Her research interests include edge computing and wireless communications in 6G.

HUANG Yansong received his BE degree in electronic engineering from Beijing University of Posts and Telecommunications (BUPT), China in 2022. He is working toward his master's degree in electronic engineering at BUPT. His research interests include integrated sensing, communication and computation in 6G.

LI Xuan (xuan.li@bupt.edu.cn) is an associate professor of Beijing University of Posts and Telecommunications, China. Her research interests include distributed learning for 6G systems, edge computing, etc.

Future Vision on Artificial Intelligence Assisted Green Energy Efficiency Network

CHEN Jiajun[1,2], GAO Yin[1,2], LIU Zhuang[2], LI Dapeng[1,2]

(1. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China；
 2. Wireless Product R&D Institute, ZTE Corporation, Shanghai 201203, China)

Abstract: To meet the key performance requirement of the 5G network and the demand of the growing number of mobile subscribers, millions of base stations are being constructed. 5G New Radio is designed to enable denser network deployments, which raises significant concerns about network energy consumption. Machine learning (ML), as a kind of artificial intelligence (AI) technologies, can enhance network optimization performance and energy efficiency. In this paper, we propose AI/ML-assisted energy-saving strategies to achieve optimal performance in terms of cell shutdown duration and energy efficiency. To realize network intelligence, we put forward the concept of intrinsic AI, which integrates AI into every aspect of wireless communication networks.

Keywords: machine learning; energy efficiency; traffic distribution; load prediction; intrinsic AI

1 Introduction

Wireless communication has witnessed rapid development, especially in terms of higher data rates, considerably smarter devices, and diverse applications. Moreover, compared with the 4G technology, 5G uses high-frequency bands, which makes the nodes denser. To achieve optimal performance in the radio access network (RAN) and meet the demand of increasing mobile subscribers, millions of base stations (BSs) are constructed. The number of BSs in developing regions has increased by over 2 million from 2007 to 2015, and data transmission rates have increased tenfold every five years[1]. However, the expected surge in traffic load requires 5G New Radio to enable denser network deployment and network densification, which results in higher energy consumption. Most of the energy is consumed by BSs in the typical RAN. However, with the deployment of more base stations with massive multiple-input multiple-output (MIMO), energy efficiency in NR becomes more urgent and challenging.

One of the energy-saving schemes that have received extensive attention from academia and the industry is cell activation/deactivation based on load prediction[2 – 3]. In the third Generation Partnership Project (3GPP), energy-saving standard cases have been specified in Releases 15 and 16, such as the intra-radio access technology (RAT) case with the central unit-distributed unit (CU-DU) split, the intra-system inter-RAT case, and multiple radio access technology-dual connectivity (MR-DC)[4]. An approach has also been recently developed to optimize wireless communications and introduced into self-organizing networks (SON) to allow for smarter operation and maintenance of operators' daily tasks[5]. The inclusion of AI-based tools enables a more proactive approach to exploiting the vast number of data available and incorporating additional dimensions, such as end-user experience and behavior characterization[6 – 8]. The cell providing capacity booster can be switched off autonomously according to its cell traffic load status. Ref. [9] leverages AI/ML methods to predict load and achieve energy efficiency performance through dynamic threshold configuration.

In this paper, we introduce and provide related works on AI/ML based energy efficiency, simulation and evaluation in real environments, and future vision on AI/ML based wireless networks. The main contributions of this work can be summarized as follows:

1）The benefits of AI/ML enabled wireless networks and the deployment of RAN intelligence are provided.

2）Compared with no-energy saving schemes and traditional energy saving strategies, the proposed AI/ML based energy saving scheme achieves great performance on power consumption and energy efficiency.

3）Based on further consideration of future wireless communication networks, we propose to integrate AI into every aspect of wireless communication systems to depict a vision of the intrinsic AI through intelligent data perception, intelligent modeling, distributed architecture, and intelligent monitoring.

The rest of this paper is organized as follows. Section 2 introduces AI/ML assisted wireless networks and describes their benefits. The definition of typical energy-saving features and the AI/ML based energy efficiency strategy is provided in Section 3, followed by the simulation and results in Section 4. Future vision on AI/ML based wireless communication networks is provided in Section 5. Section 6 shows the conclusions.

2 Machine Learning Assisted Wireless Networks

As an important research direction of AI technologies, machine learning takes advantage of the depth of the neural network's non-linear processing capability, which successfully solves a series of previously intractable problems. In image recognition, speech processing, and natural language processing, AI shows greater performance than humans and has better ability than traditional algorithms[10]. It therefore has been successfully applied in a variety of technologies, services and applications, including telecommunications. Many optimization issues in the wireless network, such as network energy saving, mobility optimization, and load balancing, can be resolved through powerful tool-machine learning, analyzing the data pattern and historical information to predict the trends or generate optimization decisions. Fig. 1 illustrates the functional framework for RAN intelligence[11].

1) Data collection is the function that provides input data to model training and model inference functions. Training data represent the data needed as input for the AI/ML model training function, while inference data are those needed as input for the AI/ML model inference function.

2) Model training is the function that performs the AI/ML model training, validation, and testing, which may generate model performance metrics as part of the model testing procedure. The model training function is also responsible for data preparation (e. g., data pre-processing and cleaning, format-

ting, and transformation).

3) Model inference is the function that provides AI/ML model inference output (e. g., predictions or decisions). The model inference function may provide model performance feedback to the model training function when applicable. The model inference function is also responsible for data preparation (e. g., data pre-processing and cleaning, formatting, and transforming). The inference output of the AI/ML model is produced by the model inference function.

4) Actor is the function that receives the output from the model inference function and triggers or performs corresponding actions. The actor may trigger actions directed to other entities or at itself.

5) Model deployment/update is used to initially deploy a trained, validated, and tested AI/ML model to the model inference function or to deliver an updated model to the model inference function.

6) Model performance feedback is used for monitoring the performance of the AI/ML model. After model inference is executed, the model performance is generated and returned to the model training function to evaluate whether the model performance is good or not. If the performance is not good, the model training function can trigger model retraining and reselection.

7) Feedback is the information needed to derive training data and inference data, or to monitor the performance of the AI/ML model and its impact on the network through updating key performance indicators (KPIs) and performance counters.

The 5G network and even future networks require the introduction of AI/ML to achieve automated and intelligent operations. For network energy saving, ML algorithms may predict the energy efficiency and load state of the next period, which can be used to make better decisions on cell activation/deactivation for energy saving. Based on the predicted load, the system may dynamically configure the energy-saving strategy. For mobility optimization, many radio resource management (RRM) actions related to mobility (e.g., selecting handover target cells) can benefit from the predicted user equipment (UE) location/trajectory. For load balancing, based on a collection of various measurements and feedback from UE, network nodes, historical data, etc., AI/ML model-based solutions and predicted load could improve load balancing performance, in order to provide a higher quality user experience and to improve the system capacity.

3 Strategies for Energy Efficiency

3.1 Energy Efficiency Features

Typical energy efficiency strategies used for wireless networks include symbol shutdown, channel shutdown, carrier shutdown and deep sleep, which can be categorized into a symbol level, a physical channel level, and a machine level. Following are the definitions of each energy efficiency strategy.
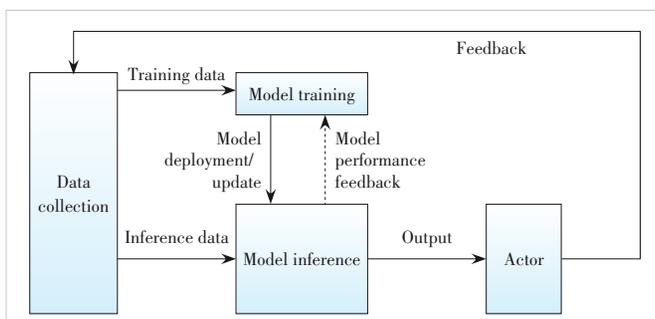
▲ Figure 1. Framework of artificial intelligence/machine learning (AI/ML) enabled wireless networks

1) Symbol shutdown: A base station detects that some downlink symbols have no data to send, and thus it turns off the PA and other analogue components, thereby reducing the power consumption of the base station, which basically has no effect on the user's latency. Moreover, by adjusting the number of synchronization signal block (SSB) beams when cells have no traffic or light traffic in a specified period of time, the proportions of symbols for which symbol power saving can take effect can be increased.

2) Channel shutdown: It refers to the technology of multichannel base stations such as 64/32 channels by muting some RF channels of the base station with low traffic, thereby reducing the power consumption of the base station. But for services with higher throughput requirements, when the user channel environment deteriorates, it is necessary to consider that the coverage cannot be reduced.

3) Carrier shutdown: When the service volume of the entire BS is low during off-peak hours at night , the BS energy consumption can be reduced by retaining only the coverage-layer cells and shutting down the capacity-layer cells. If the service load is lower than a specified threshold, the capacity layers are dynamically shut down. When the load of the carrier providing basic coverage is higher than a specified threshold, the base station dynamically turns on the carriers that have been shut down for service provisioning.

4) Deep sleep: The power requirements of the radio base station vary with the cell traffic load. As the service load of the cell increases, the power amplifier gradually becomes the most energy-consuming component of the base station. Additionally, in the scenario of no traffic load, the power demand of the wireless base station mainly comes from the digital intermediate frequency module. It is worth noting that in the absence of traffic load, and between control signaling transmissions, the BS part consumes energy even when transmissions are not required. Thus, a strategy arises to reduce unnecessary radio BS energy consumption by gradually deactivating components when they remain unused for transmission.

Energy efficiency strategies can be adopted based on the various wireless network environment, and different energy efficiency strategies produce different energy-saving results. In addition, AI/ML technologies can be used to help choose which energy-saving strategy for a certain scenario. In this paper, the channel shutdown and symbol shutdown are mainly used in simulation and evaluation in Section 4.

### 3.2 AI/ML Assisted Energy Efficiency Strategy

Cell activation/deactivation is an energy saving scheme in the spatial domain that exploits traffic offloading in a layered structure to reduce the energy consumption of the whole RAN. When the expected traffic volume is lower than a fixed threshold, the cells may be switched off, and the served UE may be offloaded to a new target cell. Efficient

energy consumption can also be achieved by other means such as reduction of load, coverage modification, or other RAN configuration adjustments. The optimal energy saving decision depends on many factors including the load situation at different RAN nodes, RAN nodes capabilities, KPI/quality of service (QoS) requirements, number of active user devices, UE mobility, cell utilization, etc. AI/ML techniques could be utilized to optimize the energy saving decisions by leveraging the data collected in the RAN network. AI/ML algorithms may predict the energy efficiency and load state of the next period, which can be used to make better decisions on cell activation/deactivation for energy saving. Based on the predicted load, the system can dynamically configure the energy-saving strategy (such as the switch-off timing and granularity and offloading actions) to keep a balance between system performance and energy efficiency and to reduce energy consumption.

Moreover, using statistics of past and current cell traffic and mobility management events, radio resource and mobility management strategies can be optimized to devise appropriate green strategies to minimize network energy consumption, while avoiding the degradation of network performance in terms of coverage quality, user rate, and handover failures. Since both the load of the serving cell and those of the neighboring ones play a role in the carrier shutdown procedure, optimized traffic distribution calls for centralized cell load predictions or exchange of predictions across cells. Only when the predicted load of the source and target cells is low, the source cell may be deactivated and the target cell is handed over to its UE to avoid QoS degradation. An optimized traffic distribution should also account for hard-to-predict traffic fluctuations and future cell loads as well as the change in signal quality of the UE in the neighborhood of a shutdown carrier, which in turn may also significantly affect cell loads in the area. A new target cell may be handed over to UE, but the load of this target cell may rapidly increase soon after, making the target cell fail to meet the required QoS, and require an immediate reactivation of the recently shutdown carrier, shown in Fig. 2.

4 Evaluation and Performance

The energy-saving strategy for the 5G system adopted in this paper comprises three components: all-day symbol shutdown, all-day AI-based channel shutdown, and AI-based deep sleep from 0 a.m. to 6 a.m. The test area is located in the Panyu District of Guangzhou, China, and involves 54 active an-
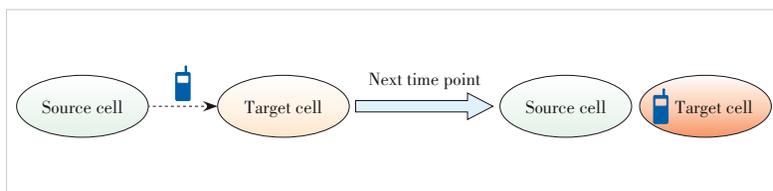
▲Figure 2. Potential scenarios cause deterioration of energy efficiency

tenna units (AAU). The energy-saving strategy is triggered by the predicted load threshold, as described in Section 3. Three-time ranges are set to compare the performance of no energy-saving, traditional energy-saving, and AI/ML-based energy-saving. T0 represents the time without energy-saving strategies, T2 represents the time when the AI/ML technology is used, and T1 represents the time with traditional energy-saving strategies. The simulation's configuration is illustrated in Table 1.

Table 2 displays statistical information on the duration of energy-saving strategies using AI/ML techniques compared with traditional methods. The data indicates a significant increase in the duration of deep sleep and channel shutdown, which has resulted in even greater energy savings. Specifically, the use of AI/ML techniques has extended the duration of shutdown for these strategies, resulting in a reduction of power consumption of 452.18 W and a 2.48% improvement in energy efficiency, compared with traditional methods. This information highlights the potential benefits of using AI/ML techniques for energy-saving purposes, particularly in extending the duration of energy-saving strategies, bringing significant energy savings. Fig. 3 shows the average power supply

▲Figure 3. Average power supply

▲Figure 4. Energy efficiency improvement

and a 23.87% reduction in power consumption with the use of AI/ML techniques. Fig. 4 shows the energy efficiency improvement of 23.40% achieved by using AI/ML techniques.

Overall, by utilizing AI/ML techniques to determine energy-saving strategies, energy efficiency can be significantly improved while simultaneously reducing energy consumption, leading to an increase in the energy-saving duration of base stations.

5 Future Vision on AI/ML Assisted Wireless Networks

The current 5G communication system is designed as a service-based architecture, providing a modular framework for meeting stringent latency and reliability requirements. Barely introducing AI technologies in wireless networks to solve a certain network optimization problem does not enable networks to be intelligent. Continuously changing radio environment requires retraining and updating the fixed ML models, resulting in repetitive work and hindering the intelligence of the wireless system. Future communication systems are not only considered to apply AI to enhance a certain function,

▼Table 1. Configuration information of evaluation

Time	Range	Type	Energy Saving Strategy
T0	2022-06-09 ~ 2022-06-15	None	W/O channel shutdown W/O symbol shutdown W/O deep sleep
T1	2022-05-27 ~ 2022-06-22	Tradition	Channel shutdown Symbol shutdown Deep sleep
T2	2022-06-25 ~ 2022-07-01	AI/ML assisted	AI/ML channel shutdown AI/ML symbol shutdown Deep sleep

AI: artificial intelligence   ML: machine learning

▼Table 2. Time statistics of the duration of shutdown

Phase	Deep Sleep/h	Channel Shutdown/h	Symbol Shutdown/h	Power Consumption/ W	Improvement (Compared with T0)	Improvement (Compared with T1)
T0-W/O ES	0	0	0	593.97	-	-
T1-Traditional ES	2.61	0.58	9.63	466.92	21.39%	2.48%
T2-AI/ML ES	3.48	5.94	9.03	452.18	23.87%	-

AI: artificial intelligence   ES: energy saving   ML: machine learning

▼Table 3. Time statistics of the duration of energy efficiency

Phase	5G Energy Efficiency /GB•(kW•h)$^{-1}$	Improvement (Compared with T0)
T0-W/O ES	2.78	-
T1-Tranditional ES	2.97	6.55%
T2-AI/ML ES	3.44	23.40%

AI: artificial intelligence   ES: energy saving   ML: machine learning

such as energy saving, mobility management, etc., but also designed to integrate AI into every aspect of wireless communication systems to depict a vision of the intrinsic AI.

The key requirements of the 5G network are stringent latency and reliability in user scenarios, e.g, ultra-reliable low-latency communications (URLLC), enhanced mobile broadband (eMBB), and massive machine type of communication (mMTC), while future wireless networks (6G) seamlessly integrates with AI, communication ne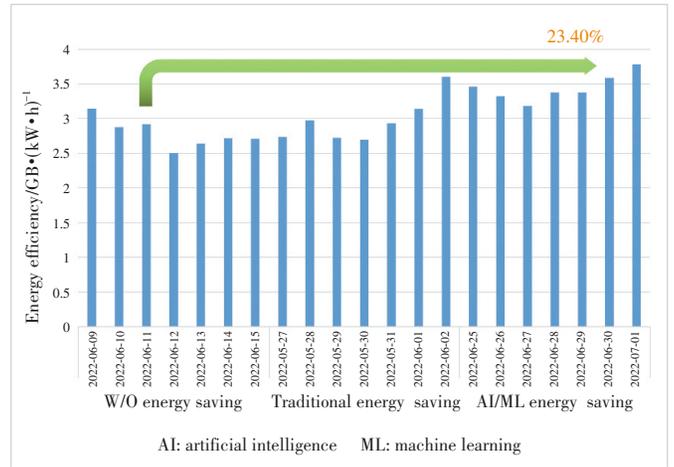tworks, and edge computing. To support AI function, 5G networks preset functional modules to monitor and enhance the performance of service-based architecture (SBA), but in the future wireless network, it involves self-sensing, self-learning, self-decision and self-evolution to support self-capability and realize AI/ML integration with communication networks. The transition in the AI/ML structure from 5G networks to future wireless networks is shown in Fig. 5.

Therefore, to bring autonomous learning, autonomous decision-making, self-optimization, and self-evolution, AI-native radio networks will be an intelligent loop including intelligent data perception, intelligent modeling, distributed architecture, and intelligent monitoring.

1) Intelligent data perception

A large quantity of data transportation will bring burdens to the current interface. On the other hand, data sensed from the radio environment sometimes do not have the corresponding labels. Now with the generative adversarial networks (GANs), it will avoid transferring a large number of data between various nodes in the network and protect the data privacy. GAN can generate the required data to simulate real data and improve the performance of models to a certain extent.

2) Intelligent modeling

The implementation of machine learning usually requires a lot of manual intervention, such as data pre-processing, feature selection, model selection, hyper-parameter, and adjustment. Each ML mode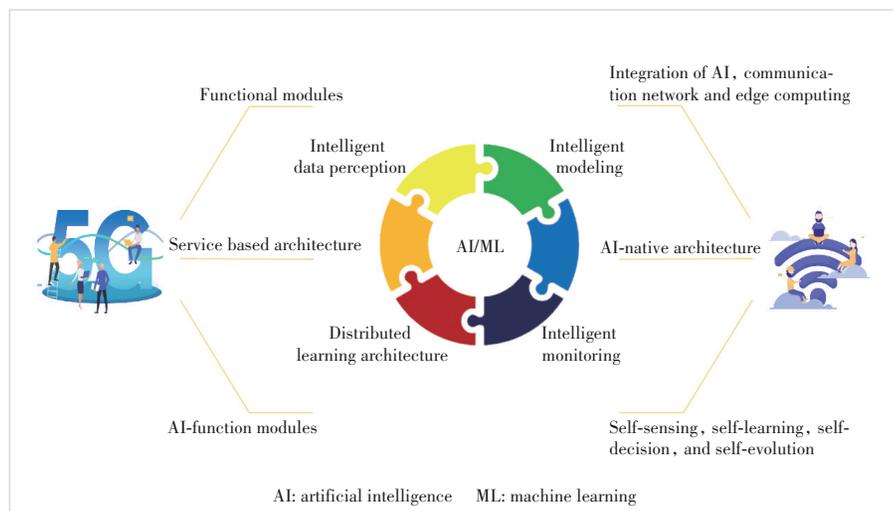l or algorithm has a specific structure and usually comes with a set of strategies or rules for model construction. The purpose of intelligent modeling is to reduce manual intervention so that the radio network can automatically generate and train the AI/ML model to achieve autonomous learning capabilities. Currently, auto-ML technologies can automatically select a machine learning model on given data and tasks, and automatically select an optimization algorithm, so that the model has the characteristics of high performance and low computational complexity of the task.

3) Distributed architecture

The centralized AI server collects the data of each network element and each node in the network, which will bring problems such as time delay and a large number of data transmission. The distributed AI server architecture can effectively solve this problem. The AI units are distributed on each network node to jointly perform the calculation tasks of the same set of AI models. Distributed AI network architecture, where each network node can be used as a part of AI training/execution and a large number of related devices can jointly build a common model based on locally collected data sets, will be the trend of intelligent AI network architecture in the future. It will reduce data transmission load and data privacy leaks on the radio interface, improve model performance, and alleviate delay problems. Distributed AI will play an important role in the subsequent evolution of the network architecture.

4) Intelligent monitoring

Intelligent monitoring is the introduction of human control into the decision-making process of the network to improve the decision-making ability of AI algorithms and help the machine better understand user preferences and make more user-preferred decisions. For example, when the AI model itself cannot make the correct decision or the cost of making the wrong decision is high, the AI algorithm can decide with human intelligence. With reinforcement learning, the agent obtains reward through interaction with the environment or feedback from users, learns the characteristics of the external environment, and improves decision-making strategies to adapt to the external environment changes.

6 Conclusions

In this paper, we introduce the benefits of AI/ML enabled wireless networks and provide the deployment of RAN intelligence. Compared with no energy saving schemes and traditional energy saving strategies, our proposed AI/ML based energy saving schemes achieve great performance on power consumption and energy efficiency. Moreover, we put forward further consideration on future wireless communication networks, which integrate AI into every aspect of wireless

▲Figure 5. Transition on AI/ML structure from 5G to future wireless network

(Figure labels: Functional modules; Integration of AI, communication network and edge computing; Intelligent data perception; Intelligent modeling; Service based architecture; AI/ML; AI-native architecture; Distributed learning architecture; Intelligent monitoring; AI-function modules; Self-sensing, self-learning, self-decision, and self-evolution; AI: artificial intelligence; ML: machine learning)

communication systems to depict a vision of the intrinsic AI through intelligent data perception, intelligent modeling, distributed architecture, and intelligent monitoring.

Reference

[1] ELKAZZI R, KHALIL A. Energy-saving solution for future cellular systems [C]//The 4th International Conference on Renewable Energies for Developing Countries (REDEC). IEEE, 2019: 1 – 6. DOI: 10.1109/REDEC.2018.8597671

[2] TONG E, WANG Y, DING F, et al. A practical eNB off/on based energy saving scheme for real LTE networks [C]//The 17th International Conference on Advanced Communication Technology (ICACT). IEEE, 2015: 12 – 17

[3] HAN T, ANSARI N. On greening cellular networks via multicell cooperation [J]. IEEE wireless communications, 2013, 20(1): 82 – 89. DOI: 10.1109/MWC.2013.6472203

[4] 3GPP. Eutran overall description: TS 36.300 [S]. 2019

[5] 3GPP. Study on ran-centric data collection and utilization for LTE and NR: TS 37.816 [S]. 2019

[6] FU Y, WANG S, WANG C X, et al. Artificial intelligence to manage network traffic of 5G wireless networks [J]. IEEE network, 2018, 32(6): 58 – 64. DOI: 10.1109/MNET.2018.1800115

[7] LI R P, ZHAO Z F, ZHOU X, et al. Intelligent 5G: when cellular networks meet artificial intelligence [J]. IEEE wireless communications, 2017, 24(5): 175 – 183. DOI: 10.1109/MWC.2017.1600304WC

[8] P'EREZ-ROMER J O, SALLENT O, FERR'US R, et al. Knowledge-based 5G radio access network planning and optimization [C]//The International Symposium on Wireless Communication Systems (ISWCS). IEEE, 2016: 359 – 365

[9] GAO Y, CHEN J J, LIU Z, et al. Machine learning based energy saving scheme in wireless access networks [C]//Proceedings of 2020 International Wireless Communications and Mobile Computing (IWCMC). IEEE, 2020: 1573 – 1578. DOI: 10.1109/IWCMC48107.2020.9148536

[10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84 – 90. DOI: 10.1145/3065386

[11] 3GPP. Study on enhancement for data collection for NR and EN-DC: TR 37.817 [S]. 2022

Biographies

Chen Jiajun (chen.jiajun1@zte.com.cn) received his master's degree in communication engineer from Shanghai University, China in 2019. He is currently a 5G research engineer at the R&D center of ZTE Corporation and State Key Laboratory of Mobile Network and Mobile Multimedia Technology, China. His research interests include 5G wireless communications and artificial intelligence.

GAO Yin received her master's degree in circuit and system from Xidian University, China in 2005. Since then she has been engaged in 4G/5G technologies and is currenly working as a wireless expert and project manager at the research center of ZTE Corporation and the State Key Laboratory of Mobile Network and Mobile Multimedia Technology, China. She has (co)-authored hundreds of proposals for 3GPP meeting and journal papers in wireless communications and has filed more than 200 patents. She has been served as RAN3 Vice Chair of 3GPP RAN3 since 2017 and was elected as Chair of 3GPP RAN3 in May 2021.

LIU Zhuang received his master's degree in computer science from Xidian University, China in 2003. He is currently a senior 5G research engineer at the R&D center of ZTE Corporation and the State Key Laboratory of Mobile Network and Mobile Multimedia, China. His research interests include 5G wireless communications and signal processing. He has filed more than 100 patents and submitted several hundreds of 3GPP contributions.

LI Dapeng received his MS degree in computer science from University of Electronic Science and Technology of China in 2003. He is currently a senior researcher in ZTE corporation and mainly focuses on research and implementation of the wireless access network system.

Machine Learning Driven Latency Optimization for Internet of Things Applications in Edge Computing

Uchechukwu AWADA[1], ZHANG Jiankang[2],

CHEN Sheng[3,4], LI Shuangzhi[1], YANG Shouyi[1]

(1. Zhengzhou University, Zhengzhou 450001, China；
 2. Bournemouth University, Poole BH12 5BB, UK；
 3. University of Southampton, Southampton SO17 1BJ, UK；
 4. Ocean University of China, Qingdao 266100, China)

Abstract: Emerging Internet of Things (IoT) applications require faster execution time and response time to achieve optimal performance. However, most IoT devices have limited or no computing capability to achieve such stringent application requirements. To this end, computation offloading in edge computing has been used for IoT systems to achieve the desired performance. Nevertheless, randomly offloading applications to any available edge without considering their resource demands, inter-application dependencies and edge resource availability may eventually result in execution delay and performance degradation. We introduce Edge-IoT, a machine learning-enabled orchestration framework in this paper, which utilizes the states of edge resources and application resource requirements to facilitate a resource-aware offloading scheme for minimizing the average latency. We further propose a variant bin-packing optimization model that co-locates applications firmly on edge resources to fully utilize available resources. Extensive experiments show the effectiveness and resource efficiency of the proposed approach.

Keywords: edge computing; execution time; IoT; machine learning; resource efficiency

1 Introduction

The Internet of Things (IoT) describes physical devices that are connected to the Internet or networks for the purpose of exchanging and sharing data. IoT enables direct fusion of physical devices into computer systems, resulting in efficiency, more reliable services and economic benefits without human intervention. However, most IoT devices have limited or no computing capability to meet some application-specific requirements. For example, emerging IoT technologies such as the smart city[1], healthcare-IoT[2], Internet of Vehicles (IoV)[3 – 5], connected and autonomous vehicles (CAVs)[6], and industry 4.0[7], require substantial resources to execute their applications. In addition, most of these applications are structured as a collection of loosely-coupled services that communicate with one another and are often latency-sensitive. A conventional approach is to offload these applications to a cloud computing (CC)[8] data center for execution. CC provides an on-demand availability of compute resources over multiple locations, each of which is a data center. However, a CC data center could be hundreds or thousands of miles away from the data sources, thereby jeopardizing the application performance through longer response time. A recent innovative distributed computing paradigm referred to as edge computing (EC)[9] brings computation and storage resources closer to the locations where they are needed, to reduce response time and save bandwidth. This enabling architecture deploys computation and storage resources at the edge of a network, and even beyond the edge of the network. It is important to note that EC computational resources are also limited compared to CC resources, but EC benefits IoT systems by deploying computing resources closer to end devices, thus reducing network traffic and latency to enable real-time insights. To this end, existing research works have exploited
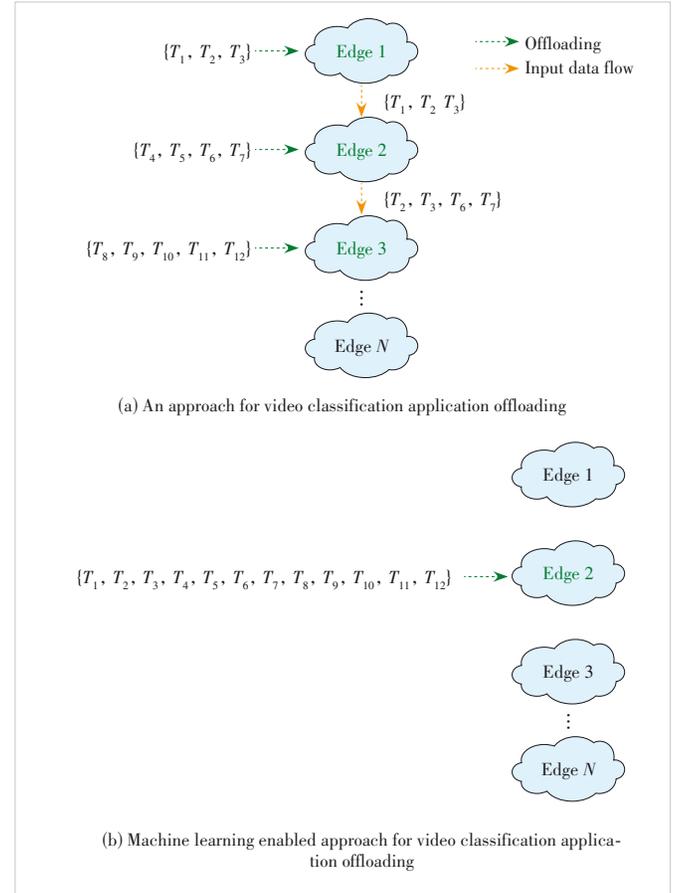
EC for task offloading in various IoT systems[3–5, 10–11]. Nevertheless, one fundamental challenge is where and how to offload and schedule complex applications so that their average latency is minimized and high resource efficiency is achieved. A common practice is to randomly offload applications or tasks individually to available edges without jointly considering task resource demands, task dependencies and edge resource availability. Such a disjointed approach would result in execution delays due to insufficient resource availability or tasks unable to communicate with their dependent tasks. Hence, it is not suitable for latency-sensitive tasks.

For example, the video classification application shown in Fig. 1(a) consists of 12 sub-applications $T_1, \cdots, T_{12}$, where T_1, T_2 and T_3 are independent tasks, whereas T_4 and T_5 require inputs from T_1 to be able to complete their executions. Similarly, T_6, T_7 and T_8 depend on the completion of T_4, T_5 and T_2, respectively. These make the execution of complex IoT applications very challenging. It is naturally important to offload and schedule such applications, to minimize their average latency. For instance, suppose each sub-application or tasks $T_1, \cdots, T_n$ of the application in Fig. 1(a) are randomly offloaded to different EC deployments, and then each dependent task would require the execution result(s) or input data from other task(s) to be transmitted back to its host edge deployment to complete its execution, as shown in Fig. 2(a). This transfer of input data is referred to as an input data flow, and such transmission would incur additional delay, thereby further affecting the average latency, given the rate and number of transmissions that could occur.

More specifically, assuming the video classification application in Fig. 1(a) is to be executed, the work in Ref. [5] proposed an approach as shown in Fig. 2(a), which offloads tasks T_1, T_2 and T_3 to Edge 1, tasks T_4, T_5, T_6 and T_7 to Edge 2, and the remaining tasks T_8, T_9, T_{10}, T_{11} and T_{12} to Edge 3. Since these tasks are interdependent tasks, the execution result of task T_1 needs to be transmitted from Edge 1 to Edge 2, to serve as the input data to tasks T_4 and T_5, while the execution results of tasks T_6 and T_7 need to be transmitted from Edge 2

to edge Edge 3, to serve as the input data to task T_{10}. Finally, the execution results of tasks T_2 and T_3 need to be transmitted $\langle c_1, m_1 \rangle$ from Edge 1 to Edge 3 to complete the video classification application execution.

In this paper, we show that machine learning (ML) techniques enable effective IoT task offloading and scheduling in edge computing systems. We propose an ML linear regression model to predict or estimate the resource requirements and

(a) An approach for video classification application offloading

(b) Machine learning enabled approach for video classification application offloading

▲Figure 2. Application offloading strategies

(a) Video classification application

(b) Video classification application, with each sub-application's CPU and memory resource requirements denoted as $\langle c, m \rangle$ and execution time denoted as E_{ex}
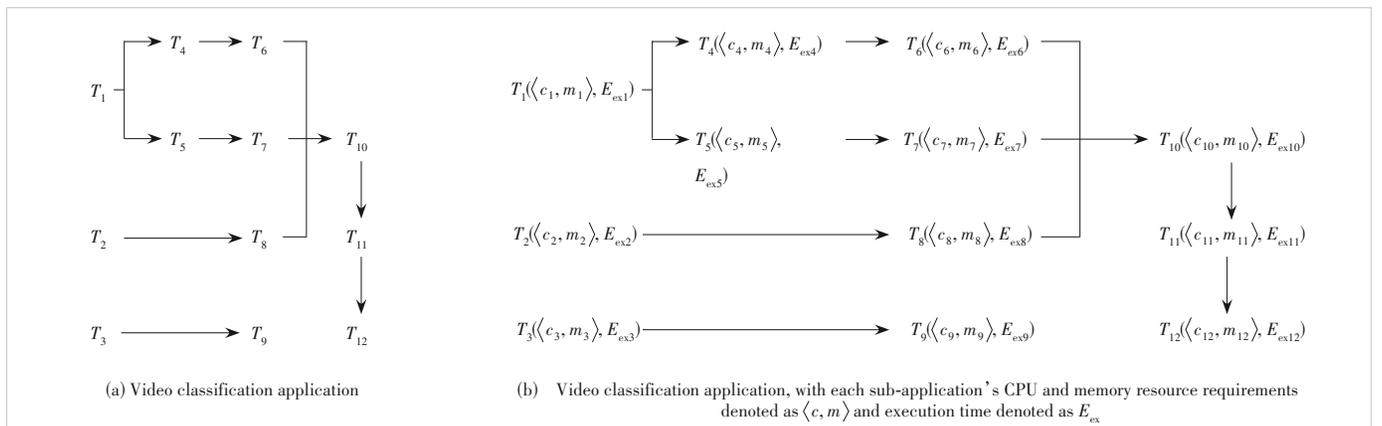
▲Figure 1. Directed acyclic graphs (DAG) of representative application

execution time of an application, as shown in Fig. 1(b), and intelligently offload them to an edge with sufficient resource availability, as shown in Fig. 2(b). This approach eliminates the need of input data flow, as sub-applications can communicate and share data quickly. However, upon arrival of an application in a suitable edge, the application may perform poorly if the sub-applications are scheduled naively, e.g., in an edge deployment that can only execute one task at any time, where each task is scheduled individually. Therefore, we further propose a variant bin-packing optimization that gang-schedules[12 – 13] and co-locates applications firmly on EC resources to fully utilize available resources. We aim to schedule and execute all the tasks by considering dependencies and resource demands, such that the actual scheduling and execution time is minimized. In summary, to achieve our Edge-IoT implementation, we address the following critical issues:

• We investigate a situation whereby multiple IoT systems can intelligently offload their complex applications to an edge deployment with sufficient resource availability to meet the resource-level demands of the applications, thus facilitating a resource-aware offloading scheme by enabling faster interactions among the applications to maximize their performance.

• Specifically, we derive a multi-task ML resource requirement and execution time estimation, so as to aid the selection of edge deployment with suitable resource availability.

• To guarantee optimal usage of edge resources and faster execution of tasks, we further propose a variant bin-packing optimization approach through gang scheduling of multi-dependent tasks, which co-schedules and co-locates tasks firmly on available nodes to avoid resource wastage.

• We show that Edge-IoT is capable of minimizing the response time of IoT applications using minimum resources, and conduct extensive experiments to compare the performance of our Edge-IoT with several existing approaches using real-world data-trace from Alibaba Cluster Trace Program, which provides information on task dependencies.

2 Related Work

Edge computing has been proven to make the IoT smarter by implementing smart connections and operation of IoT devices[14]. Emerging IoT technologies, such as the smart city[1], healthcare-IoT[2], Internet of Vehicles (IoV) [3 – 5], connected and autonomous vehicles (CAVs)[6], and industry 4.0[7], are utilizing EC for data analysis, processing and monitoring within their networks to improve both the efficiency and response speed. There are a huge number of existing works that have addressed the use of EC for IoT applications. For example, in Ref. [15], the authors studied multi-user IoT application offloading for a mobile edge computing (MEC) system and both the resources of computation and communication were cooperatively allocated. The proposed system focuses on minimizing both the weighted overhead of local IoT devices and

the offload measured by the delay and energy consumption. The authors in Ref. [16] formulated two novel optimization problems for delay-sensitive IoT applications, i. e., the total utility maximization problems under both static and dynamic offloading task request settings, to maximize the accumulative user satisfaction on the use of the services provided by an MEC system and show the non-deterministic polynomial time (NP)-hardness of the defined problems. Aiming to maximize the number of IoT devices through jointly optimizing the unmanned aerial vehicle (UAV) trajectory and service indicator as well as resource allocation and computation offloading, the authors in Ref. [17] formulated the optimization problem as a mixed integer nonlinear programming (MINLP) problem, where the chosen IoT devices would complete their computation tasks on time under given energy budgets and co-channel interference was taken into account. In Ref. [18], the authors studied the service home identification problem of service provisioning for multi-source IoT applications in an MEC network, by identifying a service home (cloudlet) of each multi-source IoT application for its data processing, querying and storage. They considered two novel service home identification problems. The work in Ref. [19] presented a joint optimization objective to evaluate the unavailability level, communication delay and resource wastage while allocating the same batch of IoT applications to multiple edge clouds. Then, the authors proposed an approach to minimizing the joint optimization objective under the condition of certain communication delays. In Ref. [20], the authors investigated the issue of joint cooperative edge caching and recommender systems to achieve additional cache gains by the soft caching framework. To measure the cache profits, they formulated the optimization problem as an Integer Linear Programming (ILP) problem, which is NP-hard.

The above methods leverage EC to offload IoT applications. They promise efficiency and better performance, but lack the consideration of a learning-based resource-aware offloading scheme with joint optimization of task resource demands and edge deployment resource availability. Therefore, we propose a joint optimization solution that guarantees faster offloading and execution of IoT applications in edge computing systems.

3 System Model and Problem Formulation

3.1 System Model

We consider an urban vehicular network environment where the IoV applications are offloaded from vehicles to EC deployments across various EC-enabled roadside units (RSUs), EC-enabled base stations (BSs), etc. We focus on V2I application offloading as illustrated in Fig. 3, where each vehicle is equipped with a powerful wireless interface that can be used to connect with RSUs, BSs, etc. We also consider the possibility that each vehicle is equipped with in-vehicle edge devices or deployment. For example, an in-vehicle EC deploy-
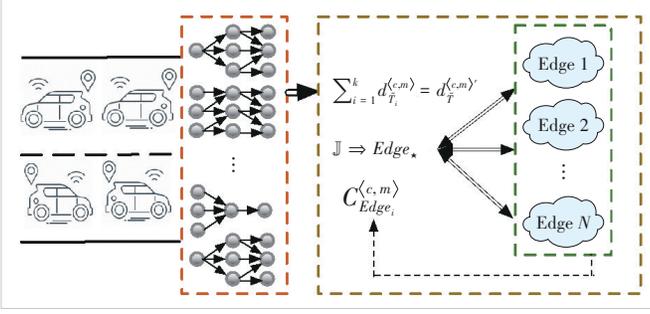
▲Figure 3. An example architecture of Internet of Vehicles (IoV) multi-task offloading

ment may not be as large as the deployments of RSUs, while those of the RSUs may not be as large as the deployments of BSs, etc., in terms of resource capacity. Therefore, IoV applications can be packaged in containers, i.e., Docker container provides a task offloading solution for isolation, portability and lightweight from devices to edge clusters, or to deploy it to the closest edge deployment with sufficient resource availability whenever it is needed. For such applications, let $\langle c, m \rangle$ represent the CPU and memory requirements.

Let $\mathbb{E} = \{ Edge_1, \cdots, Edge_M \}$ represent the set of individual participating edge deployments (i. e., in-vehicle, RSU, BS, etc.), as a cluster of container-instances (such as an edge device with virtualized container-optimized nodes). With the resource availability of each participating edge deployment $C_{Edge_i}^{\langle c,m \rangle}$, an informed decision on multi-task offloading can be made. Let $\mathbb{V} = \{ \mathcal{V}_1, \cdots, \mathcal{V}_M \}$ represent the index set of vehicles. A vehicle $\mathcal{V}_q$ can choose to execute its ready application locally in its in-vehicle edge device installation if there is sufficient resource availability or it is offloaded to the closest edge deployment $Edge_{i^*} \in \mathbb{E}$ with sufficient resource availability. Let $\vartheta \left[\mathcal{V}_q(t) \right]$ denote the offloading decision variable, which is measured by

$$\vartheta \left[\mathcal{V}_q(t) \right] = \begin{cases} 1, & \text{tasks are offloaded,} \\ 0, & \text{tasks are processed locally.} \end{cases} \quad (1)$$

A multi-task set $\mathbb{C} = \{ T_1, \cdots, T_N \}$ from the vehicles at time t requires much CPU and memory for execution. Such resource requirement, along with its execution time, is first predicted or estimated by a linear regression ML model. The multi-task features, $f_{mt}(\omega, \epsilon, \gamma)$ where ω is the number of instances, ϵ is the type of tasks, and γ is the dependency depth, are fed into the model $\Theta^\star$ to estimate the values of the resource requirement and execution time according to

$$f_{mt} \cdot \Theta^\star = \left[\tilde{E}_{ex_1} \tilde{T}_1^{\langle c, m \rangle} \tilde{E}_{ex_2} \tilde{T}_2^{\langle c, m \rangle} \cdots \tilde{E}_{ex_N} \tilde{T}_N^{\langle c, m \rangle} \right], \quad (2)$$

where $\tilde{T}_i^{\langle c, m \rangle}$ and $\tilde{E}_{ex_i}$ are the estimated resource requirement (in terms of CPU and memory $\langle c, m \rangle$) and estimated execution

time for task i, respectively. With these estimated values, a suitable edge deployment can be selected and multi-dependent tasks can be intelligently scheduled with the aim of minimizing their actual response time, while maximizing available resources. Assuming that $f_{mt} \in R^{1 \times d}$ is a d-dimensional vector (tensor), the predictor Θ is a $(d \times \epsilon)$-dimensional parameter matrix. We use historical data from previously executed tasks/jobs based on Keras to train the predictor Θ. Keras is a library that wraps TensorFlow complexity into a simple and user-friendly application programming interface (API). Dataset $\mathcal{DS} = \{ (\boldsymbol{x}_i, \boldsymbol{y}_i) \}_{i=1}^n$ contains d-dimensional tensors of data features $\boldsymbol{x}_i \in R^{1 \times d}$ and ϵ-dimensional tensors of labels (the actual execution times) $\boldsymbol{y}_i \in R^{1 \times \epsilon}$. The learning problem is to solve the following optimization:

$$\Theta^\star = \arg \min_{\Theta \in R^{d \times \epsilon}} \frac{1}{2n} \sum_{i=1}^n \| \mathbf{x}_i \Theta - \mathbf{y}_i \|_2^2 + \frac{\lambda}{2} \| \Theta \|_F^2, \quad (3)$$

where λ is the regularization parameter and $\| \cdot \|_F$ denotes the Frobenius norm. Optimization (3) is solved using gradient-descent, where the model is updated iteratively until convergence, i. e., $\Theta^{t+1} = \Theta^t - \eta (\frac{1}{n} g(\Theta^t) + \lambda \Theta^t)$, in which η is the learning rate, $\mathbf{g}(\Theta) = \frac{1}{n} X^T (X\Theta - Y)$ denotes the gradient of the loss function, $X = [\boldsymbol{x}_1^T \cdots \boldsymbol{x}_n^T]^T$ and $Y = [\boldsymbol{y}_1^T \cdots \boldsymbol{y}_n^T]^T$ are the feature set and label set, respectively. To guarantee the accuracy of the proposed model, we introduce the normalized absolute estimate error (NAEE), defined as:

$$\text{NAEE} = \frac{|\text{estimated value} - \text{actual value}|}{\text{actual value}}, \quad (4)$$

for both the resource requirement and execution time estimation, which serves as the estimation accuracy measure for the trained linear regression model.

At time t, while $\vartheta \left[\mathcal{V}_q(t) \right] = 0$, the multi-task set $\mathbb{C} \in \mathcal{V}_q$ is decided to perform local execution procedure in the vehicle $\mathcal{V}_q$; while $\vartheta \left[\mathcal{V}_q(t) \right] = 1$, $\mathbb{C} \in \mathcal{V}_q$ is otherwise to be offloaded to the edge deployment ($Edge_{i^*}$) with sufficient resources closest to $\mathcal{V}_q$. Multi-task set $\mathbb{C}$ is a loosely coupled inter-dependent application, as shown in Fig. 1, where each task $T \in \mathbb{C}$ has two resource requirements: CPU and memory, as the total number of estimated resources needed for its execution is denoted as $d_{\tilde{T}}^{\langle c, m \rangle}$. For each task $T \in \mathbb{C}$, let E_{sh}, E_{st} and E_{cp} denote its scheduling time, starting time and completion time, respectively. Therefore, the execution time of a task is thus:

$$E_{ex} = E_{cp} - E_{st}. \quad (5)$$

Existing offloading strategies (i.e., Refs. [4], [5], [21], etc.,) allow subtasks of an application or a job to be offloaded separately across different edge deployments, thus creating addi-

tional delay in the application's response time, as explained in Section 1. For example, when a vehicle in such an approach begins to offload its tasks, the delay includes three parts: 1) the time for offloading subtasks from the vehicle to different edge deployments, given as E_{of}, 2) the time for transmitting the results of executed subtasks (known as input data flow) from one edge deployment to another edge deployment, given as E_{sub}, and 3) the time for transmitting the final result from EC deployment to the vehicle, given as E_{rst}. Therefore, the response time of the vehicle's job is given as:

$$E_{rsp} = \sum_{T \in \mathbb{C}} (E_{of} + E_{sub} + E_{sh} + E_{ex}) + E_{rst}. \qquad (6)$$

In this paper, we aim to offload or dispatch a set of applications $\mathbb{C}$ belonging to a parked or moving vehicle $\mathcal{V}_q$ directly to a single and the closest edge deployment $Edge_i$, having sufficient resource capacity or availability to accommodate the tasks such that E_{of} is minimized, E_{sub} is avoided, as well as the overall E_{sh} and E_{ex} are minimized, namely,

$$\mathbb{C} \Rightarrow Edge_{\star}. \qquad (7)$$

Hence, the response time of the vehicle's job changes to:

$$E_{rsp} = E_{of} + \sum_{T \in \mathbb{C}} (E_{sh} + E_{ex}) + E_{rst}. \qquad (8)$$

Once $\mathbb{C}$ has been offloaded to $Edge_{\star}$, Edge-IoT utilizes the gang-scheduling[12–13] strategy to co-schedule all the applications at a time in $Edge_{\star}$. Given a cluster of container instances or nodes $I_i \in Edge_{\star}$, let $I_{Edge_{\star}}^{\langle c, m \rangle}$ denote each node's resource capacity or availability. In a real scenario where multi-vehicle set $\mathcal{V} \in \mathbb{V}$ offload multi-job tasks at t, these applications are offloaded as a multi-job set $\mathbb{J}$, i.e., $\mathbb{J} \Rightarrow Edge_{\star}$, where its collective estimated resource demand denoted as $\sum_{i=1}^{k} d_{\tilde{T}_i}^{\langle c, m \rangle} = d_{\tilde{T}}^{\langle c, m \rangle'}$. Hence, we can offload $\mathbb{J}$ to $Edge_{\star}$ with suitable resource availability. Therefore, the aggregate scheduling time and execution time of multi-job set $\mathbb{J}$ is given as:

$$\sum_{J \in \mathbb{J}} \sum_{i=1}^{k} \frac{E_{sh_i}}{k} = E_{sh}', \qquad (9)$$

$$\sum_{J \in \mathbb{J}} \sum_{i=1}^{k} \frac{E_{ex_i}}{k} = E_{ex}'. \qquad (10)$$

The estimated resource utilization of the edge for multi-job tasks is thus

$$\tilde{\mathcal{U}}_{Edge_i}^{\langle c, m \rangle} = \frac{\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c, m \rangle'}}{C_{Edge_i}^{\langle c, m \rangle}}. \qquad (11)$$

Similarly, $\tilde{\mathcal{U}}_{Edge_i}^{\langle c, m \rangle}$ includes CPU utilization $\tilde{\mathcal{U}}_{Edge_i}^{\langle c \rangle}$ and memory utilization $\tilde{\mathcal{U}}_{Edge_i}^{\langle m \rangle}$, which are defined respectively by

$$\tilde{\mathcal{U}}_{Edge_i}^{\langle c \rangle} = \frac{\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c \rangle'}}{C_{Edge_i}^{\langle c \rangle}}, \qquad (12)$$

$$\tilde{\mathcal{U}}_{Edge_i}^{\langle m \rangle} = \frac{\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle m \rangle'}}{C_{Edge_i}^{\langle m \rangle}}, \qquad (13)$$

where $\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c \rangle'}$ and $\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle m \rangle'}$ are the total collective estimated CPU and memory, respectively. After completing the multi-job executions, the final execution results are immediately and deterministically transmitted back to the vehicles.

3.2 Problem Formulation

The basic notations adopted are described in Table 1. The objectives are to minimize the response time, E_{rsp} in Eq. (8) for all $J \in \mathbb{J}$ and to maximize the computation or cluster resource utilization $\mathcal{U}_{Edge_i}^{\langle c, m \rangle}$ in Eq. (11), subject to certain constraints. The response time E_{rsp} in Eq. (8) comprises the dispatching or offloading time E_{of}, the scheduling time E_{sh}', the execution time E_{ex}', and the transmission time of final execution results E_{rst}. The closest computation offload-

▼Table 1. Notations

Notation	Description	Notation	Description
$\mathbb{E}$	A set of edge deployments	$\mathcal{V}, \mathbb{V}$	A vehicle, a set of vehicles
T	Individual application or task	I_i	Container-instance or node in a cluster
$\langle c, m \rangle$	CPU and memory resources	$I_i^{\langle c, m \rangle}$	Resource capacity or availability of a node
$\mathbb{C}$	A set of containerized applications	$C_{Edge_i}^{\langle c, m \rangle}$	Resource capacity/availability in an edge
$d_T^{\langle c, m \rangle}$	Application resource requirements	$U_{Edge_i}^{\langle c, m \rangle}$	Resources used for execution
$Edge_i$	Individual edge deployment or cluster	$U_{Edge_i}^{\langle c \rangle}, U_{Edge_i}^{\langle m \rangle}$	CPU, memory resource used for execution
$Edge_{\star}$	Closest edge deployment or cluster	$RU_{Edge_i}^{\langle c, m \rangle}$	Actual resources usage of jobs
$RU_{Edge_i}^{\langle c \rangle}, RU_{Edge_i}^{\langle m \rangle}$	Actual CPU, memory resources usage	E_{st}, E_{cp}	Application/task start, completion time
E_{ex}	Application or task execution time	$\mathcal{U}_{Edge_i}^{\langle c, m \rangle}$	Cluster resource utilization
$\mathcal{U}_{Edge_i}^{\langle c \rangle}, \mathcal{U}_{Edge_i}^{\langle m \rangle}$	Cluster CPU, memory resource utilization	$J, \mathbb{J}$	A job, a set of jobs

ing policies are jointly adopted in E_{of}, thus enabling faster offloading time.

1) Constraints

The collective resource demand or request of multi-job set $\mathbb{J}$ at any given time t cannot exceed the collective resource capacity or available in the selected EC deployment:

$$\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c, m \rangle '} \leqslant C_{Edge_\star}^{\langle c, m \rangle}, \quad \forall_{c, m}, \tag{14}$$

and the unused or inactive nodes $I_i \in Edge_\star$ would be shut down. All the nodes are in active or inactive states. An active node is a node that is running and currently considered for allocation or has at least a job being started, executing or completing. An inactive node is a node that is not running and is not currently considered for allocation or has no job. These two states can be expressed as follows:

$$\forall c, m \; \beta(I_i) = \begin{cases} 1, & \text{Active if } J_i \in [E_{\text{st}}, E_{\text{cp}}, E_{\text{ex}}], \\ 0, & \text{Inactive if } J_i \notin [E_{\text{st}}, E_{\text{cp}}, E_{\text{ex}}], \end{cases} \tag{15}$$

where indicator $\beta(I_i) = 1$ indicates that node I_i is ready to accept new jobs, and at least job J_i is being started, executing or completing, i.e., $J_i \in [E_{\text{st}}, E_{\text{cp}}, E_{\text{ex}}]$, on I_i; otherwise $\beta(I_i) = 0$.

2) Optimization formulation

Hence, maximizing utilization of the selected edge deployment or cluster depends on application orchestration:

$$\text{Maximize} \quad \tilde{\mathcal{U}}_{Edge_i}^{\langle c, m \rangle} = \frac{\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c, m \rangle '}}{C_{Edge_i}^{\langle c, m \rangle}}, \tag{16}$$

$$\text{subject to} \quad \mathbb{J} \Rightarrow Edge_\star, \quad \exists, \tag{17}$$

$$\beta(I_i) \in \{0, 1\}, \quad \exists, \tag{18}$$

$$\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c, m \rangle '} \leqslant C_{Edge_\star}^{\langle c, m \rangle}, \quad \forall_{c, m}. \tag{19}$$

The constraints in Eqs. (17) to (19) indicate the dispatching of multi-job set $\mathbb{J}$ to the closest edge having sufficient resource capability or availability. More specifically, Eq. (17) is the constraint for $\mathbb{J}$ offloading, guaranteeing that $\mathbb{J}$ is dispatched to a cluster such that dependent tasks within each $J \in \mathbb{J}$ can communicate and execute faster. Condition (18) guarantees that active nodes $(\beta(I_i) = 1)$ are used for execution and that inactive nodes $(\beta(I_i) = 0)$ are shut down. The constraint in Eq. (19) guarantees that $d_{\tilde{T}}^{\langle c, m \rangle '}$ of $\mathbb{J}$ does not exceed $C_{Edge_i}^{\langle c, m \rangle}$ any selected cluster. The details of our multi-job dispatching principle will be discussed in Section 4.1 and Algorithm 1. We aim to minimize the number of active nodes used for execution by co-locating jobs tightly on each node to maximize resource

utilization. The details of our co-location strategy will be discussed in Section 4.2 and Algorithm 2.

On the other hand, the overall scheduling time and execution time can be minimized depending on orchestration:

$$\text{Minimize} \quad \sum_{J \in \mathbb{J}} \sum_{i=1}^{k} \frac{E_{\text{sh}_i}}{k} = E_{\text{sh}}', \tag{20}$$

$$\text{subject to} \quad \mathbb{J} \Rightarrow Edge_\star, \quad \forall_{c, m}. \tag{21}$$

$$\text{Minimize} \quad \sum_{J \in \mathbb{J}} \sum_{i=1}^{k} \frac{E_{\text{ex}_i}}{k} = E_{\text{ex}}', \tag{22}$$

$$\text{subject to} \quad \mathbb{J} \Rightarrow Edge_\star, \quad \forall_{c, m}. \tag{23}$$

The constraints in Eqs. (21) and (23) guarantee that $\mathbb{J}$ is dispatched to the same cluster such that dependent tasks within each $J \in \mathbb{J}$ can communicate and execute faster. The details of our multi-job dispatching principle are given in Section 4.1 and Algorithm 1.

4 Edge-IoT Algorithm Framework

The proposed Edge-IoT solution in this paper is focused on offloading and scheduling. The offloading strategy is based on the orchestration of ready multi-job tasks to the closest edge deployment with sufficient available resources to accommodate the tasks, as expressed in Eq. (17), while the scheduling strategy involves packing or co-location of these tasks tightly on container instances to fully utilize the available resources. These components aim at providing optimal performance for vehicular multi-task execution in EC systems such that the optimizations in Eqs. (16), (20) and (22) are achieved.

4.1 Offloading Policy

When sets of vehicular multi-job tasks $\mathbb{J} = J_1, \cdots, J_N$ are ready to be offloaded, our policy is to offload them to the closest edge $Edge_\star$ with sufficient resource capacity or availability, i.e., $\mathbb{J} \Rightarrow Edge_\star$, while $\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c, m \rangle '} \leqslant C_{Edge_\star}^{\langle c, m \rangle}$. For the rationale of this strategy, consider the Ericsson Connected Vehicle Platform (CVP), which serves about 5.5 million active vehicles across more than 150 countries. Assuming that there are 0.1% of these vehicles at a location $\mathcal{L}$ and at time t deciding to offload their multiple tasks i.e., $\vartheta[\mathcal{V} \in \mathbb{V}] = 1$, we would see a total load of 4 000 requests. Executing these loads would require an edge deployment with 40 nodes or container instances if we assume that a container instance can co-locate 100 containerized tasks. To serve these vehicles efficiently, it is better to dispatch these tasks as units to a closest edge deployment, i.e., $\mathbb{J} \Rightarrow Edge_\star$, having sufficient resource capacity or availability. The closest heuristic given in Eq. (17) is to minimize the offloading time E_{of} and to further minimize the

overall response time E_{rsp}. Algorithm 1 describes the offloading procedure.

Algorithm 1. Edge-IoT: multi-job offloading

Input: $\mathbb{J}$ arrived at time t; $Edge_i \in \mathbb{E}$; $\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c, m \rangle'}$

Output: Offload $\mathbb{J}$ to $Edge_\star$ with matching $C_{Edge_\star}^{\langle c, m \rangle}$ such that
$$\mathbb{J} \Rightarrow Edge_\star$$
1: **for** $Edge_i \in \mathbb{E}$ **do**
2:     **if** $\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c, m \rangle'} \leqslant C_{Edge_i}^{\langle c, m \rangle}$ **then**
3:         $\mathbb{J} \Rightarrow Edge_i = Edge_\star$
4:     **else**
5:         Offload $\mathbb{J}$ to next $Edge_\star$
6:     **end if**
7: **end for**
8: **if** $\mathbb{J}$ cannot be offloaded as a whole **then**
9:     **for** $Edge_i \in \mathbb{E}$ **do**
10:       **for** $J \in \mathbb{J}$ **do**
11:         **if** $\sum_{J \in \mathbb{J}} d_{\tilde{T}}^{\langle c, m \rangle'} \leqslant C_{Edge_i}^{\langle c, m \rangle}$ **then**
12:           $J \Rightarrow Edge_i = Edge_\star$
13:         **else**
14:           Dispatch J to next $Edge_\star$
15:         **end if**
16:       **end for**
17:     **end for**
18: **end if**

4.2 Scheduling Policy

Once $\mathbb{J}$ is offloaded to $Edge_\star$, our scheduling algorithm uses the resource availability $I_i^{\langle c, m \rangle}$ of each container-instance in $Edge_\star$, and the resource demand $d_T^{\langle c, m \rangle'}$ of each $J \in \mathbb{J}$ to provide efficient co-location such that fewer container-instances are used for execution in $Edge_\star$. Specifically, the gang scheduling approach is adopted alongside our bin-packing optimization to co-schedule and co-locate all $J \in \mathbb{J}$ at a time. Bin-packing is one of the most popular packing problems. The goal is to minimize the number of nodes used as given in optimization in Eq. (31). Unlike other approaches, such as the first fit bin packing problem (FFBPP)[22], it requires the next J_i to be placed on the active node; otherwise, it is placed on a new node. Our scheduling strategy co-locates multi-dependent tasks firmly on nodes (Algorithm 2) such that for any given job, resource wastage is avoided and fewer nodes are used for execution. It takes the resource demand of multi-job tasks and resource availability of nodes as input, then scans all $J \in \mathbb{J}$ and maps them to active nodes in full utilization. Our approach scans all $J \in \mathbb{J}$ and maps J_i to active nodes in full utilization (Line 2 in Algorithm 2). All $J \in \mathbb{J}$ are co-located firmly on active nodes, so that resource wastage is avoided and fewer nodes are used to execute all jobs concurrently (Lines 4–9 in Algorithm 2).

Algorithm 2. Edge-IoT: multi-job co-location

Input: $\mathbb{J}$ offloaded to $Edge_\star$, resource demand of each $J \in \mathbb{J}$: $d_{\tilde{T}}^{\langle c, m \rangle'}$, resource availability of each node $I_i \in Edge_\star$: $I_i^{\langle c, m \rangle}$

Output: $\mathbb{J}$ is co-located, such that
$$\text{Minimize} \sum_{I_i \in Edge_\star} I_i \equiv \text{Minimize } RU_{Edge_\star}^{\langle c, m \rangle}$$
1: **for** $I_i \in Edge_\star$ **do**
2:     **if** $\beta(I_i) = 1$ **then**
3:       $I_i^{\langle c, m \rangle} = \langle c, m \rangle$, i.e., initial resource available
4:       **for** $J \in \mathbb{J}$ **do**
5:         **if** $\Gamma[J, I_i] = 0$ and $d_{\tilde{T}}^{\langle c, m \rangle'} \leqslant I_i^{\langle c, m \rangle}$ **then**
6:           $J \Rightarrow I_i$
7:           $\Gamma[J, I_i] = 1$
8:           $I_i^{\langle c, m \rangle} = I_i^{\langle c, m \rangle} - d_{\tilde{T}}^{\langle c, m \rangle'}$
9:         **end if**
10:         **if** $I_i^{\langle c, m \rangle}$ close to zero **then**
11:           **break**
12:         **end if**
13:       **end for**
14:     **end if**
15: **end for**

Hence, for every $\mathbb{J}$ offloaded to $Edge_\star$, our co-location strategy is to find the solution to the problem:

$$\text{Minimize} \sum_{I_i \in Edge_\star} I_i \equiv \text{Minimize } RU_{Edge_\star}^{\langle c, m \rangle} = \frac{U_{Edge_\star}^{\langle c, m \rangle}}{C_{Edge_\star}^{\langle c, m \rangle}}, \quad (24)$$

$$\text{subject to } \mathbb{J} \Rightarrow Edge_\star, | \exists , \quad (25)$$

$$\sum_{J \in \mathbb{J}} \Gamma[J, I_i] \cdot d_{\tilde{T}}^{\langle c, m \rangle'} \leqslant I_i^{\langle c, m \rangle}, \quad \forall c, m , \quad (26)$$

where

$$\Gamma[J, I_i] = \begin{cases} 1, & \text{if } J \Rightarrow I_i, \\ 0, & \text{otherwise.} \end{cases} \quad (27)$$

We aim to minimize the number of nodes used for executing $\mathbb{J}$, which is equivalent to minimizing the actual resource usage in $Edge_\star$, given as $RU_{Edge_\star}^{\langle c, m \rangle}$, which is the ratio of the resources used for execution $U_{Edge_\star}^{\langle c, m \rangle}$ over the edge's resource capacity $C_{Edge_i}^{\langle c, m \rangle}$. The metric $RU_{Edge_\star}^{\langle c, m \rangle}$ includes the actual CPU resource usage $RU_{Edge_\star}^{\langle c \rangle}$ and the actual memory resource usage $RU_{Edge_\star}^{\langle m \rangle}$, which are defined respectively as

$$RU_{Edge_\star}^{\langle c \rangle} = \frac{U_{Edge_\star}^{\langle c \rangle}}{C_{Edge_\star}^{\langle c \rangle}} , \quad (28)$$

$$RU_{Edge_\star}^{\langle m \rangle} = \frac{U_{Edge_\star}^{\langle m \rangle}}{C_{Edge_\star}^{\langle m \rangle}} , \quad (29)$$

where $U_{Edge_\star}^{\langle c \rangle}$ and $U_{Edge_\star}^{\langle m \rangle}$ are the used CPU and memory re-

sources, respectively, while $C_{Edge_\star}^{\langle c\rangle}$ and $C_{Edge_\star}^{\langle m\rangle}$ are the edge's CPU and memory resource capacity, respectively. Then the actual CPU utilization $\rho_{\mathcal{DR}_i}^{\langle c\rangle}$ and the actual memory utilization $\rho_{\mathcal{DR}_i}^{\langle m\rangle}$ are defined respectively by

$$\mathcal{U}_{Edge_i}^{\langle c\rangle} = \frac{\sum_{J\in\mathbb{J}} d_T^{\langle c, m\rangle'}}{U_{Edge_\star}^{\langle c\rangle}}, \tag{30}$$

$$\mathcal{U}_{Edge_i}^{\langle m\rangle} = \frac{\sum_{J\in\mathbb{J}} d_T^{\langle c, m\rangle'}}{U_{Edge_\star}^{\langle c\rangle}}. \tag{31}$$

Algorithms 1 and 2 are directly connected with minimizing E_{sh}', minimizing E_{ex}' as well as maximizing $\tilde{\mathcal{U}}_{Edge_i}^{\langle c, m\rangle}$. Therefore, Eq. (25) is the constraint for multi-job set $\mathbb{J}$ deployment, guaranteeing that $\mathbb{J}$ is offloaded to the closest cluster such that dependent tasks within each $J\in\mathbb{J}$ can communicate and execute faster. As we have stated previously that if $\mathbb{J}$ cannot be dispatched as a whole to a cluster, the dispatcher will allow fractional dispatching of each $J\in\mathbb{J}$ to the closest member edge. The constraint in Eq. (26) indicates that the total estimated resource requirements of co-located jobs $d_T^{\langle c, m\rangle'}$ cannot exceed $I_i^{\langle c, m\rangle}$, the node resource availability. The condition in Eq. (27) means that $\Gamma[J_i, I_i] = 1$ if job J_i is placed on the node I_i; otherwise, $\Gamma[J_i, I_i] = 0$. This is to guarantee that each $J\in\mathbb{J}$ is placed in exactly one node. To solve this multi-job packing problem, we have adopted the solving Constraint Integer Programs (SCIP) solver, which is currently one of the fastest mathematical programming (MP) solvers for this problem.

4.3 Connection with Optimization Objectives

Our objectives are to minimize the total response time of multiple IoV applications as stated in Eqs. (20) and (22) and maximize the edge cluster resource utilization in Eq. (26). Algorithms 1 and 2 together achieve these objectives. By offloading multi-job tasks to an edge having sufficient resource availability, Algorithm 1 ensures that any edge deployment selected has sufficient resources $C_{Edge_\star}^{\langle c, m\rangle}$ needed for multi-job execution such that the dependent tasks can be executed faster, ultimately leading to a smaller aggregate scheduling time E_{sh}' and execution time E_{ex}'. By intelligently packing dependent tasks tightly on nodes, Algorithm 2 is capable of fully utilizing available resources at EC clusters, ultimately leading to the resource assigned for the execution of jobs $U_{Edge_\star}^{\langle c, m\rangle}$ to be fewer while guaranteeing it is sufficient for multi-job tasks. More specifically, the resource usage (RU) of the cluster for multi-job tasks is given in Eqs. (28) and (29).

5 Experiment Setup

Our experiment setup consists of six edge deployments distributed across RSUs, BSs and vehicles, as summarized in

Table 2. These platforms consist of large resource capacity EC devices. The input data flow time, final result transmission time, vehicle's speed, and road area were drawn from a uniform distribution range of $(0.2, 0.4]$ s, $(0.4, 4]$ s, $(40, 80]$ km/h and $[2\text{ km} \times 2\text{ km}]$, respectively[23]. Therefore, we conduct extensive experiments with orchestrated sets of multi-dependent tasks with heterogeneous resource requests across the EC resources. For each deployment, we compare the performance of our Edge-IoT with the existing state of the art.

As for applications, the v-2018 version of Alibaba cluster trace is used, which records the activities of about 4 000 machines in a period of eight days. The entire trace contains more than 14 million tasks with more than 12 million dependencies and more than four million jobs, among which we deploy a total of 48 jobs with total of 204 tasks (including dependencies) for our experiments. The task dependency depth among the jobs is in the range of $(1, 17]$. Table 3 lists the details of our multi-job sets.

5.1 Heuristics and Baselines

In our experiments, we assume that all tasks are of high priority. The proposed Edge-IoT utilizes the closest heuristic and adopts the gang-scheduling strategy and a variant bin-packing optimization to efficiently co-schedule and co-locate multi-job tasks in a cluster or edge to minimize the overall response time. We consider Edge-IoT as a full dependency and full packing (FDFP) approach.

We compare the scheduling approach of Edge-IoT with the following three existing schemes, fixing their dispatching policy to that of Edge-IoT, as follows:

1) Full dependency and partial packing (FDPP)[5] is an ap-

▼Table 2. Edge deployments and their resource capacities

Edge Deployment	Edge Device	CPU Capacity	Memory Capacity/GiB
Edge 1	Acer aiSage (x2)	12 Cores	4
Edge 2	AWS Snowcone (x10)	20 Cores	40
Edge 3	Huawei AR502H Series (x6)	24 Cores	12
Edge 4	HIVECELL (x6)	36 Cores	48
Edge 5	NVIDIA Jetson Xavier NX (x3)	36 Cores	24
Edge 6	INTELLIEDGE G700 (x5)	48 Cores	80

▼ Table 3. Multi-job execution, where the actual resources consumed for multi-job execution $d_T^{\langle c, m\rangle}$ are taken from the original Alibaba data and the estimated resource demands $d_{\tilde{T}}^{\langle c, m\rangle'}$ are calculated by linear regression model

Multi-Job $\mathbb{J}$	$\mathbb{C}$	T	$d_{\tilde{T}}^{\langle c, m\rangle'}$	$d_T^{\langle c, m\rangle'}$	NAEE
1	5	22	$\langle 1\,195.24, 4.35\rangle$	$\langle 1\,135, 3.77\rangle$	$\langle 0.1, 0.15\rangle$
2	7	29	$\langle 1\,501.5, 5.81\rangle$	$\langle 1\,325, 4.23\rangle$	$\langle 0.13, 0.37\rangle$
3	9	38	$\langle 2\,011.55, 7.57\rangle$	$\langle 1\,820, 5.76\rangle$	$\langle 0.1, 0.3\rangle$
4	12	52	$\langle 2\,762.25, 10.4\rangle$	$\langle 2\,560, 8.2\rangle$	$\langle 0.1, 0.26\rangle$
5	15	63	$\langle 3\,369.68, 12.58\rangle$	$\langle 3185, 10.17\rangle$	$\langle 0.1, 0.23\rangle$

NAEE: normalized absolute estimate error

proach that executes subtasks of a job locally in the vehicle and offloads subtasks to the cloud server and the remaining tasks to the RSU for execution at the same time.

2) Full dependency and no packing (FDNP)-1[3] is an approach that offloads all tasks of a job to the same EC deployment, but assumes that at any EC deployment, a node can only execute one task at a time, and FDNP-1 schedules one task at a time. Therefore, unscheduled tasks must wait in a queue until resources become available for the next task(s). Such a queue is constructed based on the application priority, where it keeps multiple applications in decreasing order of their priority.

3) FDNP-2[4] is an approach that offloads different subtasks of a job to different EC deployments, where each node at the selected EC deployment can only schedule and execute one task at a time, and the task with the highest priority is first selected for scheduling.

4) No dependency and partial packing (NDPP)[23] is an approach that offloads different multi-job subtasks to available EC deployment, by considering the completion deadline of each task. However, this approach does not respect inter-task dependencies, but co-locates tasks on a node.

5.2 Comparison of Offloading and Execution Results

The investigation focuses on the IoV multi-task response time, which includes the multi-job offloading, resource utilization/usage, scheduling, execution and response time. The multi-job execution information across the edge deployments, obtained according to Alibaba data, are listed in Table 3, where the actual resources consumed for the multi-job execution $d_T^{\langle c, m \rangle \prime}$ are taken from the original data. NAEE defined in Eq. (4) and listed in Table 3 for resource consumed serves as the estimation accuracy measure for the trained linear regression model. The average NAEE across six deployments is 0.12 for CPU and 0.23 for memory. Note that we only focus on the resource demand estimation for multi-job tasks, as the execution time estimation is not required to select suitable on-premise edge deployments given in Table 2. The results obtained by Edge-IoT (FDFP), FDPP, FDNP-1, FDNP-2 and NDPP are compared.

1) Resource usage and resource utilization

Fig. 4 shows the task deployment ratio of Edge-IoT with four baseline schemes. It can be seen that for each multi-job task

offloaded, Edge-IoT is able to deploy its constituent tasks to a single edge. This is because Edge-IoT selects the closest edge with sufficient resource availability to accommodate all the tasks, and co-locates them tightly in each node. Recall that some of the baseline schemes, i.e., FDNP-1 and FDNP-2, do not co-locate tasks on each node, but assume each node can only execute one task at a time. Therefore, FDNP-1 can neither offload all its subtasks nor execute them at a time, given the number of nodes at each edge. For example, Multi-Job 1 that consists of five jobs is deployed and co-located on edge Edge-1 by Edge-IoT, and in turn, allows for faster input data flow transmissions. For the same Multi-Job 1, FDPP, FDNP-2 and NDPP deploy the jobs across two edge deployments. Although FDPP and NDPP can partially co-locate tasks at each of the edges, the three schemes incur additional execution delays due to input data flow transmissions across the two edge
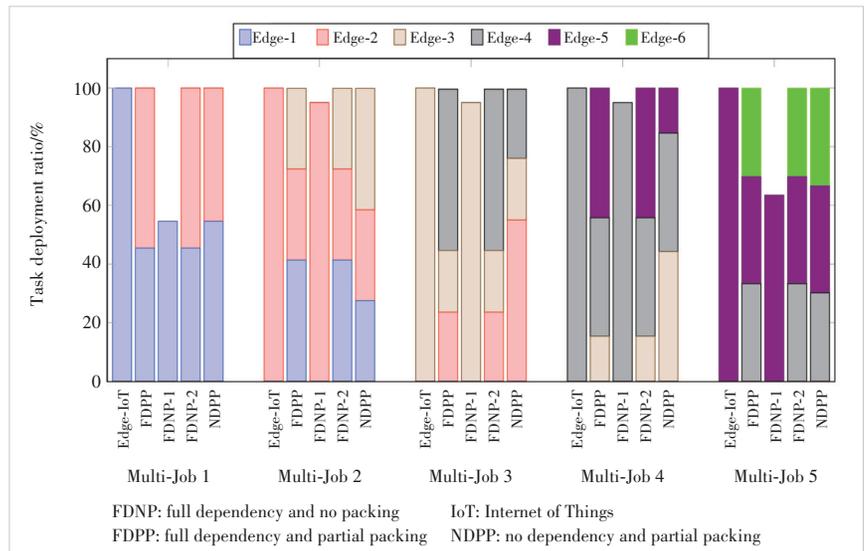
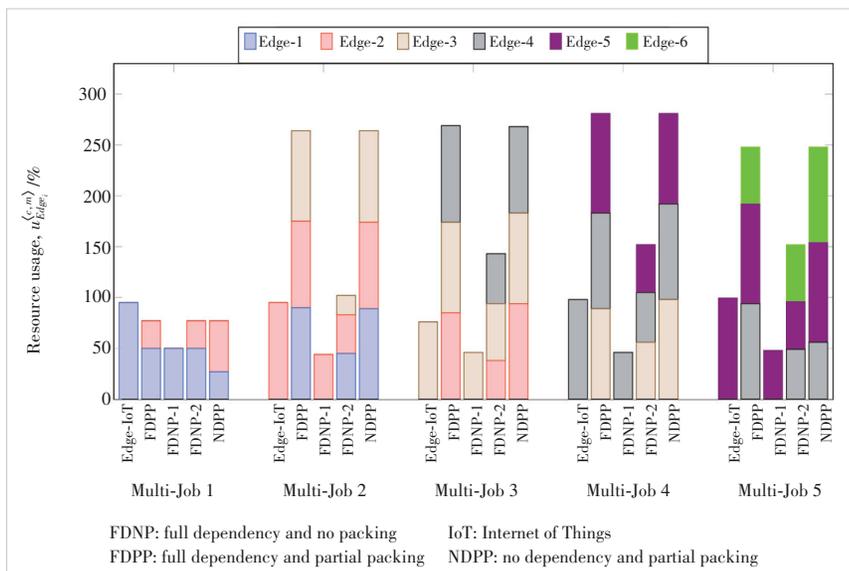▲Figure 4. Tasks deployment ratio across the edge deployments
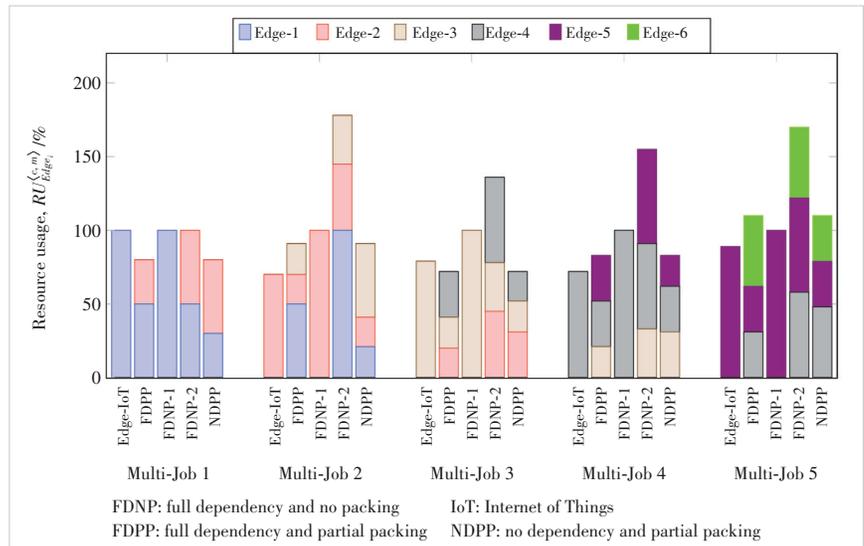
▲Figure 5. Average resource usage across the edge deployments

deployments. On the other hand, FDNP-1 is not able to deploy all the jobs on edge Edge-1, because it executes a task on each node at a time. Hence, it can only execute several tasks at a time, given the number of nodes available in the edge cluster, while the remaining tasks wait in a queue. Fig. 5 shows the average resource usage of the multi-job tasks deployed by Edge-IoT with those of the four baseline schemes across the edge clusters. It can be seen that Edge-IoT consumes the fewest resources by using a single edge for each multi-job task, while FDNP-2 uses the highest resources (up to three edge deployments) for the same multi-job task. The average resource utilization comparison is shown in Fig. 6. Again, Edge-IoT achieves the highest resource utilization compared with the four baseline schemes. We now examine the performance of Edge-IoT compared with the baseline schemes for each multi-job offloaded (as shown in Table 3) in detail.

• Multi-Job 1: Edge-IoT dispatches 100% of the tasks in a single-hop offloading to Edge-1. It first optimizes the deployment by gang-scheduling and co-locating as many tasks in a node as possible to fully utilize the available resources in the node. These tasks are tightly packed on nodes using the packing algorithm, which uses all of Edge-1 resources to execute the tasks, and achieves 95% resource utilization. For the same Multi-Job 1, some of the baseline schemes such as FDPP, FDNP-2 and NDPP offload the tasks across two edge clusters (Edge-1 and Edge-2), using up to two times more resources than Edge-IoT. FDNP-1 schedules one task on a node at a time using a single edge deployment (Edge-1). Thus, it uses all available resources (100%) at the edge deployment and keeps the unscheduled tasks on a task queue until resources become available. Overall, Edge-IoT achieves better resource usage and utilization compared to the four baseline schemes, as shown in Figs. 5 and 6.

• Multi-Job 2: This multi-job task consists of seven jobs

with a total of 29 tasks, where each job has a task dependency in the range of (1, 5]. Edge-IoT optimizes the deployment to ensure that the resources are fully utilized. Containers provide isolation to running applications, making it possible to co-locate multiple applications on the same node without any interference. A single container-optimized node can execute more containerized applications, given that there are sufficient available resources. For scheduling, Edge-IoT deploys all the tasks at a time on edge cluster Edge-2, using 70% of the resources, while with three edge deployments, FDPP, FDNP-2 and NDPP use 50%, 20% and 21% on Edge-1, 100%, 45% and 33% on Edge-2, and 21%, 20% and 50% on Edge-3. Edge-IoT and FDNP-1 utilize 95% and 55% of resources, respectively. Although FDNP-1 uses all available resources in the cluster, it achieves low resource utilization due to its inability to co-locate tasks on nodes, which results in resource under-utilization. Again Edge-IoT outperforms all the four baseline schemes in terms of task deployment ratio, resource usage and utilization.

• Multi-Job 3: Edge-IoT offloads all tasks of Multi-Job 3 to edge Edge-3. This edge deployment is made up of six Huawei AR502H Series edge devices, with CPU and memory capacity of 24 vCPU and 12 GiB, respectively. The multi-job task consists of nine jobs, with a total of 38 tasks, where each job has a task dependency range (1, 8]. Edge-IoT improves resource usage by using a single edge and up to three times fewer resources compared with the four baseline schemes, as can be seen from Fig. 5. It also achieves 76% resource utilization in a single cluster. On the other hand, with three edge deployments, FDPP and NDPP achieve 85% and 89% resource utilization on Edge-2; 94% and 94% on Edge-3; and 89% and 85% on Edge-4). FDNP-1 and FDNP-2 perform worst with the highest resource consumption and the lowest resource utilization.

• Multi-Job 4 and Multi-Job 5: These multi-job tasks are offloaded by Edge-IoT to Edge-4 and Edge-5, respectively. Among all the schemes, Edge-IoT uses the least resources for each multi-job execution across the two edge clusters. Specifically, Edge-IoT consumes 72% and 89% of resources at Edge-4 and Edge-5, respectively. It also achieves the highest resource utilization of 98% and 99% across the two clusters, compared to the four baseline schemes. FDPP consumes 21%, 31% and 31% of resources across Edge-3, Edge-4 and Edge-5, and NDPP consumes 31%, 31% and 21% of resources across Edge-4, Edge-4 and Edge-6. FDNP-1 consumes all available resources at Edge-3 and Edge-4 for Multi-Job4 and Multi-Job5, respectively, while recording the lowest resource utilization at each cluster. FDNP-2 consumes the second highest resources and achieves the

▲Figure 6. Average resource utilization across the edge deployments

FDNP: full dependency and no packing
FDPP: full dependency and partial packing
IoT: Internet of Things
NDPP: no dependency and partial packing

Uchechukwu AWADA, ZHANG Jiankang, CHEN Sheng, LI Shuangzhi, YANG Shouyi

second lowest resource utilization for the same multi-job task execution.

2) Multi-Task Scheduling, Execution and Response Time

The aggregate job scheduling time E_{sh}' defined in Eq. (9), which is the time for placing multi-job tasks on the nodes in a cluster, is an important performance metric to assess the integrated edge clusters. Another important performance metric is the aggregate job execution time E_{ex}' defined in Eq. (10). The response time E_{rsp}' defined in Eq. (8) is even more important. Figs. 7, 8 and 9 compare the scheduling time, execution time and response time, respectively, attained by the five schemes.

It can be seen that the scheduling time is typically very small, and the execution time and response time by contrast are significantly larger. Across the edge clusters, Edge-IoT consistently achieves the fastest scheduling, execution and response, compared to the other four benchmark strategies. Note that we focus on the scheduling time, execution time and result transmission time components of the response time. This is because the offloading time E_{of}' is relatively small due to our offloading policy which ensures that jobs are offloaded to the closest edge cluster and within a single-hop offloading. Specifically, for Multi-Job 1, Edge-IoT achieves a very fast scheduling, which is 11.6 times faster than FDPP and NDPP, and 16 times faster than FDNP-1 and FDNP-2. For Multi-Job 2 scheduling, Edge-IoT achieves significantly shorter scheduling time than the four benchmark strategies, i. e., Edge-IoT is 12 times faster than FDPP and NDPP, and 29 times faster than FDNP-1 and FDNP-2. For Multi-Job 3, FDNP-1 and FDNP-2 attain the lowest scheduling time, while FDPP and NDPP attain the second lowest scheduling time. Edge-IoT achieves the best performance with up to 38 times faster than the other four schemes. For Multi-Job 4 and Multi-Job 5, Edge-IoT again achieves the fastest scheduling, followed by FDPP and NDPP, while FDNP-1 and FDNP-2 have the worst scheduling performance.

In terms of the execution time, it is important to note that the input data flow time also contributes to the total execution time of a job. FDPP, FDNP-2 and NDPP incur additional time due to their approaches of task offloading across multiple clusters, which leads to input data flows (which is in the range of (0.2, 0.4] s) across the clusters. Edge-IoT is 111.4, 22.3, 112
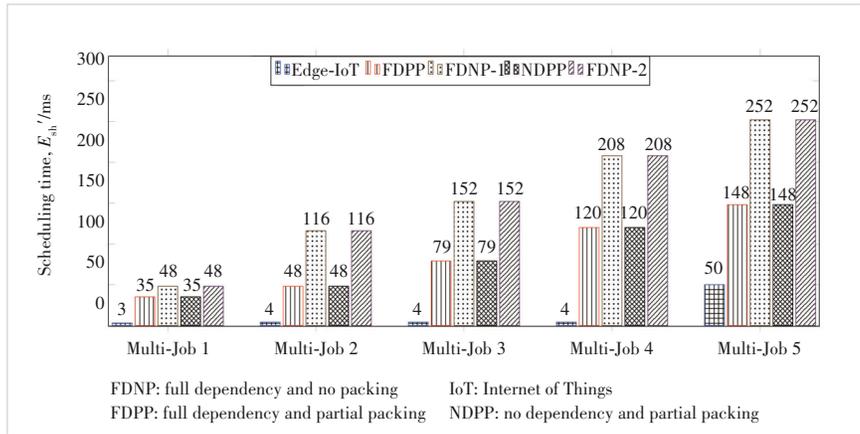
▲Figure 7. Task scheduling time across edge deployments
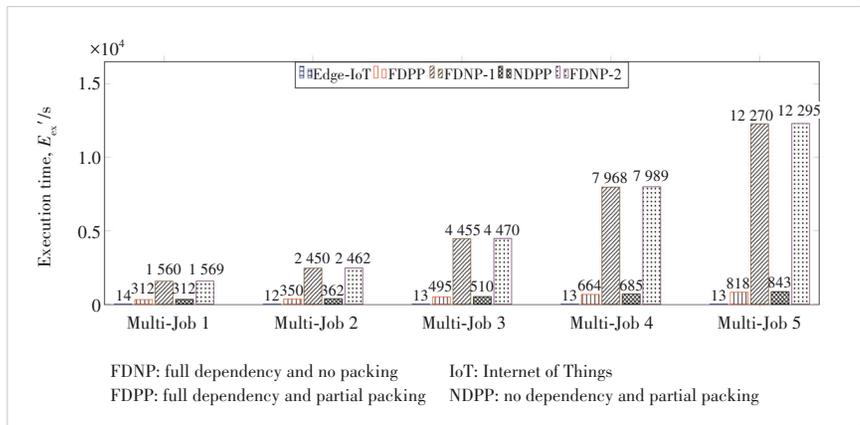
▲Figure 8. Task execution time across edge deployments
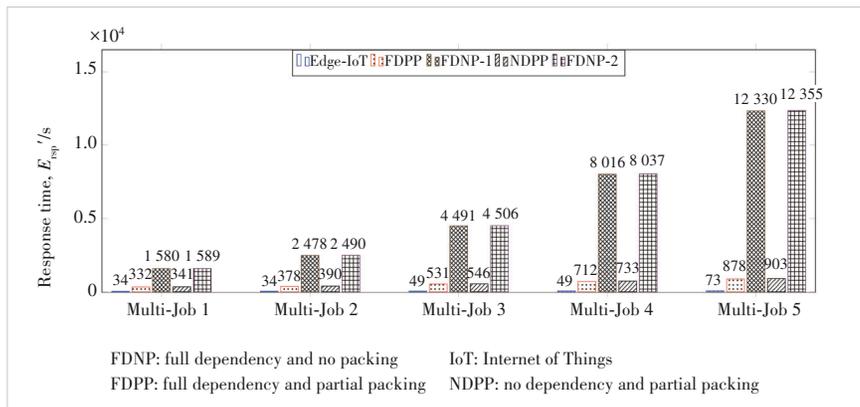
▲Figure 9. Task response time across edge deployments

and 23 times faster than FDNP-1, FDPP, FDNP-2 and NDPP, respectively, for executing Multi-Job 1, while for Multi-Job 2 execution, it is approximately 204, 29, 205 and 30 times faster, respectively. Similarly, for Multi-Job 3, Multi-Job 4 and Multi-Job 5 executions, Edge-IoT achieves approximately up to 943.8, 63, 945.7 and 64.8 times shorter execution time than FDNP-1, FDPP, FDNP-2 and NDPP, respectively. The significant advantage of Edge-IoT in terms of the aggregate job execution time can be explained as follows. It deploys sets of

multi-job tasks as a unit through the gang scheduling strategy in a single-edge deployment. These applications are deployed and executed concurrently. By contrast, the benchmark approaches schedule and execute the given DAGs individually and in parts across multiple edge deployments, resulting in input data flow transmission delays and longer time to execute the overall tasks.

Recall that the response time of a job defined in Eq. (8) is the addition of its offloading time, scheduling time, execution time and final result transmission time. Therefore, the ultimate aim is to minimize the response time of IoV applications offloaded to EC. Fig. 9 compares the response time of Edge-IoT and the four benchmark schemes. Edge-IoT outperforms the four benchmark schemes by achieving shorter response time for all the multi-job tasks, and up to 169, 12, 169.2 and 12.4 times faster than FDNP-1, FDPP, FDNP-2 and NDPP, respectively.

6 Conclusions

Edge-IoT, a machine learning-enabled IoT application orchestration in an EC system proposed in this paper, has demonstrated superior QoS in resource management and IoT multi-task orchestration in edge clusters. Unlike Edge-IoT, the existing methods do not deploy all the ready tasks at a time or in a single edge cluster or do not respect task dependencies, leading to more edge resource usage and cluster under-utilization as well as causing longer task execution time. This paper has presented Edge-IoT to improve edge resource efficiency and performance. We have utilized a resource-aware offloading strategy that selects the closest edge cluster suitable for a given job, and a container-based bin packing optimization strategy that packs or co-locates tasks tightly on nodes to fully utilize available resources. To evaluate our approach, we have illustrated use cases of real-world CPU and memory-intensive tasks from Alibaba cluster trace, which records the activities of both long-running containers (for Alibaba's e-commerce business) and batch jobs across eight days. We have compared our approach with the state-of-the-art dependency-aware IoV task orchestration baseline strategies. Our proposed algorithm achieves both the highest edge cluster resource utilization and the minimum scheduling, execution and response time for IoV multi-job tasks compared to the baseline strategies. The gains achieved by Edge-IoT as observed from our experiments include faster response time of the overall tasks and improved usage of edge resources.

References

[1] KHAN L U, YAQOOB I, TRAN N H, et al. Edge-computing-enabled smart cities: a comprehensive survey [J]. IEEE Internet of Things journal, 2020, 7(10): 10200 – 10232. DOI: 10.1109/JIOT.2020.2987070

[2] AMIN S U, HOSSAIN M S. Edge intelligence and Internet of Things in healthcare: a survey [J]. IEEE access, 2020, 9: 45 – 59. DOI: 10.1109/ACCESS.2020.3045115

[3] LIU Y J, WANG S G, ZHAO Q L, et al. Dependency-aware task scheduling in vehicular edge computing [J]. IEEE Internet of Things journal, 2020, 7(6): 4961 – 4971. DOI: 10.1109/JIOT.2020.2972041

[4] SHEN Q Q, HU B J, XIA E J. Dependency-aware task offloading and service caching in vehicular edge computing [J]. IEEE transactions on vehicular technology, 2022, 71(12): 13182 – 13197. DOI: 10.1109/TVT.2022.3196544

[5] REN H, LIU K, JIN F, et al. Dependency-aware task offloading via end-edge-cloud cooperation in heterogeneous vehicular networks [C]//25th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2022: 1420 – 1426. DOI: 10.1109/ITSC55140.2022.9922334.

[6] LIU S S, LIU L K, TANG J, et al. Edge computing for autonomous driving: opportunities and challenges [J]. Proceedings of the IEEE, 2019, 107(8): 1697 – 1716. DOI: 10.1109/jproc.2019.2915983

[7] MAHMUD R, TOOSI A N, RAMAMOHANARAO K, et al. Context-aware placement of industry 4.0 applications in fog computing environments [J]. IEEE transactions on industrial informatics, 2020, 16(11): 7004 – 7013. DOI: 10.1109/TII.2019.2952412

[8] OTHMAN M M, EL-MOUSA A. Internet of Things & cloud computing Internet of Things as a service approach [C]//11th International Conference on Information and Communication Systems (ICICS). IEEE, 2020: 318 – 323. DOI: 10.1109/ICICS49469.2020.239503

[9] REN J, ZHANG D Y, HE S W, et al. A survey on end-edge-cloud orchestrated network computing paradigms: transparent computing, mobile edge computing, fog computing, and cloudlet [J]. ACM computing surveys, 2020, 52(6): 1 – 36. DOI: 10.1145/3362031

[10] HWANG J, NKENYEREYE L, SUNG N, et al. IoT service slicing and task offloading for edge computing [J]. IEEE Internet of Things journal, 2021, 8 (14): 11526 – 11547. DOI: 10.1109/jiot.2021.3052498

[11] ALMUTAIRI J, ALDOSSARY M. A novel approach for IoT tasks offloading in edge-cloud environments [J]. Journal of cloud computing, 2021, 10(1): 1 – 19. DOI: 10.1186/s13677-021-00243-9

[12] AWADA U, ZHANG J K, CHEN S, et al. Air-to-air collaborative learning: a multi-task orchestration in federated aerial computing [C]//14th International Conference on Cloud Computing (CLOUD). IEEE, 2021: 671 – 680. DOI: 10.1109/CLOUD53861.2021.00086

[13] AWADA U, ZHANG J K, CHEN S, et al. AirEdge: a dependency-aware multi-task orchestration in federated aerial computing [J]. IEEE transactions on vehicular technology, 2022, 71(1): 805 – 819. DOI: 10.1109/TVT.2021.3127011

[14] TU Y F, DONG Z J, YANG H Z. Key Technologies and application of edge computing [J]. ZTE communications, 2017, 15(2): 26-34. DOI: 10.3969/j.issn.1673-5188.2017.02.004

[15] LI X W, ZHAO L, YU K P, et al. A cooperative resource allocation model for IoT applications in mobile edge computing [J]. Computer communications, 2021, 173: 183 – 191. DOI: 10.1016/j.comcom.2021.04.005

[16] LI J, LIANG W F, XU W Z, et al. Maximizing user service satisfaction for delay-sensitive IoT applications in edge computing [J]. IEEE transactions on parallel and distributed systems, 2022, 33(5): 1199 – 1212. DOI: 10.1109/TPDS.2021.3107137

[17] ZHAN C, HU H, LIU Z, et al. Multi-UAV-enabled mobile-edge computing for time-constrained IoT applications [J]. IEEE Internet of Things journal, 2021, 8 (20): 15553 – 15567. DOI: 10.1109/JIOT.2021.3073208

[18] LI J, LIANG W F, XU W Z, et al. Service home identification of multiple-source IoT applications in edge computing [J]. IEEE transactions on services computing, 2023, 16(2): 1417 – 1430. DOI: 10.1109/TSC.2022.3176576

[19] LIU J L, LIU C H, WANG B, et al. Optimized task allocation for IoT application in mobile-edge computing [J]. IEEE Internet of Things journal, 2022, 9 (13): 10370 – 10381. DOI: 10.1109/JIOT.2021.3091599

[20] HAN S N, LI X H, SUN C, et al. RecCac: Recommendation-empowered cooperative edge caching for internet of things [J]. ZTE communications, 2021, 19 (2): 2 – 10. DOI: 10.12142/ZTECOM.202102002

[21] LIU C H, LIU K, GUO S T, et al. Adaptive offloading for time-critical tasks in heterogeneous Internet of vehicles [J]. IEEE Internet of Things journal, 2020, 7 (9): 7999 – 8011. DOI: 10.1109/JIOT.2020.2997720

[22] RAMPERSAUD S, GROSU D. Sharing-aware online virtual machine packing

in heterogeneous resource clouds [J]. IEEE transactions on parallel and distributed systems, 2017, 28(7): 2046 – 2059. DOI: 10.1109/TPDS.2016.2641937

[23] HONG Z C, CHEN W H, HUANG H W, et al. Multi-hop cooperative computation offloading for industrial IoT-edge-cloud computing environments [J]. IEEE transactions on parallel and distributed systems, 2019, 30(12): 2759 – 2774. DOI: 10.1109/TPDS.2019.2926979

Biographies

Uchechukwu AWADA is currently working toward a PhD degree at the School of Information Engineering, Zhengzhou University, China. His current research interests include edge computing, cloud computing, aerial computing, distributed systems, IoT, IoV and wireless communications. He is a student member of the ACM.

ZHANG Jiankang (jzhang3@bournemouth.ac.uk) is a senior lecturer at Bournemouth University, UK. Prior to joining Bournemouth University, he was a senior research fellow at the University of Southampton, UK. Dr. ZHANG was a lecturer from 2012 to 2013 and then an associate professor from 2013 to 2014 at Zhengzhou University, China. His research interests are in the areas of aeronautical communications, aeronautical networks, evolutionary algorithms and edge computing.

CHEN Sheng received his BE degree from the East China Petroleum Institute, China in 1982 and his PhD degree from City, University of London, UK in 1986, both in control engineering. In 2005, he was awarded the higher doctoral degree, Doctor of Sciences (DSc), from the University of Southampton, UK.

From 1986 to 1999, He held research and academic appointments at the Universities of Sheffield, Edinburgh and Portsmouth, all in UK. Since 1999, he has been with the School of Electronics and Computer Science, the University of Southampton, where he holds the post of Professor in Intelligent Systems and Signal Processing. His research interests include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, and evolutionary computation methods and optimization. He has published over 600 research papers. He has 18 500+ Web of Science citations with h-index of 59, and 36 700+ Google Scholar citations with h-index of 81. Dr. CHEN is a Fellow of the United Kingdom Royal Academy of Engineering, a Fellow of the Asia-Pacific Artificial Intelligence Association, and a Fellow of IET. He is one of the original ISI's highly cited researchers in engineering (March 2004). He is named a 2023 Electronics and Electrical Engineering Leader in the UK by Research.com.

LI Shuangzhi received his BS and PhD degrees from the School of Information Engineering, Zhengzhou University, China in 2012 and 2018, respectively. From 2015 to 2017, he was a visiting student with the Department of Electrical and Computer Engineering, McMaster University, Canada. He is currently a lecturer with the School of Information Engineering, Zhengzhou University, China. His research interests include noncoherent space-time coding and ultra-reliable low-latency communications.

YANG Shouyi received his PhD degree from the Beijing Institute of Technology, China in 2002. He is currently a full professor with the School of Information Engineering, Zhengzhou University, China. He has authored or co-authored various articles in the field of signal processing and wireless communications. His current research interests include signal processing in communications systems, wireless communications, and cognitive radio.

Multi-User MmWave Beam Tracking via Multi-Agent Deep Q-Learning

MENG Fan[1], HUANG Yongming[2], LU Zhaohua[3],

XIAO Huahua[3]

(1. Purple Mountain Laboratories, Nanjing 211111, China；
 2. School of Information Science and Engineering, Southeast University, Nanjing 210096, China；
 3. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, ZTE Corporation, Shenzhen 518057, China)

Abstract: Beamforming is significant for millimeter wave multi-user massive multi-input multi-output systems. In the meanwhile, the overhead cost of channel state information and beam training is considerable , especially in dynamic environments. To reduce the overhead cost, we propose a multi-user beam tracking algorithm using a distributed deep Q-learning method. With online learning of users' moving trajectories, the proposed algorithm learns to scan a beam subspace to maximize the average effective sum rate. Considering practical implementation, we model the continuous beam tracking problem as a non-Markov decision process and thus develop a simplified training scheme of deep Q-learning to reduce the training complexity. Furthermore, we propose a scalable state-action-reward design for scenarios with different users and antenna numbers. Simulation results verify the effectiveness of the designed method.

Keywords: multi-agent deep Q-learning; centralized training and distributed execution; mmWave communication; beam tracking; scalability

1 Introduction

Millimeter wave (mmWave) communications have gained extensive attention due to vast bandwidth resources. The beamforming technique with large antenna arrays can improve the mmWave communication network coverage and make up for severe free-space path loss. MmWave signals are highly directional with beamforming, and thus beam tracking is needed to ensure the stability and quality of connected links in mobile scenarios. Currently, mmWave systems typically use hybrid analog-digital architectures to reduce the hardware cost and power consumption.

Traditional beam alignment exhaustively scans the whole beam space, and the introduced high overhead is unacceptable for mobile scenarios. The efficiency of beam training can be improved by the hierarchical searching method with a multi-resolution codebook. Refs. [1 – 3] have reduced the beam training overhead by exploiting prior knowledge of the mmWave channel such as the angle of departure (AoD) or the angle of arrival (AoA), and a low-resolution codebook is further considered in fast-varying scenarios[3]. As a heuristic solution, a deep learning based fast beamforming design method is introduced, without complex operations and iterations in conventional methods[4].

To better utilize implicit prior information embedded in the practical environments, data-driven approaches are feasible[5 – 7]. A fingerprint database is used in Ref. [8] to access historical training records according to the user's location. In Ref. [9], a data-driven data fusion module is developed to combine AoD and time of arrival (ToA) positioning, and positioning-based beam tracking methods are introduced for high-speed railway scenarios. In general, offline learning requires a large number of collected samples in advance, and recollection is needed once the environment changes, leading to difficulties in deployment. Meanwhile, reinforcement learning can realize online learning without offline data, and optimize the policy through interactions with the environment. To reduce beam training overhead, Ref. [10] proposes a multi-armed bandit (MAB) based approach where the training beams are selected by the upper confidence bound strategy. However, the simple MAB model has limited ability to learn from the surroundings, furthermore, a centralized deep Q-learning (DQL) method is proposed in Ref. [11], where the beam training problem is modeled as a Markov decision process (MDP). However, due to its multi-user single-agent model, the action space exponentially explodes with the growth of the user number and lacks scalability to different

user and antenna numbers.

In this paper, under the centralized training and distributed execution (CTDE) framework, we propose a beam tracking method with distributed DQL for the beam tracking problem in dynamical mmWave scenes. Specifically, a distributed beam tracking algorithm is designed to adapt to the changing environments, where each user is regarded as an agent. We also propose several enhancements on the vanilla DQL, including simplified deep Q-network (DQN) training and scalable state-action-reward designs. The main contributions are summarized as follows:

• We develop multi-agent DQL for simultaneous multi-user beam tracking, and the DQL method follows the CTDE framework, where all users share the same policy learned with collected data from all the users.

• We prove that the beam tracking problem is a quasi-static optimization problem instead of an MDP, and a simplified DQL training scheme is proposed to reduce the complexity.

• We propose scalable state-action-reward designs for the DQL which can work in scenarios with different BS antenna and user numbers. In comparison, the existing centralized DQL methods cannot be transferred to a different scenario due to a mismatch of input and output.

The rest of this paper is organized as follows. Section 2 presents the system model. Section 3 describes the beam tracking design with a distributed DQL method. Section 4 gives the simulation results. Section 5 draws the conclusions.

2 System Model and Problem Formulation

2.1 System Model

We consider the downlink transmission in a link-level mmWave communication system composed of one base station (BS) and U mobile users (MU). The BS is equipped with M transmit antennas and N_{rf} radio frequency (RF) chains which are fully-connected, and each MU has a single receiving antenna. One data stream is simultaneously allocated to each user, and thus $U = N_{rf}$. On the BS side, the hybrid analog-digital precoding is considered. The analog precoding matrix is denoted by $A \in \mathbb{C}^{M \times N_{rf}}$, where the u-th column, i.e., $A[:,u]$, is the analog precoding vector of user u, and it is selected from the discrete Fourier transformation (DFT) codebook $F \in \mathbb{C}^{M \times M}$. Similarly, the digital precoder is $V = \left[v_1, \cdots, v_{N_{rf}} \right]$, where the u-th column $H = \left[h_1, \cdots, h_{N_{rf}} \right]^H$ denotes the digital precoding vector, and s is the independent and identical distributed (i. i. d.) data stream. The received signal can be written as:

$$y = HAVs + w, \tag{1}$$

where $w \sim \mathcal{CN}\left(0, \sigma_n^2 I_{N_{rf}}\right)$ denotes zero mean additive white Gaussian noise (AWGN) with variance σ_n^2, and the channel matrix is denoted by $v_u \in \mathbb{C}^{N_{rf} \times 1}$, where h_u is the downlink

channel vector from the BS to MU u.

Without sacrificing generality, the DFT codebook F is constructed by evenly sampling the beam space, and thus the i-th column is:

$$F_i = a(\phi_i)|_{\phi_i = \frac{\pi}{2}\left(\frac{2i}{M} - 1\right)}, \tag{2}$$

where the array response with azimuth being ϕ is:

$$a(\phi) = \frac{1}{\sqrt{M}}\left[1, \cdots, e^{jkdm\sin(\phi)}, \cdots, e^{jkd(M-1)\sin(\phi)}\right], \tag{3}$$

where $k = \frac{2\pi}{\lambda}$, and λ is the wavelength.

Instead of directly estimating high-dimensional channel state information (CSI) $\{h_u\}$, we use low-dimensional equivalent CSI $\{\bar{h}_u\}$ obtained by beam scanning. Specifically, the equivalent channel is a multiplication of the channel matrix H and the analog precoding matrix A, and then the BS can be considered as a transmitter with N_{rf} ports. The equivalent channel vector between the BS and MU u is $\bar{h}_u = A^H h_u$, which will be used for digital precoding.

2.2 Problem Formulation

As illustrated in Fig. 1, the precoded signal is transmitted within a correlation block which is divided into three phases: beam scanning, hybrid precoding, and data transmission. After the beam scanning in time slot t, we can obtain the equivalent channel vectors $\{\bar{h}_u\}$. Then, the digital precoding problem is modeled as:

$$\max_{\{v_u\}} \quad \sum_{u \in \mathcal{U}} \log\left(1 + \frac{\left|\bar{h}_u^H v_u\right|^2}{\sum_{v \neq u}\left|\bar{h}_u^H v_v\right|^2 + \sigma_n^2}\right),$$

$$\text{s.t.} \qquad \sum_{u \in \mathcal{U}}(v_u)_F^2 \leq P_m, \tag{4}$$

where P_m denotes the maximal transmit power of the BS. Eq. (4) can be solved by minimum mean square error
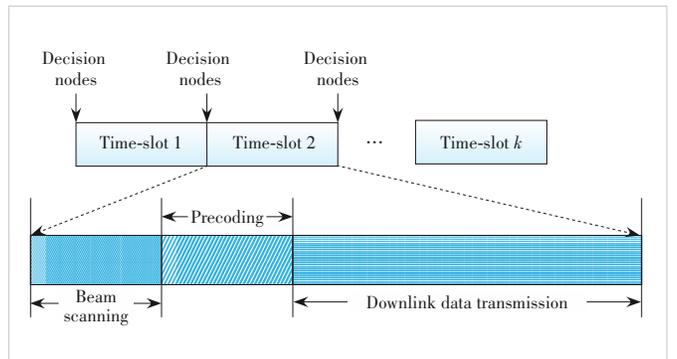
▲Figure 1. Three phases of a time slot

(MMSE) precoding. We adopt a classical linear MMSE to derive the transmitter digital precoder as follows:

$$D = \xi \bar{H} \left(\bar{H}^H \bar{H} + \sigma_n^2 I_{N_{rf}} \right)^{-1}, \tag{5}$$

where $\bar{H} = \left[\bar{h}_1, \cdots, \bar{h}_{rf} \right]^T$, and ξ is a factor to control the BS maximum transmit power.

We evaluate the system performance by an effective sum-rate. Let f_t be the optimal value of an objective function in Eq. (4). Considering the beam training overhead, the effective achievable sum rate during time slot t is defined as

$$R_t = \left(1 - \frac{|\mathcal{F}_t| t_S + t_P}{t_C} \right) f_t, \tag{6}$$

where $\mathcal{F}_t$ is a subset of the codebook $\mathcal{F}$ and its elements are the training beams to be scanned, $|\mathcal{F}_t|$ denotes the corresponding cardinal number, t_S is the duration of one training beam, t_P denotes the duration of precoding and online learning, and t_C denotes the duration of one time slot. With previous known experience, the investigated problem is to design a beam tracking algorithm to maximize time average of Eq. (6), where the digital precoding vectors $\{ v_u \}$ in Eq. (4) are derived from the beam scanning results.

3 Beam Tracking with Deep Q-Learning

3.1 Preliminary of Deep Q-Learning

Without loss of generality, single-agent DQL is developed for a problem modeled as a process of continuous interactions between an intelligent agent and the environment, i.e., MDP. In each interaction, the agent conducts an action a by a policy π with an observed state s, then receives a feedback reward r from the environment, and enters a new state s'. The goal is to learn a strategy for cumulative reward maximization. In a value-based algorithm, the action is selected by the values of state-action pairs, i.e., Q-values. The Q-value is defined as follows:

$$Q^\pi(s,a) = \mathbb{E}_\pi \left[\sum_{k=0}^\infty \gamma^k r_{k+1} |_{s,a} \right], \tag{7}$$

where $\gamma \in [0,1)$ is the discount factor. The mapping from the state to the action values is realized by a learnable network DQN.

3.2 Centralized Deep Q-Learning

Intuitively, the investigated multi-user beam tracking problem can be modeled as an MDP, and the centralized DQL method is considered in Ref. [11]. Specifically, during time-slot t, the modulus of the channel vector of MU u in beam space is given as：

$$I_u^t = \text{abs} \left(F^H h_u^t \right). \tag{8}$$

Stacking $\{ I_u^t \}$ into a matrix I^t, we have

$$I^t = \left[I_1^t, \cdots, I_U^t \right] \in \mathbb{R}^{M \times U}, \tag{9}$$

where we can obtain an "image" I^t as the state s^t, which describes the distribution of effective paths or beam directions. Since mmWave channels are sparse in the beam domain and the training beam set is a subset of the DFT codebook, I^t is a sparse image and most elements of I^t are near zero.

To achieve the goal of sensing the environment, an action is defined based on the difference in the indices between two adjacent beams. An action for a single MU is defined by a pair of integers (μ, σ), where μ denotes the difference of the indices of the optimal beams in two adjacent time-slots, i.e.,

$$\mu^t = \left(b^t - b^{t-1} \right) \text{mod} M, \tag{10}$$

where b denotes the beam index, "mod" denotes the modular arithmetic, and σ denotes the number of beams used to sweep the beam space, respectively. The action space corresponding to MU u is denoted by

$$\mathcal{A}_u = \left\{ \begin{array}{l} \left(\mu_1 - \left\lfloor \frac{\sigma_1}{2} \right\rfloor, \mu_1 + \left\lfloor \frac{\sigma_1}{2} \right\rfloor \right), \cdots, \\ \left(\mu_L - \left\lfloor \frac{\sigma_L}{2} \right\rfloor, \mu_L + \left\lfloor \frac{\sigma_L}{2} \right\rfloor \right) \end{array} \right\} \text{mod} M, \tag{11}$$

where L is the size of the action space. The action space for all MUs is a product of $\{ \mathcal{A}_u \}$, i.e., $\mathcal{A} = \prod_u \mathcal{A}_u$. Finally, the immediate reward in time-slot t is given in Eq. (6), i.e., $r^t = R_t$. The scanned beams are $\mathcal{F}_t = a^t \in \mathcal{A}$.

In DQL[12], a separate target network is introduced to stabilize DQN training, the weights of which change slowly compared with the primary network.

However, several shortcomings of the centralized framework must be observed. Firstly, as the user number U increases, the cardinality of DQN input space $|\mathcal{S}| = M \times U$ grows linearly, and cardinality of output space $|\mathcal{A}| = L^U$ grows exponentially. The training is difficult for such a DQN since the state-action space increases exponentially with user numbers. Additionally, exploration in high-dimensional space is inefficient, and thus the learning can be impractical. Secondly, the DQL lacks scalability in changing user number U and the BS antenna number M.

3.3 Simplicated DQN Training

3.3.1 Centralized Training and Distributed Execution Framework

Single-agent DQL for multi-user beam tracking can lead to action space explosion[13]. To address this issue, we propose

the multi-agent DQL with CTDE. Specifically, each MU is regarded as one agent. All agents are synchronized and distributed, and they share the same policy for online training and inference. The collected data from all agents are aggregated to form a centralized training set, and the shared policy is trained with the centralized training set. The shared policy is then executed by all agents. Thus, the training is centralized and the execution is distributed. The CTDE framework can solve the space explosion problem, and also improve network scalability and reduce training difficulty.

3.3.2 Non-MDP Problem

To adapt to the dynamic environment, low computational complexity is significant for online training, therefore we propose to simplify the vanilla DQN method, i. e., reducing the beam tracking problem as a static optimization problem and solving it in a greedy manner[13]. From the perspective of MDP, the following conclusion can be drawn.

Theorem 1. When the state transition function is independent of the current action and the reward is independent of the state to be transferred to, the maximized cumulative reward under the optimal policy is equivalent to the combination of single-step rewards.

The description of the assumed conditions can be mathematically formulated as:

$$P^a_{s \to s'} = P_{s \to s'}, \tag{12}$$

$$r^a_{s \to s'} = r^a_s, \tag{13}$$

where P denotes state transition probability. The proof of Theorem 1 is given in the Appendix.

In practice, the beam alignment success rate reaches a certain extent p_{thr} close to 100%. Once the misalignment occurs, the BS instantly realigns and a partial observation is obtained. This observation is very similar to the one observed when the beam is successfully tracked. Thus, the new state observed from the environment is mainly determined by the moving users and the fading channels, and is weakly related to the taken action. Additionally, when the reward is sum-rate R_t in Eq. (6), the current reward is irrelevant to the new state. Therefore, we can regard that the system satisfies Eqs. (12) and (13), and we set the discount factor γ as 0. Formally, the Q-value function in Eq. (7) can be simplified as follows:

$$Q(s_t, a_t) = r_t. \tag{14}$$

In summary, when Theorem 1 holds, we can replace the above Eq. (7) with Eq. (14) for DQN training, which has the following benefits:

1) With no need for target networks, the training complexity is reduced;

2) The variance of Q-value estimation is reduced, and thus the training is more efficient.

3.4 State, Action and Reward Design

To make the choice of action in each state logical, the design of the state must reflect the state of the user's interaction with the environment. Since the irregular movements of the user are the main cause of the dynamic changes in the environment, the state can be defined according to the movement of the user. We propose to use the index difference of optimal beams measured in successive time slots as the state. This state design reflects changes in the direction and rate of the user's motion over a period of time.

In the state design of centralized DQL, the state space grows linearly with the user number U. To achieve scalability against U, firstly we propose to decouple the centralized state I^t as a bunch of distributed states $\{I^t_u\}$. Thus, once training is finished, the distributed state can be extended to the scenarios with any user number.

The mmWave channel is sparse in the beam domain, thus most of the element values in I^t_u are equal or close to zero. Besides, in the beam tracking period t for user u, only a small subset of training beams is scanned. Therefore, except for σ^t_u scanned beams, $(M - \sigma^t_u)$ elements in I^t_u are zero, indicating that we can retain the scanned beams as the distributed state and leave out the others. Secondly, as shown in Fig. 2, we propose to cap I^t_u with a mask $m^t_u \in \mathbb{R}^{M \times 1}$, the center of which is b^t_u and the half width is $B = \max^L_{l=1} \sigma_l$, to achieve scalability against the BS antenna number M. Formally, the masked distributed state is:

$$s^t_u = I^t_{u,i} \Big|^{(b^t_u + B)\bmod M}_{i = (b^t_u - B)\bmod M}. \tag{15}$$

The other elements in I^t_u are left out. Similarly, we decouple the centralized action $a \in \mathcal{A}$ as $\{a_u \in \mathcal{A}_u\}$. Thus, the state space is $2B$ and the action space is L which are fixed and irrelevant to the user number U. This indicates the proposed distributed design is scalable to changing user numbers and BS antenna numbers.

Reward acquisition requires completion of analog and digital precoding. At time-slot t, the effective achievable rate of user u in Eq. (6) is defined as the reward

Note: The modular arithmetic is neglected.
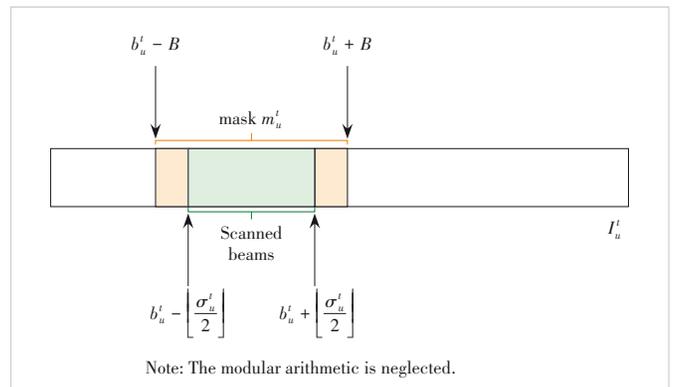
▲Figure 2. State of user *u* after masking

$$r_u^t = \left(1 - \frac{|\mathcal{F}_t| t_S + t_P}{t_C}\right) \log\left(1 + \frac{\left|\overline{\boldsymbol{h}}_u^H \boldsymbol{v}_u\right|^2}{\sum_{v \neq u} \left|\overline{\boldsymbol{h}}_u^H \boldsymbol{v}_v\right|^2 + \sigma_n^2}\right). \quad (16)$$

The scanned beams is $\mathcal{F}_t = \bigcup_{u=1}^{U} a_u^t$.

In summary, compared with the centralized DQL introduced in Section 3.2, the proposed distributed DQL has the following benefits:

1) The input and output of the DQN are greatly reduced, and thus the DQN is simplified.

2) The DQN is scalable to the changing user number U and the BS antenna number M.

3) The sample number is U times higher than that of the centralized DQL.

We give two instances of DQNs in Table 1, and they both have a three-layer neural network (NN). The activation function $f(\cdot)$ and the neuron number of each DNN layer are listed on the left and the right sides, respectively. The activation functions are rectified linear unit (ReLU): $f(\boldsymbol{x}) = \max(0, \boldsymbol{x})$, and linear: $f(\boldsymbol{x}) = \boldsymbol{x}$.

3.5 Distributed Beam Tracking Algorithm Procedure

For clarity, the flow of the distributed beam tracking algorithm is summarized in Algorithm 1. At the beginning of each episode, the entire codebook is scanned to obtain the initial state s_u^t for every single user. Then, the agent selects action a^t by the ε-greedy strategy, and ε for the ε-greedy strategy varies as

$$\varepsilon = \frac{n_{\text{to}} - n_{\text{cur}}}{n_{\text{to}}}, \quad (17)$$

where n_{to} is a fixed value. We set another fixed value n_{thr} that is less than n_{to}, and n_{cur} varies as:

$$n_{\text{cur}} = \begin{cases} n, & n < n_{\text{thr}} \\ n_{\text{thr}} & n \geq n_{\text{thr}}. \end{cases} \quad (18)$$

In Algorithm 1, by performing steps (1), r^t and digital precoding vectors are obtained. Then downlink data transmission is executed in step (2). In step (3), the parameters of DQN are updated.

The ε-greedy strategy is used to explore the environment, the existence of which can lead to the failure to find the optimal beam at each time-slot, i.e., misalignment. Five consecu-

▼Table 1. Deep Q-network (DQN) setting

DQL type	Centralized DQL	Distributed DQL (proposed)
Output layer	linear, L^U	linear, L
Hidden layer	ReLU, 32	ReLU, 32
Input layer	linear, $M \times U$	linear, $2B$

DQL: deep Q-learning   ReLU: rectified linear unit

tive moments of mis-alignment are defined as an incident. Once an incident occurs, the optimal beam initialization process starts immediately from this moment, which is called "calibration" and is achieved via exhaustive search.

Algorithm 1: Distributed beam tracking algorithm

1: **Initialize**: 1) DFT codebook $\boldsymbol{F}$; 2) DQN with random weights θ; 3) replay memory D;
2: for each episode do
3:     **scan** optimal beam in codebook to obtain U initial states
4:     **while** $t \geq k$ and $t \leq$ snapshot **do**
5:         (1) **obtain** analog and digital precoding
6:             a) choose action a^t according to the ε-greedy strategy
7:             b) execute action a^t and observe the next state s^{t+1}
8:             c) compute reward r^t and obtain (s^t, a^t, r^t)
9:             4) obtain precoding vectors
10:         (2) **transmit** data during the remaining of time-slot t
11:         (3) **update** parameters θ of DQN
12:             a) store transition (s^t, a^t, s^{t+1}, r^t) in D
13:             b) sample batch of transitions from D
14:             c) update θ with the gradient descent optimizer
15:         let $t \leftarrow t + 1$
16:     **end while**
17: **end for**

---

4 Simulation Results

In this section, we evaluate the performance of the proposed beam training algorithm via numerical results. The MUs are assumed to move along a circle and the BS is located at the center. To reflect dynamical changes of the distances between the BS and the MUs, the time-varying path-loss of the MUs is incorporated into the mmWave channel model. The movement velocity of the MUs is assumed to be stochastic, and obeys a known probability law. Accordingly, switching to another beam in the next time-slot is also stochastic and obeys some probability law.

For each MU, the probability that the optimal beam of the MU switches to the i-th beam of the next S beams is denoted by $p_{S,i}(i = 0, \cdots, S)$, where $p_{S,0}$ is the probability that the optimal beam of the MU in the next time-slot is still the current beam. For example, two probability distributions are considered, where $p_{S,i}$ is given by:

$$p_{S,i} = e^{-\eta i} \left(\sum_{k=0}^{S} e^{-\eta k}\right)^{-1}. \quad (19)$$

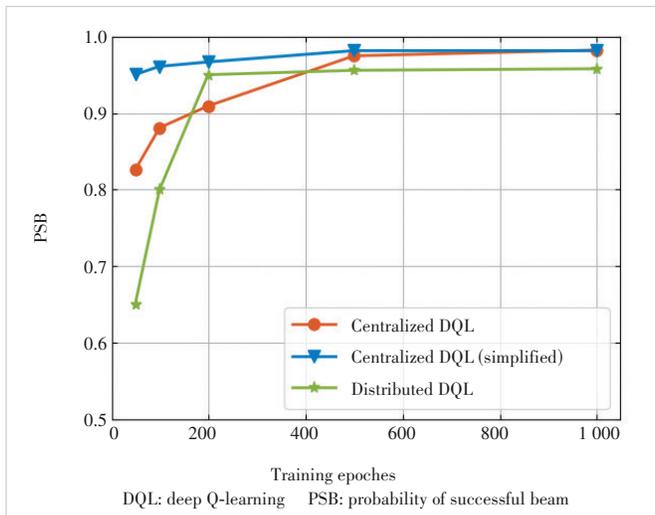The parameter $\eta > 0$ defines the "decay" rate. Specifically,

we consider $S = 4$ and $\eta = 1.0$. For each MU u, the action space $\mathcal{A}_u$ is given by

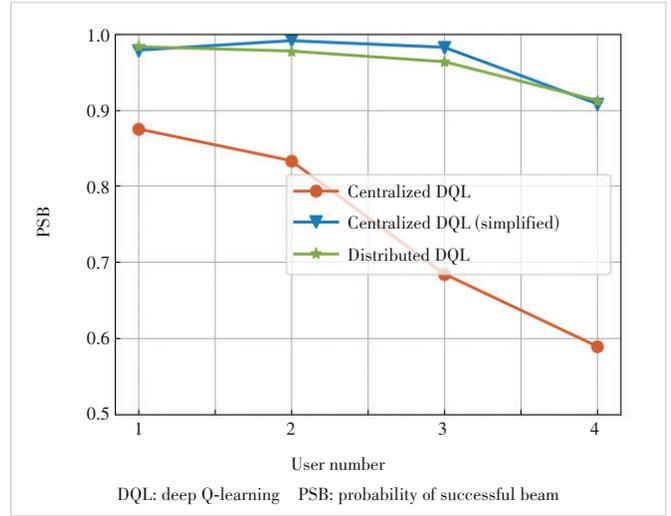$$\mathcal{A}_u = \left\{ (a,b)|a = \{0,1,2,3\}; b = \{1,3,5\} \right\}. \tag{20}$$

Next, we evaluate the performance of the designed DQL algorithm. The simulation results of the centralized DQL in Ref. [11], the proposed simplified DQL in Section 3.3.2 named centralized DQL (simplified) and the proposed distributed DQL are provided for comparison. Besides, the exhaustive search beam tracking, the bandit learning based beam tracking, Q-learning based beam tracking and the centralized DQL algorithm are studied in Ref. [11], and the simulation results show that the centralized DQL algorithm is the best. We use the average effective sum-rate (AESR) and the probability of successful beam (PSB) alignment as the two metrics for performance evaluation. The simulation platform is presented as Python 3.9, Tensorflow 2.9.0, CPU Intel i7-9700K and GPU Nvidia GTX-1070Ti.

The PSB for different beam tracking algorithms with $M = 32$, $U = 2$ is shown in Fig. 3. We have noticed that the proposed centralized DQL (simplified) has the fastest convergence speed and the highest PSB performance. Meanwhile, the centralized DQL converges slowly. The proposed distributed DQL converges fastly when the epoch number is up to 200, but it cannot work well with small epoch numbers lower than 100.

The PSB for different beam tracking algorithms with $M = 64$ is shown in Fig. 4, and the user number $U \in \{1,2,3,4\}$. We have observed that the proposed centralized DQL (simplified) and distributed DQL have similar PSB performance, with varying user numbers. However, the centralized DQL cannot work well and has a poor PSB performance. Besides, the training time costs are listed in Table 2. The time cost increases significantly for centralized methods and remains fixed for the proposed distributed method (the cost time rises due to interactions

▲Figure 4. PSB performance versus user numbers

▼Table 2. Training time cost

User Number	Centralized DQL/s	Centralized DQL (simplified)/s	Distributed DQL (proposed)/s
$U = 1$	10.61	8.82	10.09
$U = 2$	11.91	10.18	11.51
$U = 3$	16.32	14.10	12.82
$U = 4$	87.55	61.97	14.40

DQL: deep Q-learning

with the environment). As U increases, the action space grows exponentially for the centralized methods, and the training is very difficult for $U > 4$. This indicates the proposed distributed method is computationally efficient.

The AESR of the proposed distributed DQL for different beam tracking algorithms with $U = 2$ is shown in Fig. 5, and the user number $M \in \{32, 64, 128, 256\}$. Similar to the case with different user numbers, the proposed centralized DQL (simplified) and distributed DQL have similar AESR perfor-

▲Figure 3. PSB performance versus training epochs
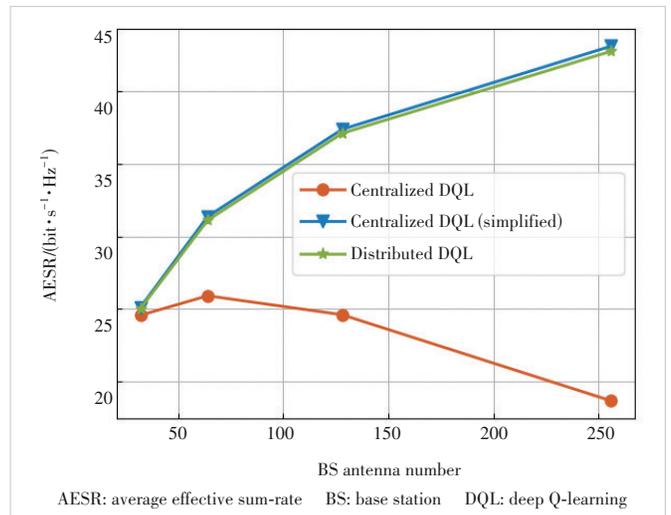
▲Figure 5. AESR performance versus BS antenna numbers

mance, with varying BS antenna numbers, and the centralized DQL has a poor AESR performance.

The scalability is studied in Fig. 6. The distributed DQL indicates the training data and the test data are independent and identically distributed (i.i.d). The distributed DQL (generalized) indicates the training data has a fixed user number/BS antenna number, meanwhile the test data have changing user numbers/BS antenna numbers. The results show that the learned DQN in changing scenarios has the same AESR performance as those with fixed scenarios, and the scalability and generalization ability is verified. The centralized methods cannot work in this test due to mismatched input/output.
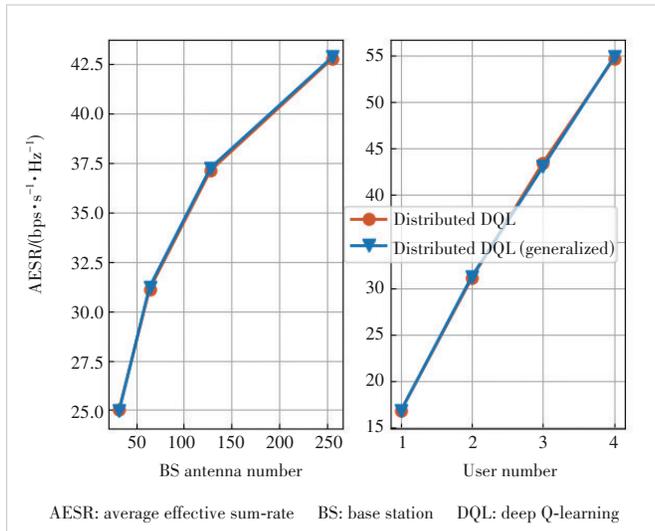
5 Conclusions

In this paper, we investigate multi-user beam tracking in dynamical mmWave scenes, and a multi-agent DQL method under centralized training and distributed execution framework is proposed for online learning. The vanilla DQL is improved in many aspects, such as distributed architecture, rational simplification of training, and state-action-reward designs. Moreover, the proposed method is adaptable to the environment, and is scalable for different BS antenna numbers and user numbers. Simulation results demonstrate the effectiveness of the proposed algorithm.

Appendix

Proof. With the policy π and the initial state s_1, the T-step cumulative reward is defined as:

$$V_\pi^T(s_1) = \mathbb{E}_\pi\left[\frac{1}{T}\sum_{t=1}^{T} r_t | s_1\right] =$$
$$\sum_{a_1 \in A} \pi(a_1|s_1) \sum_{s_2 \in S} P_{s_1 \to s_2}^{a_1} \times \left(\frac{1}{T} r_{s_1 \to s_2}^{a_1} + \frac{T-1}{T} V_\pi^{T-1}(s_2)\right). \quad (19)$$

According to Eqs. (12) and (13), the state value function in Eq. (19) can be rewritten as:

$$V_\pi^T(s_1) = \sum_{a_1 \in A} \pi(a_1|s_1) \sum_{s_2 \in S} P_{s_1 \to s_2} \times$$
$$\left(\frac{1}{T} r_{s_2}^{a_1} + \frac{T-1}{T} V_\pi^{T-1}(s_2)\right) =$$
$$\frac{1}{T}\sum_{a_1 \in A} \pi(a_1|s_1) r_{s_1}^{a_1} + \frac{T-1}{T} \sum_{s_2 \in S} P_{s_1 \to s_2} V_\pi^{T-1}(s_2). \quad (20)$$

The full unrolling of Eq. (20) is given as:

$$V_\pi^T(s_1) = \frac{1}{T}\sum_{a_1 \in A} \pi(a_1|s_1) r_{s_1}^{a_1} +$$
$$\frac{1}{T}\sum_{t=2}^{T}\sum_{a_t \in A} \pi(a_t|s_t) \sum_{s_t \in S}\prod_{t'=1}^{t-1} P_{s_{t'} \to s_{t'+1}} r_{s_t}^{a_t}. \quad (21)$$

As the state transfer is independent of the action and the state can be independently represented as $s = < s_1, \cdots, s_{T+1}>$. Therefore, the maximization of Eq. (21) with respect to a_t, $\forall t$ can be decomposed into the subproblem:

$$\max_{a_t} V_\pi^T(s) \Leftrightarrow \max_{a_t} r_{s_t}^{a_t}. \quad (22)$$

$$\max_{a_t} V_\pi^T(s) = \max_{a_t} \frac{1}{T}\sum_{t'=1}^{T}\sum_{a_{t'} \in A} \pi(a_{t'}|s_{t'}) r_{s_{t'}}^{a_{t'}} \Leftrightarrow$$
$$\max_{a_t} \sum_{a_t \in A} \pi(a_t|s_t) r_{s_t}^{a_t} = \max_{a_t} r_{s_t}^{a_t}. \quad (23)$$

In summary, it can be proved that the maximization of Eq. (21) with respect to $\{a_t|\forall t\}$ can be decomposed into T subproblems:

$$\max_{\{a_t|\forall t\}} V_\pi^T(s) \Leftrightarrow \left\{\max_{a_t} r_{s_t}^{a_t}|\forall t\right\}. \quad (24)$$

The equivalence proof of γ-discounted cumulative reward is similar.

▲Figure 6. Scalability of proposed distributed DQL

AESR: average effective sum-rate   BS: base station   DQL: deep Q-learning

References

[1] KIM Y J, CHO Y S. Beam-tracking technique for millimeter-wave cellular systems using subarray structures [J]. IEEE transactions on vehicular technology, 2018, 67(8): 7806 – 7810. DOI: 10.1109/TVT.2018.2834346

[2] DUAN Q Y, KIM T, HUANG H, et al. AoD and AoA tracking with directional sounding beam design for millimeter wave MIMO systems [C]//The 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC). IEEE, 2015: 2271 – 2276. DOI: 10.1109/PIMRC.2015.7343676

[3] ZHANG D Y, LI A, SHIRVANIMOGHADDAM M, et al. Codebook-based train-

ing beam sequence design for millimeter-wave tracking systems [J]. IEEE transactions on wireless communications, 2019, 18(11): 5333 – 5349. DOI: 10.1109/twc.2019.2935731

[4] HUANG H, PENG Y, YANG J, et al. Fast beamforming design via deep learning [J]. IEEE transactions on vehicular technology, 2020, 69(1): 1065 – 1069. DOI: 10.1109/tvt.2019.2949122

[5] HE H T, JIN S, WEN C K, et al. Model-driven deep learning for physical layer communications [J]. IEEE wireless communications, 2019, 26(5): 77 – 83. DOI: 10.1109/mwc.2019.1800447

[6] HE W L, ZHANG C, HUANG Y M, et al. Intelligent optimization of base station array orientations via scenario-specific modeling [J]. IEEE transactions on communications, 2022, 70(3): 2117 – 2130. DOI: 10.1109/TCOMM.2021.3135532

[7] SU J Y, MENG F, LIU S H, et al. Learning to predict and optimize imperfect MIMO system performance: Framework and application [C]//The IEEE Global Communications Conference. IEEE, 2023: 335 – 340. DOI: 10.1109/GLOBECOM48099.2022.10001369

[8] VA V, CHOI J, SHIMIZU T, et al. Inverse multipath fingerprinting for millimeter wave V2I beam alignment [J]. IEEE transactions on vehicular technology, 2018, 67(5): 4042 – 4058. DOI: 10.1109/TVT.2017.2787627

[9] MENG F, LIU S H, HUANG Y M, et al. Learning-aided beam prediction in mmWave MU-MIMO systems for high-speed railway [J]. IEEE transactions on communications, 2022, 70(1): 693 – 706. DOI: 10.1109/TCOMM.2021.3124963

[10] ZHANG J J, HUANG Y M, ZHOU Y, et al. Beam alignment and tracking for millimeter wave communications via bandit learning [J]. IEEE transactions on communications, 2020, 68(9): 5519 – 5533. DOI: 10.1109/TCOMM.2020.2988256

[11] ZHANG J J, HUANG Y M, WANG J H, et al. Intelligent interactive beam training for millimeter wave communications [J]. IEEE transactions on wireless communications, 2021, 20(3): 2034 – 2048. DOI: 10.1109/TWC.2020.3038787

[12] SUTTON R S, BARTO A G. Reinforcement learning: an introduction [M]. Cambridge, USA: MIT Press. 2018

[13] MENG F, CHEN P, WU L N, et al. Power allocation in multi-user cellular networks: Deep reinforcement learning approaches [J]. IEEE transactions on wireless communications, 2020, 19(10): 6255 – 6267. DOI: 10.1109/TWC.2020.3001736

Biographies

MENG Fan (mengfan@pmlabs.com.cn) received his BS degree from the School of Electronic Engineering, University of Electronic Science and Technology of China in 2015 and the PhD degree with the School of Information Science and Engineering, Southeast University, China in 2020. He is now working in the Purple Mountain Laboratories. His research mainly focuses on applying machine learning techniques in the wireless communication systems. His research interests include machine learning in physical layer, model- and data-driven design, beamforming, beam alignment and tracking, and AI-enhanced positioning.

HUANG Yongming received his BS and MS degrees from Nanjing University, China in 2000 and 2003, respectively, and PhD degree in electrical engineering from Southeast University, China in 2007. Since March 2007, he has been with the School of Information Science and Engineering, Southeast University, where he is currently a full professor. During 2008 – 2009, he visited the Signal Processing Lab, Royal Institute of Technology, Sweden. He has authored or coauthored more than 200 peer-reviewed papers, and holds more than 80 invention patents. He has submitted 20 technical contributions to IEEE standards. His research interests include intelligent 5G/6G mobile communications and millimeter wave wireless communications.

LU Zhaohua received his PhD degree from Tianjin University, China in 2006. He is currently a senior wireless communication system research expert at ZTE Corporation and has long been engaged in the field of wireless communication system design and the key technologies of the physical layer. He has many technical contributions, papers, and patents in interference mitigation in the MIMO field.

XIAO Huahua received his MS degree in computer software and theories from Sun Yat-Sen University, China. He is currently a senior engineer in the field of antenna algorithm pre-research with ZTE Corporation. He has applied for more than 150 patents in the multi-antenna field home and abroad.

RIS-Assisted UAV-D2D Communications Exploiting Deep Reinforcement Learning

YOU Qian[1], XU Qian[1], YANG Xin[1], ZHANG Tao[2], CHEN Ming[3]

(1. School of electronics and information, Northwestern Polytechnical University, Xi'an 710072, China；
2. China Academy of Launch Vehicle Technology, Beijing 100076, China；
3. Hangzhou Hikvision Digital Technology Co., Ltd., Hangzhou 310051, China)

Abstract: Device-to-device (D2D) communications underlying cellular networks enabled by unmanned aerial vehicles (UAV) have been regarded as promising techniques for next-generation communications. To mitigate the strong interference caused by the line-of-sight (LoS) air-to-ground channels, we deploy a reconfigurable intelligent surface (RIS) to rebuild the wireless channels. A joint optimization problem of the transmit power of UAV, the transmit power of D2D users and the RIS phase configuration are investigated to maximize the achievable rate of D2D users while satisfying the quality of service (QoS) requirement of cellular users. Due to the high channel dynamics and the coupling among cellular users, the RIS, and the D2D users, it is challenging to find a proper solution. Thus, a RIS softmax deep double deterministic (RIS-SD3) policy gradient method is proposed, which can smooth the optimization space as well as reduce the number of local optimizations. Specifically, the SD3 algorithm maximizes the reward of the agent by training the agent to maximize the value function after the softmax operator is introduced. Simulation results show that the proposed RIS-SD3 algorithm can significantly improve the rate of the D2D users while controlling the interference to the cellular user. Moreover, the proposed RIS-SD3 algorithm has better robustness than the twin delayed deep deterministic (TD3) policy gradient algorithm in a dynamic environment.

Keywords: device-to-device communications; reconfigurable intelligent surface; deep reinforcement learning; softmax deep double deterministic policy gradient

1 Introduction

Current communication systems and applications are pursuing higher and higher transmission rates, which brings greater challenges to the scarce spectrum resources. Thus, spectrum-efficient communications become increasingly important, which promotes the development of the next-generation cellular networks. Among the various spectrum-efficient techniques, the device-to-device (D2D) communication underlying the cellular network has been considered a promising technique for boosting the communication rates between two neighbor nodes, since it allows the two users to transmit signals directly without passing through a base station (BS)[1]. To maximize the performance of the D2D and cellular network, the location of the BS usually needs to be optimized, which is difficult to realize for the traditional terrestrial cellular network. Fortunately, unmanned aerial vehicles (UAVs) have played a critical role in 6G networks due to their flexibility. For instance, UAVs can work as the aerial BS to improve the network capacity and expand the coverage area, and thus help overcome the limitations of the terrestrial wireless communication at the physical layer[2].

With the UAV aerial BS, the dominant links are usually line-of-sight (LoS) links that benefit the intended receivers while causes strong interference to the unintended users. In this case, reconfigurable intelligent surface (RIS) can be employed to reconstruct the transmission environment and thus reach a compromise between the performances of the intended and other users[3 – 5]. RIS consists of many low-cost passive reflection elements, where each element can adaptively adjust its reflection amplitude and/or phase to control the intensity and the direction of the electromagnetic wave. In this way, RIS can enhance and/or weaken the strength of the reflected signal for different users[3]. For the D2D communication system with plenty of low-power terminal devices,

RIS can be deployed to improve the quality of the communication links for cellular user equipment (CUE) and mitigate the co-channel interference between CUE and D2D users[6]. There have been some studies on RIS-assisted D2D communications. For example, the authors in Ref. [7] proposed a relaxation-based algorithm called Riemannian manifold based alternating direction multiplier (RM-ADMM) to optimize the system configuration, which is a quadratic constraint quadratic optimization (QCQP) problem. This kind of proposal adopts traditional optimization methods, which may converge to the local optima and cause system performance loss.

Recently, artificial intelligence (AI) has been regarded as a powerful tool to solve complicated non-linear optimization problems. In Ref. [8], a deep learning-based method is proposed for the effective online configuration of the smart surface, where the proposed deep neural network (DNN) model maps the target user's information and the optimal phase matrix to maximize the user's received signal strength by calculating the measurement coordinates. It is worth noting that the deep learning method requires large-scale data sets, which is impractical for some applications. To overcome the limitations of deep learning, deep reinforcement learning (DRL), which combines deep learning and reinforcement learning, has been widely used in wireless communication systems. In Ref. [9], the non-convex optimization problem consisting of beamforming design, power control, and interference coordination is jointly optimized by DRL. In Ref. [10], the authors investigated the simultaneous wireless information and power transfer network where the UAV and the RIS are deployed. By exploiting the DRL to optimize the RIS passive beamforming, the total harvested energy is maximized while meeting the quality of service (QoS) requirements for communications. Ref. [11] is a very early attempt to develop a framework for integrating DRL techniques into optimization designs with no need to understand explicit models or specific mathematical formulas of the wireless environment to solve large-dimensional optimization problems.

At present, the commonly used algorithms for processing continuous action space in DRL are deep deterministic policy gradient (DDPG) and its improved version, the twin delayed deep deterministic (TD3) policy gradient. But the introduction of the underestimation bias by the TD3 algorithm will affect the performance. Studies have shown that softmax's smoothing effect can help learn and reduce the number of local optima[12]. Thus, the authors in Ref. [13] proposed a softmax deep double deterministic (SD3) policy gradients algorithm. The analyses show that the error between the value function and the optimal value under the softmax operator is bounded.

To overcome the complex problem of traditional algorithm calculation, we exploit the SD3 algorithm to jointly design the transmit power of the UAV, the transmit power of the D2D users, and the RIS phase configuration. The main contributions of this paper are summarized as follows:

1) Firstly, we formulate a RIS-assisted UAV-D2D communication system model. In our considered system, the UAV is used as an aerial BS to overcome the limitations of conventional terrestrial BSs. Besides, to investigate the impact of the time-varying channels on the system performance, the motion state of the UAV moving from the CUE to the D2D users is taken into consideration.

2) Secondly, we propose a RIS-SD3 algorithm to solve the complex optimization problem involved in the RIS-assisted UAV-D2D communication system. Unlike the TD3 algorithm, SD3 merges the softmax operator into the action key of continuous control, which makes the optimization environment smoother and thus is conducive to empirical learning.

3) Finally, unlike previous studies that exploit alternating methods to optimize the transmit power and the RIS phase, the proposed algorithm optimizes the transmit power and the phase of the RIS simultaneously. To be more specific, the sum rate of the D2D users is adopted as an immediate reward for training the RIS-SD3 algorithm. The sum rate is gradually maximized by iteratively adjusting the parameters of the RIS-SD3 according to the reward.

The remainder of this paper is organized as follows. The system model is described in Section 2. In Section 3, the RIS-SD3 algorithm is introduced to optimize the phase shift and the transmit power. In Section 4, simulation results are presented to evaluate the performance of the proposed algorithm. The conclusions are given in Section 5.

2 System Model

We consider a practical RIS-assisted UAV-D2D communication network. For example, in a dense urban environment with tall buildings, the primary user, like CUE, is close to the RIS, while the D2D user is located at the edge of the cell. The detailed system description is as follows.

2.1 System Descriptions

The system model is depicted in Fig. 1. We consider a downlink cellular transmission assisted by UAV and RIS. The system consists of one UAV serving as the BS, one RIS, K CUE, and D D2D pairs. To simplify the following analysis, only one CUE is considered in this paper, and a scenario with multiple CUE will be studied in future work. The BS, CUE, D2D transmitter (DT), and the associated D2D receiver (DR) are all single antenna devices. Besides, the RIS is equipped with M reflecting elements and the reflection coefficient matrix $\boldsymbol{\Theta}$ can be described as $\boldsymbol{\Theta} = \text{diag}\left(\beta_1 e^{j\theta_1}, \beta_2 e^{j\theta_2}, \cdots, \beta_M e^{j\theta_M}\right)$.

The CUE receives the desired signals including the signals sent by the BS and the signals reflected from the RIS. In addition, it will receive the interference signals from all the D2D pairs. Therefore, the signal received by the CUE can be

written as:

$$y_c = \left(\boldsymbol{h}_{r,c}^H \Theta \boldsymbol{h}_r + h_c \right) \sqrt{P_c}\, s + \sum_{d=1}^{D} h_{d,c} \sqrt{P_d}\, \mu_d + n_c, \quad (1)$$

where $\boldsymbol{h}_r \in \mathbb{C}^{M \times 1}, \boldsymbol{h}_{r,c} \in \mathbb{C}^{M \times 1}, h_c \in \mathbb{C}$, and $h_{d,c} \in \mathbb{C}$ represent the channel gains of UAV-RIS, RIS-CUE, UAV-CUE, and the d-th DT to CUE, respectively; $P_c \in \mathbb{R}$ and $s \in \mathbb{C}$ denote the transmit power and the transmit signal of the BS-CUE link, respectively; $P_d \in \mathbb{R}$ and $\mu_d \in \mathbb{C}$ are the transmit power of the d-th DT and the data transmitted to the d-th DR, respectively; $n_c \sim \mathcal{CN}\left(0, \sigma_c^2 \right)$ is the additive Gaussian white noise at the CUE.

The wireless transmission link between the user and the UAV can be either LoS or NLoS. Thus, the received signal power at each user's location is given by Ref. [14].

$$P_r = \begin{cases} P_c d^{-\alpha_0}, & \text{LoS} \\ \eta P_c d^{-\alpha_0}, & \text{NLoS}, \end{cases} \quad (2)$$

where d is the distance between the user and the UAV, α_0 is the path loss exponent over the user-UAV link, and η is an additional factor related to the NLoS link. The LoS probability can be expressed as $P_{\text{LoS}} = \dfrac{1}{1 + A \exp\left(-B(\theta - A) \right)}$, where A and B are constant values that depend on the environment. In this paper, we set $A = 9.6, B = 0.15$, and $\eta = 20$ dB ; $\theta = \dfrac{180}{\pi} \sin^{-1}\left(\dfrac{h}{d} \right)$ is the elevation angle where h is the altitude between the user and UAV. The probability of NLoS is $P_{\text{NLoS}} = 1 - P_{\text{LoS}}$[14].

For the terrestrial links, we assume that they follow the Rayleigh distribution where the path-loss is given by $\rho \left(\dfrac{d}{d'} \right)^{-v}$, where ρ, d and v represent the path loss at the refer-
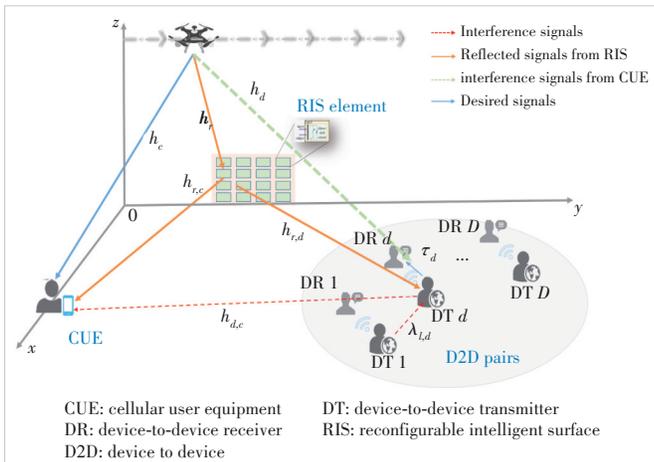
▲ **Figure 1. System model of a practical RIS-assisted unmanned aerial vehicle (UAV)-D2D communication network**

ence distance of $d' = 1$, the individual link distance, and the corresponding path loss exponent, respectively.

Note that the m-th element of the diagonal matrix can be written as $\phi_m = \beta_m e^{j\theta_m}$, where $\theta_m \in [\, 0, 2\pi)$ is the phase shift. Generally speaking, phase-shift control achieves better passive beamforming performance than amplitude control, so we assume ideal reflection by the RIS so that the signal power is lossless from each reflection element, e.g., the amplitude reflection coefficient $\beta_m = 1$[15].

The Signal to Interference plus Noise Ratio (SINR) for the received signal of CUE can be calculated as:

$$\text{SINR}_c = \frac{\left| \boldsymbol{h}_{r,c}^H \Theta \boldsymbol{h}_r + h_c \right|^2 P_c}{\displaystyle\sum_{D}^{d=1} \left| h_{d,c} \right|^2 P_d + \sigma_c^2}. \quad (3)$$

Thus, the achievable rate of CUE is:

$$R_c = \log_2\left(1 + \text{SINR}_c \right) = \log_2\left(1 + \frac{\left| \boldsymbol{h}_{r,c}^H \Theta \boldsymbol{h}_r + h_c \right|^2 P_c}{\displaystyle\sum_{D}^{d=1} \left| h_{d,c} \right|^2 P_d + \sigma_c^2} \right). \quad (4)$$

The signals received at the d-th DR consists of the desired signal received from the d-th DT, the interference signal from the UAV, and the reflected signal from the RIS, in addition to the interference signal received from the other D2D pairs. Thus, the signal received at the d-th DR is given by:

$$y_d = \tau_d \sqrt{P_d}\, \mu_d + \left(\boldsymbol{h}_{r,d}^H \Theta \boldsymbol{h}_r + h_d \right) \sqrt{P_c}\, s + \sum_{D}^{l \ne d} \lambda_{l,d} \sqrt{P_l}\, \mu_l + n_d, \quad (5)$$

where $\boldsymbol{h}_{r,d} \in \mathbb{C}^{M \times 1}, h_d \in \mathbb{C}, \tau_d \in \mathbb{C}$, and $\lambda_{l,d} \in \mathbb{C}$ denote the channel gains of RIS-DR d, UAV-DR d, DT d-DR d, and DT l-DR d, respectively; P_l and μ_l are the transmit power of the l-th DT and the transmit data of D2D to the l-th DR, respectively; $n_d \sim \mathcal{CN}\left(0, \sigma_d^2 \right)$ denotes the additive Gaussian white noise at the d-th DR.

Similarly, the received SINR for the d-th DR is given by:

$$\text{SINR}_d = \frac{\left| \tau_d \right|^2 P_d}{\displaystyle\sum_{l \ne d}^{D} \left| \lambda_{l,d} \right|^2 P_l + \left| \boldsymbol{h}_{r,d}^H \Theta \boldsymbol{h}_r + h_d \right|^2 P_c + \sigma_d^2}. \quad (6)$$

The achievable rate of the d-th DR is

$$R_d = \log_2\left(1 + \text{SINR}_d \right) =$$

$$\log_2\left(1 + \frac{\left| \tau_d \right|^2 P_d}{\displaystyle\sum_{D}^{l \ne d} \left| \lambda_{l,d} \right|^2 P_l + \left| \boldsymbol{h}_{r,d}^H \Theta \boldsymbol{h}_r + h_d \right|^2 P_c + \sigma_d^2} \right). \quad (7)$$

Accordingly, the sum rate of all the D2D pairs is

$$R_{\text{total}} = \sum_{d=1}^{D} R_d.$$ (8)

2.2 Problem Formulation

In order to increase the sum rate of the D2D pairs while limiting the amount of interference to the CUE, the problem is formulated as a non-convex optimization problem as follows

$$\max_{p,\Theta,\varphi_c} \sum_{D}^{d=1} \log_2(1 + \text{SINR}_d)$$ (9)

$$\text{s.t.} \sum_{D}^{d=1} \left| h_{d,c} \right|^2 P_d \leq I_T,$$ (9a)

$$0 \leq P_d \leq P_t, \forall d \in \{1,2,\cdots,D\},$$ (9b)

$$0 \leq P_c \leq P_{\max},$$ (9c)

$$\text{SINR}_c \geq \text{SINR}_{\text{thr}},$$ (9d)

$$R_d \geq R_{d\_\text{thr}}, \forall d \in \{1,2,\cdots,D\},$$ (9e)

$$\theta_m \in [0, 2\pi), \forall m \in \{1,2,\cdots,M\},$$ (9f)

$$\left| \phi_m \right| = 1, \forall m \in \{1,2,\cdots,M\},$$ (9g)

where $P = \{P_1, P_2, \cdots, P_D\}$ is the transmit power vector for D2D pairs; P_t is the maximum transmit power of DT and $P_{\max}$ is the maximum transmit power of UAV. I_T in Constraint (9a) indicates the maximum allowable interference to the cellular transmission. Constraints (9b) and (9c) denote the transmit power limit for each DT and the maximum power limit for the UAV BS. Constraints (9d) and (9e) denote the QoS requirements for CUE and D2D pairs. Constraints (9f) and (9g) specify the phase shift and the amplitude constraint of the RIS.

Due to the non-convexity of the objective function and the performance loss of the traditional successive convex approximation (SCA) method[16], we propose a DRL-based framework to solve the non-convex optimization problem.

3 Proposed RIS-SD3 Algorithm

3.1 Description of SD3

SD3 is the abbreviation for the deep double deterministic policy gradient al-

gorithm, which enables a better value estimation by reducing the overestimation bias in DDPG and smoothing the optimized environment, thus contributing to experiential learning[13].

The process of the SD3 algorithm is shown in Fig. 2. SD3 includes an actor network $\mu(\cdot)$ and a critic network $Q(\cdot)$. The actor network consists of two online and two target policy networks with the different parameters $\theta_i^\mu, \theta_i^{\mu'} (i = 1,2)$. Similarly, the critic network consists of two online and two target Q-networks with different parameters $\theta_i^Q, \theta_i^{Q'} (i = 1,2)$.

According to Fig. 2, we can see that at the time step t, the agent selects an action a_t based on the actor network $\mu(s_t; \theta^\mu)$. Meanwhile, a random noise $N_t \sim \mathcal{N}(0,\sigma)$ is added to interact with the environment to more fully explore the policy. Thus, the action a_t can be written as

$$a_t = \mu(s_t; \theta^\mu) + N_t.$$ (10)

After the agent executes the action, it will return a reward defined as below

$$Q_{t+1}(s, a) = r_t(s,|a) + \gamma \mathbb{E}_{s' \sim p(\cdot|s,a)} [V_t(s')],$$ (11)

where $V_{t+1}(s) = \text{softmax}_\beta(Q_{t+1}(s, \cdot))$ is the softmax operator, which is used to update the value function $Q_{t+1}(s,a)$ iteratively. Since the softmax operator itself involves integrals and thus is difficult to handle in continuous action space, we use the following unbiased estimation to replace the term $V_t(s')$ as in Ref. [13]:

$$\mathbb{E}_{a' \sim p} \left[\frac{\exp(\beta\hat{Q}(s',a'))\hat{Q}(s',a')}{p(a')} \right] / \mathbb{E}_{a' \sim p} \left[\frac{\exp(\beta\hat{Q}(s',a'))}{p(a')} \right],$$ (12)

where $p(a')$ is the probability density function of Gaussian distribution.
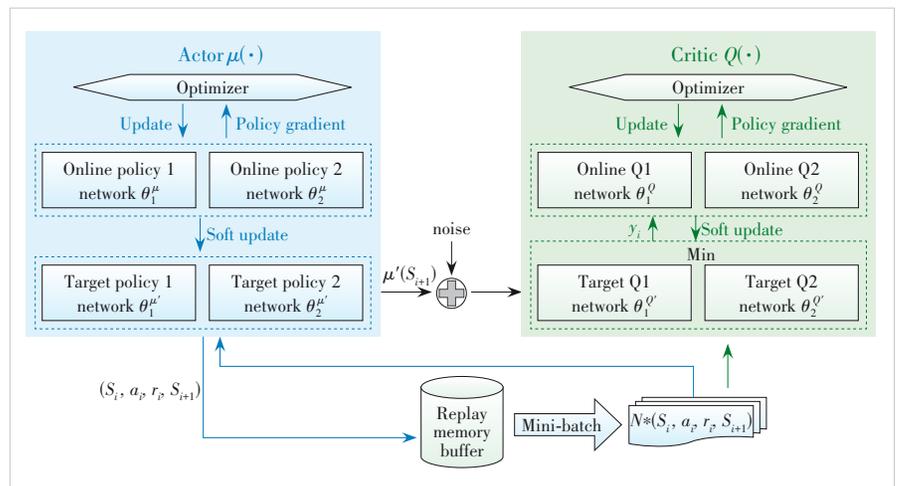
▲Figure 2. Workflow of the deep double deterministic (SD3) policy gradient algorithm

For ease of representation, we introduce

$$\hat{Q}_i(s',a') = \min\left(Q_i(s',a';\theta_i^{Q'}),Q_{-i}(s',a';\theta_{-i}^{Q'})\right), \quad (13)$$

and

$$Y_{SD_3}^{-i}(s') = \text{soft max}_\beta\left(\hat{Q}_i(s',\cdot)\right). \quad (14)$$

With Eqs. (13) and (14), the target values for the critic network in Fig. 2 can be estimated as:

$$y_i = r + \gamma Y_{SD_3}^{-i}(s'), i = 1, 2. \quad (15)$$

Then, the critic network optimizes its parameters θ_i^Q by minimizing the loss function given by:

$$L = \frac{1}{N_B}\left(y_i - Q_i(s,a;\theta_i^Q)\right)^2. \quad (16)$$

After the critic network updates its parameters, the actor network is updated by θ_i^μ following the applying the chain rule.

$$\nabla_{\theta_i^\mu} = \frac{1}{N_B}\nabla_a Q_i(s,a;\theta_i^Q)\Big|_{s=s_t,a=\mu(s_t;\theta_i^\mu)}\nabla_{\theta^\mu}\mu(s;\theta_i^\mu)\Big|_{s=s_t}, \quad (17)$$

To make the learning process more stable, the SD3 also uses a soft target update approach:

$$\begin{aligned}\theta_i^{Q'} &\leftarrow \tau\theta_i^Q + (1-\tau)\theta_i^{Q'}, \\ \theta_i^{\mu'} &\leftarrow \tau\theta_i^\mu + (1-\tau)\theta_i^{\mu'},\end{aligned} \quad (18)$$

where τ is the learning rate for updating the target critic network and the target actor network.

Algorithm 1: Learning algorithm of RIS-SD3

Input: $h_{r,c}, h_r, h_{r,d}, h_c, h_{d,c}, h_d, \tau_d, \lambda_{l,d}$
Output: the optimal action
$\quad a = \left\{p_1^{\text{opt}},p_2^{\text{opt}},\theta_1^{\text{opt}},\theta_2^{\text{opt}},\cdots,\theta_M^{\text{opt}},p_c^{\text{opt}}\right\}$

1   Initialize actor networks μ_1,μ_2 and critic networks Q_1,Q_2
     with random parameters $\theta_1^\mu, \theta_2^\mu, \theta_1^Q, \theta_2^Q$;
2   Initialize the size of experience replay N_R, the size of mini-batches N_B and replay buffer R;
3   **for** $t = 1,\cdots,T$ **do**
4   Select an action with exploration noise $N_t \sim \mathcal{N}(0,\sigma)$ based on executing action a, obtained reward r, new state s' and done;
5   Store transition tuple $(s, a, r, s', \text{done})$ in R;
6   **for** $i = 1, 2$ **do**
7   Sample a random minibatch of N from R;
8   Sample K noises $N_t \sim \mathcal{N}(0,\sigma)$;
9   Set $\hat{a}' = \mu_i'(s_{i+1}) + \text{clip}(N_t, -c,c)$;
10   Set $\hat{Q}(s',\hat{a}') = \min_{j=1,2}\left(Q_j(s',\hat{a}';\theta_j^{Q'})\right)$;
11   Set $\text{softmax}_\beta\left(\hat{Q}(s',\hat{a}')\right)$ as Eq. (12)
12   Update critic net via minimizing Eq. (16);
13   Update actor net by policy gradient in Eq. (17);
14   Update the target networks
       $\theta_i^{Q'} \leftarrow \tau\theta_i^Q + (1-\tau)\theta_i^{Q'}$;
       $\theta_i^{\mu'} \leftarrow \tau\theta_i^\mu + (1-\tau)\theta_i^{\mu'}$
15   **end**
16   **end**

3.2 Details of RIS-SD3

In this paper, the environment depends on our proposed system model. At the time step t, the agent can collect the current channel information, and combined with the current state, the agent selects the action and calculates the reward according to the current policy. There are E episodes in the whole training process, and each episode is iterated by T times. The detailed workflow of the proposed RIS−SD3 algorithm is shown in Algorithm 1. The state space, action space and reward function are given as follows.

1) State: The state s_t at the t-th time step is constructed by the received signal of CUE, the UAV's location at the t-th time step, and the SINR of D2D pairs. So the total number of the state is $D + M + K + 1$.

2) Action: The action is constructed by the transmit power vector $\boldsymbol{P} = \{P_1,P_2,\cdots,P_D\}$, the transmit power of BS P_c and the phase $\theta_i(i = 1,2,\cdots,M)$ of RIS. In order to reduce the complexity of the action space, we convert both phase and power into one-dimensional vectors, i. e., action $= \{P_1,P_2,\cdots P_D,\theta_1,\theta_2,\cdots\theta_M,P_c\}$. So the total number of the action is $D + M + 1$.

3) Reward: In the proposed RIS-SD3 algorithm, the sum rate of the D2D pairs is taken as the reward. Furthermore, in order to satisfy the minimum signal-to-noise ratio and the maximum interference requirements for CUE users and the QoS requirements for D2D users, the reward can therefore be set as:

$$R_t = \begin{cases} R_{\text{total}}, & \begin{aligned}&\text{if } \sum_{d=1}^{D}\left|h_{d,c}\right|^2 P_d \leqslant I_T \\ &\text{SINR}_c \geqslant \text{SINR}_{\text{thr}} \\ &R_d \geqslant R_{d\_\text{thr}}\end{aligned} \\ 0, & \text{else} \end{cases} \quad (19)$$

The reward for each episode is:

$$R = \sum_{t=1}^{T} R_t. \quad (20)$$

4 Numerical Results

In order to facilitate the analysis, we consider *D=2*, and the other parameters used in the algorithm are shown in

Table 1, and the establishment of the coordinate system is shown in Fig. 1. After continuous training tests, we then find the training work the best when the main hyper-parameters in the RIS-SD3 are set as follows: $E = 10\,000$, $T = 61$, $\gamma = 0.99$, and $\tau = 0.005$.

4.1 Impact of Parameters Settings of RIS-SD3

In our proposed RIS-SD3 algorithm, we use a constant learning rate and batch size for all networks to investigate their effects on the performance and convergence speed for the DRL-based approach. Fig. 3 demonstrates the average rewards versus time episodes at different learning rates. It can be seen that different learning rates have a great impact

▼Table 1. Parameters of the proposed system

	Parameter	Value
Location	UAV	From (0, 0, 1) m to (0, 60, 1) m
	RIS	(0, 10, 2) m
	CUE	(20, 0, 1) m
	DT1	(20, 60, 1) m
	Distance of D2D	5 m
	Size area of D2D	10 m
$SINR_{thr}$	Minimum SINR of CUE	12 dB
R_{d_thr}	Minimum achievable rate of D2D	2 dB
I_T	Maximum interference of CUE	−30 dB
P_{max}	Max transmit power of UAV	30 W
P_t	Max transmit power of DT	10 W, 20 W, 30 W
β	Path loss coefficient	−30 dB
α_0	Path loss exponent over the user-UAV link	3
ν	Path loss exponent	2.5
ρ	The path loss at the reference distance	0.01

CUE: cellular user equipment
D2D: device-to-device
DT: D2D transmitter

RIS: reconfigurable intelligent surface
SINR: Signal to Interference plus Noise Ratio
UAV: unmanned aerial vehicle
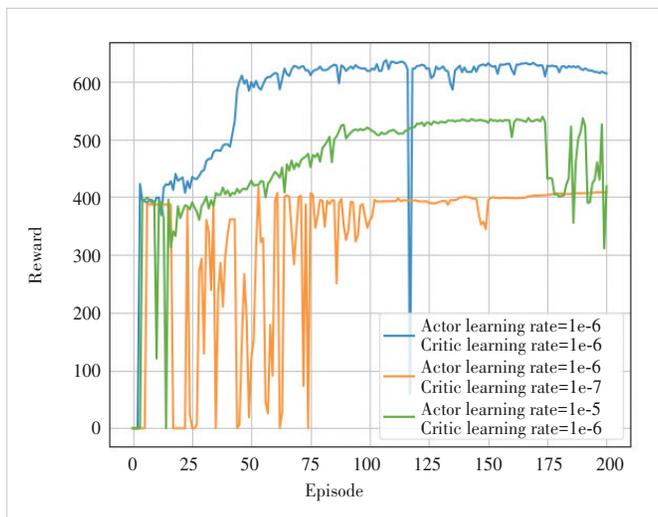
▲Figure 3. Effect of the learning rate
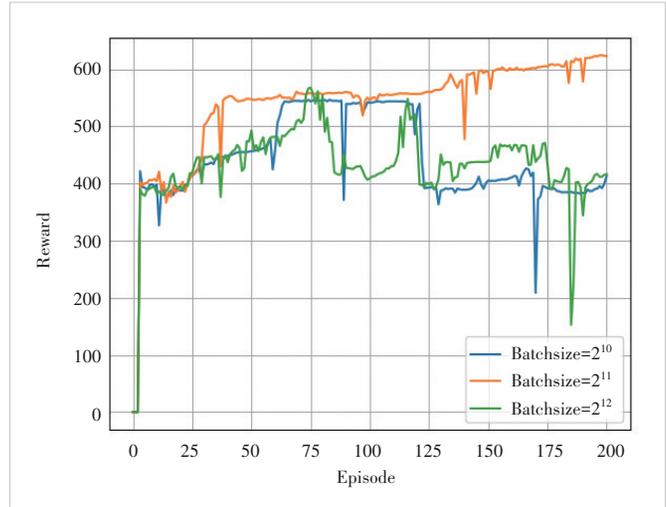
▲Figure 4. Effect of batchsize on the training model

on the performance of the proposed RIS-SD3 algorithm. As shown in Fig. 3, RIS-SD3 with actor and critic learning rates of 1e-6 performs best. Specifically, when the learning rate is too large, the algorithm will be unstable and even cannot converge. On the contrary, when the learning rate is too small, the convergence rate will be slow or even incapable to learn, and thus the training time is wasted.

Batchsize is the number of data used for each update when using the optimizer. In short, it is how many data we want to put into the model at a time to train. This value is between 1 and the total number of training samples.

As shown in Fig. 4, we explore the impact of batchsize on the training model. If the batchsize is too small, time-consuming and training efficiency is low, the training data will be very difficult to converge, resulting in a state of under-fitting. In a certain range, generally speaking, the larger the batchsize, the more accurate the determined descending direction, and the smaller the training shock. The batchsize increases to a certain extent, and its determined decline direction has basically not changed. Therefore, the larger the batch size is, the more stable the gradient will be, while the smaller the batch size is, the higher the randomness of the gradient will be. However, if the batch size is too large, the the demand for memory will be higher, and it is not conducive to the network jumping out of the local minimum. We can see that batchsize = 2^{11} is the best, so this value is used in the following simulations.
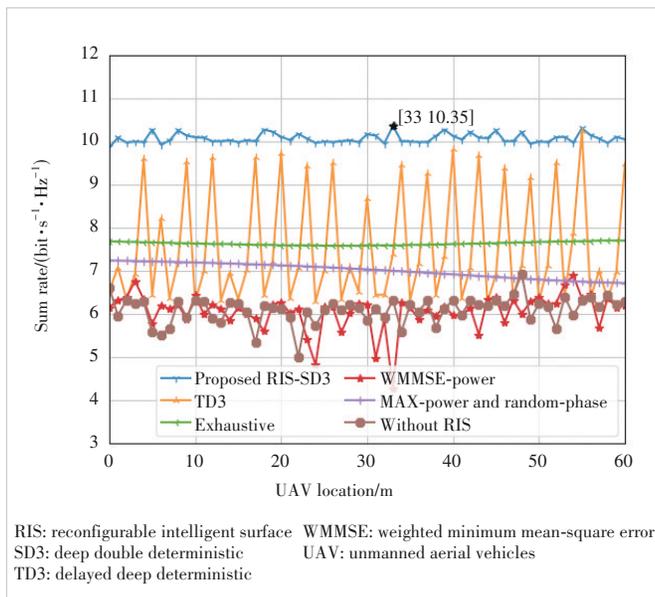
4.2 Comparisons with Benchmarks

To further demonstrate the performance and the time complexity of the proposed RIS-SD3 algorithm, we consider the following baseline schemes. Firstly, we use the exhaustive searching approach to find the approximate optimal value, where the transmit power and the phase are limited to ten equally spaced values. Then, for the weighted minimum

mean-square error (WMMSE)-power" baseline scheme, we use the WMMSE algorithm in Ref. [17] to optimize the transmit power of D2D. For the "max-power and random-phase" baseline scheme, we assume that the RIS configures the phase shifts in a random manner with the maximum D2D transmit power. For the "without RIS" baseline scheme, we assume that the D2D transmit power is random without the deployment of the RIS. Moreover, the TD3 algorithm is also introduced. Unless otherwise specified, the learning rate for the RIS-SD3 algorithm is set as 1e-6.

As shown in Fig. 5, compared with the TD3 algorithm for continuous actions, the proposed RIS-SD3 algorithm is more robust under dynamic channel conditions because it refers to the soft operator in Ref. [13]. Compared with the more accurate exhaustive searching results and the WMMSE algorithm, the proposed RIS-SD3 algorithm can obtain a larger sum rate. In addition, it can be seen from the figure that the proposed algorithm and the exhaustive searching algorithm are more robust to the position change of the UAV. Finally, by comparing the results of the proposed algorithm with the "without RIS" scheme, we improve the system performance by introducing RIS, since the RIS provides the additional degrees of freedom (DoF) to improve the sum rate. In addition, since exhaustive search only considers partially discrete values, its effect is slightly lower than that of the RIS-SD3 algorithm that considers continuous values.

Moreover, it can be seen from Fig. 5 that the sum rate fluctuates as the position of UAVs changes, especially for the TD3 and the WMMSE algorithms. Actually, due to the introduction of RIS, the system performance is not that sensitive to the position of the UAV. The up and down phenomena indicate that the performance of the TD3 algorithm is poor for the considered scenario, which motivates us to propose the RIS-SD3 algorithm. As for the WMMSE algorithm, the reason for the fluctuation is that this algorithm only optimizes the D2D user's transmit power, while the phase is random.

To evaluate the time complexity of the proposed method, the time consumption of the proposed scheme and the baseline schemes are shown in Table 2, where the device we use is NVIDIA GPU RTX 3090. It can be observed that the time consumption of the proposed algorithm is less than most of the baselines, but a little bit more than the TD3 algorithm.

However, it should be noted that the TD3 cannot adapt to the change of the UAV location, as observed in Fig. 5.
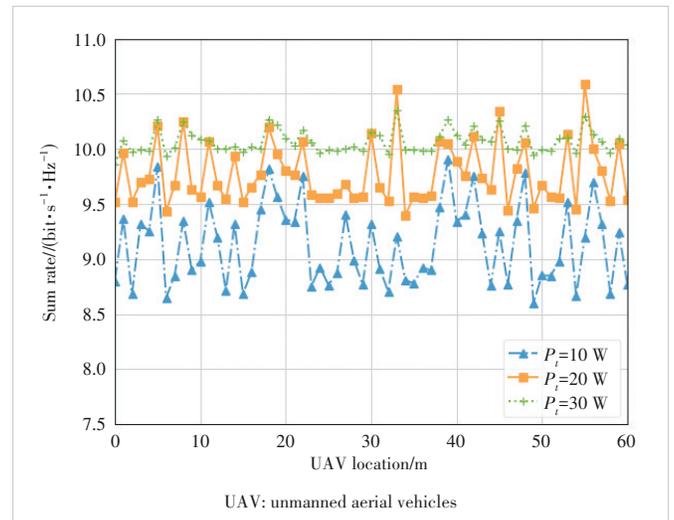
4.3 Impact of Parameter Settings on System

To get a better understanding of the RIS-SD3 method, we investigate the impact of the max power of DT. When more transmitting power is allocated to D2D users, the proposed RIS-SD3 algorithm can obtain a higher sum rate. This observation is consistent with the results in the traditional multi-input single-output (MISO) system. Through the joint design of transmit beamforming and phase shift, the common channel interference of multi-user MISO systems can be effectively reduced, thereby improving performance.

▼Table 2. Time consumption comparison

Scheme	Time Consumption/s
Proposed RIS-SD3	1.74
TD3	1.08
Exhaustive	3.79e+05
WMMSE-power	1.68
Max-power and random-phase	3.67
Without RIS	3.98

RIS: Reconfigurable intelligent surface
SD3: softmax deep double deterministic
TD3: delayed deep deterministic
WMMSE: weighted minimum mean-square error

RIS: reconfigurable intelligent surface   WMMSE: weighted minimum mean-square error
SD3: deep double deterministic   UAV: unmanned aerial vehicles
TD3: delayed deep deterministic

▲Figure 5. RIS-SD3 in comparison with other baseline schemes

UAV: unmanned aerial vehicles

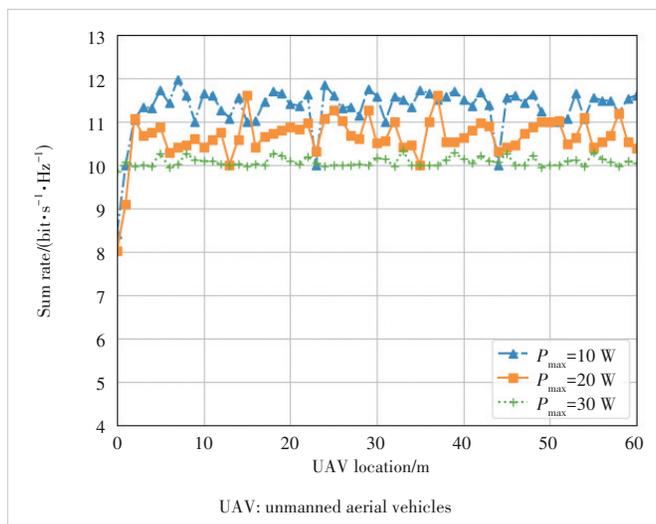▲Figure 6. Sum Rate under different P_t

As can be seen from Fig. 6, during the movement of the UAV from (0, 0, 30) m to (0, 60, 30) m, there are also certain fluctuations in the system about the sum rate. In fact, the UAV only changes continuously by 1 m, while the receiver may not have time to feel this change, and then the UAV moves to the next position. Therefore, it can be seen from Fig. 6 that the undulating change is random.

In addition, we simulate the effect of the maximum transmitting power of UAV P_{max} on the D2D sum rate. It can be seen from Fig. 7 that, as the maximum transmitting power of the UAV increases, the D2D sum rate decreases. This is because with the transmit power of the UAV increases, the interference of the cellular user to the D2D user increases, so the D2D sum rate decreases.

It can also be seen from Fig. 7 that in the process of the drone moving from $(0, 0, 30)$ m to $(0, 60, 30)$ m, the sum rate has certain ups and downs. The specific reason may be that the UAV changes less, so the fluctuations are more random, but the overall change is not very significant.

5 Conclusions

Based on the latest progress in DRL for continuous action space, a RIS-SD3 optimization algorithm is proposed to solve the joint power allocation and phase optimization problem in a dynamic RIS-assisted UAV-D2D communication network. With the RIS-SD3 algorithm, the sum rate of the D2D users is maximized while meeting the QoS requirement for the cellular user. Specifically, by introducing softmax operators, the proposed algorithm learns about the environment more efficiently, and thus has better robustness to the change of the environment. Simulation results show that the proposed RIS-SD3 method can learn from the environment by observing the instantaneous reward got from the time-varying wireless channels, and then gradually improves its behavior to the optimal

result. Compared with the baseline schemes, the proposed scheme can increase the sum rate as well as improve the robustness of the transmission environment.

References

[1] DANG S P, CHEN G J, COON J P. Multicarrier relay selection for full-duplex relay-assisted OFDM D2D systems [J]. IEEE transactions on vehicular technology, 2018, 67(8): 7204 – 7218. DOI: 10.1109/TVT.2018.2829401

[2] WU Q Q, ZENG Y, ZHANG R. Joint trajectory and communication design for multi-UAV enabled wireless networks [J]. IEEE transactions on wireless communications, 2018, 17(3): 2109 – 2121. DOI: 10.1109/TWC.2017.2789293

[3] WU Q Q, ZHANG R. Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network [J]. IEEE communications magazine, 2020, 58(1): 106 – 112. DOI: 10.1109/MCOM.001.1900107

[4] WU Q Q, ZHANG R. Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design [C]//IEEE Global Communications Conference (GLOBECOM). IEEE, 2018: 1 – 6. DOI: 10.1109/GLOCOM.2018.8647620

[5] HUANG C W, ZAPPONE A, ALEXANDROPOULOS G C, et al. Reconfigurable intelligent surfaces for energy efficiency in wireless communication [J]. IEEE transactions on wireless communications, 2019, 18(8): 4157 – 4170. DOI: 10.1109/TWC.2019.2922609

[6] YANG G, LIAO Y, LIANG Y C, et al. Reconfigurable intelligent surface empowered device-to-device communication underlaying cellular networks [J]. IEEE transactions on communications, 2021, 69(11): 7790 – 7805. doi: 10.1109/TCOMM.2021.3102640

[7] CAO Y S, LV T J, NI W, et al. Sum-rate maximization for multi-reconfigurable intelligent surface-assisted device-to-device communications [EB/OL]. [2023-01-20]. https://arxiv.org/abs/2108.07091

[8] HUANG C W, ALEXANDROPOULOS G C, YUEN C, et al. Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces [C]//IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2019: 1 – 5. DOI: 10.1109/SPAWC.2019.8815412

[9] SHAFIN R, CHEN H, NAM Y H, et al. Self-tuning sectorization: Deep reinforcement learning meets broadcast beam optimization [J]. IEEE transactions on wireless communications, 2020, 19(6): 4038 – 4053. DOI: 10.1109/TWC.2020.2979446

[10] PENG H R, WANG L C, YE L G, et al. Long-lasting UAV-aided RIS communications based on SWIPT [C]//2022 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2022: 1844 – 1849. DOI: 10.1109/WCNC51071.2022.9771999

[11] HUANG C W, MO R H, YUEN C. Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning [J]. IEEE Journal on selected areas in communications, 2020, 38(8): 1839 – 1850. DOI:10.1109/JSAC.2020.3000835

[12] CESA-BIANCHI N, GENTILE C, LUGOSI G, et al. Boltzmann exploration done right [C]//The 31st International Conference on Neural Information Processing Systems. NIPS, 2017: 6287 – 6296

[13] PAN L, CAI Q P, HUANG L B. Softmax deep double deterministic policy gradients [C]//The 34th International Conference on Neural Information Processing Systems. NIPS, 2020: 11767 – 11777

[14] AL-HOURANI A, KANDEEPAN S, LARDNER S. Optimal LAP altitude for maximum coverage [J]. IEEE wireless communications letters, 2014, 3(6): 569 – 572. doi: 10.1109/LWC.2014.2342736

[15] WU Q Q, ZHANG S W, ZHENG B X, et al. Intelligent reflecting surface-aided wireless communications: A tutorial [J]. IEEE transactions on communications, 2021, 69(5): 3313 – 3351. DOI: 10.1109/TCOMM.2021.3051897

[16] WANG W H, YANG L, MENG A Q, et al. Resource allocation for IRS-aided JP-CoMP downlink cellular networks with underlaying D2D communications [J]. IEEE transactions on wireless communications, 2022, 21(6): 4295 – 4309. DOI: 10.1109/TWC.2021.3128711

[17] SHI Q J, RAZAVIYAYN M, LUO Z Q, et al. An iteratively weighted MMSE

▲Figure 7. Sum rate under different P_{max}

approach to distributed sum-utility maximization for a MIMO interfering broadcast channel [C]//IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2011: 3060 – 3063. DOI: 10.1109/ICASSP.2011.5946304

Biographies

YOU Qian received her BS degree from Yangzhou University, China in 2021. She is currently pursuing her MS degree at the school of electronics and information, Northwestern Polytechnical University, China. Her research interests include machine learning for communications, IRS-assisted communications, and UAV-assisted Communications.

XU Qian (qianxu@nwpu.edu.cn) received her BS and PhD degrees both from Xi'an Jiaotong University, China. She is currently an associate professor at the school of electronics and information, Northwestern Polytechnical University, China. Her current research interests include mobile wireless communications with emphasis on physical layer security and QoS provisioning. She has published more than 20 technical papers.

YANG Xin received his BS degree in communication engineering and MS degree in electronics and communication engineering from Xidian University, China in 2011 and 2014, respectively, and PhD degree in information and communication engineering from Northwestern Polytechnical University, China in 2018. He is an associate professor with school of electronics and information, Northwestern Polytechnical University. His research interests include wireless communications and ad hoc networks.

ZHANG Tao received his MS degree from Nankai University, China. He is currently a senior engineer in China Academy of Launch Vehicle Technology. His current research interests mainly focus on wireless communications.

CHEN Ming received his BS degree from Xidian University, China. He is currently a senior engineer in Hangzhou Hikvision Digital Technology Co., Ltd. His current research interests mainly focus on intelligent signal processing.

SST-V: A Scalable Semantic Transmission Framework for Video

LIU Chenyao[1], GUO Jiejie[2], ZHANG Yimeng[1],

XU Wenjun[1,3], LIU Yiming[1]

(1. State Key Laboratory of Network and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China；
 2. School of Artificial Intelligence, Beijing University of Posts and Tele-communications, Beijing 100876, China；
 3. Department of Mathematics and Theories, Peng Cheng Laboratory, Shenzhen 518066, China)

Abstract: The emerging new services in the sixth generation (6G) communication system impose increasingly stringent requirements and challenges on video transmission. Semantic communications are envisioned as a promising solution to these challenges. This paper provides a highly-efficient solution to video transmission by proposing a scalable semantic transmission algorithm, named scalable semantic transmission framework for video (SST-V), which jointly considers the semantic importance and channel conditions. Specifically, a semantic importance evaluation module is designed to extract more informative semantic features according to the estimated importance level, facilitating high-efficiency semantic coding. By further considering the channel condition, a cascaded learning based scalable joint semantic-channel coding algorithm is proposed, which autonomously adapts the semantic coding and channel coding strategies to the specific signal-to-noise ratio (SNR). Simulation results show that SST-V achieves better video reconstruction performance, while significantly reducing the transmission overhead.

Keywords: scalable coding; semantic communication; video transmission

1 Introduction

The wireless communication paradigm is envisioned to shift from connecting things to connecting intelligence, imposing new challenges on the developing sixth generation (6G) communication systems. On the one hand, new intelligent applications, such as the digital twin and the smart city, emerge with a surging number of terminals and explosively increasing data[1], bringing a great burden to existing communication systems. At the same time, to achieve real-time intelligent decision and control, communications are expected to be extremely low-delay and reliable. These challenges become more stringent when it comes to video data, which accounts for more than 80% of Internet traffic[2] and is further rising driven by the demand for ultra-high definition (HD) video. For example, a 1 080P HD video with 50 frames per second requires a bandwidth of 60 – 70 Mbit/s in the advanced H.265 format encoding. As a result, existing coding and transmission strategies, aiming at transmitting every bit, face the dual challenges of bandwidth and delay and are not capable enough of meeting the future demands of ultra-low delay and even real-time video transmission. It is urgent to develop a more efficient video compression and transmission paradigm.

In recent years, the semantic communication driven by artificial intelligence (AI), which is regarded as one of the potential technologies of 6G, has shown great potential due to its superior performance in data compression and transmission. WEAVER and SHANNON[3] divided the communication problems into three levels, namely the technical problem, semantic problem, and effectiveness problem, which corresponds to the definition of syntactic, semantic, and pragmatic in the theory of signs[4]. Based on the syntactic level of information, existing communication systems are developed, aiming at achieving complete and correct transmission of every symbol. Differently, semantic communication systems focus on the semantic level of information and aim at delivering the goal-related

parts of messages, which reduces the communication overhead to a great extent. Recently, semantic communications have been widely studied for the text[5], speech[6], and image[7], and demonstrated to be a powerful solution to high-efficiency compression and transmission.

For the video, existing source coding methods, like H.264, take the numerical difference between pixels as the evaluation index of the reconstruction quality, and all pixels are supposed to be transmitted completely. When researchers are coping with the phenomenon of lag due to the limited bandwidth, the clarity of the reconstructed video often gives place to fluency. However, for the human visual perception system, eyes tend to be attracted by some specific parts, e.g., the dynamic information in the foreground of the picture, and are not very sensitive to others like absolute errors between pixels of background. Inspired by this, the trade-off between visual experience and communication overhead can be better achieved in video compression and transmission by preserving the informative and significant parts, i.e., semantic information. Some researchers have investigated the effectiveness of semantic-enabled video compression and transmission. LU et al. proposed the first semantic compression framework for video, named Deep Video Compression (DVC)[8], which implements the video compression modules with convolutional neural networks (CNN) and optimizes them in an end-to-end manner. In Ref. [9], the effect of channel transmission is taken into consideration by jointly designing the semantic and channel coding for video transmission, and the coding rate of semantic information is determined by the entropy model, which is further expanded to frame-level control in Refs. [10] and [11]. Although these works provide efficient solutions to channel-aware semantic transmission, they ignore the change of transmission environments and are suboptimal under dynamic channel conditions.

To improve the efficiency in semantic communications of videos under dynamic channels, this paper proposes a scalable semantic transmission framework for video (SST-V), which achieves adaptive control of the coding rate towards dynamic channels. Specifically, a semantic importance estimation (SIE) module is proposed to evaluate the importance of different semantic features, where the semantic features of higher significance are given higher weights in the successive coding. To improve the efficiency and robustness of semantic transmission, we design a scalable multi-level joint semantic-channel (S-JSC) coding algorithm, where the coding rate of semantic feature is adaptively adjusted according to the corresponding importance level and the specific channel condition. In addition, a cascade-learning-based training strategy is applied for S-JSC, which greatly reduces the training and storage overhead.

The rest of the paper is as follows. Section 2 summarizes the existing video compression standard and semantic transmission methods for video. In Section 3, a basic framework for SST-V is proposed, including an SIE module and an S-JSC coding algorithm. Section 4 gives specific implementation details, simulation results, and performance analysis. Finally, Section 5 concludes the paper.

2 Overview of Video Transmission

In this section, we introduce the research status and progress of relevant fields, including existing video compression, semantic communication, and semantic transmission of video. Meanwhile, some private opinions are given about problems of current research and possible directions for improvement.

2.1 Existing Video Compression Coding Method

Relevant standards in the field of video coding are mainly formulated by two major organizations: the International Organization for Standardization/the International Electrotechnical Commission (IOS/IEC) and the International Telecommunication Union (ITU-T). The Moving Picture Experts Group (MPEG) of ISO/IEC has formulated the MPEG series of video coding standards for motion image compression. ITU-T has formulated the H.26x series of video coding standards, which is mainly used for low-bit-rate video telephony. There are also some standards that are formulated jointly by IOS/IEC and ITU-T, such as H.264/MPEG-4 part10 and H.265/HEVC.

Jaswant. R. JAIN and Anil. K. JAIN proposed a hybrid coding framework based on block motion compensation (MC) and transform coding such as discrete cosine transform (DCT) in Picture Coding Symposium (PCS) in 1979, which has become the main framework of almost all later video coding standards. The framework is also known as the MC/DCT hybrid coding. The core modules in MC/DCT hybrid coding include predictive coding and transform coding. In order to reduce the complexity of coding and make the operation of video coding easy to execute, each frame is divided into fixed-size blocks first, and then the blocks are compressed and encoded. In addition, there are quantization, entropy coding, and other modules. For example, the H.264 video coding framework is shown in Fig. 1.

In order to reduce the time redundancy among video frames, inter-frame prediction and motion compensation are usually used. MPEG series compression coding standards di-

▲Figure 1. H.264 coding framework[12]

vide different video frames into intra-frames (I-frame), predictive frames (P-frame), and bi-directional interpolated prediction frames (B-frame). For different types of frames, different compression ratios are used to achieve a balance between compression efficiency and video quality. The I-frame is the key frame. In the H.264 standard, video frames of fixed length are divided into a Group of Pictures (GOP) to prevent error propagation among reconstructed frames. The first frame in each GOP, as the key frame, is independently compressed by the Joint Photographic Experts Group (JPEG) or other image encoding methods to maximize the preservation of frame information. The previous I-frame or P-frame is used as a reference frame for the subsequent P-frame. Huffman encoding is performed on the motion vector, and a higher degree of compression is performed on the residual by using approximate JPEG encoding. The B-frame uses the front and back frames for bidirectional interpolation prediction, which has the highest degree of compression, but the decoding complexity and distortion are higher.

However, existing video compression methods have trouble dealing with increasing video data and the reasons are as follows. Firstly, existing standards take minimizing pixel error as the reconstruction goal and ignore the semantic information contained in the video. Secondly, they adopt fixed modular designs in which each module is independent of the others, such as DCT transform and entropy coding. As a result, they cannot obtain an overall performance gain. Benefiting from the development of deep learning, which has a strong nonlinear characterization ability, the latest evolution schemes of mainstream coding methods such as H.265, AVS2, and AVS3 have taken it into consideration to improve the coding performance.

2.2 Video Semantic Communication

The development of semantic information theory[3, 13 – 18] has supported the rapid growth of semantic communications in recent years. Studies on different modalities[19 – 21] have shown that joint source-channel coding can improve the overall performance of the system, and efficiently handle wireless channel fading and interference. Inspired by the joint design, joint semantic-channel coding schemes are widely adopted in semantic communication systems, which combines the semantic representation of the source with the link state of the physical layer.

2.2.1 Semantic Compression for Video

Video semantic compression and reconstruction can be divided into two categories: optimizing the existing video compression framework and extracting key information from videos to be compressed. The optimization of the existing compression framework mainly considers the method with a delay constraint, which means the reference frame is only from

previous frames. This makes it more suitable for actual application scenarios like streaming media. Specific modules in the existing video coding framework have been considered to be replaced by neural networks (NN). In Ref. [22], the existing video compression algorithm based on DCT is combined with video frame interpolation based on deep learning. According to the threshold of peak signal-to-noise ratio (PSNR), the encoded data can be selected to provide adjustable compression for residuals. In Ref. [23], four kinds of deformation of attention mechanisms are proposed, which respectively use the I-frame, motion vector, residual error, audio signal, and other mode information for video action recognition. Different types of information are processed in different ways and compared with each other. In 2019, LU et al. proposed an end-to-end deep video compression (DVC) model for the first time[8], which combines the optimization of video compression modules and uses CNN to optimize the network. The CNN network implements an encoder, a decoder, and a motion compensation network and uses a highly nonlinear transformation to represent residuals, which improves compression efficiency. Based on DVC, network modules such as feature prediction, loop filter, and discriminator[24 – 28] are added to further improve compression performance, making end-to-end video compression an important research trend. Multiple frames prediction for learned video compression (MLVC)[24] calculates relative motion using multiple previous frames, thus reducing coding residuals. A deep contextual video compression (DCVC) is proposed in Ref. [25], which uses the feature domain context as a condition and performs conditional coding instead of sub-optimal residuals. Considering the similarity of spatial dependencies, advanced learned video compression (ALVC)[28] predicts the current frame from previous frames without consuming any bits, further reducing coding overhead.

For the second category, key semantic features are extracted from original videos. There are relatively mature attempts for certain types of data sets. Based on the semantic segmentation technology, frames are divided into different semantic units, each of which has a specific spatial arrangement and visual characteristics, and is coded separately[29]. ZHANG et al. extract the semantic features of football video games from the elements of foreground, background, and the relationship between different objects[30]. The encoded sequences of
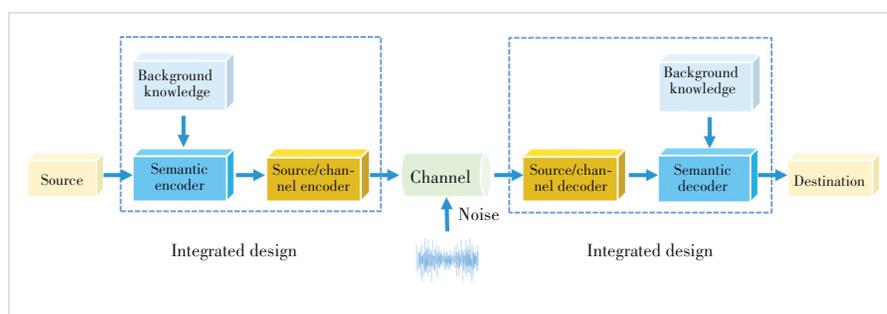
▲Figure 2. Semantic communication systems[17]

these features are decoded at the receiving end correspondingly, and then fused by a U-net network to generate a complete video. A semantic video conferencing (SVC)[31] network is proposed to extract key points of speakers to realize the semantic transmission of the video conference. CHEN et al. propose a framework for Interactive Face Video Coding (IFVC)[32] where each talking frame is expressed by highly-independent facial features such as mouth motion and eye blinking, achieving superior performance for face videos.

2.2.2 Semantic Transmission for Video

The above works consider video semantic compression schemes under the condition of sufficient bandwidth and ideal channels. Considering practical communication scenarios, new video semantic transmission schemes have been developed to facilitate joint optimization of semantic coding and channel coding. To achieve the balance between video quality and transmission delay in real-time video transmission scenarios, CUI et al.[33] use reinforcement learning (RL) to generate inferencing models based on playback and cache information when network throughput fluctuates. ELGAMAL et al.[34] manage to carry out targeted video coding according to specific downstream tasks in edge computing and cloud computing scenarios. They focus on capturing the target object in the picture. When the target object has a violent change (e.g., vehicle entering or leaving the picture), a new I-frame is selected, and only the I-frame is retrieved under specific tasks to reduce computing and transmission overhead. For surveillance videos[35], only the salient zones are encoded with high resolution, therefore the calculation ability of fog nodes can be reasonably allocated to obtain low delay while maintaining the video quality.

The joint design under specific scenarios considers schemes with a fixed rate, and there is still room for performance improvement for dynamic coding. Therefore, some works have studied variable-length semantic coding methods for the video to further improve the efficiency of semantic encoding. In Ref. [36], the resource allocated to different video frames is determined by the position in GOPs. According to the distance between the video frame and the key frame, a hierarchical learned video compression (HLVC) method is established with three hierarchical quality layers and a recurrent enhancement network. However, this allocation strategy treats the frame as the rate control unit and does not go deep into the level of semantic features, which ignores the importance of different semantic information and is difficult to further compress the content redundancy within frames. In Ref. [9], an entropy model is used to obtain the rate-adaptive

transmission strategy, which has several parallel autoencoders with a different number of output channel symbols. Although the proposed framework achieves adaptive control towards the importance of semantic features, which, however, is evaluated with a syntactic information entropy, leading to suboptimal rate control strategies for semantic coding. In Ref. [29], a semantic bit allocation model based on RL is proposed, which aims at improving the rate-semantic-sensing performance by encoding a certain semantic concept. Different features are put into the semantic decoder at different quantization levels according to the allocated resources and then reconstruct the image. These two works achieve a variable length encoding of semantic features, while the bit rates are changed through multiple parallel encoders, which greatly increases the complexity of neural network architecture. In Ref. [37], a rate allocation network is introduced to analyze the semantic information and anti-noise capability of the frame features. Features are coded and transmitted in a descending order of semantic importance according to the mask generated following the rate allocation network, and features with lower importance may be discarded to achieve video transmission of different bit rates. The above works mainly study variable-length coding schemes with different semantic features from the perspective of reconstruction. However, it is still an unsolved problem how to implement a flexible and scalable video semantic transmission scheme when the channel dynamically changes.

3 Proposed Framework of Scalable Semantic Transmission

In this section, we first present the basic framework of scalable video semantic transmission, and then introduce the proposed SIE module and S-JSC coding algorithm. The cascaded training strategy for the proposed system is finally presented.

3.1 Proposed Framework

The total framework of the proposed SST-V is shown in Fig. 3. The transmitter consists of a semantic feature extrac-
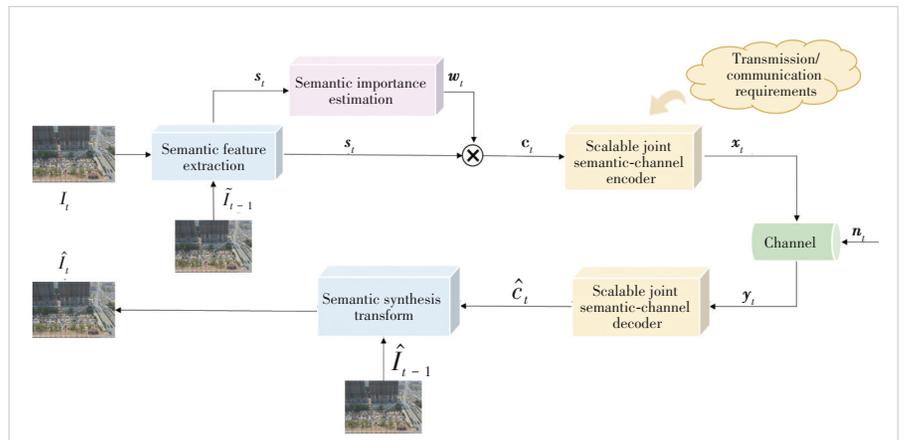
▲Figure 3. Framework of scalable semantic transmission for video

tion module $S(\cdot)$, a SIE module $A(\cdot)$ and a S-JSC encoder $E(\cdot)$. The receiver mainly consists of a S-JSC decoder $D(\cdot)$ and a semantic synthesis transform $T(\cdot)$. The video $I = \{I_1, I_2, \cdots, I_T\}$ is input into the semantic feature extraction module by frames, while the input image can be expressed as $I_t \in \mathbb{R}^{W \times H}$, where W and H are the width and height of the images, respectively. The semantic feature extraction module compares the semantic information changes between the current image I_t and the previous reference frame $\tilde{I}_{t-1}$ and calculates the semantic information required to reconstruct the video $s_t = S(I_t, I_{t-1}) \in \mathbb{R}^{N \times M_W \times M_H}$, where N is the number of the semantic features and $M_W \times M_H$ is the size of each feature map. The weight of different semantic information $w_t = A(s_t) \in \mathbb{R}^{N \times M_W \times M_H}$ is given by SIE. The semantic information to be transmitted can be expressed as $c_t = s_t \times w_t$. The S-JSC encoder encodes c_t as $x_t = E(c_t) \in C^l$, where l is the number of the transmission symbols, the value of which is chosen among the set $L = \{l_1, l_2, \cdots, l_C\}$ according to the channel state or different communication demands.

In this paper, we consider both the additive white Gaussian noise (AWGN) channel and the Rayleigh fading channel. As for the AWGN channel, a received sequence can be expressed as $y_t = x_t + n_t \in C^l$. The noise vector n_t consists of independent and equally distributed cyclic symmetric complex Gaussian random variables n_i which follows $n_i \sim \mathcal{CN}(0, \sigma^2)$, $i = 1, \cdots, l$, and σ^2 is the average noise power. For Rayleigh fading channels, the single point fading model is considered in this paper and all transmission symbols experience the same channel response. Clarke Model[38] shows that a flat fading channel is composed of several multipath signals under a rich-scattering electromagnetic environment. According to the central limit theorem, both the I-path and Q-path of the channel response can be approximated as Gaussian random processes when the number of paths is large enough. Similar to Refs. [5] and [7], the received sequence can be expressed as $y_t = h_t x_t + n_t \in C^l$, where $h_t \sim \mathcal{CN}(0, \sigma_t^2)$ is a random variable satisfying a cyclic symmetric complex Gaussian distribution.

At the receiver side, the received sequence is decoded as $\hat{c}_t = D(y_t) \in \mathbb{R}^{N \times M_W \times M_H}$ by the S-JSC decoder, which selects different decoder structures according to the number of symbols received. Finally, the decoded sequence of semantic information is transmitted to the semantic synthesis transform module to reconstruct the original video frame $\hat{I}_t \in \mathbb{R}^{W \times H} = T(c_t)$.

3.2 Semantic Feature Extraction

The framework of the semantic feature extraction module $S(\cdot)$ is shown in Fig. 4, which is referred to the end-to-end
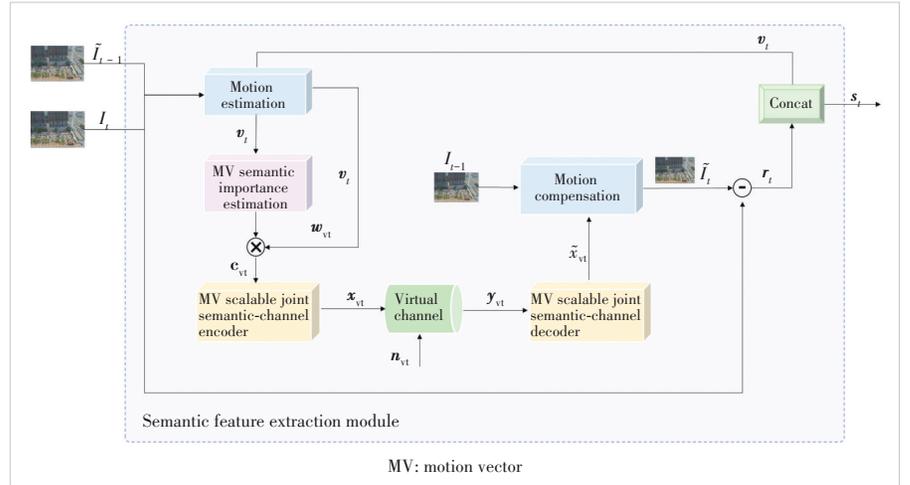
▲Figure 4. Framework of semantic feature extraction module

MV: motion vector

video compression structure in Ref. [8]. The extracted semantic feature s_t is mainly obtained from the motion vector (MV) v_t and the residual vector (RES) r_t.

The current frame I_t and the reference frame $\tilde{I}_{t-1}$ are first fed into the motion estimation module to get the motion vector v_t[39]. A semantic feature map of the motion vector with new distribution is obtained in MV SIE and further encoded in an MV S-JSC encoder. In particular, the semantic feature map needs to pass through a virtual channel at the transmitter and then the decoding process is simulated. The decoded semantic sequence $\tilde{x}_{vt}$ is considered to be approximately consistent with the semantic sequence obtained by the receiver. Therefore, the semantic synthesis transform module can predict the received frame from $\tilde{x}_{vt}$ and I_{t-1}. By comparing the error between the simulated reconstructed frame $\tilde{I}_t$ and the original frame I_t, the residual error vector r_t of the current motion vector can be calculated. The encoded motion vector and residual vector are spliced together to get the final sequence $s_t = \{v_t, r_t\}$. Decoded semantic sequence $\hat{s}_t$ after joint semantic channel decoder can be divided into decoded motion vector $\hat{v}_t$ and residual error $\hat{r}_t$ correspondingly. The semantic synthesis firstly carries out motion estimation based on the previous reference frame $\hat{I}_{t-1}$ and $\hat{v}_t$ to generate frame I_{MC-t}, followed by correction using $\hat{r}_t$ to obtain the final reconstructed frame $\hat{I}_t$.

3.3 Semantic Importance Estimation

For video frames, semantic feature vectors indicate what is present in each frame and what has changed compared with the previous frame. However, it is worth noting that these semantic feature vectors have different importance. In the case of street surveillance video, we care more about the movement of cars and pedestrians that dominate the video than the swaying leaves in the wind.

Therefore, we design the SIE module $A(\cdot)$ based on squeeze-and-excitation networks (SE-Net) shown in Fig. 5. Following the semantic features extraction, SIE is firstly used to compre-
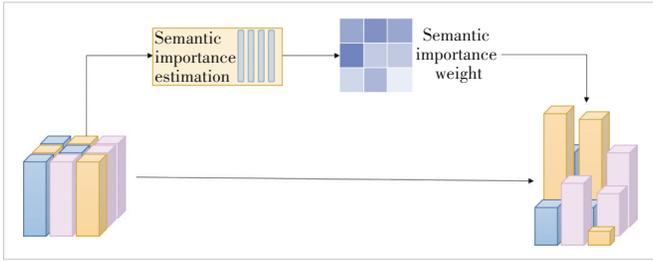
▲Figure 5. Semantic importance estimation (SIE) module structure

hensively analyze the relationship between different feature maps to estimate the importance degree of different features, and provide different weights for each feature. Since the semantic information s_t consists of two parts, SIE estimates the semantic importance of v_t and r_t respectively, and the output is $w_t = \{ w_{vt}, w_{rt} \}$. Then, the extracted semantic feature is multiplied by the weights to produce a new feature map. On the one hand, more power can be allocated to important information during an actual communication process to reduce the effect of noise. On the other hand, important semantic information needs more strict protection by the S-JSC coding algorithm explained below. When the channel conditions are severe with a low SNR, the correct transmission of important semantic features can be guaranteed with the same number of channel symbols to realize the reconstruction of basic semantic information despite the interference of noise.

3.4 Scalable Joint Semantic-Channel Coding Algorithm

Joint semantic-channel coding is supposed to be used for end-to-end overall optimization, which further enhances the accuracy of the semantic reconstruction of transmitted videos and protects the semantic information obtained in Sections 3.2 and 3.3. In particular, existing video semantic coding methods based on deep learning cannot adjust the code rate flexibly. To solve this problem, the S-JSC coding algorithm is designed, which can adjust the code rate adaptively according to the actual transmission requirements.

According to the training strategy of cascade learning[40], several different source channel coding rates are designed. With the increase in coding level, the output dimension of the S-JSC coding algorithm is continuously reduced while the compression ratio is continuously improved. Higher-level algorithms with fewer symbols manage to maintain the maximum transmission quality within limited resources. Instead of indiscriminately compressing the encoding output of the upper level, semantic features of different importance obtained from SIE are protected with different degrees. The redundancy of semantic information with less importance may be decreased greatly to realize reliable transmission of the most important information, therefore achieving more efficient video semantic transmission.

The coding level is used as the control parameter and input into the S-JSC encoder/decoder together with the information to be encoded. According to the coding level, the scalable au-

toencoder layer specifies the neural network architecture to change the dimensions of output. The training and storage overhead is greatly reduced with multilevel coding algorithms stored in a serial structure, improving the deployment efficiency of the model.

3.5 Training Strategy

During the training process, it is of limited significance for the subsequent layers to participate in the training before the semantic feature extraction module becomes basically stable, due to the relatively complex network architecture and the large correlation between the front and back layers. Therefore, the semantic feature extraction module with relatively stable performance can be obtained by separating the training first. After it is basically stable, SIE and levels of the S-JSC coding algorithm are added successively. Due to the random distribution of parameters of the newly added layer, the semantic feature extraction module is frozen, which makes the subsequent structure converge quickly. When lower levels of S-JSC are trained, there are epochs where all preceding components are updated to achieve an end-to-end gain.

With the gradual addition of the new JSC coding layer, the parameters of the trained S-JSC coding layers and SIE will also be frozen to ensure that the output of the lower coding level is not damaged as much as possible during the training of the higher coding level. The update of parameters will not be carried out in the back propagation, and only the new coding layers will be trained. SIE and the low-level S-JSC encoder layers get the overall gain through the joint training, which improves the anti-interference ability from the perspective of semantic information protection. The high-level S-JSC coding layers mainly further reduce the transmission symbol error rate from the perspective of transmission under insufficient channel carrying capacity.

4 Experiments and Results

In this section, we present the details of training, including data set processing, optimization objectives, and evaluation indicators. Simulation results are analyzed, which proves the ef-
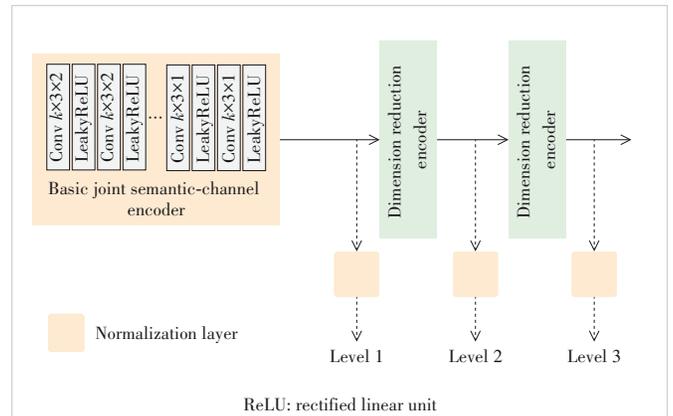
▲Figure 6. Scalable multi-level joint semantic-channel (S-JSC) coding architecture based on cascade learning

fectiveness of the proposed framework.

4.1 Datasets

The proposed system is trained with the Vimeo-90k dataset[41], which is one of the most commonly used video datasets built for evaluating different video processing tasks. The complete dataset is about 82 G in size, and according to Ref. [8], it would take about 7 days to train a single similar semantic feature extraction module using two Titan X GPUs. The dataset contains 89 800 independent short films, each of which consists of 7 frames. During each epoch, according to the time relativity, the relative motion for each frame with respect to its reference frame is considered to be approximate and the coding strategy is therefore similar among the 7 frames of each video clip. Only one frame of each video is selected randomly for our training, together with the previous frame as references. Compared with training each group of the frame and its reference frame, the volume of training details decreases to 1/6 of the original volume, which greatly reduces the time cost while the performance in subsequent validation is basically unchanged.

4.2 Optimization Objective and Metrics

The optimization objective of SST-V is to improve the video reconstruction performance with fewer channel symbols. The entropy of preliminary semantic information s_t from the semantic feature extraction module is supposed to decrease and therefore reduce the difficulty of subsequent module coding. Therefore, the rate-distortion (RD) function is adopted as the loss function, i.e.,

$$\text{loss} = \lambda l_D + l_R = \lambda d(x_t, \hat{x}_t) + (H(v_t) + H(r_t)), \tag{1}$$

where λ is the Lagrange multiplier that represents the tradeoff between bit overhead and video distortion, l_R is the coding bit rate of the semantic feature extraction module, represented by entropies of the moving vector $\boldsymbol{v}_t$ and the residual vector $\boldsymbol{r}_t$, and l_D is the distortion constraint of video reconstruction quality that consists of mean square error (MSE) of the original video frame. Since the reconstruction process includes both motion compensation using motion vector and correction using residual vector, we need to minimize the errors after motion compensation l_{mc} and the overall distortion after reconstruction l_{re}, i.e.,

$$l_d = l_{mc} + l_{re} = $$
$$w * \text{MSE}(I_{MC-t}, I_t) + \text{MSE}(I_t, \hat{I}_t), \tag{2}$$

where w is weight of the distortion of motion-vector-based reconstruction that decreases with the training process.

PSNR and multi-scale structural similarity index (MS-SSIM) are used to measure the distortion degree of reconstructed frames. PSNR is calculated as:

$$\text{PSNR} = 10 \log_{10}(\text{MAX}^2/\text{MSE}), \tag{3}$$

where MSE is the mean square error of the reconstructed image and original image, and MAX represents the maximum pixel value possible for a frame and is set as 1 during the experiment. Compared with PSNR, MS-SSIM is closer to the real perception of human eyes that ranges from 0 to 1, where a higher value indicates lower distortion. It considers that visual distortion is composed of brightness, contrast and structure, and the influence of the distance from the viewer to the image and the density of pixel information on subjective visual experience are further considered. The detailed calculation of MS-SSIM can be found in Refs. [42] and [43].

Similar to k/n in Ref. [21], channel symbols per pixel (CPP) is defined to measure the coding rate of the system. For a fixed resolution $W \times H$, the number of channel input symbols is R, and then CPP is calculated as:

$$\text{CPP} = R/(W \times H). \tag{4}$$

4.3 Simulation Results

We set up three coding levels, namely Level 1, Level 2, and Level 3, and the value of CPP is 0.110, 0.055, and 0.014, respectively. The ideal channel environment is first considered, where the channel capacity is assumed to always be enough to serve the transmission of all the channel input symbols. Fig. 7 shows the reconstruction performance of the schemes with and without SIE under different coding levels. "SIE-L1" indicates the scheme at Level 1 with SIE, while "L1" means the scheme without SIE.

Since most values of MS-SSIM are distributed densely, we will use both raw values and the form of n dB for better visual effect, which is calculated by

$$\text{MS} - \text{SSIM(dB)} = -10\log_{10}(1 - \text{MS} - \text{SSIM}). \tag{5}$$

Under different channel conditions and metrics, the two curves at the top are both the results of schemes with SIE, which achieve higher performance gain under the metric of MS-SSIM. Note that benefitting from SIE, the scheme at Level 2 performs even better than the non-SIE scheme at Level 1 in most cases. Therefore, SIE helps greatly in extracting important semantic information from the source video, reducing the communication cost significantly. The advantage of SIE is also shown in Fig. 8, which gives several visual results of the schemes with or without SIE at 0 dB with Level 3 under the ideal channel.

However, in the practical communication system, the channel capacity is not always sufficient and deteriorates severely under bad channel conditions. Hence, the proposed schemes are also evaluated under capacity-limited cases. During the test process, we make SNR vary uniformly from 0 dB to 10 dB. The average PSNR and MS-SSIM of different schemes are shown in Table 1, where schemes 1 and 2 perform better than schemes 3 and 4 under non-ideal channels, respectively, which further demonstrates the effectiveness of the proposed
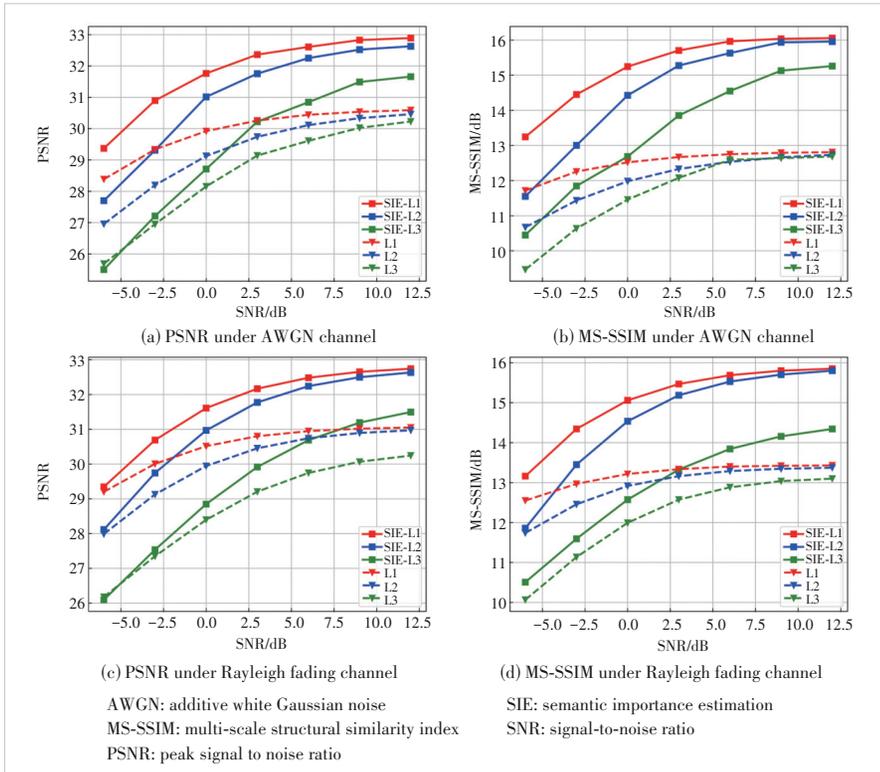
▲Figure 7. Reconstruction performance of the schemes with or without SIE at different coding levels

▲ Figure 8. Examples of the reconstructed frames of the schemes with or without semantic importance estimation (SIE)

SIE. Scheme 3 performs worse than scheme 4 even if the former uses more symbols for transmission. The reason is that all the semantic information is transmitted without considering the variation of channel capacity and the important information is deteriorated with higher probability, exposing the limitation of the fixed level coding scheme.

For the proposed scalable multilevel coding scheme, the coding level increases for lower SNR. Level 3, Level 2 and Level 1 are selected respectively for channel conditions ranging in $[0,3)$, $[3,6)$ and $[6,10]$. The number of three levels is assumed to be the maximum number of symbols that can be accurately transmitted under the corresponding channel conditions. For the proposed SST-V, each frame can automatically

switch to a different coding model according to the SNR, while the non-scalable baselines (schemes 1, 2, 3 and 4) use a fixed encoding level at all SNRs.

As shown in Table 1, compared with schemes 3 and 4, the proposed SST-V with both SIE and S-JSC, i.e., scheme 6, achieves performance gains of 2.3 dB and 4.6 dB in terms of PSNR and MS-SSIM, respectively. It shows that when the channel capacity is limited, the SST-V can adapt to the dynamic channel environment, significantly improving the transmission efficiency. Note that schemes 5 and 6 further evaluate the effectiveness of SIE under the scalable multilevel coding rate. With the proposed SIE, scheme 6 achieves gains of 1.3 dB in PSNR and 2.7 dB in MS-SSIM. It proves that SIE helps the SST-V focus on the semantic information of higher importance, and hence improves the reconstruction performance under practical dynamic channels.

▼Table 1. PSNR and MS-SSIM of different schemes

Number	Scheme	PSNR	MS−SSIM/dB
1	Without SIE, fixed Level 1 (L1)	23.937 6	6.704 7
2	Without SIE, fixed Level 2 (L2)	27.307 8	8.743 6
3	With SIE, fixed Level 1 (SIE-L1)	26.099 3	7.454 1
4	With SIE, fixed Level 2 (SIE-L2)	28.900 34	9.754 9
5	Scalable multilevel coding without SIE	29.935 9	11.682 3
6 (SST-V)	Scalable multilevel coding with SIE	31.190 78	14.349 9

MS-SSIM：multi-scale structural similarity index
PSNR：peak signal-to-noise ratio
SIE：semantic importance estimation
SST-V：scalable semantic transmission framework for video

5 Conclusions

In this paper, we discuss the video transmission problem in the future 6G mobile communication scenarios and review the existing video coding and semantic-based video coding transmission methods. To achieve efficient and robust video transmission under dynamic channel conditions, this paper proposes a scalable semantic transmission framework for video, namely SST-V. Besides semantic information extraction, SST-V estimates the importance of different semantic features with the proposed SIE, and obtains a more compact and robust rep-

resentation of semantic information. An S-JSC coding algorithm based on cascading learning is designed, where the coding rate can be adjusted adaptively according to dynamic channel states. The simulation results show that SST-V has better video reconstruction performance in terms of PSNR and MS-SSIM compared with the baseline schemes, and provides a more efficient solution to video transmission under bandwidth constraints.

References

[1] TU Y, CHEN W. A deep learning-based semantic communication system [J]. Mobile communications, 2021, 45(4): 91-94. DOI: 10.3969/j. issn. 1006-1010.2021.04.015

[2] CISCO. 2020 global networking trends report [EB/OL]. (2019-11-17) [2023-04-01]. https://www. cisco. com/c/dam/m/en_us/solutions/enterprise-networks/networking-report/files/GLBL-ENG_NB-06_0_NA_RPT_PDF_MOFU-no-NetworkingTrendsReport-NB_rpten018612_5.pdf

[3] WARREN W, SHANNON C E. Recent contributions to the mathematical theory of communication [EB/OL]. [2023-02-01]. http://www. sietmanagement. fr/wp-content/uploads/2016/04/Weaver1949.pdf

[4] MORRIS C W. Foundations of the theory of signs [M]. Chicago, USA: The University of Chicago Press, 1938

[5] XIE H Q, QIN Z J, LI G Y, et al. Deep learning enabled semantic communication systems [J]. IEEE transactions on signal processing, 2021, 69: 2663 – 2675. DOI: 10.1109/TSP.2021.3071210

[6] WEI H, XU W J, WANG F Y, et al. SemAudio: semantic-aware streaming communications for real-time audio transmission [C]//IEEE Global Communications Conference. IEEE, 2022: 3965 – 3970. DOI: 10.1109/GLOBECOM48099.2022.10001043

[7] XU W J, ZHANG Y M, WANG F Y, et al. Semantic communication for the Internet of vehicles: a multiuser cooperative approach [J]. IEEE vehicular technology magazine, 2023, 18(1): 100 – 109. DOI: 10.1109/MVT.2022.3227723

[8] LU G, OUYANG W L, XU D, et al. DVC: an end-to-end deep video compression framework [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020: 10998 – 11007. DOI: 10.1109/CVPR.2019.01126

[9] WANG S X, DAI J C, LIANG Z J, et al. Wireless deep video semantic transmission [J]. IEEE journal on selected areas in communications, 2023, 41(1): 214 – 229. DOI: 10.1109/JSAC.2022.3221977

[10] TUNG T Y, GÜNDÜZ D. DeepWiVe: deep-learning-aided wireless video transmission [J]. IEEE journal on selected areas in communications, 2022, 40(9): 2570 – 2583. DOI: 10.1109/JSAC.2022.3191354

[11] HUANG B W, YAN X, ZHOU J J, et al. CSMCNet: scalable video compressive sensing reconstruction with interpretable motion estimation [EB/OL]. (2021-08-03) [2023-02-01]. arXiv: 2108.01522. https://arxiv.org/abs/2108.01522

[12] WIEGAND T, SULLIVAN G J, BJONTEGAARD G, et al. Overview of the H.264/AVC video coding standard [J]. IEEE transactions on circuits and systems for video technology, 2003, 13(7): 560 – 576. DOI: 10.1109/TCSVT.2003.815165

[13] CARNAP R, BAR-HILLEL Y. An outline of a theory of semantic information [EB/OL]. [2023-02-01]. https://courses.cs.tau.ac.il/0368-4341/shared/Papers/CARNAP-HILLEL.pdf

[14] BAR-HILLEL Y, CARNAP R. Semantic information [J]. The British journal for the philosophy of science, 1953, 4(14): 147 – 157. DOI: 10.1093/bjps/iv.14.147

[15] FLORIDI L. Outline of a theory of strongly semantic information [J]. Minds and machines, 2004, 14(2): 197 – 221. DOI: 10.1023/B: MIND.0000021684.50925.c9

[16] KOLCHINSKY A, WOLPERT D H. Semantic information, autonomous agency and non-equilibrium statistical physics [J]. Interface focus, 2018, 8(6): 20180041. DOI: 10.1098/rsfs.2018.0041

[17] ZHANG P, XU W J, GAO H, et al. Toward wisdom-evolutionary and primitive-concise 6G: a new paradigm of semantic communication networks [J]. Engineering, 2022, 8: 60 – 73. DOI: 10.1016/j.eng.2021.11.003

[18] ZHONG Y X. A theory of semantic information [J]. China communications, 2017, 14(1): 1 – 17. DOI: 10.1109/CC.2017.7839754

[19] RAO M, FARSAD N, GOLDSMITH A. Variable length joint source-channel coding of text using deep neural networks [C]//IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2018: 1 – 5. DOI: 10.1109/SPAWC.2018.8445924

[20] BOURTSOULATZE E, BURTH KURKA D, GÜNDÜZ D. Deep joint source-channel coding for wireless image transmission [J]. IEEE transactions on cognitive communications and networking, 2019, 5(3): 567 – 579. DOI: 10.1109/TCCN.2019.2919300

[21] KURKA D B, GÜNDÜZ D. DeepJSCC-f: deep joint source-channel coding of images with feedback [J]. IEEE journal on selected areas in information theory, 2020, 1(1): 178 – 193. DOI: 10.1109/JSAIT.2020.2987203

[22] JALALPOUR Y, WANG L Y, FENG W C, et al. FID: frame interpolation and DCT-based video compression [C]//IEEE International Symposium on Multimedia (ISM). IEEE, 2021: 218 – 221. DOI: 10.1109/ISM.2020.00045

[23] CHEN J W, HO C M. MM-ViT: multi-modal video transformer for compressed video action recognition [C]//IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). IEEE, 2022: 786 – 797. DOI: 10.1109/WACV51458.2022.00086

[24] LIN J P, LIU D, LI H Q, et al. M-LVC: multiple frames prediction for learned video compression [C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020: 3543 – 3551. DOI: 10.1109/CVPR42600.2020.00360

[25] LI J, LI B, LU Y. Deep contextual video compression [EB/OL]. [2023-04-01]. https://arxiv.org/pdf/2109.15047.pdf

[26] LIU C, SUN H M, ZENG X Y, et al. Learned video compression with residual prediction and feature-aided loop filter [C]//IEEE International Conference on Image Processing (ICIP). IEEE, 2022: 1321 – 1325. DOI: 10.1109/ICIP46576.2022.9897989

[27] ZHANG S P, MRAK M, HERRANZ L, et al. DVC-P: deep video compression with perceptual optimizations [C]//Proceedings of 2021 International Conference on Visual Communications and Image Processing (VCIP). IEEE, 2022: 1 – 5. DOI: 10.1109/VCIP53242.2021.9675350

[28] YANG R, TIMOFTE R, VAN GOOL L. Advancing learned video compression with In-loop frame prediction [J]. IEEE transactions on circuits and systems for video technology, 2023, 33(5): 2410 – 2423. DOI: 10.1109/TCSVT.2022.3222418

[29] HUANG D L, GAO F F, TAO X M, et al. Toward semantic communications: deep learning-based image semantic coding [J]. IEEE journal on selected areas in communications, 2023, 41(1): 55 – 71. DOI: 10.1109/JSAC.2022.3221999

[30] DUAN Y P, LI M Z, WEN L J, et al. From object-attribute-relation semantic representation to video generation: a multiple variational autoencoder approach [C]//Proceedings of 2022 IEEE 32nd International Workshop on Machine Learning for Signal Processing (MLSP). IEEE, 2022: 1 – 6. DOI: 10.1109/MLSP55214.2022.9943394

[31] JIANG P W, WEN C K, JIN S, et al. Wireless semantic communications for video conferencing [J]. IEEE journal on selected areas in communications, 2023, 41(1): 230 – 244. DOI: 10.1109/JSAC.2022.3221968

[32] CHEN B, WANG Z, LI B, et al. Interactive face video coding: a generative compression framework [EB/OL]. [2023-02-20]. https://arxiv.org/abs/2302.09919

[33] CUI L Z, SU D Y, YANG S, et al. TCLiVi: transmission control in live video streaming based on deep reinforcement learning [J]. IEEE transactions on Multimedia, 2020, 23: 651-663. DOI: 10.1109/TMM.2020.2985631

[34] ELGAMAL T, SHI S, GUPTA V, et al. SiEVE: semantically encoded video analytics on edge and cloud [C]//Proceedings of 2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2021: 1383 – 1388. DOI: 10.1109/ICDCS47774.2020.00182

[35] WANG Y Q, XU J C, JI W. A feature-based video transmission framework for visual IoT in fog computing systems [C]//Proceedings of 2019 ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS). IEEE, 2019: 1 – 8. DOI: 10.1109/ANCS.2019.8901872

[36] YANG R, MENTZER F, VAN GOOL L, et al. Learning for video compression with hierarchical quality and recurrent enhancement [C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

IEEE, 2020: 6627 – 6636. DOI: 10.1109/CVPR42600.2020.00666

[37] ZHANG B, QIN Z, LI Y. Semantic communications with variable-length coding for extended reality [EB/OL]. [2023-03-11]. https://arxiv.org/abs/2302.08645.

[38] RAPPAPORT T S. Wireless communications: principles and practice [M]. Upper Saddle River, USA: Prentice Hall PTR, 1996

[39] RANJAN A, BLACK M J. Optical flow estimation using a spatial pyramid network [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 2720 – 2729. DOI: 10.1109/CVPR.2017.291

[40] MARQUEZ E S, HARE J S, NIRANJAN M. Deep cascade learning [J]. IEEE transactions on neural networks and learning systems, 2018, 29(11): 5475 – 5485. DOI: 10.1109/TNNLS.2018.2805098

[41] XUE T F, CHEN B A, WU J J, et al. Video enhancement with task-oriented flow [J]. International journal of computer vision, 2019, 127(8): 1106 – 1125. DOI: 10.1007/s11263-018-01144-2

[42] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE transactions on image processing, 2004, 13(4): 600 – 612. DOI: 10.1109/TIP.2003.819861

[43] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment [C]//The 37th Asilomar Conference on Signals, Systems & Computers. IEEE, 2004: 1398 – 1402. DOI: 10.1109/ACSSC.2003.1292216

Biographies

LIU Chenyao received her BE degree from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunication (BUPT), China in 2022. She is currently pursuing her PhD degree at the School of Artificial Intelligence, BUPT. Her research interests include semantic communication, video coding, and machine learning.

GUO Jiejie is currently pursuing her BE degree from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, China. Her research interests include semantic communication, video coding, and artificial intelligence.

ZHANG Yimeng received her BE degree from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunication (BUPT), China in 2018. She is currently pursuing her PhD degree at the School of Artificial Intelligence, BUPT. Her research interests include semantic communication and intelligent resource allocation in emerging wireless applications. She is a graduate student Member of IEEE.

XU Wenjun (wjxu@bupt.edu.cn) is a professor with the State Key Laboratory of Network and Switching Technology, Beijing University of Posts and Telecommunications, China, and with Peng Cheng Laboratory, China. He received his PhD degree from Beijing University of Posts and Telecommunications in 2008. His research interests include artificial intelligence-driven networks, semantic communications, unmanned aerial vehicle communications and networks, and green communications and networking. He is an editor of *China Communications* and a senior member of IEEE.

LIU Yiming received her BE degree in communication engineering from Shanghai University, China in 2014, and PhD degree in information and communication engineering from Beijing University of Posts and Telecommunications (BUPT), China in 2019. She was a visiting PhD student with The University of British Columbia, Canada in 2017 and 2018. She is currently an associate researcher with the School of Information and Communication Engineering, BUPT. Her research interests include next-generation wireless networks, semantic communication, edge intelligence, blockchain, and the distributed ledger technology.

UAV Autonomous Navigation for Wireless Powered Data Collection with Onboard Deep Q-Network

LI Yuting[1], DING Yi[2], GAO Jiangchuan[1], LIU Yusha[1], HU Jie[1], YANG Kun[3]

(1. School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China；
2. China Mobile Communications Group Jilin Co., Ltd., Changchun 130061, China；
3. School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, United Kingdom)

Abstract: In a rechargeable wireless sensor network, utilizing the unmanned aerial vehicle (UAV) as a mobile base station (BS) to charge sensors and collect data effectively prolongs the network's lifetime. In this paper, we jointly optimize the UAV's flight trajectory and the sensor selection and operation modes to maximize the average data traffic of all sensors within a wireless sensor network (WSN) during finite UAV's flight time, while ensuring the energy required for each sensor by wireless power transfer (WPT). We consider a practical scenario, where the UAV has no prior knowledge of sensor locations. The UAV performs autonomous navigation based on the status information obtained within the coverage area, which is modeled as a Markov decision process (MDP). The deep Q-network (DQN) is employed to execute the navigation based on the UAV position, the battery level state, channel conditions and current data traffic of sensors within the UAV's coverage area. Our simulation results demonstrate that the DQN algorithm significantly improves the network performance in terms of the average data traffic and trajectory design.

Keywords: unmanned aerial vehicle; wireless power transfer; deep Q-network; autonomous navigation

1 Introduction

Wireless sensor networks (WSNs) have been widely used in various scenarios, like environment monitoring[1]. However, the energy of the sensors in WSN is usually limited and recharging sensors is challenging[2]. When a WSN is deployed in remote areas, it is not realistic for traditional terrestrial communication networks to charge sensors. In this situation, unmanned aerial vehicles (UAVs) can be used to charge ground sensors and complete tasks such as traffic monitoring, autonomous driving complement, flying relay and data collection[3-6]. In addition, various types of natural disasters, such as earthquakes, wildfires, hurricanes, etc., have caused serious damage to communication infrastructure. UAVs as mobile stations can help quickly establish emergency communication and maintain real-time communication to obtain post-disaster situational awareness, which can significantly improve the efficiency of rescue missions.

The UAV is used as a mobile access point (AP) to charge a set of sensors via wireless power transfer (WPT) in the downlink, and the ground sensors leverage the harvested energy to transmit data back to the UAV via wireless information transfer (WIT) in the uplink[7]. However, the resource allocation problem in this scenario is non-convex and difficult to solve directly. Therefore, we formulate the problem as a Markov decision process (MDP), which will be optimally solved with a deep reinforcement learning (DRL) approach[8].

The deep Q-network (DQN) framework has been widely applied in UAV-assisted wireless communication systems. In Ref. [9], the authors investigated UAV-assisted WPT and data collection and employed the DQN to optimize the UAV's instantaneous patrolling velocity as well as plan the flight trajectory, in order to minimize the packet loss. TANG et al.[10] designed a DRL strategy for maximizing the minimum throughput, where the sparse reward was used to ensure that the UAV could complete the optimization task. In Ref. [11], the authors

proposed to dynamically adjust the flying trajectory of the UAV based on the changes of point of interest (PoI) in the coverage range of the UAV, in order to cover as many PoIs as possible, and to improve the fairness of ground users. In Ref. [12], the authors investigated UAV-aided mobile networks, where multiple ground mobile users (GMUs) desired to upload data to a UAV, and maximized the uplink throughput by optimizing the UAV's trajectory. ABEDIN et al.[13] designed a navigation policy for multi-UAVs to improve the data freshness and connectivity to the Internet of Things (IoT) devices, which incorporated different contextual information such as energy and age of information (AoI) constraints. In Ref. [14], the authors investigated a UAV-based emergency communication network, in which UAVs could collect information from ground users in post-disaster scenarios, and transformed the problem into a constrained Markov decision-making process (CMDP). LI et al.[15] formulated a joint optimization of flight cruise control and data collection schedule to minimize network data loss as a partially observable Markov decision process (POMDP), where the states of individual IoT nodes could be obscure to the UAV.

However, the above-mentioned works are all based on the condition that the UAV has partial or full prior knowledge of the environment or waypoints. Investigation into UAVs' navigation with no prior knowledge of sensors' location is still a blank in the literature. Therefore, motivated to fill this gap, we study the UAV navigation problem under the assumption of no prior knowledge about sensor positions on the UAV side. To complete the autonomous navigation task, which is difficult to solve by convex optimization, we propose a novel DRL-based framework to optimize the UAV's trajectory as well as the ground sensors selection with the objective of maximizing the average data traffic of all sensors in a UAV-assisted WSN. The problem is formulated as MDP with a large state and action space. To obtain the up-to-date knowledge about the state information of ground sensors, DRL is used for the UAV to autonomously navigate to the next position. Numerical results show that the proposed algorithm significantly improves the network performance while ensuring the UAV trajectory is optimized.

The rest of this paper is organized as follows. In Section 2, we describe the system model and problem formulation. The DRL-aided algorithm for UAV-assisted networks is presented in Section 3. Section 4 shows the simulation results of the proposed algorithms. Finally, the conclusion of this paper is presented in Section 5.

2 System Model and Problem Formulation

2.1 System Model

We consider a single-UAV-assisted WSN consisting of K sensors shown in Fig. 1, where the UAV is responsible for charging and collecting data, and the ground sensors harvest

energy from the UAV in the downlink and then send collected data in the uplink. Without loss of generality, we assume that the flying height of the UAV is fixed at H m. The UAV has a coverage range of R m^2, i.e., only sensors with a distance of less than R from the UAV can communicate with the UAV. The UAV starts from the origin and flies back to it within the specified flight time. The notations used in this paper are summarized in Table 1.

In actual scenarios, the UAV has no prior knowledge of the location and status of sensors. Therefore, how the UAV performs autonomous navigation also becomes a problem to be solved in our paper. Given the flight height, the UAV coverage area R can be expressed as:

$$R = \pi \left(h \tan \frac{\phi}{2} \right)^2, \tag{1}$$

where ϕ denotes the antenna beam width, and only the sensors in the coverage area can communicate with the UAV.

At any time slot, the UAV has three operation modes: the uplink data collection (DC) mode, the downlink WPT mode and the state listening mode. We employ $\rho_k(t) \in (0,1)$, $\forall k \in \mathcal{K}$ to denote the UAV operational mode selection at the

UAV: unmanned aerial vehicle
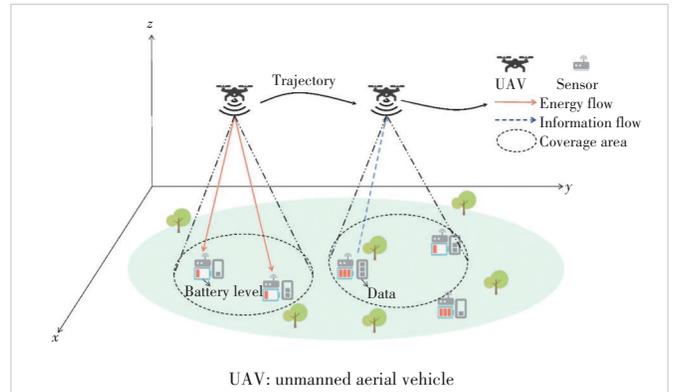
▲Figure 1. A UAV-assisted rechargeable wireless sensor network (WSN)

▼Table 1. Notation list

Notation	Definition
k, K	Sensor index, number of ground sensors
t, T	Time slot index, total UAV flight time slots
τ, T	Time slot duration, total UAV flight time
$\boldsymbol{w}_k$	Coordinate of the k-th sensor
$v_{\mathrm{UAV}}(t)$	Velocity of the UAV at the t-th time slot
$\boldsymbol{q}(t)$	Position of the UAV at the t-th time slot
$d_k(t)$	Distance between sensor k and the UAV
$h_k(t)$	Channel gain between sensor k and the UAV
R	The coverage range of the UAV
$\bar{P}_k(t)$	Battery level of the k-th sensor
$\rho_k(t)$	Operation mode factor of sensor k
P_{UAV}, P_s	Transmit power of the UAV and ground sensors

UAV: unmanned aerial vehicle

time slot t. In the uplink DC mode, we have $\rho_k(t) = 1$, and sensor k is selected to send its data to the UAV by consuming its energy storage. In order to avoid mutual interference, only a single sensor is allowed to send data to the UAV, which yields $\sum_{k=0}^{K} \rho_k(t) = 1$. In contrast, in the downlink WPT mode, we have $\rho_k(t) = 0$, and the UAV will charge the ground sensors within the coverage area R. These sensors may harvest energy from the downlink radio frequency (RF) signals of the UAV to replenish their energy storage. In the state listening mode, the UAV obtains the status information of the sensors within its coverage area through its beacons, thus making a partial observation of the UAV. The state information includes the sensor's battery level, data traffic, and instantaneous channel conditions. This state information is then used to execute the actions of the UAV. Note that, since the state listening mode occupies a much shorter time duration compared with the other two operation modes, it can be reasonably omitted from the mode selection of the UAV.

Note that the UAV can only obtain state information for sensors within their coverage area R, therefore, the UAV must autonomously navigate to the vicinity of all sensors without fully knowing their locations and should cover as many sensors as possible based on local observations.

Fig. 2 illustrates the operation modes of the UAV over five consecutive time slots. In the time slots t_1 and t_3, the UAV operates in a downlink WPT mode. Since only sensors 1, 3, and k are in their coverage area, they may harvest energy from the UAV's downlink WPT signal. In the t_2, t_4 and t_5, the UAV operates in an uplink DC mode, while sensors 3, 1 and k upload data to the UAV, respectively.

2.2 Problem Formulation

The locations of sensor k and the UAV at time slot t are denoted as $\boldsymbol{w}_k = (x_k, y_k)$ and $\boldsymbol{q}(t) = (x(t), y(t))$, respectively. Accordingly, the distance from the UAV to sensor k is given by

$$d_k(t) = \sqrt{\left\| \boldsymbol{q}(t) - \boldsymbol{w}_k \right\|^2 + H^2}, \tag{2}$$

where $\|\cdot\|$ denotes the Euclidean norm of a vector.

UAV: unmanned aerial vehicle   WPT: wireless power transfer
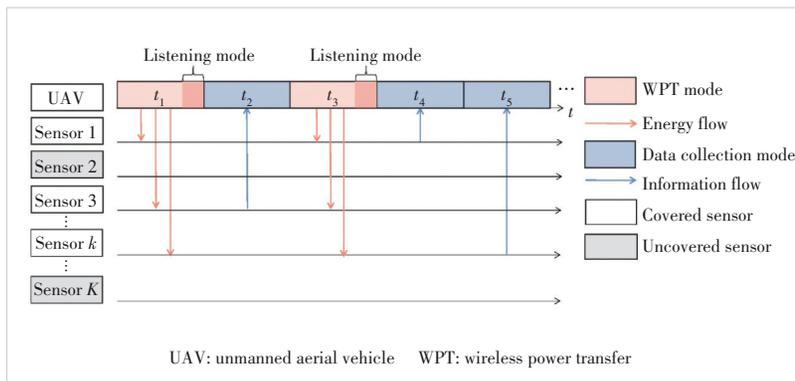
▲Figure 2. Communication protocol for a UAV-assisted wireless sensor network

The UAV communicates with sensors via the line-of-sight (LoS) communication links. The channel power gain between the UAV and sensor k at time slot t is given by:

$$h_k(t) = \beta_0 d_k^{-2}(t) = \frac{\beta_0}{\left\| \boldsymbol{q}(t) - \boldsymbol{w}_k \right\|^2 + H^2}, \forall k \in \mathcal{K}, \tag{3}$$

where β_0 denotes the channel power gain at a reference distance of 1 m.

First, we consider the WPT mode, where $\rho_k(t) = 0$. Let P_{UAV} denote the transmit power of the UAV, while all the K sensors have the same transmit power of P_s. Accordingly, at time slot t, the energy harvested by sensor k can be expressed as:

$$\hat{P}_k(t) = \left(1 - \rho_k(t)\right) \eta P_{\text{UAV}} h_k(t) \tau = \frac{\left(1 - \rho_k(t)\right) \eta \beta_0 P_{\text{UAV}}}{\left\| \boldsymbol{q}(t) - \boldsymbol{w}_k \right\|^2 + H^2} \tau, \tag{4}$$

where $0 < \eta \leq 1$ denotes the RF-to-direct current energy conversion efficiency.

Then, in the DC mode, we have $\rho_k(t) = 1$. The achievable uplink throughput of sensor k in bits per second can be expressed as:

$$\hat{r}_k(t) = \rho_k(t) B \log_2 \left(1 + \frac{P_s h_k(t)}{\sigma^2}\right) =$$
$$\rho_k(t) B \log_2 \left(1 + \frac{P_s h_k(t)}{\left\| q(t) - w_k \right\| + h^2}\right), \tag{5}$$

where σ^2 is the noise power, and $\lambda \triangleq \beta_0 / \sigma^2$ is the reference signal-to-noise ratio (SNR). Thus, the total number of data $r_k(t)$ collected from sensor k at the end of the time slot t is given by:

$$r_k(t) = r_k(t-1) + \hat{r}_k(t)\tau, \tag{6}$$

where we reasonably assume $r_k(0) = 0$. By jointly considering energy harvesting and energy consumption at the time slot t, the remaining energy on sensor k is given by

$$P_k(t) = P_k(t-1) + \hat{P}_k(t) - P_s\tau \geq 0, \text{ for } \forall k. \tag{7}$$

Appropriate actions $\left\{\alpha(t), v_{\text{UAV}}(t), \rho_k(t)\right\}$ must be carefully chosen for the UAV to ensure that the energy consumption of sensor k should not exceed the energy stored, which constitutes the energy causality constraint on all sensors of Eq. (7).

For simplicity, in the state listening mode, sensor k may report its energy state to the UAV in the

form of a single binary bit. Therefore, the energy state information of sensor k is quantified as:

$$
\bar{P}_k(t) = \begin{cases} 1, & \text{if } P_k(t) \geqslant P_s\tau \\ 0, & \text{if } P_k(t) < P_s\tau. \end{cases} \tag{8}
$$

Since the flight time is stipulated, we need to predict whether the UAV can fly back to the origin within the specified time at the current position. Accordingly, we express the judgment basis as:

$$
\frac{\sum_{t'=1}^{t} v_{\text{UAV}}(t')}{t}(T-t) \geqslant \sqrt{\|q(t+1)-q(0)\|^2}, \tag{9}
$$

where $\dfrac{\sum_{t'=1}^{t} v_{\text{UAV}}(t')}{t}$ represents the UAV's average flying velocity so far. Additionally, $q(t+1)$ represents the next position obtained by the UAV according to the DQN algorithm that will be introduced in Section 3.

Our objective is to maximize the average data traffic of all sensors during a finite flight time. Therefore, the data traffic optimization problem can be mathematically formulated as follows:

$$
(\text{P1}): \max_{\{q(t),(v_{\text{UAV}}),\rho_k(t)\}} \frac{\sum_{k=0}^{K} r_k(T)}{K}, \tag{10}
$$

$$
\text{s. t. } \rho_k(t) = \{0,1\}, \forall k \in \mathcal{K}, \tag{10a}
$$

$$
\sum_{k=0}^{K} \rho_k(t) = 1, \forall t \in T, \tag{10b}
$$

$$
P_k(t) - P_s T_{\text{DC}} \geqslant 0, \forall k \in \mathcal{K}, \tag{10c}
$$

$$
r_k(T) \geqslant r_{\text{QoS}}, \forall k \in \mathcal{K}, \tag{10d}
$$

Eqs. (7) and (9).

In addition to the flight Constraints (9) and (10c), the operating mode Constraints (10a) and (10b) and the energy causality Constraint (7), we also need to guarantee that the average data collection of all sensors should be above the minimum quality of service (QoS) requirement r_{QoS}, as expressed in Constraint (10d).

3 Deep Reinforcement Learning for UAV-Assisted Power Transfer and Data Collection

Due to the non-convex optimization problem and the large action space, in this section, we employ a DQN-based algorithm to solve Problem (P1). The UAV will periodically select the best action based on the network status while maximizing the average data traffic.

3.1 Deep Q-Network

We first provide a brief review of the DQN framework. The DQN approach can be described as a MDP, which is defined by a 4-tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P} \rangle$, where $\mathcal{S}$ is the set of states, $\mathcal{A}$ is the set of all possible actions, $\mathcal{R}$ represents the reward when an action is taken, and $\mathcal{P}$ denotes the transition probability from one state to another. The DQN structure is illustrated in Fig. 3, where the agent observes the environment, obtains the state $s_t \in \mathcal{S}$, chooses an action $a_t \in \mathcal{A}$, and then receives a reward $r_t \in \mathcal{R}$ according the observation and the next state.

The DQN obtains an optimal policy π by maximizing the long-term expected accumulated rewards. The expected accumulated reward for each state-action pair is defined as:

$$
Q(s_t,a_t) = \mathbb{E}\left[\sum_{t=1}^{T} \gamma^{t-1} r_t \middle| s_t,a_t\right], \tag{11}
$$

where $\gamma \in [0,1]$ is the discounted factor. Then, we can get the optimal policy: $\pi(s_t) = \arg\max_{a_t} Q(s_t,a_t)$. By selecting the optimal action a^*, we have the optimal action-value function:

$$
Q^*(s_t,a_t^*) = (1-\varpi)Q^*(s_t,a_t^*) + \varpi\left[r(s_t,a_t^*) + \gamma \max_{a_t \in \mathcal{A}} Q(s_t,a_t)\right], \tag{12}
$$

where $\varpi \in (0,1]$ is the learning rate. A DQN uses a deep neural network (DNN) as the approximator of the $Q(\cdot)$ function and is trained by minimizing the following loss function:

$$
L(\theta) = \mathbb{E}[y_t - Q(s_t,a_t|\theta)^2], \tag{13}
$$

where θ^Q is the weight vector of the DQN, and y_t is the target value, which is denoted by:

$$
y_t = r_t + \gamma \max_{a_t} \hat{Q}(s',a';\theta^-). \tag{14}
$$

In conventional Q-learning, the Q-value is updated by both the return at the current slot and the value estimated at the
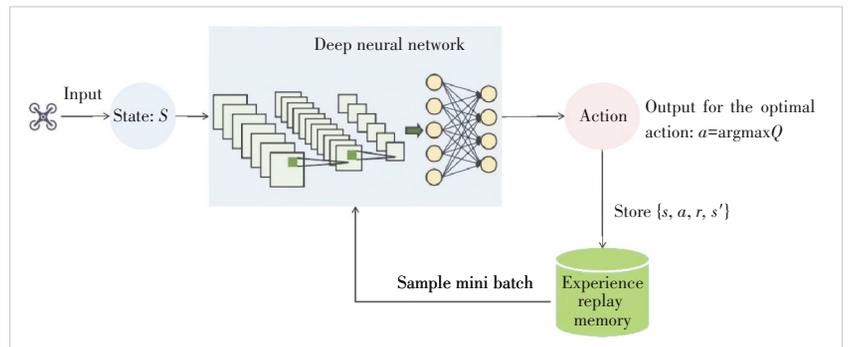
▲Figure 3. Schematics of the proposed algorithm

next slot. Due to the instability of the training samples, some fluctuations may occur in each iteration, which will be immediately reflected in the next iteration. In order to reduce the impact of related issues, it is necessary to decouple the two parts as much as possible. Therefore, the DQN introduces a dual neural network and replay buffer mechanism. It uses another network to generate the target Q-value, which is used to calculate the evaluated network and loss function. After M iterations, the parameters of the evaluated network will be copied to the target network. The model for calculating the target value will be fixed for a period of time, hence reducing the volatility of the model. In addition, a replay buffer is applied to store the transition samples (s,a,r,s') that are generated in each iteration.

3.2 DQN-Based Solution

In the DQN algorithm we used, the UAV implements the ground sensor selection and operation mode selection, and calculates the next location and flying speed. Each action depends on the network state. Let us explain the definition of state space, action space and reward function of the UAV in our proposed DQN-based algorithm, as summarized in Algorithm 1.

Algorithm 1: DQN-based algorithm

1: Initialize replay memory $\mathcal{D}$ to capacity N;

2: Initialize action-value function Q with random weights θ;

3: Initialize action-value function $\hat{Q}$ with random weights $\theta^- = \theta$;

4: **for** episode = 1, M, **do**

5:     Start state $s_t \rightarrow s_1$ and accordingly update θ;

6:     **for** $t = 1, \cdots, T$ **do**

7:         Execute action a_t in an emulator and observe reward r_t and the next state s_{t+1};

8:         Store transition (s_t, a_t, r_t, s_{t+1}) in $\mathcal{D}$;

9:         Sample M random minibatch of transitions (s_j, a_j, r_j, s_{j+1}), $j = 1, \cdots, M$ from $\mathcal{D}$;

10:         Calculate y_j according to Eq. (14);

11:         Calculate the MSE loss function $\frac{1}{M} \sum_{j=1}^{M} \left(y_j - Q(s_j, a_j; \theta) \right)^2$ and update the network parameters θ;

12:         Update $(x(t+1), y(t+1))$;

13:         Every δ steps reset $\hat{Q} = Q$;

14:     **end for**

15: **end for**

---

1) The UAV obtains the status information of the sensors in the coverage area through autonomous navigation. The state $\mathcal{S}$ of the network at the time slot t contains four parts:

- $\bar{P}_k(t) | k \in \mathcal{K}(t)$: the battery level of the sensors within the coverage area;

- $r_k(t) | k \in \mathcal{K}(t)$: the current data traffic of all sensors, where $\mathcal{K}(t) = \{k | d_k(t) \leqslant R\}$, indicates the sensors within the

coverage area R of UAV;

- $q(t)$: the UAV location;

- $h_k(t) | k \in \mathcal{K}(t)$: the channel gain between the UAV and the sensors within the coverage area.

2) The steering angle α_t of the UAV is assumed to be selected from D directions in the angular domain. The action $\mathcal{A}$ contains four parts:

- k: the ground sensor selection, where $k \in \mathcal{K}(t)$;

- $\rho_k(t)$: the operation mode selection, by which the UAV chooses to charge or receive data, where $\rho_k(t) \in \{0,1\}$;

- $\alpha(t)$: the steering angle;

- $v_{UAV}(t)$: the next flying speed.

As it can be seen from the above, the action space $\mathcal{A}$ has a cardinality of $3 \cdot K \cdot D$, where 3 represents that the flying speed of the UAV can have three values. Thus we can get the next location of the UAV:

$$\begin{cases} x(t+1) = x(t) + v_{UAV} \cos \alpha(t), \\ y(t+1) = y(t) + v_{UAV} \sin \alpha(t). \end{cases} \quad (15)$$

3) The reward $\mathcal{R}$: since our objective is to maximize the average data traffic while ensuring that the battery level $\bar{P}_k(t)$ is not lower than 1, our reward function consists of two parts:

- $r_{data}(t)$ is the change of the average data traffic of all sensors after selecting an action:

$$r_{data}(t) = \Delta \frac{\sum_{k=1}^{K} r_k(t)}{K}. \quad (16)$$

- $r_{penalty}(t)$ is an action penalty when none of the constraints in (P1) is satisfied.

To summarize, we give the final reward function:

$$r(t) = r_{data}(t) + r_{penalty}(t). \quad (17)$$

4 Simulation Results

In this section, we present network configurations and illustrate numerical results including the trajectory of the UAV, energy, battery level and data traffic of all sensors to validate the proposed DQN-based algorithm. Our experiments are performed on Tensorflow 1.11.0 (the symbolic math library for numerical computation) and Python 3.6.

4.1 Experiments Settings

As it is shown in Tables 2 and 3, we assume that $K = 10$ sensors are uniformly distributed within a 200×200 m^2 district in our environments. The time-slot duration is fixed at $t = 1$ s. The transmit power of the UAV and ground sensors are $P_{UAV} = 30$ dBm and $P_s = 0$ dBm, respectively. The channel power gain at the reference distance $d_0 = 1$ m is set as $\beta_0 = -30$ dB. The QoS requirement is set as 0 bit/s. And we set

▼Table 2. Simulation parameters: DNN

DNN Parameters	Value
Learning rate	0.000 1
Discount factor	0.9
Replay memory size	10 000
Batch size	32
ReLu hidden neurons	20
Number of neural network layers	2

DNN：deep neural network

▼Table 3. System parameters

System Parameters	Value
Bandwidth	1 MHz
Energy conversion efficiency	0.9
Noise power	−60 dBm
Flying height	10 m
Coverage area	70 m^2
Steering angle $(\alpha(t))$	$\left\{\frac{k}{4}\pi \mid k = 0,\cdots,7\right\}$
Flying velocity $(v_{\text{UAV}}(t))$	$\{0,5,10\}$ m/s

r_{penalty} as 0.01 when a constraint is not satisfied.

4.2 Performance Evaluation

The optimized trajectory of the UAV within different specified flight time is shown in Fig. 4. The red dots represent 10 sensors, and the asterisks indicate the start and end of the UAV. The black trajectories represent the UAV trajectories obtained by DQN. We assume that when the distance between the UAV and the origin is less than 10 m and all sensors satisfy the QoS requirement, the episode is ended. It can be observed that the UAV can complete autonomous navigation tasks and adjust its trajectory either to transfer energy or collect data. As it can be seen from Fig. 4, the UAV needs to cover as many sensors as possible for charging in order to receive data later, so the UAV will not fly straightly over a specific UAV. Similarly, we can see from Figs. 4(a) and 4(b) that when the flight time increases, the UAV's flight trajectory is closer to the sensor.

Furthermore, we compare the situation of the UAV flying along the preset circular trajectory with the same start points. As shown in the colored trajectories in Fig. 4(a), the UAV will fly at a constant speed along circular trajectories with the radius of 75 m, 100 m, and 125 m respectively. At this time, the action of the UAV will only include the selection of the operation mode and the specific sensor, namely $a_t = \{k, \rho_k(t)\}$.

We then investigate the DQN-based algorithm performance by analyzing the battery and data traffic of sensors. Figs. 5(a) and 5(b) respectively show the energy and data traffic of the first, fifth and tenth sensors, where a training step is equivalent to the time slot duration. The rising part of the curve in Fig. 5(a) indicates that the corresponding sensor is being charged, which corresponds to the level part of the curve in

DQN: deep Q-network   UAV: unmanned aerial vehicle
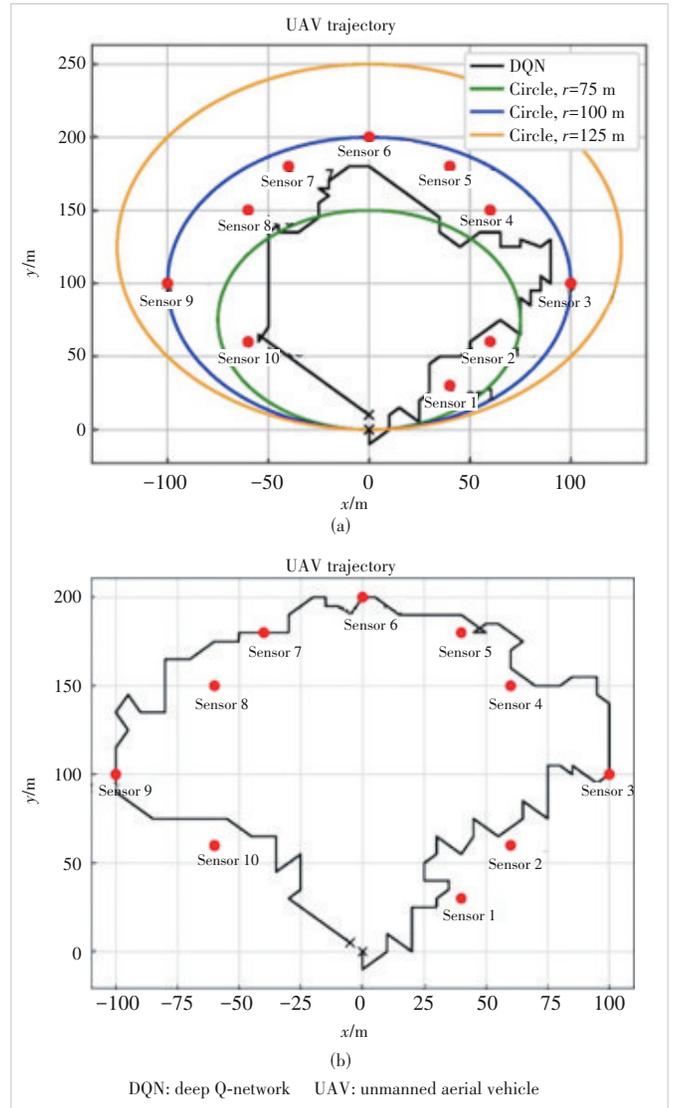
▲ **Figure 4. UAV's trajectory under different flight time: (a) 2 000 s and (b) 4 000 s**

Fig. 5(b). In contrast, the falling part of Fig. 5(a) indicates that the sensor k transmits data to the UAV, which also corresponds to the rising curve in Fig. 5(b).

For the purpose of testing the performance of the algorithm, we then simulate the scene with different sensor numbers and different flight time by DQN and three circular trajectories with different radii in Fig. 6. It can be seen from Fig. 6(a) that when the flight time of UAV increases, the average data traffic on the sensor increases. This is because the flight time increases, and the UAV will have more time to receive data from the sensors. And it can be seen that the performance obtained by the DQN algorithm is the best, followed by the circular trajectory with $r = 100$ m, and the circular trajectory with $r = 125$ m is the worst. Because the DQN algorithm can choose the speed and the next position of the UAV, it can hover around the sensor more. The circu-
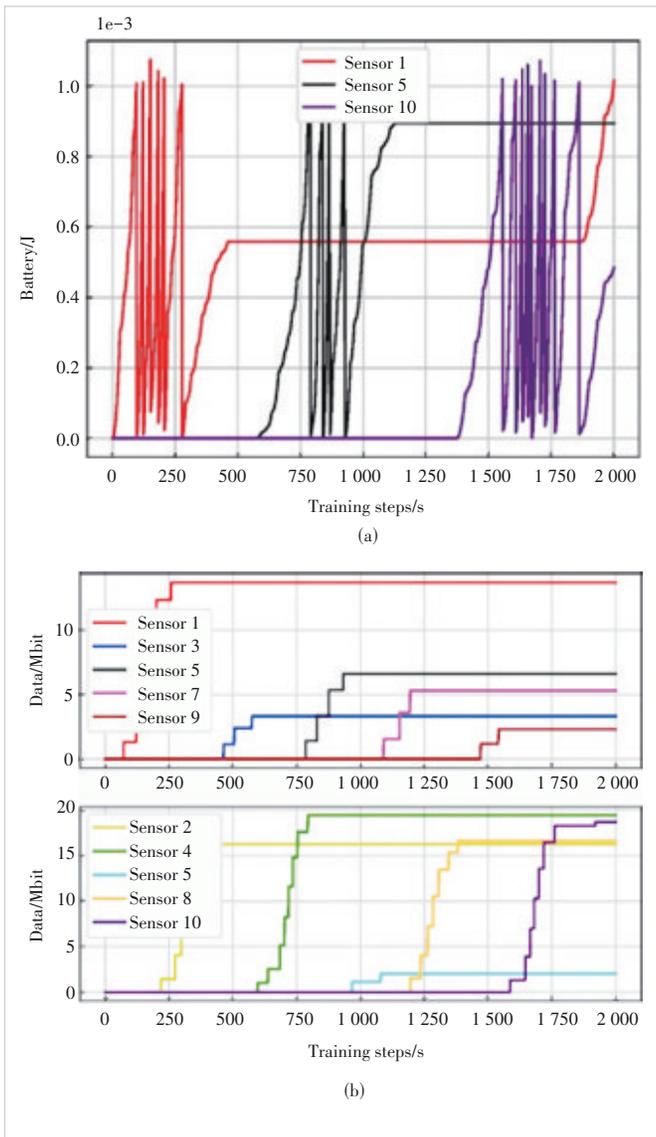
▲Figure 5. Energy and data traffic of *k*=1.5 and 10 ground sensors: (a) battery and (b) data traffic

▲Figure 6. Comparison of average data traffic per sensor by DQN and three circular trajectories

lar trajectory with *r* = 100 m is the closest to the sensor among all circular trajectories shown in Fig. 4(a), so the average data traffic obtained is the largest among the three circular trajectories.

Fig. 6(b) shows that under the same flight time, the fewer the number of sensors, the more the UAV will choose to perform the data collection mode, which results in higher average data traffic. After the number of sensors increases to 16, the average data traffic per sensor by DQN becomes stabilized. This is because the total distribution area of the sensors and the coverage of the UAV remains unchanged. When the number of sensors increases, the distribution of sensors will be denser, and the UAV can charge more sensors with the same coverage and the same charging time. Additionally, it can be observed from Fig. 6(b) that the number of sensors does not

have much impact on the circular trajectories, since the UAV always flies at a constant speed. There is no situation where the UAV has a higher probability of hovering over certain sensors when the number of sensors is small. This shows that the DQN algorithm is more suitable for sparsely distributed and nonuniform networks.

5 Conclusions

In this paper, we formulate the problem of power transfer and data collection in a UAV-assisted WSN, while ensuring autonomous navigation of the UAV. We then propose a DQN-based algorithm to solve the problem in order to maximize the average data traffic. Given the states of the battery level, the current data traffic of all sensors, the position of the UAV, and channel conditions, the UAV takes actions according to the

proposed DQN-based algorithm. Numerical results illustrate that the proposed algorithm significantly improves the network performance. However, the proposed algorithm also has high complexity. Due to the real-time interaction with the environment, the UAV needs significant energy consumption, which can be solved by the digital twin network.

References

[1] HU J, WANG Q, YANG K. Energy self-sustainability in full-spectrum 6G [J]. IEEE wireless communications, 2021, 28(1): 104 – 111. DOI: 10.1109/MWC.001.2000156

[2] WU D P, HE J, WANG H G, et al. A hierarchical packet forwarding mechanism for energy harvesting wireless sensor networks [J]. IEEE communications magazine, 2015, 53(8): 92 – 98. DOI: 10.1109/MCOM.2015.7180514

[3] GHARIBI M, BOUTABA R, WASLANDER S L. Internet of drones [J]. IEEE access, 2016, 4: 1148 – 1162. DOI: 10.1109/ACCESS.2016.2537208

[4] ZHU S C, GUI L, CHENG N, et al. Joint design of access point selection and path planning for UAV-assisted cellular networks [J]. IEEE Internet of Things journal, 2020, 7(1): 220 – 233. DOI: 10.1109/JIOT.2019.2947718

[5] BEN GHORBEL M, RODRÍGUEZ-DUARTE D, GHAZZAI H, et al. Joint position and travel path optimization for energy efficient wireless data gathering using unmanned aerial vehicles [J]. IEEE transactions on vehicular technology, 2019, 68(3): 2165 – 2175. DOI: 10.1109/TVT.2019.2893374

[6] HU J, CAI X P, YANG K. Joint trajectory and scheduling design for UAV aided secure backscatter communications [J]. IEEE wireless communications letters, 2020, 9(12): 2168 – 2172. DOI: 10.1109/LWC.2020.3016174

[7] XIE L F, XU J, ZHANG R. Throughput maximization for UAV-enabled wireless powered communication networks [J]. IEEE Internet of Things journal, 2019, 6(2): 1690 – 1703. DOI: 10.1109/JIOT.2018.2875446

[8] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing atari with deep reinforcement learning [EB/OL]. [2023-03-10]. https://arxiv.org/abs/1312.5602

[9] LI K, NI W, TOVAR E, et al. On-board deep Q-network for UAV-assisted online power transfer and data collection [J]. IEEE transactions on vehicular technology, 2019, 68(12): 12215 – 12226. DOI: 10.1109/TVT.2019.2945037

[10] TANG J, SONG J R, OU J H, et al. Minimum throughput maximization for multi-UAV enabled WPCN: A deep reinforcement learning method [J]. IEEE access, 2020, 8: 9124 – 9132. DOI: 10.1109/ACCESS.2020.2964042

[11] LIU C H, CHEN Z Y, TANG J, et al. Energy-efficient UAV control for effective and fair communication coverage: a deep reinforcement learning approach [J]. IEEE journal on selected areas in communications, 2018, 36(9): 2059 – 2070. DOI: 10.1109/JSAC.2018.2864373

[12] LU Y P, XIONG G, ZHANG X, et al. Uplink throughput maximization in UAV-aided mobile networks: a DQN-based trajectory planning method [J]. Drones, 2022, 6(12): 378. DOI: 10.3390/drones6120378

[13] ABEDIN S F, MUNIR M S, TRAN N H, et al. Data freshness and energy-efficient UAV navigation optimization: a deep reinforcement learning approach [J]. IEEE transactions on intelligent transportation systems, 2021, 22(9): 5994 – 6006. DOI: 10.1109/TITS.2020.3039617

[14] ZHANG T K, LEI J Y, LIU Y W, et al. Trajectory optimization for UAV emergency communication with limited user equipment energy: a safe-DQN approach [J]. IEEE transactions on green communications and networking, 2021, 5(3): 1236 – 1247. DOI: 10.1109/TGCN.2021.3068333

[15] LI K, NI W, TOVAR E, et al. Joint flight cruise control and data collection in UAV-aided Internet of Things: an onboard deep reinforcement learning approach [J]. IEEE Internet of Things journal, 2021, 8(12): 9787 – 9799. DOI: 10.1109/JIOT.2020.3019186

Biographies

LI Yuting is with the University of Electronic Science and Technology of China. Her current research interests include data and energy integrated communication networks and machine learning.

DING Yi received his bachelor's degree in communication and master's degree in communication and information systems engineering both from Beijing University of Posts and Telecommunications, China. After graduation, he has been engaged in the field of wireless network planning and optimization, demonstrating expertise in the 5G network technology and its evolving trends.

GAO Jiangchuan received his BS degree from the Southwest Jiaotong University, China in 2022. He is currently pursuing his master's degree in the University of Electronic Science and Technology of China. His current research interests include data and energy integrated communication networks and unmanned aerial vehicle trajectory planning.

LIU Yusha (yusha.liu@uestc.edu.cn) received her PhD degree from the University of Southampton, UK, and is currently with the University of Electronic Science and Technology of China. Her current research interests include wireless communications, signal processing and deep learning.

HU Jie received his BE and MS degrees from Beijing University of Posts and Telecommunications, China in 2008 and 2011, respectively, and received his PhD degree from the School of Electronics and Computer Science, University of Southampton, UK in 2015. Since 2016, he has been working with the School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC). He is now a research professor and PhD supervisor. He won UESTC's Academic Young Talent Award in 2019. Now he is supported by the "100 Talents" program of UESTC. He is an editor for *IEEE Wireless Communications Letters*, *IEEE/CIC China Communications* and *IET Smart Cities*. He serves for *IEEE Communications Magazine*, *Frontiers in Communications and Networks* as well as *ZTE communications* as a guest editor. He is a program vice-chair for IEEE TrustCom 2020, a technical program committee (TPC) chair for IEEE UCET 2021 and a program vice-chair for UbiSec 2022. He also serves as a TPC member for several prestigious IEEE conferences. He has won the best paper award of IEEE SustainCom 2020 and the best paper award of IEEE MMTC 2021. His current research focuses on wireless communications and resource management for B5G/6G, wireless information and power transfer as well as integrated communication, computing and sensing.

YANG Kun received his PhD from the Department of Electronic & Electrical Engineering of University College London (UCL), UK. He is a Chair Professor in the School of Computer Science & Electronic Engineering, University of Essex, UK, and is leading the Network Convergence Laboratory (NCL), UK. He is also an affiliated professor at UESTC, China. Before joining in the University of Essex at 2003, he had worked at UCL on several European Union (EU) research projects for several years. His main research interests include wireless networks and communications, IoT networking, data and energy integrated networks and mobile computing. He manages research projects funded by various sources such as UK EPSRC, EU FP7/H2020, etc. He has published more than 400 journal papers and filed 20 patents. He is an IEEE ComSoC Distinguished Lecturer (2020 – 2021) and a member of Academia Europaea (MAE). Professor YANG is an IEEE Fellow.

Research Towards Terahertz Power Amplifiers in Silicon-Based Process

CHEN Jixin[1], ZHOU Peigen[1], YU Jiayang[1], LI Zekun[1], LI Huanbo[1], PENG Lin[2]

(1. Southeast University, Nanjing 211189, China；
2. ZTE Corporation, Shenzhen 518057, China)

Abstract: In view of the existing design challenges for Terahertz (THz) power amplifiers (PAs), the common design methods and the efforts of the State Key Laboratory of Millimeter Wave, Southeast University, China in the development of silicon-based THz PAs, mainly including silicon-based PAs with operating frequencies covering 100−300 GHz, are summarized in this paper. Particularly, we design an LC-balun-based two-stage differential cascode PA with a center frequency of 150 GHz and an output power of 14 dBm. Based on a Marchand balun, we report a 220 GHz three-stage differential cascode PA with a saturated output power of 9.5 dBm. To further increase the output power of THz PA, based on a four-way differential power combining technique, we report a 211−263 GHz dual-LC-tank-based broadband PA with a recorded 14.7 dBm Psat and 16.4 dB peak gain. All the above circuits are designed in a standard 130 nm silicon germanium (SiGe) BiCMOS process.

Keywords: power amplifier; power combining; SiGe; silicon-based; Terahertz

1 Introduction

In recent years, the importance of the Terahertz (THz) spectrum in electronics, radio astronomy and other fields, such as biological imaging, high-speed communication, and high-resolution radar, has gradually emerged mainly due to its spectral resolution safety, perspective and broadband characteristics[1]. However, due to the large loss of free space propagation in the THz frequency band (especially when frequencies are above 100 GHz), combined with the effects of atmospheric attenuation, a transmitter needs to have sufficient effective isotropic radiated power (EIRP) to achieve long-distance wireless transmission. Fig. 1 shows a block diagram of a typical THz transmitter based on a communication application scenario. The power amplifier (PA), as the final active stage of the transmitter, directly drives the post-antenna to transmit the radio frequency (RF) signal into free space. Therefore, its performance indicators, including output power, RF bandwidth, etc., directly determine the performance of the entire transmitter and then restrict the wireless transmission distance of the entire transceiver system.

The III-V compound semiconductor processes are often used in THz PA designs due to their higher cutoff frequency (f_T), higher maximum oscillation frequency (f_{MAX}), higher cutoff voltage, and smaller substrate losses. However, the III-V process is not suitable for integrating large-scale digital and analog circuits, so multiple discrete chips implemented by different processes need to be integrated into the transceiver system. As a result, the channel module size tends to be large, making it difficult to meet the half-wavelength spacing requirements for THz array integration, which limits the performance of THz RF array systems.

With the development of silicon-based integrated circuit

IBB: I-baseband   PA: power amplifier   RF: radio frequency
LO: local oscillator   QBB: Q-baseband   Tx: transmitter
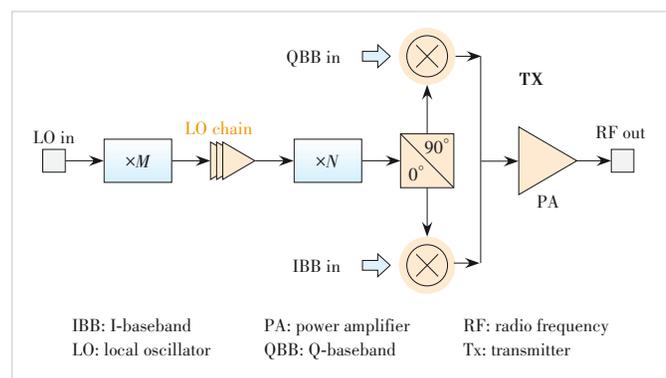
▲Figure 1. Block diagram of a typical Terahertz（THz）transmitter

processes, the f_T/f_{MAX} of its transistors has been covered to the 300 GHz band[2], making it possible to design low-cost, highly integrated and small-size THz PAs based on advanced silicon-based processes. However, based on existing advanced silicon-based processes, there are still many challenges in designing high-performance PAs in the frequency band above 100 GHz[3]. First, the shrinking of the characteristic size of the transistor causes the withstand voltage value of the transistor to decrease, so the supply voltage is limited, thereby limiting the output power of the PA. Second, when the operating frequency increases to more than 100 GHz, the loss of the passive structure increases due to the influence of skin effect and radiation effect, which further deteriorates the performance of silicon-based THz PA. Fig. 2 summarizes the output power of currently representative reported PAs, mainly based on bulk complementary metal oxide semiconductor (CMOS) and silicon germanium (SiGe) processes[3]. It can be seen that the output power of silicon-based PAs drops sharply at above 100 GHz, and there is an urgent need to study high-performance silicon-based THz PAs.

To overcome the above difficulties, we have conducted research on THz broadband low-loss single-ended-to-differential signal conversion baluns and gain-boosted PA cores, and developed a variety of 100 – 300 GHz silicon-based differential PAs, including 150 GHz, 220 GHz and 250 GHz PAs[5 – 7]. In addition, aiming at the development of broadband low-loss power combining techniques and high-efficiency miniaturized multi-way power combining structures, we design a 240 GHz three-stage cascode structured PA based on four-way zero-degree transmission line power combining techniques[8].

2 Silicon-Based Differential THz PA

The development of high-performance silicon-based THz PAs is a key link in the practical process of THz communica-
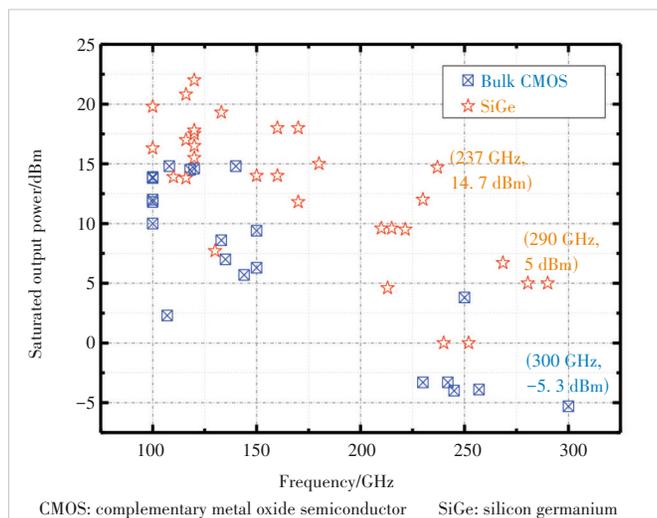
▲ Figure 2. Research status of Terahertz (THz) power amplifiers (PAs) based on bulk CMOS and SiGe process[3]

tion and radar systems. To ensure that the transmitter has sufficient transmit power, it is critical to study the design of PAs with high output power, high gain, and miniaturization. When the operating frequency rises to the THz frequency band, the influence of circuit parasitic parameters is intensified, coupled with the influence of skin effect and others, the quality factor of passive devices is reduced, the circuit loss increases, and the working bandwidth deteriorates sharply. Conventional single-ended PAs have limited gain, output power, and efficiency in the THz band. Therefore, scholars have conducted extensive research on how to design and produce differential structured silicon-based THz PAs with high performance. According to Ref. [3], PAs based on silicon-based processes currently can operate at frequencies up to 300 GHz. In typical reports like Ref. [9], the PA shows a 3 dB bandwidth of 63 GHz (239 – 302 GHz) in small-signal operation and 94 GHz (223 – 317 GHz) when saturated, and the PA is fabricated in an experimental 130 nm SiGe BiCMOS process with f_T/f_{MAX} of 470/650 GHz.

Differential combining is the commonest power combining method in THz PA designs, and differential power combining structures include LC baluns, transformer baluns, and Marchand baluns. LC baluns consist of lumped component inductors and capacitors, making them easy to design. In the microwave and millimeter wave bands, LC baluns are less used due to the larger inductor size. When operating frequencies are above 100 GHz, the inductor size is greatly reduced as the wavelength decreases, making the LC balun suitable for on-chip differential power combining. Transformer baluns are widely used in the millimeter wave circuit design, by changing the size, spacing, and linewidth of the primary and secondary coils that make up the transformer, and the operating band, coupling coefficient, and bandwidth can be changed to achieve single-ended-to-differential signal conversion. Transformer baluns have the characteristics of small size, high order, wide bandwidth, many adjustable parameters, and strong design flexibility. However, in the THz frequency band, limited by the design rules of the chip processing process, the transformer balun is small in size, and the parasitic parameters and losses are often larger. The coils that make up Marchand baluns are half a wavelength in size and are larger in the lower frequency bands, usually reduced in area by spiral layouts. In the THz band, Marchand baluns, benefiting from a drastically reduced operating wavelength, often have the advantage of a moderate size and excellent RF performance.

2.1 150 GHz PA Based on LC Balun

The THz band has become a popular candidate for next-generation wireless communications due to its extremely rich spectrum resources. The D-band (110 – 170 GHz) is located in the THz low-frequency band, and there is an atmospheric window in the frequency range of 120 – 160 GHz. The wireless propagation path loss is small, so it is suitable for prelimi-

nary verification of THz communication systems. As one of the important components in the communication transceiver, the PA directly restricts the performance of the entire transceiver.

Fig. 3 shows a 150 GHz differential structured PA based on a commercial 130 nm SiGe BiCMOS process[6]. Common-emitter topology amplifiers have high power-added efficiency, but in the 150 GHz band, differential amplifiers based on common-emitter structures have lower single-stage gain due to the large decrease in the intrinsic gain of the transistor, and usually need to cascade four to five stages to obtain a small-signal gain of more than 20 dB. Therefore, the PA usually adopts a cascode structure. In order to further increase the gain of a single-stage cascode amplifier core, inductors L1 and L2 shown in Fig. 3 are introduced to the base of the common-base structure transistor. The base decoupling capacitor of the common-base transistor in a single-ended cascode amplifier will introduce an additional parasitic signal low impedance path to the RF ground, and the introduction of this pole also brings about stability issues, so the PA uses a differential structure. In addition, the size of the transistors in the PA is multiplied in the output direction to increase the power handling capability of the last stage and guarantee the low power performance of the PA.

For easy testing and easy connection to the pre-mixer and post-antenna, the amplifier has single-ended ground-signal-ground (GSG) RF PADs for both input and output. The core of the amplifier is differential, so a low-loss single-ended-to-differential signal conversion structure is required at the input and output ports of the amplifier. As shown in Fig. 3, LC baluns are used for single-ended and differential signal conversion at the input and output ports of the amplifier, where the inductor consists of microstrip lines. When operating at 150 GHz, the differential ports of the on-chip LC balun will have a phase imbalance problem, so the phase compensation lines TL4 and TL11 shown in Fig. 3 are introduced to achieve a good balance

performance at the differential ports.

The chip micrograph of the 150 GHz PA is shown in Fig. 4, and the simulation and measured results for the small-signal gain and saturated output power of this PA are shown in Fig. 5.
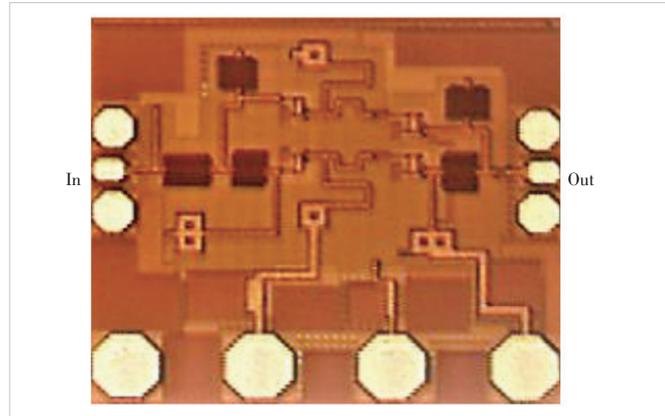
▲Figure 4. Die photo of the 150 GHz power amplifier (PA)
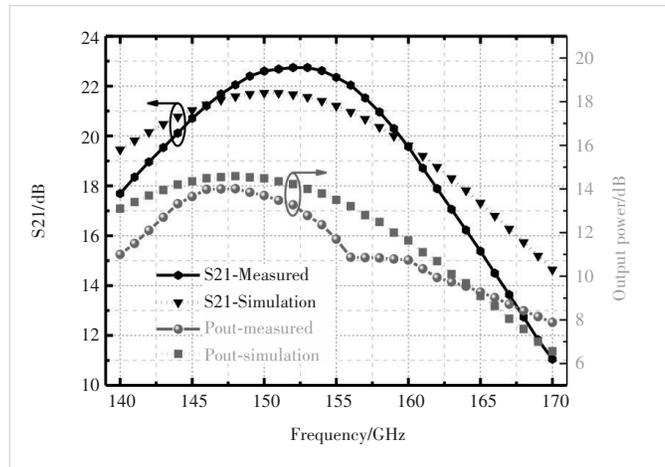
▲Figure 5. Simulated and measured results of the 150 GHz power amplifier (PA)[6]
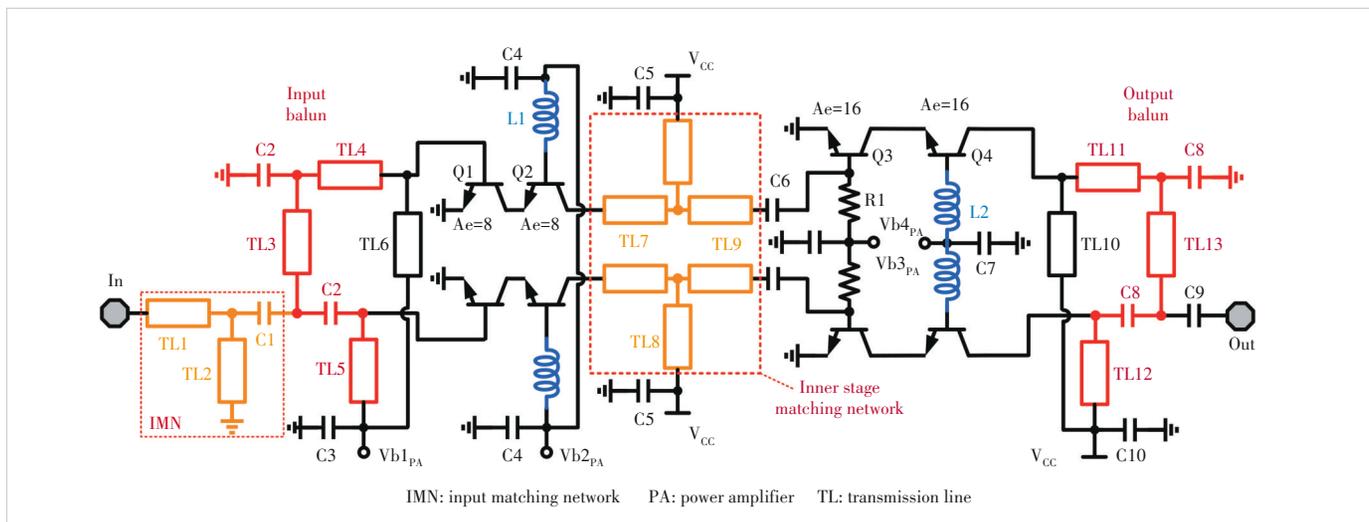
▲Figure 3. Schematic of the 150 GHz PA[6]

In the 140 – 160 GHz band, the small-signal gain of the amplifier is measured to be larger than 15 dB, reaching a maximum of 22.7 dB at 152 GHz. Meanwhile, the test shows that the PA obtains a maximum output power at the frequency of 148 GHz, where the power value is 14 dBm, and the 3 dB bandwidth of the output power reaches 16 GHz.

2.2 220 GHz PA Based on Marchand Balun

The 220 GHz band is currently a hot spot band, and there are many studies on this frequency band, including ultra-high-speed communications, high-resolution radar, etc. Among them, the PA is one of the key components that restrict the performance of the 220 GHz RF system. Achieving high performance PA in the 220 GHz band faces two major difficulties: one is that the operating frequency is close to the f_T/f_{MAX} of the transistor, resulting in a steep gain drop in the active device, and the other is the low quality factor and large insertion loss of the passive device in the THz band.

As shown in Figs. 6 and 7, the PA is a fully differential three-stage cascode structure[7], and the amplification unit using the cascode structure can reduce the number of stages, thereby reducing the passive loss caused by the inter-stage matching network, and the cascode topology also has better reverse isolation. The input and output networks are based on stacked Marchand baluns for single-ended-to-differential signal conversion. DC feed is performed through the center tap of the Marchand balun's secondary coil. In addition, some series transmission lines and parallel capacitors participate in impedance matching of the input and output networks. In the inner-stage impedance matching network, thanks to the fully differential structure, the intersection of the parallel transmission lines is an RF virtual point that can be used for DC feeding.

The measured results show the small-signal gain of the amplifier exceeds 20 dB, and the 3 dB bandwidth covers 204 – 239 GHz. Fig. 8 shows the simulated and measured large-signal performance of this PA, with saturated output power exceeding 9 dBm in the 200 – 216 GHz band.

3 Silicon-Based THz PA Based on Multi-way Power Combining Techniques

In silicon-based THz systems, due to the large connection loss of the RF interface (the connection between the RF output port and the antenna), the solution of integrated antenna-on-chip is often used. Limited by the thickness of the silicon process's metal layer, the on-chip antenna has a limited gain of about a few dBi. Therefore, as the device is directly con-
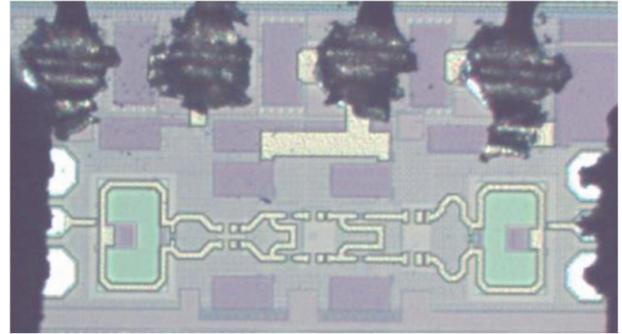
▲Figure 7. Die photo of the 220 GHz power amplifier (PA)[7]
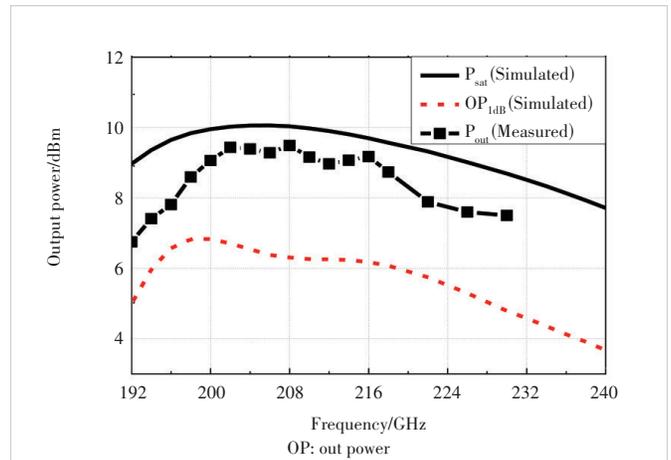
▲Figure 8. Simulated and measured results of the 220 GHz power amplifier (PA)[7]
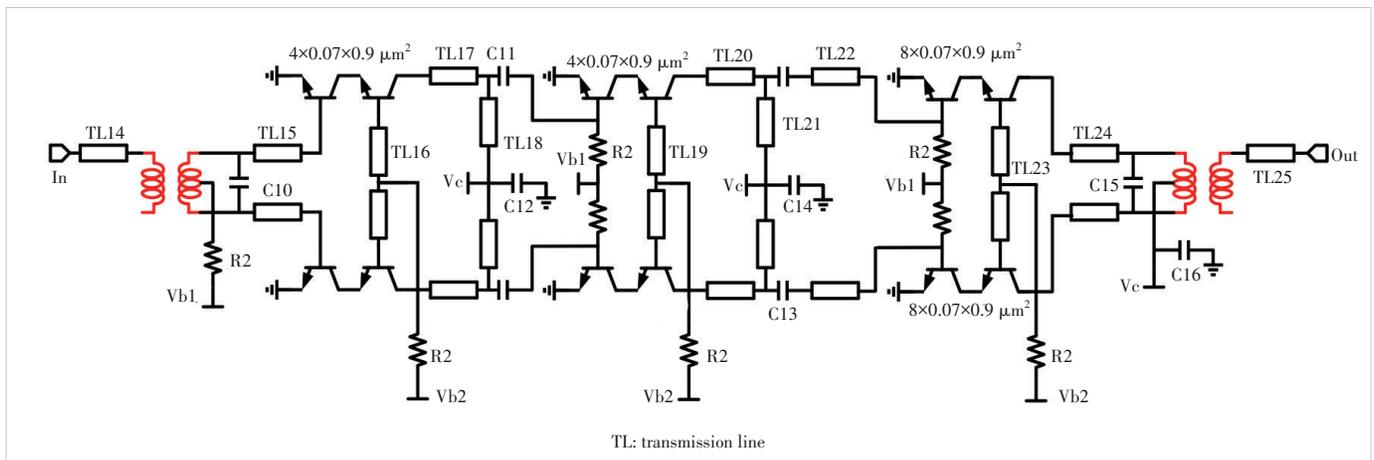
▲Figure 6. Schematic of the 220 GHz power amplifier (PA)[7]

CHEN Jixin, ZHOU Peigen, YU Jiayang, LI Zekun, LI Huanbo, PENG Lin

nected to the on-chip antenna in the RF system, the output power of the PA directly determines the transmission distance of the system[10]. However, the ability of transistors to deliver power in the THz band in silicon-based processes is limited, and power combining techniques must be employed to achieve higher output power.

Power combining techniques in THz bands include free-space power combining and on-chip power combining. Free-space power combining techniques can be co-designed with the on-chip antenna to generate high radiated power in equivalent space through array design[11]. However, on-chip antennas based on silicon-based processes are less efficient, and on-chip antennas tend to be larger in area and higher in cost. On-chip power combining techniques include the direct use of multiple transistors in parallel or multiple transistors stacking, and passive power synthesis techniques based on multi-way PAs. Transistor parallel connection is a current-based power combining method and transistor stacking is a voltage-based power combining method, and a better power combining efficiency can only be achieved when the current and voltage through each transistor are in phase. While the THz frequency is high and the wavelength is short, the connection line of tens of microns will introduce a large phase difference, so the number of transistors connected in parallel or stacked is limited. Passive multi-way power combining is to design a passive power combining network inside the chip and add the output power of the multi-way PAs with high efficiency to improve the final output power.

There are three main types of passive on-chip power combining techniques, namely power combining techniques based on a Wilkinson power divider, transformer-based power combining techniques, and zero-degree power combining techniques based on transmission lines. The power combining technique based on Wilkinson power dividers requires a pair of quarter-wavelength transmission lines that occupy a large area and are not suitable for on-chip power combining. At the same time, limited by the quarter-wavelength line, its bandwidth is relatively narrow. When the operating frequency deviates from the design frequency or there is a large parasitic, the isolation and matching between the ports will deteriorate significantly. Power synthesis based on transformers mainly includes voltage combining and current combining. For the transformer, common-mode signals and noise can be transmitted to the output due to parasitic capacitance between coils. When frequencies rise above 200 GHz, capacitance parasitic increases significantly, severely limiting transformers to be used in this frequency band, where Marchand balun's common-mode conversion is much lower than that of transformers, resulting in a smaller insertion loss. Zero-degree power combining techniques directly synthesize the current of multiple PAs to obtain high output power. And in the frequency band above 200 GHz, the T-type zero-degree power combining network has a compact size and low insertion loss.

Combining the characteristics of the above-mentioned on-chip power combining techniques, we have designed a three-stage differential cascode structure PA based on a four-way zero-degree power combining architecture, as depicted in Fig. 9[8]. As illustrated in Fig. 10, the PA's power combining network consists of an improved four-way zero-degree power combining network and a three-conductor Marchand balun, which
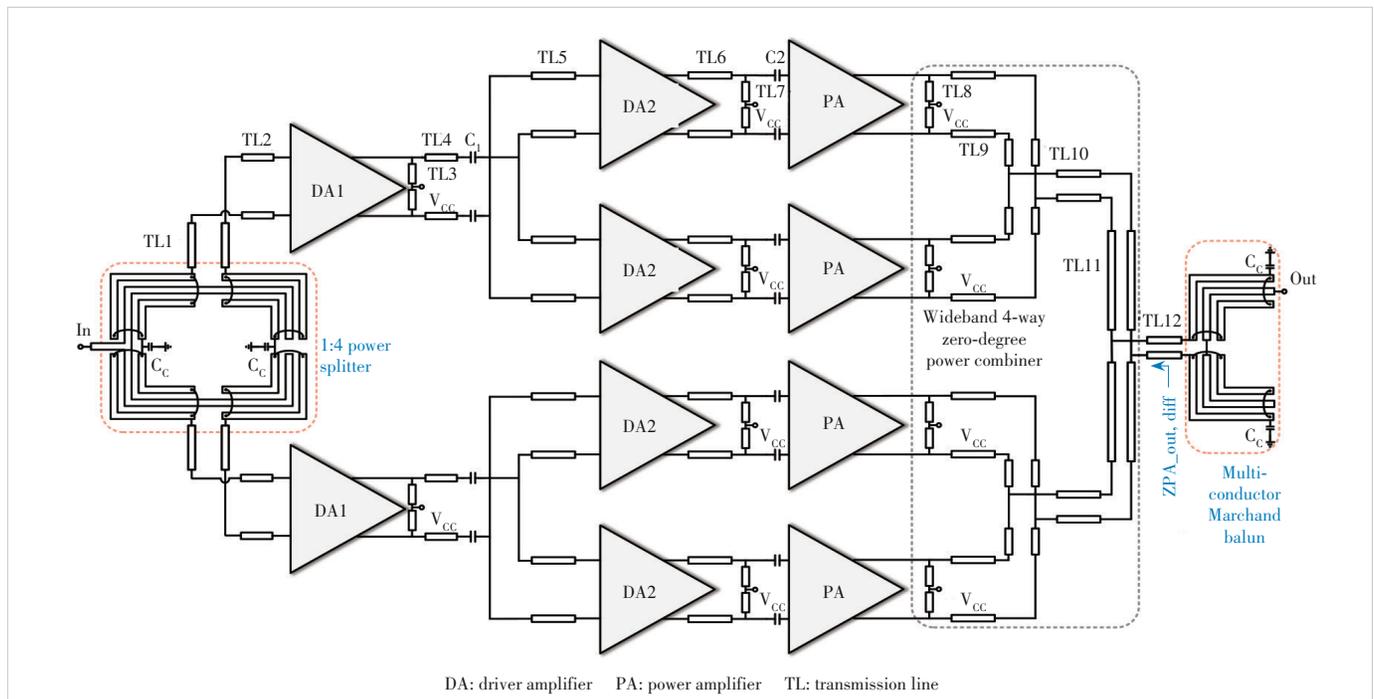
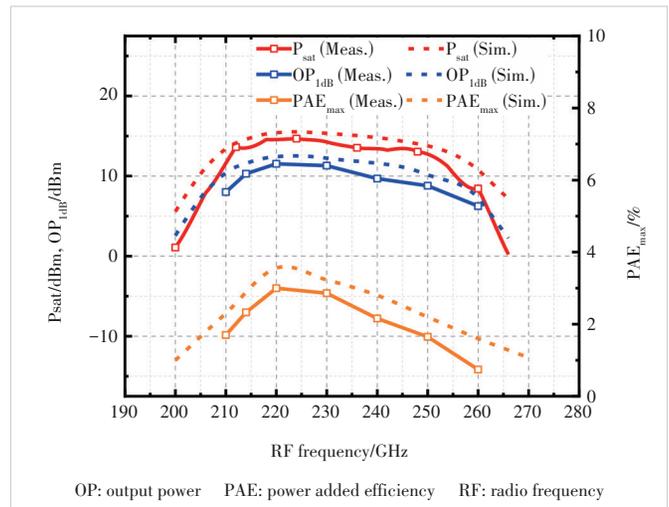▲Figure 9. Schematic of the 240 GHz four-way power combining PA[8]
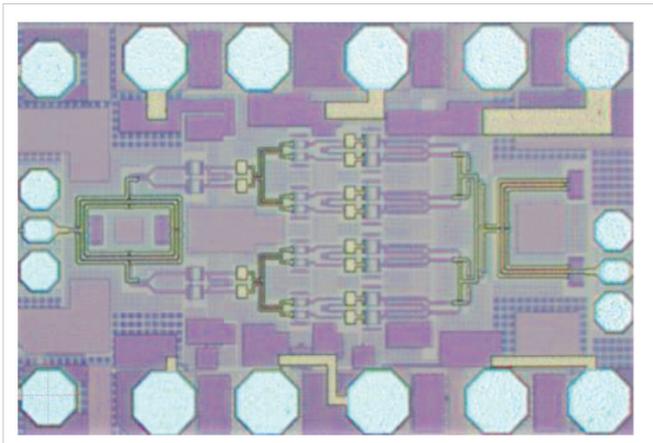
▲Figure 10. Die photo of the 240 GHz four-way power combining PA[8]

can achieve wideband impedance matching while completing high-efficiency power synthesis. Traditional zero-degree power combining networks use quarter-wavelength lines and bypass capacitors for DC feeding, which limits the operating bandwidth of the PA.

The zero-degree power combing network used by the PA contains RF virtual points for DC feed, as shown by the microstrip lines TL3 in Fig. 9, which is formed by the central point of the microstrip lines connected in parallel between differential ports. The advantage of such a design is that the broadband matching characteristics of the output network are guaranteed. In addition, to further increase the bandwidth of the output power combining network, we use a wideband Marchand balun based on a three-conductor coupling line as shown by the orange dotted box in Fig. 9 in the output differential-to-single-ended conversion network. When the operating frequency rises to 200 GHz, the lumped components with low figure of merit result in high insertion loss, and a broadband matching network with multi-resonant cavities can only be achieved by distributed networks. The three-conductor Marchand balun in the output impedance matching network achieves wideband matching with dual-LC resonances by absorbing two resonant networks into the coupling structure.

At the input of the PA, unlike the common passive or active power division structure, the PA uses a folded 1-to-4 power distribution network as shown in Fig. 9, where the RF ground of the secondary coil that makes up the balun is achieved by the metal-oxide-metal (MOM) capacitors Cc. The metal-insulator-metal (MIM) capacitors provided in the process have a low quality factor in the frequency band above 200 GHz, so the capacitors used in the PA design are all MOM capacitors. MOM capacitors are made of multilayer metal crossings and stacks for high design flexibility.

The chip micrograph of the designed PA is shown in Fig. 10. Due to the folded input power distribution network and the compact output power combining network, the core area of the PA is only 770×280 μm². As shown in Fig. 11, the measured results show that the small-signal gain of the PA ex-

▲Figure 11. Die photo of the 240 GHz four-way power combining PA[8]

ceeds 15 dB and the 3 dB bandwidth is larger than 50 GHz. In the 210 – 250 GHz band, the PA is tested to have saturated output power of more than 13 dBm.

4 Conclusions

Due to the abundant spectrum resources of the THz frequency band, application scenarios, such as high-speed communication and high-resolution radar with this frequency band, have become research hotspots. The rapid development of silicon-based processes provides new low-cost solutions beyond compound processes for these studies. The above systems have high demands on the performance of PAs such as output power and RF bandwidth. Combined with several PAs designed in our lab in the THz band, this article summarizes the common structure and design methods of PAs in the THz band.

References

[1] KISSINGER D, KAHMEN G, WEIGEL R. Millimeter-wave and terahertz transceivers in SiGe BiCMOS technologies [J]. IEEE transactions on microwave theory and techniques, 2021, 69(10): 4541 – 4560. DOI: 10.1109/TMTT.2021.3095235

[2] RÜCKER H, HEINEMANN B, FOX A. Half-terahertz SiGe BiCMOS technology [C]//The 12th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems. IEEE, 2012: 133 – 136. DOI: 10.1109/SiRF.2012.6160164

[3] WANG H, WANG F, NGUYEN H T, et al. Power Amplifiers Performance Survey 2000-Present [EB/OL]. [2023-04-16]. https://gems.ece.gatech.edu/PA_survey.html

[4] LI X C, CHEN W H, ZHOU P G, et al. A 250 – 310 GHz power amplifier with 15 dB peak gain in 130 nm SiGe BiCMOS process for terahertz wireless system [J]. IEEE transactions on terahertz science and technology, 2022, 12(1): 1 – 12. DOI: 10.1109/TTHZ.2021.3099057

[5] LI H B, CHEN J X, HOU D B, et al. A 250 GHz differential SiGe amplifier with 21.5 dB gain for sub-THz transmitters [J]. IEEE transactions on terahertz sci-

ence and technology, 2020, 10(6): 624 – 633. DOI: 10.1109/TTHZ.2020.3019361

[6] ZHOU P G, CHEN J X, YAN P P, et al. A 150 GHz transmitter with 12 dBm peak output power using 130 nm SiGe: C BiCMOS process [J]. IEEE transactions on microwave theory and techniques, 2020, 68(7): 3056 – 3067. DOI: 10.1109/TMTT.2020.2989112

[7] LI Z K, CHEN J X, LI H B, et al. A 220 GHz power amplifier with 22.5 dB gain and 9 dBm Psat in 130 nm SiGe [J]. IEEE microwave and wireless components letters, 2021, 31(10): 1166 – 1169. DOI: 10.1109/LMWC.2021.3105611

[8] YU J Y, CHEN J X, ZHOU P G, et al. A 211-to-263 GHz dual-LC-tank-based broadband power amplifier with 14.7 dBm PSAT and 16.4 dB peak gain in 130 nm SiGe BiCMOS [J]. IEEE journal of solid-state circuits, 2023, 58(2): 332 – 344. DOI: 10.1109/JSSC.2022.3192043

[9] BÜCHER T, GRZYB J, HILLGER P, et al. A broadband 300 GHz power amplifier in a 130 nm SiGe BiCMOS technology for communication applications [J]. IEEE journal of solid-state circuits, 2022, 57(7): 2024 – 2034. DOI: 10.1109/JSSC.2022.3162079

[10] LI X C, CHEN W H, LI S Y, et al. A high-efficiency 142 – 182 GHz SiGe BiCMOS power amplifier with broadband slotline-based power combining technique [J]. IEEE journal of solid-state circuits, 2022, 57(2): 371 – 384. DOI: 10.1109/JSSC.2021.3107428

[11] ATESAL Y A, CETINONERI B, CHANG M, et al. Millimeter-wave wafer-scale silicon BiCMOS power amplifiers using free-space power combining [J]. IEEE transactions on microwave theory and techniques, 2011, 59(4): 954 – 965. DOI: 10.1109/TMTT.2011.2108313

[12] YORK R A. Some considerations for optimal efficiency and low noise in large power combiners [J]. IEEE transactions on microwave theory and techniques, 2001, 49(8): 1477 – 1482. DOI: 10.1109/22.939929

[13] PARK D W, UTOMO D R, LAM B H, et al. A 230 – 260 GHz wideband and high-gain amplifier in 65 nm CMOS based on dual-peak Gmax-core [J]. IEEE journal of solid-state circuits, 2019, 54(6): 1613 – 1623. DOI: 10.1109/JSSC.2019.2899515

[14] SARMAH N, AUFINGER K, LACHNER R, et al. A 200 – 225 GHz SiGe power amplifier with peak Psat of 9.6 dBm using wideband power combination [C]//The 42nd European Solid-State Circuits Conference. IEEE, 2016: 193 – 196. DOI: 10.1109/ESSCIRC.2016.7598275

[15] EISSA M H, MALIGNAGGI A, KISSINGER D. A 13.5 dBm 200 – 255 GHz 4-way power amplifier and frequency source in 130 nm BiCMOS [J]. IEEE solid-state circuits letters, 2019, 2(11): 268 – 271. DOI: 10.1109/LSSC.2019.2951689

[16] LI X C, CHEN W H, ZHOU P G, et al. A 250 – 310 GHz power amplifier with 15 dB peak gain in 130 nm SiGe BiCMOS process for terahertz wireless system [J]. IEEE transactions on terahertz science and technology, 2022, 12 (1): 1 – 12. DOI: 10.1109/TTHZ.2021.3099057

[17] BUCHER T, GRZYB J, HILLGER P, et al. A 239 – 298 GHz power amplifier in an advanced 130nm SiGe BiCMOS technology for communications applications [C]//The 47th European Solid State Circuits Conference (ESSCIRC). IEEE, 2021. DOI:10.1109/ESSCIRC53450.2021.9567853

[18] YOON D, SEO M G, SONG K, et al. 260 GHz differential amplifier in SiGe heterojunction bipolar transistor technology [J]. Electronics letters, 2017, 53 (3): 194 – 196. DOI: 10.1049/el.2016.3882

[19] KUCHARSKI M, NG H J, KISSINGER D. An 18 dBm 155 – 180 GHz SiGe power amplifier using a 4-way T-junction combining network [C]//The 45th European Solid State Circuits Conference (ESSCIRC). IEEE, 2019: 333 – 336. DOI: 10.1109/ESSCIRC.2019.8902847

[20] STÄRKE P, CARTA C, ELLINGER F. High-linearity 19 dB power amplifier for 140 – 220 GHz, saturated at 15 dBm, in 130 nm SiGe [J]. IEEE microwave and wireless components letters, 2020, 30(4): 403 – 406. DOI: 10.1109/LMWC.2020.2978397

[21] PHILIPPE B, REYNAERT P. 24.7 A 15 dBm 12.8% PAE compact D-band power amplifier with two-way power combining in 16 nm FinFET CMOS [C]//IEEE International Solid State Circuits Conference (ISSCC). IEEE, 2020: 374 – 376. DOI: 10.1109/ISSCC19947.2020.9062920

[22] BAMERI H, MOMENI O. An embedded 200 GHz power amplifier with 9.4 dBm saturated power and 19.5 dB gain in 65 nm CMOS [C]//Proceedings of 2020 IEEE Radio Frequency Integrated Circuits Symposium (RFIC). IEEE, 2020: 191 – 194. DOI: 10.1109/RFIC49505.2020.9218441

Biographies

CHEN Jixin received his BS degree in radio engineering from Southeast University, China in 1998, and MS and PhD degrees from Southeast University in 2002 and 2006, respectively, all in electromagnetic field and microwave techniques. Since 1998, he has been with the Sate Key Laboratory of Millimeter Waves, Southeast University, and is currently a professor of School of Information Science and Engineering. His current research interests include microwave and millimeter-wave circuit design and MMIC design. He has authored and co-authored more than 100 papers and presented invited papers at ICMMT2016, IMWS2012, and GSMM2011. He is the winner of 2016 Keysight Early Career Professor Award. He has served as the TPC Co-chair of HSIC2012, UC-MMT2012, LOC Co-chair of APMC2015, Session Co-chair of iWAT2011, ISSSE2010, and APMC2007, and a reviewer for *IEEE MTT* and *IEEE MWCL*.

ZHOU Peigen (pgzhouseu@seu.edu.cn) received his BS degree in radio engineering in 2015 and PhD degree in electromagnetic field and microwave techniques in 2020, both from Southeast University, China. Since 2021, he has been with the State Key Laboratory of Millimeter Waves, Southeast University, and is currently an assistant researcher of the School of Information Science and Engineering. His research interests include silicon-based millimeter-wave/THz on-chip wireless communication/radar phased-array transceivers.

YU Jiayang is currently working toward his PhD degree in the School of Information Science and Engineering, State Key Laboratory of Millimeter Waves, Southeast University, China. His current research interests include millimeter-wave/terahertz integrated circuits for radar and high speed communication.

LI Zekun received his BS degree from the School of Information Science and Engineering, Southeast University, China in 2018, where he is currently pursuing his PhD degree. His current research focuses on silicon-based millimeter-wave and terahertz integrated circuits and systems for high-speed wireless communication and radar.

LI Huanbo received his BS degree from the School of Information Science and Engineering and PhD degree in electromagnetic field and microwave techniques from Southeast University, China in 2016 and 2021, respectively. His current research focuses on silicon based mm-wave and terahertz integrated circuits and systems for high-speed wireless communication and radar imaging.

PENG Lin received his BS degree in information engineering and MS degree in electromagnetic field and microwave techniques from Nanjing University of Science and Technology, China in 2004 and 2006, respectively. Since 2006, he has been with ZTE Corporation, where he is mainly engaged in wireless communications. His current research interests include beyond-5G and 6G technologies, millimeter-wave and terahertz communication, and intelligent reflecting surface for wireless applications.

The 1st Youth Expert Committee
for Promoting Industry-University-Institute Cooperation

ZTE COMMUNICATIONS
中兴通讯技术(英文版)