Research
6G Requirements, Vision, and Enabling Technologies—Review

Federated Learning for 6G: Applications, Challenges, and Opportunities

Zhaohui Yang [a], Mingzhe Chen [b,*], Kai-Kit Wong [a], H. Vincent Poor [b], Shuguang Cui [c,d,*]

[a] Department of Electronic and Electrical Engineering, University College London, London WC1E 6BT, UK
[b] Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA
[c] Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong, Shenzhen 518172, China
[d] School of Science and Engineering and Future Network of Intelligence Institute, The Chinese University of Hong Kong, Shenzhen 518172, China

ARTICLE INFO	ABSTRACT
	Standard machine-learning approaches involve the centralization of training data in a data center, where centralized machine-learning algorithms can be applied for data analysis and inference. However, due to privacy restrictions and limited communication resources in wireless networks, it is often undesirable or impractical for the devices to transmit data to parameter sever. One approach to mitigate these problems is federated learning (FL), which enables the devices to train a common machine learning model without data sharing and transmission. This paper provides a comprehensive overview of FL applications for envisioned sixth generation (6G) wireless networks. In particular, the essential requirements for applying FL to wireless communications are first described. Then potential FL applications in wireless communications are detailed. The main problems and challenges associated with such applications are discussed. Finally, a comprehensive FL implementation for wireless communications is described. © 2022 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

1.1. Motivation

Owing to the significant growth in data traffic, machine learning has gained a considerable amount of attention and are anticipated to be vital in the development of sixth generation (6G) wireless networks [1]. Centralized machine learning methods require the collection of training samples at a centralized parameter server. Hence, transmitting a large amount of data samples can cause significant transmission delay. Meanwhile, user privacy is not guaranteed in standard centralized machine learning approaches. However, low latency and privacy requirements are important in many newly emerging applications, such as unmanned aerial vehicles, extended reality (XR) services, and autonomous driving. Therefore, using centralized machine-learning approaches to optimize these emerging applications is inappropriate. Meanwhile, due to limited communication resources, it is often impossible for all edge devices to upload their data to a parameter server for centralized machine learning.

For these reasons, it is desirable to introduce distributed learning algorithms which enables devices to cooperatively build a unified learning model with local training. One of the most promising distributed machine-learning frameworks is federated learning (FL) [2–19]. In FL, edge devices collaboratively build a learning model by transmitting only locally learned models to a base station (BS) while keeping the local training data, as illustrated in Fig. 1 [20]. Note that FL can also be performed without a parameter server, where each device can communicate with neighboring devices [21]. Since the data center cannot access the local data sets at the user level, FL can improve the data privacy of the users.

In wireless communications, implementation of FL has the following advantages [15,22]: ① Exchanging local machine learning model parameters instead of voluminous training data can save energy and consume less wireless resources; ② training machine learning model parameters locally can effectively reduce transmission latency; ③ FL can help improve data privacy since the training data remains at end-user devices and only the local learning model parameters are uploaded; and ④ using different learning processes to train multiple classifiers from edge datasets increases the possibility of achieving higher learning performance.

FL can be utilized to solve complex convex and nonconvex problems in various use cases, such as interference cancelation, network control, resource allocation, and user grouping. In

* Corresponding authors.
E-mail addresses: mingzhec@princeton.edu (M. Chen), shuguangcui@cuhk.edu.cn (S. Cui).
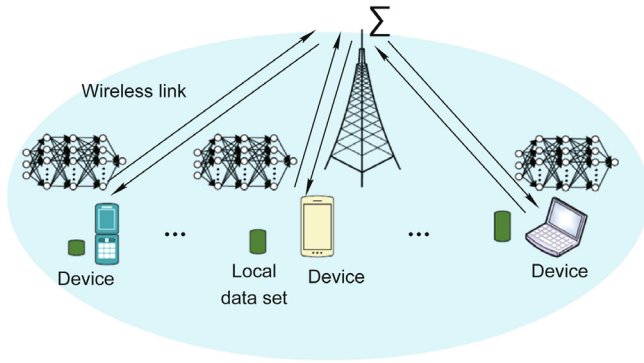
Fig. 1. FL over a wireless communication network.

addition, FL enables users to cooperatively learn unified prediction models while storing the collected data on their devices for wireless environment analysis, user movement prediction, and user identification. Based on the predicted results, the BS can efficiently allocate wireless resources to the devices.

1.2. Types of FL

There are certain common types of FL: federated reinforcement learning (FRL), federated supervised learning (FSL), FL for generative adversarial networks (GANs) (unsupervised learning), and FL for contrastive learning (self-supervised learning). In Refs. [23,24], the goal of FRL is to enable wireless devices to remember what they and the other wireless devices have learned. FRL can be used in cases where multiple wireless devices make decisions in different environments. In FRL, each wireless device builds a learning network with the help of other wireless devices.

(1) Initially, one edge device obtains its private model through reinforcement learning (RL) in its own environment. The edge device uploads its private model to the BS as a shared model.

(2) Then, the wireless devices download the common shared strategy model from the BS as the initial model for RL. Wireless devices obtain their own private learning networks through RL in new environments. When the training is completed, the wireless devices upload their private learning networks to the BS.

(3) At the BS, the private learning networks are integrated into the shared model, which produces a new shared model. The new shared model will be utilized by any other wireless device. The wireless devices will also transmit private learning networks to the data center to calculate the shared model.

The FSL technique builds a uniform learning model by iteratively updating information between the BS and wireless devices, where the local private data are fully labeled. In FSL, the devices can remember what they have learned via local learning model parameters, and the local learning model is built with the help of other devices via global model aggregation. The FSL scheme contains three procedures for each iteration: local computation at the wireless device, local FSL model parameter transmission from each wireless device, and global model generation and broadcasting at the BS.

- Every wireless device needs to compute the result by using its fully labeled dataset locally.
- All wireless devices transmit local prediction results to the center through wireless channels in the uplink.
- The BS obtains the prediction model parameters and transmits the unified prediction learning model coefficients to all wireless devices.

1.3. Relevant surveys and contributions

There are some interesting surveys on the use of FL in networks, such as those in Refs. [25–30]. The unique characteristics and challenges of FL are discussed in Ref. [25], which also provices a summary of the current approaches and outlined multiple directions for future research. Ref. [26] introduces the FL implementation challenges and reviews the current approaches to these challenges. In Ref. [27], the authors describe the challenges of machine learning systems configured on edge computer networks. For RL, the authors in Ref. [28] propose the integration of deep RL techniques and FL schemes with emerging edge systems for unified optimization of wireless communication, edge computing, and cached resources. Ref. [29] explores the key parameters of edge machine learning and various wireless architectural splits for wireless communications. Practical aspects of FL are surveyed in Ref. [30], including applications, usage scenarios, and hardware platforms. The overview of some studied about FL in wireless communications is given in Table 1 [25–30].

We aim to gather the contributions that highlight the key challenges of applying FL techniques to wireless networks. Particularly, our objectives are threefold: to provide a comprehensive description of the FL algorithm, to identify the key problems in wireless communication systems that can be solved using FL methods, and to point out the emerging FL applications in wireless communication.

2. Performance and requirements for FL

2.1. Performance metrics

Fig. 2 shows a procedure for implementing FL in a wireless communication network. The FL scheme contains three procedures at each step: local iteration at every device (with multiple local times), uploading of locally computed FL model parameters, and global model aggregation and re-broadcasting at the center. The local iteration procedure signifies that every device computes its local FL parameters by using its local data and the received global FL parameters. There are four main performance metrics for FL: delay, energy, reliability, and massive connectivity.

(1) **Delay.** According to Fig. 3, the delay of FL includes the local iteration delay of edge devices, uplink communication delay, BS aggregation delay, and downlink transmission delay. The delay of FL is also determined by the number of iterations FL needs for convergence [31]. Considering the tradeoff between the local computation delay and communication delay, it is crucial to minimize the delay for implementing FL via joint transmission and computation optimization.

(2) **Energy.** Because the total energy of each wireless device is limited both the transmission energy and local computation energy affect the FL procedure. The local computation energy of a

Table 1
Overview of some studied about FL in wireless communications.

| Subject	Contributions	Reference
FL	Introductory tutorial on characteristics and challenges of FL	[25]
FL	Challenges of FL implementation	[26]
Edge machine learning	Challenges of machine learning systems at the edge networks	[27]
FL	FL and RL for optimizing mobile edge computing and caching	[28]
Edge machine learning	Edge machine learning architectures	[29]
FL	FL applications and use-cases	[30]
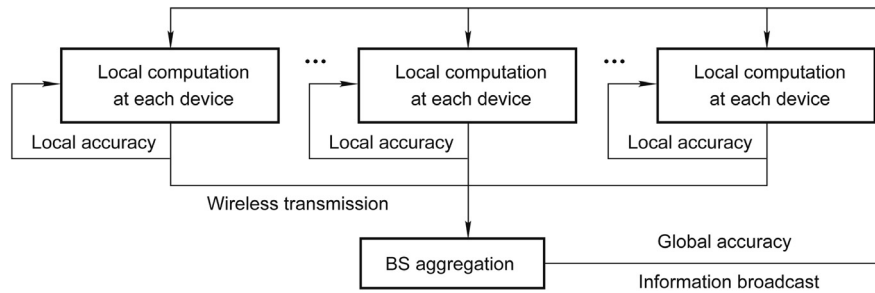
Fig. 2. FL procedures over wireless networks.
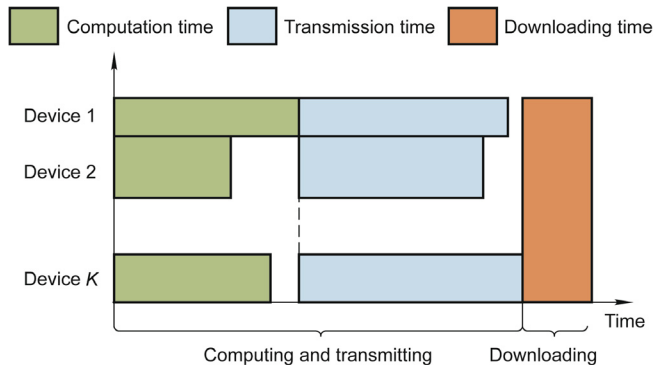
Fig. 3. Time performance of FL over wireless networks. *K*: the total number of all devices.

device depends on the number of iterations needed for the local computation procedure at that device, while the transmission energy is related to the number of iterations during the implementation of FL.

(3) **Reliability**. End-user devices must transmit their training parameters through wireless links to the aggregating device. Owing to the limited wireless resources (such as bandwidth) and inherent unreliability of wireless links, training errors may be introduced. In particular, symbol errors caused by the unreliable characteristics of wireless channels and limited resources will affect the performance and success rate of FL iterations [32,33]. The overall performance of the FL algorithms and convergence speed are affected by these factors.

(4) **Massive connectivity**. To satisfy the low latency requirement of FL, we must obtain the data from numerous edge devices efficiently and rapidly using wireless communications. However, owing to the large number of devices, traditional interference avoidance channel access schemes are infeasible because they usually cause excessive delays. To overcome this challenge, an emerging approach is over-the-air computation, which can gather wireless data quickly by using the superposition nature of wireless transmission [34,35]. Although over-the-air computation has some attractive advantages, it is not compatible with existing digital wireless communication systems. In addition, scheduling only a fraction of all devices at each round of FL uploading is a promising alternative [36,37].

2.2. Potential to meet 6G requirements

It is envisioned that 6G networks will need to accommodate 125 billion wireless devices by 2030. As a result, it is crucial to design an intelligent signal and data processing system to allow edge learning to occur. As a key technology, FL has the potential to meet the following anticipated 6G requirements [1].

(1) **Massive ultra-reliable low latency communications (mURLLCs)**. Because of the expected growth in the number of 6G wireless end-user devices, the fifth generation (5G) ultra-reliable low latency communication (URLLC) metrics must be updated to mURLLC. With FL, multiple edge computing units can be used to cooperatively learn a shared network model, which can decrease service delay and provide high reliability [38,39].

(2) **Scalable architecture**. Unlike centralized intelligence, edge intelligence, such as FL, is built in a distributed manner, which includes many edge servers with computing and communication capabilities. To serve a large number of end-user devices in future 6G communications, it is important to provide a decomposable and scalable architecture to allow simultaneous computing among multiple edge servers. Such architectures are expected to play an important role in the emerging wireless communication services and applications.

(3) **Human-centric services**. Unlike the rate-reliability-latency metrics in 5G, 6G is anticipated to involve human-centric services, which will require quality of experience levels related to the physical movement of the users. FL can be used to predict the movements and gestures of users, and the BS can use the predicted results to improve the quality of user experience.

3. FL for wireless communications: Motivation behind applications

Machine learning approaches can use data analytics to estimate the state of wireless networks and find connections between optimized variables and objective functions online, which reduces the computational complexity of solving nonconvex optimization problems in wireless systems. In addition, machine learning is powerful because it can optimize problems in which the problem description is unknown. However, given that multicell networks require global channel state information (CSI), centralized learning algorithms may require BSs to continuously upload their obtained data to a centralized processing server, which leads to high network overhead and significant delays. Consequently, using a centralized learning algorithm for resource management or network control may require many iterations to converge. As a result, traditional machine learning algorithms with centralized training may not be able to handle resource allocation, signal detection, and user behavior prediction problems in future 6G networks. As a more practical alternative, FL can enable users or BSs to manage the resources in a distributed manner and locally analyze collected data. Section 3.1 reports a summary of driving FL applications for wireless problems, and Sections 3.2–3.5 describe four applications where FL can be used to solve various wireless network problems.

3.1. Driving FL applications for wireless problems

(1) **Resource management**. Spectral efficiency and connectivity optimization of multicell networks typically leads to nonconvex

resource allocation problems. Conventional algorithms, such as matching theory can be used to solve such nonconvex resource allocation problems. However, the complexity is high. Therefore, there is a need to introduce new FL algorithms that can be used to address a variety of resource management problems, such as distributed power control for multi-cell networks, joint user association and beamforming design, as well as dynamic user clustering.

For multi-cell power control, as depicted in Fig. 4, FRL enables each BS to settle the connection between the power control schemes and utility values to find the globally optimal resource allocation scheme. In FRL, the BSs on a connected network process data locally by minimizing small optimization problems and exchange the local results among their neighbors to arrive at a global solution.

Furthermore, FRL can be used for dynamic user clustering, where end-users individually learn the clustering parameters by RL, and the BS builds unified clustering parameters based on the received clustering parameters from all end-users.

(2) **User behavior prediction**. Due to the various quality-of-service requirements of users, user behavior prediction is crucial for the optimization of wireless network performance.

User behaviors, such as mobility patterns, can be predicted using FL, where each user performs a local FL algorithm to compute its local model using private user behavior data and uploads the obtained model to the center. The center then generates and broadcasts aggregated FL parameter coefficients to all users. Based on the mobility predictions, in the uplink, the users can dynamically choose a subchannel and the users that occupy the same subchannel can perform non-orthogonal multiple access (NOMA) or full duplex to upload their models. In contrast, in the downlink, the BS can dynamically allocate multiple subchannels to several users.

The quality-of-service of users can be predicted using FL, where each BS uses the FL algorithm based on stored information such as requested data, device type, and so forth, and all BSs transmit the FL model results to a server to obtain a unified FL model.

(3) **Channel estimation and signal detection**. Channel estimation and signal detection are major challenges because of the random features of wireless channels in wireless communication networks. For downlink systems, FL algorithms are used for channel estimation and multi-user detection, where each user performs an FL scheme for channel estimation and signal detection and sends the locally obtained FL parameters to the center which computes the unified FL model. To enable channel detection via FL, each user can perform the same channel detection task; for example, obtaining CSI from the BS to a passive relay. The training convergence time scale and required number of datasets are suitable for fitting within the coherence duration, as only one common channel needs to be predicted. For multicell uplink systems, multi-user signals can be detected by iteratively transmitting individual FL model parameters from all BSs to a server and broadcasting the unified FL model parameters from the server back to all BSs. Furthermore, FL algorithms can be utilized to automatically design the BS codebooks and decoding strategy of users to minimize the bit error rate, where users upload the learned result to the corresponding BSs and the BSs forward their unified learned result to a server.

3.2. Reconfigurable intelligent surfaces

Reconfigurable intelligent surface (RIS) based wireless communication systems are regarded as a potential technology for improving the energy efficiency of communication networks [40–51], as shown in Fig. 5. An RIS is mainly composed of numerous high-efficiency hardware components, which can change the phase of the input signal. In RIS-based wireless communication systems, the RIS is usually managed by the BS via a backhaul link between the BS and RIS to determine the properties of the incident waves. Thus, the wireless environment can be controlled for various design objectives using the RIS. The RIS serves as a mirror that will not require any digital operations. Therefore, if deployed properly, RISs are expected to reduce energy consumption compared to existing amplify-and-forward (AF) relays [52–54]. However, it is challenging to jointly optimize the active beamforming at the BS and passive phase beamforming at and RIS owing to the unique constraints on the RIS coefficient matrix phases. To deal with complicated and varying electromagnetic (EM) environments and nonlinear problems of communication systems that are difficult to solve mathematically, an FL algorithm can be used as a practical alternative.
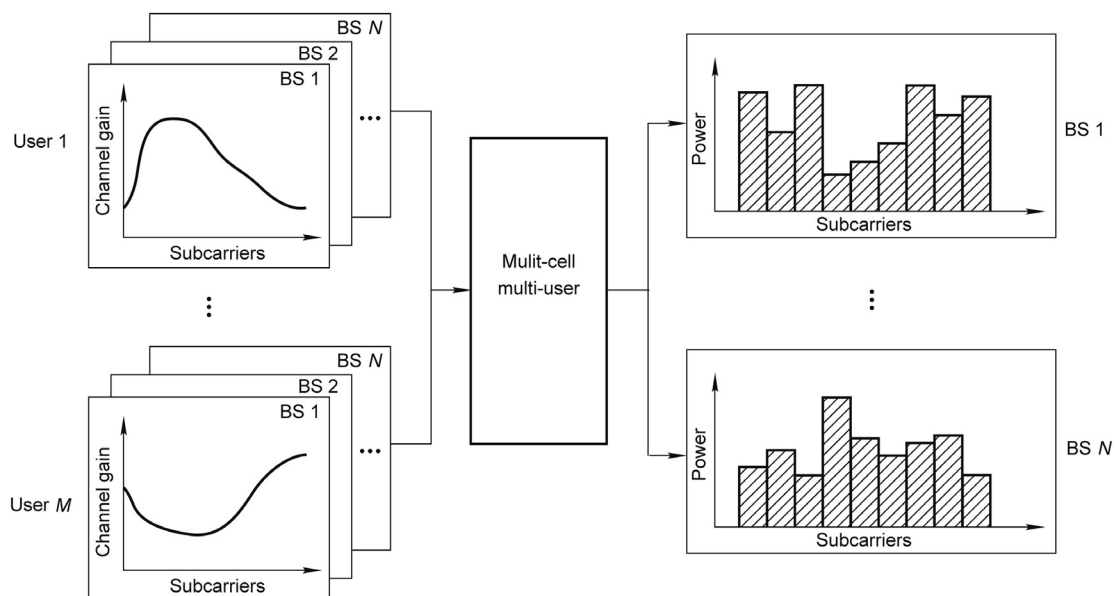

Fig. 4. Multi-cell power control scheme. *M*: the total number of all users; *N*: the total number of all BSs.
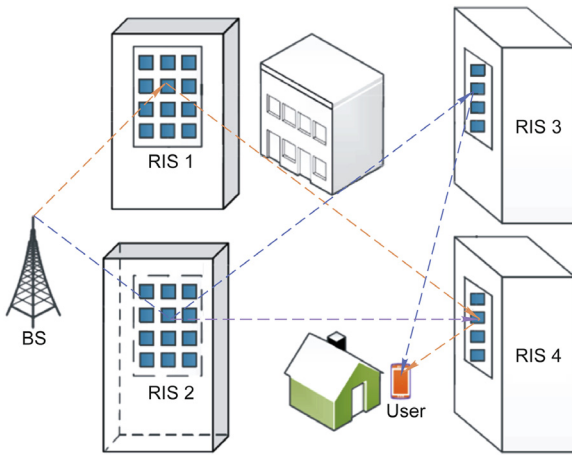
Fig. 5. Example of an RIS-enhanced communication network.

(1) **CSI detection**. In an RIS-based system, to fully exploit the advantages of the architecture, multiple high-efficiency technologies, such as energy-saving designs, resource allocation, and active and passive joint beamforming, are required. Note that all the above designs depend on perfect knowledge of CSI between the RIS and BS, and between the user and RIS. However, when the RIS is not built on a radio frequency (RF) chain or sensor, the RIS enhanced system cannot accurately estimate the CSI. To this end, it is meaningful to use FL for CSI detection in RIS-assisted wireless communications.

The FL-based model training approach can be used in RIS-assisted massive multiple-input multiple-output (MIMO) systems [55]. The FL approach mainly includes three steps: data gathering, sample training, and task prediction. In the first step, every user collects its local training dataset, where the pilot sequence is the input, and the received signal is the output. Then, each user computes the updated model by utilizing its own local data samples, and the BS generates a global model after receiving the updated models from all users. In the last step, each user estimates its own channel by inputting the received pilot data into the training model.

(2) **Distributed joint passive and active beamforming**. In an RIS-assisted wireless communication system, the phase of each element in the RIS can be controlled to improve the performance of RIS-assisted wireless communication systems. In contrast to conventional communications, it is important to optimize both passive beamforming (phase shift matrices at the RIS) and active beamforming (beamforming at the multi-antenna transmitter) [56,57]. Deep learning (DL) has been applied to solve complicated joint passive and active beamforming to optimize the reflection matrix of RIS components [58]. In practice, multiple RISs can be utilized to overcome severe signal congestion between a user and the BS, thereby achieving better service coverage, which is similar to a multi-hop relay system. A multi-hop RIS auxiliary communication scheme was proposed in Ref. [59] to deal with the increase in coverage and severe pathloss in the terahertz frequency band, where a hybrid optimization of phase shift matrices and transmitted beamforming at the BS is obtained by an advanced RL. Owing to the high complexity of using a centralized RL, FRL can be utilized to solve the joint passive and active beamforming problem, where all users can individually optimize their phase shift matrices and transmit beamforming via RL, and the BS transmits the unified learning model back to all users.

(3) **Phase shift prediction**. Owing to the randomness of wireless communication channels, the RIS phase-shift matrices must be determined as the wireless channel changes. By exploiting the time-correlated property of channel fading, the phase-shift matrices of the RIS can be predicted via FL. To predict the phase shift, each user uses a long short-term memory (LSTM) network for the prediction of future CSI and phase shift matrices using a local data set, while the BS aggregates the received results from all users.

3.3. Semantic communication

Semantic communication is similar to the communication that takes place in the human brain, where the difference between the meaning of transmitted symbols and that of recovered ones is correlated [60]. This correlation can be useful for joint encoding and decoding when the bandwidth of the system is limited, or the bit error rate is high for some typical communication systems.

(1) **Channel encoder and decoder design**. Using a semantic communication technique that enables the devices to transmit semantic information to the server, rather than traditional bits or symbols, can effectively improve the network bandwidth utility. However, the semantic communication model requires training data from multiple distributed devices, which incurs very substantial communication costs for data transmission. To solve this problem, an FL-based DL-enabled semantic communication can be used for channel encoder and decoder design. First, a DL model can be used to extract semantic information from text or audio with robustness against noise. Then, in an FL approach, the end-user devices and server obtain practicable DL models with the server aggregating the locally trained models and sending back the unified model to the devices.

(2) **Distributed semantic communication for Internet of Things (IoT)**. Emerging technologies, such as smart connectivity, IoT, and machine-to-machine (M2M) networks, require intelligent communication between different ends, such as humans and machines. For these applications, intelligent communication depends on the background and interface language models [61]. In addition, there are always numerous devices in IoT networks. These factors motivate the design of a distributed semantic communication for IoT networks with FL. The distributed semantic communication with FL includes three steps. In the first step, the center computes the semantic communication model using DL. In the second step, the center transmits the trained DL model to each device. In the third step, each user obtains the semantic features through received broadcast information. Then, each user uploads the semantic features to the BS, then, the BS calculates the semantic communication model accordingly.

3.4. Extended reality

XR refers to all computer-generated graphics in real and virtual environments that consist of mixed reality (MR), augmented reality (AR), and virtual reality (VR). Deploying XR over wireless communication networks is an essential step for realizing XR applications [1]. Owing to the seamless and immersible requirements, it is important to introduce wireless communication technologies that can meet the stringent quality-of-service requirements, such as high data rate and ultra-low latency. For XR allocation over wireless communications, the location and orientation information need to be sent to the BS, which constructs 360° images for users based on the received information.

(1) **User movement prediction**. In a wireless XR network, user body movements can heavily influence wireless resource allocation and network management [62]. FL is effective in predicting user actions and movements, which are used to deal with user movement challenges. Based on the predicted movements and actions, the BS can improve the generated XR image and optimize the wireless resource allocation of XR users.

(2) **Resource allocation.** FL can be used to design self-organizing schemes for solving dynamic resource management problems for XR networks [63]. Specifically, FL can be used to dynamically optimize wireless resources and construct the structure of XR images based on the wireless environment.

3.5. Non-orthogonal multiple access

NOMA is envisioned as a promising technique for next-generation wireless communication networks [64]. By serving several users on the same time and frequency resource, compared to the alternative orthogonal multiple access (OMA) technology, NOMA can expand the number of connected users, improve user fairness, and improve spectral efficiency. Recently, significant research effort has been focused on various challenges of NOMA implementations [65–67], including modeling, performance analysis, signal processing, and emerging NOMA applications, such as heterogeneous networks (HetNets), cognitive radio networks, and millimeter wave (mmWave) communications. The non-orthogonal resource allocation nature of NOMA necessitates the introduction of novel models and algorithms to address several challenges, including joint user clustering and resource allocation for devising a scalable multicell NOMA design, advanced channel estimation and signal detection for large-scale NOMA networks, and dynamic user behavior prediction in NOMA-based mobile networks.

Owing to the non-orthogonal resource allocation property, intra-cell interference always exists in NOMA networks, which usually leads to nonconvex resource allocation problems. Traditional optimization methods, which are used to solve the nonconvex problems for optimizing the performance of NOMA networks, mostly operate offline with extremely high computational complexity and depend on precise CSI [68–71]. Big data analysis can be used to estimate the state of the wireless network and find the relationship between the optimized variable and the objective function online via machine learning schemes [72–75] which minimize the computational complexity for solving the nonconvex problems in NOMA. However, given that multicell NOMA needs global CSI, a centralized learning algorithm may require the BSs to continuously upload their obtained data to a centralized processing server, which leads to a high network overhead and significant delays. In addition, in NOMA, each subcarrier can be occupied by multiple users. Consequently, using a centralized learning algorithm for resource management or network control may require many iterations to converge. Therefore, the conventional central machine learning methods described in Refs. [76–79] cannot handle resource allocation, signal detection, and user behavior prediction problems in NOMA. For NOMA, FL has two important applications: ① the complex convex and nonconvex optimization problems that can be solved by FRL, which include resource allocation, interference mitigation, user grouping, and network control, and ② FSL which can enable edge users to cooperatively obtain a unified learning parameter while protecting their obtained data on their devices for CSI prediction and user detection.

(1) **Resource management in NOMA**. With the superposition coding technique at the transmitter and successive interference cancellation (SIC) at the receiver, NOMA can yield higher spectral efficiency compared to OMA [80,81]. Moreover, NOMA can take advantage of user differences in the power domain to provide services for multiple users connected to the same resource. The power domain characteristics of NOMA can help support massive NOMA connections and meet a range of quality services.

The spectral efficiency and connectivity optimization of NOMA typically leads to nonconvex resource allocation problems, which are optimized using conventional algorithms [65]. Therefore, there is a need to introduce new distributed learning techniques that can

be used to address many resource management challenges, such as distributed power control for multicell NOMA [70], joint user association and beamforming design [67], and dynamic user clustering [82]. For multi-cell power control, FRL enables each BS to build a connection between the power control schemes and utility functions to find an optimal power control scheme. FRL can also be used to study user association and beamforming of a multi-antenna NOMA network [83]. Furthermore, FRL is used for dynamic user clustering in NOMA, where users individually learn the clustering parameters by RL, and the BS builds unified clustering parameters based on the received clustering parameters from all users.

(2) **Signal detection and channel estimation in NOMA**. Signal detection and channel estimation in NOMA are major challenges owing to error propagation in SIC for NOMA networks. FSL algorithms can be utilized for channel estimation and multi-user detection in downlink NOMA networks, where each user executes a supervised learning (SL) algorithm for signal detection and channel estimation of multiple users and sends its local FL model coefficients to the BS that will generate the global FL model. As reported in Ref. [84], FSL can detect multi-user signals in multi-cell uplink NOMA networks by iteratively transmitting individually learned model parameters from all BSs to a server and broadcasting the unified learning model parameters from the server back to all BSs. Furthermore, FSL can be used to automatically design the codebook of BSs and decoding strategy of users in code-domain NOMA networks to minimize the bit error rate [85], where users upload the learned result to the corresponding BS, which forwards their unified learned result to a server.

(3) **User behavior prediction in NOMA**. Owing to the heterogeneous quality-of-service requirements of users in NOMA, where devices in the same group may have diversified channel values and quality-of-service requirements, user behavior prediction is crucial for the implementation of NOMA networks. To predict certain user behaviors, such as mobility information, each user in the FSL scheme executes an SL algorithm to train the learning model, utilizing its own user behavior data, and uploads the obtained local model to the BS via NOMA. Then, the BS generates and broadcasts the unified learning model coefficients to all users using NOMA. Based on the mobility pattern predictions, the users can dynamically choose subchannels to upload data in the uplink, the BS dynamically allocates multiple subchannels to multiple users in the downlink, and multiple users that occupy the same subchannel can perform NOMA. For multiple BSs to predict the quality-of-service of users [86] in FSL, each BS uses an SL algorithm based on its stored data set, and device type. All BSs transmit the learning model results to a server via NOMA to obtain a unified FL model.

4. Research directions and open problems

4.1. Research directions and challenges

FL ensures that the resource allocation or behavior prediction problem can be solved in a distributed manner for wireless networks. The utilization of FL for wireless networks has the following five main directions and challenges:

(1) **Scalability**. FL should be scalable because an increased number of computers or processors may offset the increased amount of data and provide a solution to the complexity and memory issues in large-scale learning networks. For a large-scale learning network, it is important to investigate issues related to distributed training.

(2) **Privacy and security**. In FL, the raw data set for each user can be protected because only the locally obtained FL model is transmitted to the center. However, it is also possible for an

eavesdropper to conduct approximate reconstruction of the original data, particularly when the local and global model coefficients cannot be protected [87]. In addition, the local FL model may leak private information. In FL, privacy can be categorized into two types: global and local. The model generation at each iteration is invisible to all unknown devices except the BS in global privacy, and the model aggregation at each iteration is confidential to all unknown third parties and the BS in local privacy.

(3) **Asynchronous communication**. FL involves information exchange between the wireless devices and BS. Synchronous communication methods are simple, but they can introduce stragglers among devices. An attractive way to alleviate laggards in a heterogeneous environment is an asynchronous solution. Although asynchronous server parameters in the distributed data center are successful in dealing with stragglers, assumptions of bounded delay may be impractical in federated schemes.

(4) **Non-independent identically distribution (Non-IID) devices**. When training a joint model from differently distributed data across devices, challenges arise both in terms of data modeling and analyzing the convergence trend of the relevant training process [88]. One key aspect of FL is coping with heterogeneous settings and competing and distributed decision-making environments.

(5) **Joint communication and computation design**. To deploy FL in a wireless communication network, each device needs to transmit its multimedia data or local training results through an unreliable wireless link. It is important to consider the multicell and multi-hop FL implementations for real scenarios [89]. In addition, the performance of FL learning schemes is degraded by limited radio resources. Thus, it is important to consider the joint management of communications and computing resources to achieve efficient and effective FL.

4.2. Open problems and future directions

This section presents several open problems based on the above issues to reveal future research opportunities. Although FL has been extensively researched, there are still several key issues to be studied regarding wireless communication and FL.

(1) **Convergence**. Because of the limited wireless resources in communication networks, only a fraction of users can be activated in each learning step to upload their local model parameters to the center. However, owing to the diversity of training data samples of different users, the center would like to involve the local FL models of all users to determine the best overall global FL model. So, user upload scheduling is a key issue and affects the FL performance and convergence time. Many studies of FL convergence are based on the assumption of a convex loss function [90,91]. However, the loss functions for many learning problems are non-convex, and there are challenges associated with investigating the convergence rate of FL with non-convex loss functions [92]. Moreover, there are still some key problems for the FL convergence rate as well, even for convex loss. For example, there is a need for an exact/more accurate convergence formulation with fewer assumptions and approximations [90] in order to be consistent with real FL experiment data. Although there are some studies in this area, most of them are based on convex loss functions. Furthermore, owing to the heterogeneous property of the quality-of-service, it is necessary to simultaneously conduct multi-task FL. In addition, for large-scale systems, multicell and multi-hop FL should be considered, which both require greater insights into the FL convergence analysis. Moreover, a particular challenge is to study the mobility of wireless devices for FL convergence. Owing to such mobility, the channel gains between the devices and BS are dynamically changing; thus, it is possible that some devices will exit the FL process

owing to serious CSI, which affects the convergence of the entire FL process.

(2) **Privacy and security**. There are a number of open problems associated with privacy and security in FL: privacy protection for each user, privacy preservation of the BS, and security for the entire FL algorithm. Regarding the privacy protection for each user and the BS, a promising approach is to use differential privacy, which introduces a tradeoff between privacy and FL performance [93].

To ensure the security of the entire FL algorithm, traditional methods such as encryption can be considered, as well as more recent developments such as secure multi-party computation and physical layer security, which can provide security in situations (such as massively deployed IoT) where more conventional methods cannot be applied.

(3) **Performance evaluation**. One of the main challenges is to investigate the effects of communication bandwidth on FL delay performance. Although the computing resources of mobile phones are becoming increasingly powerful, the bandwidth of wireless communication has not increased significantly. Consequently, the bottleneck has shifted from computing to communication capabilities. Therefore, the limited communication bandwidth may cause a longer communication delay, which can result in long convergence times for FL. Communication-efficient FL is thus an important area of current and future study [94–96].

(4) **FL for emerging technologies**. The interplay between FL and emerging technologies introduces new challenges. For instance, a very high propagation attenuation in the terahertz band can affect the convergence analysis. Moreover, in satellite communications, FL can be used to optimize the beam and location of the satellite [97–99]. Another example is in quantum communication, where there is a need to use FL to optimize parameters (such as base probability) for quantum key distribution.

5. Conclusions

In this study, we have considered FL applications for wireless communications. Two main classifications of FL are have been introduced, namely, FRL and FSL. In addition, we have discussed the motivations behind using FL for wireless communication applications. Furthermore, we have identified some of the techniques required to meet the challenges of using FL in practical wireless communications situations. Therefore, it is hoped that this study on FL for wireless communications will provide insights useful for the operation, design, and optimization of FL-based wireless networks.

Acknowledgments

Compliance with ethics guidelines

Zhaohui Yang, Mingzhe Chen, Kai-Kit Wong, H. Vincent Poor, and Shuguang Cui declare that they have no conflict of interest or financial conflicts to disclose.

References

[1] Saad W, Bennis M, Chen M. A vision of 6G wireless systems: applications, trends, technologies, and open research problems. IEEE Netw 2020;34 (3):134–42.
[2] Chen M, Yang Z, Saad W, Yin C, Poor HV, Cui S. A joint learning and communications framework for federated learning over wireless networks. IEEE Trans Wirel Commun 2021;20(1):269–83.

[3] Konečný J, McMahan HB, Ramage D, Richtárik P. Federated optimization: distributed machine learning for on-device intelligence. 2016. arXiv: 1610.02527.

[4] Bennis M, Debbah M, Huang K, Yang Z. Guest editorial: communication technologies for efficient edge learning. IEEE Commun Mag 2020;58(12):12–3.

[5] Zhu G, Wang Y, Huang K. Broadband analog aggregation for low-latency federated edge learning. IEEE Trans Wirel Commun 2020;19(1):491–506.

[6] Zhu G, Du Y, Gunduz D, Huang K. One-bit over-the-air aggregation for communication-efficient federated edge learning: design and convergence analysis. 2020. arXiv: 2001.05713.

[7] Zeng Q, Du Y, Huang K, Leung KK. Energy-efficient resource management for federated edge learning with CPU–GPU heterogeneous computing. 2020. arXiv: 2007.07122.

[8] Mohammadi Amiri M, Gunduz D. Machine learning at the wireless edge: distributed stochastic gradient descent over-the-air. IEEE Trans Signal Process 2020;68:2155–69.

[9] Gunduz D, Kurka DB, Jankowski M, Amiri MM, Ozfatura E, Sreekumar S. Communicate to learn at the edge. IEEE Commun Mag 2020;58(12):14–9.

[10] Amiri MM, Gunduz D. Federated learning over wireless fading channels. IEEE Trans Wirel Commun 2020;19(5):3546–57.

[11] Hosseinalipour S, Brinton CG, Aggarwal V, Dai H, Chiang M. From federated learning to fog learning: towards large-scale distributed machine learning in heterogeneous wireless networks. 2020. arXiv: 2006.03594.

[12] Liu Y, Yuan X, Xiong Z, Kang J, Wang X, Niyato D. Federated learning for 6G communications: challenges, methods, and future directions. China Commun 2020;17(9):105–18.

[13] Hosseinalipour S, Azam SS, Brinton CG, Michelusi N, Aggarwal V, Love DJ, et al. Multi-stage hybrid federated learning over large-scale wireless fog networks. 2020. arXiv: 2007.09511.

[14] Jin R, He X, Dai H. On the design of communication efficient federated learning over wireless networks. 2020. arXiv: 2004.07351.

[15] Liu D, Simeone O. Privacy for free: wireless federated learning via uncoded transmission with adaptive power control. IEEE J Select Areas Commun 2021;39(1):170–85.

[16] Kassab R, Simeone O. Federated generalized Bayesian learning via distributed stein variational gradient descent. 2020. arXiv: 2009.06419.

[17] Kairouz P, McMahan HB, Avent B, Bellet A, Bennis M, Bhagoji AN, et al. Advances and open problems in federated learning. 2019. arXiv: 1912.04977.

[18] Samarakoon S, Bennis M, Saad W, Debbah M. Distributed federated learning for ultra-reliable low-latency vehicular communications. IEEE Trans Commun 2020;68(2):1146–59.

[19] Kang J, Xiong Z, Niyato D, Xie S, Zhang J. Incentive mechanism for reliable federated learning: a joint optimization approach to combining reputation and contract theory. IEEE Internet Things J 2019;6(6):10700–14.

[20] Yang Z, Chen M, Saad W, Hong CS, Shikh-Bahaei M. Energy efficient federated learning over wireless communication networks. IEEE Trans Wirel Commun 2021;20(3):1935–49.

[21] Chen M, Poor HV, Saad W, Cui S. Wireless communications for collaborative federated learning. IEEE Commun Mag 2020;58(12):48–54.

[22] Kang J, Xiong Z, Niyato D, Zou Y, Zhang Y, Guizani M. Reliable federated learning for mobile networks. IEEE Wirel Commun 2020;27(2):72–80.

[23] Liu B, Wang L, Liu M, Xu C. Lifelong federated reinforcement learning: a learning architecture for navigation in cloud robotic systems. 2019. arXiv: 1901.06455.

[24] Wang X, Wang C, Li X, Leung VCM, Taleb T. Federated deep reinforcement learning for internet of things with decentralized cooperative edge caching. IEEE Internet Things J 2020;7(10):9441–55.

[25] Li T, Sahu AK, Talwalkar A, Smith V. Federated learning: challenges, methods, and future directions. IEEE Signal Process Mag 2020;37(3):50–60.

[26] Lim WYB, Luong NC, Hoang DT, Jiao Y, Liang YC, Yang Q, et al. Federated learning in mobile edge networks: a comprehensive survey. IEEE Commun Surv Tutor 2020;22(3):2031–63.

[27] Murshed MG, Murphy C, Hou D, Khan N, Ananthanarayanan G, Hussain F. Machine learning at the network edge: a survey. 2019. arXiv: 1908.00080.

[28] Wang X, Han Y, Wang C, Zhao Q, Chen X, Chen M. In-edge AI: intelligentizing mobile edge computing, caching and communication by federated learning. IEEE Netw 2019;33(5):156–65.

[29] Park J, Samarakoon S, Bennis M, Debbah M. Wireless network intelligence at the edge. Proc IEEE 2019;107(11):2204–39.

[30] Aledhari M, Razzak R, Parizi RM, Saeed F. Federated learning: a survey on enabling technologies, protocols, and applications. IEEE Access 2020;8:140699–725.

[31] Sun Y, Shi W, Huang X, Zhou S, Niu Z. Edge learning with timeliness constraints: challenges and solutions. IEEE Commun Mag 2020;58(12):27–33.

[32] Wei X, Shen C. Federated learning over noisy channels: convergence analysis and design examples. 2021. arXiv: 2101.02198.

[33] Zheng S, Shen C, Chen X. Design and analysis of uplink and downlink communications for federated learning. IEEE J Sel Areas Commun 2021;39 (7):2150–67.

[34] Yang K, Jiang T, Shi Y, Ding Z. Federated learning via over-the-air computation. IEEE Trans Wirel Commun 2020;19(3):2022–35.

[35] Xu C, Liu S, Yang Z, Huang Y, Wong KK. Learning rate optimization for federated learning exploiting over- the-air computation. IEEE J Sel Areas Commun. In press.

[36] Shi W, Zhou S, Niu Z, Jiang M, Geng L. Joint device scheduling and resource allocation for latency constrained wireless federated learning. IEEE Trans Wirel Commun 2021;20(1):453–67.

[37] Amiri MM, Gunduz D, Kulkarni SR, Poor HV. Convergence of update aware device scheduling for federated learning at the wireless edge. IEEE Trans Wirel Commun 2021;20(6):3643–58.

[38] Park J, Samarakoon S, Shiri H, Abdel-Aziz MK, Nishio T, Elgabli A, et al. Extreme URLLC: vision, challenges, and key enablers. 2020. arXiv: 2001.09683.

[39] Li Z, Uusitalo MA, Shariatmadari H, Singh B. 5G URLLC: design challenges and system concepts. In: Proceedings of 15th International Symposium on Wireless Communication Systems; 2018 Aug 28–31; Lisbon, Portugal; 2018.

[40] Basar E, Di Renzo M, De Rosny J, Debbah M, Alouini MS, Zhang R. Wireless communications through reconfigurable intelligent surfaces. IEEE Access 2019;7:116753–73.

[41] Zhang S, Zhang R. Capacity characterization for intelligent reflecting surface aided MIMO communication. IEEE J Sel Areas Commun 2020;38(8):1823–38.

[42] Hu S, Chitti K, Rusek F, Edfors O. User assignment with distributed large intelligent surface (LIS) systems. In: Proceedings of 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications; 2018 Sep 9–12; Bologna, Italy; 2018.

[43] Pan C, Ren H, Wang K, Xu W, Elkashlan M, Nallanathan A, et al. Intelligent reflecting surface for multicell MIMO communications. 2019. arXiv: 1907.10864.

[44] Nadeem QUA, Kammoun A, Chaaban A, Debbah M, Alouini MS. Asymptotic analysis of large intelligent surface assisted MIMO communication. 2019. arXiv: 1903.08127.

[45] Wei L, Huang C, Alexandropoulos GC, Yang Z, Yuen C, Zhang Z. Joint channel estimation and signal recovery in RIS-assisted multi-user MISO communications. In: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC); 2021 Mar 29–Apr 1; Nanjing, China; 2021.

[46] Huang C, Zappone A, Alexandropoulos GC, Debbah M, Yuen C. Reconfigurable intelligent surfaces for energy efficiency in wireless communication. IEEE Trans Wirel Commun 2019;18(8):4157–70.

[47] Huang C, Mo R, Yuen C. Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning. IEEE J Sel Areas Commun 2020;38(8):1839–50.

[48] Huang C, Hu S, Alexandropoulos GC, Zappone A, Yuen C, Zhang R, et al. Holographic MIMO surfaces for 6G wireless networks: opportunities, challenges, and trends. 2019. arXiv: 1911.12296.

[49] Yu X, Xu D, Sun Y, Ng DWK, Schober R. Robust and secure wireless communications via intelligent reflecting surfaces. IEEE J Sel Areas Commun 2020;38(11):2637–52.

[50] Zheng B, You C, Zhang R. Double-IRS assisted multi-user MIMO: cooperative passive beamforming design. IEEE Trans Wirel Commun 2021;20(7):4513–26.

[51] Chaccour C, Soorki MN, Saad W, Bennis M, Popovski P. Risk-based optimization of virtual reality over terahertz reconfigurable intelligent surfaces. 2020. arXiv: 2002.09052.

[52] Hum SV, Perruisseau-Carrier J. Reconfigurable reflect arrays and array lenses for dynamic antenna beam control: a review. IEEE Trans Antenn Propag 2014;62(1):183–98.

[53] Huang J, Li Q, Zhang Q, Zhang G, Qin J. Relay beamforming for amplify-and-forward multi-antenna relay networks with energy harvesting constraint. IEEE Signal Process Lett 2014;21(4):454–8.

[54] Di Renzo M, Ntontin K, Song J, Danufane FH, Qian X, Lazarakis F, et al. Reconfigurable intelligent surfaces vs. relaying: differences, similarities, and performance comparison. IEEE Open J Commun Soc 2020;1:798.

[55] Elbir AM, Coleri S. Federated learning for channel estimation in conventional and IRS-assisted massive MIMO. 2020. arXiv: 2008.10846.

[56] Yang K, Shi Y, Zhou Y, Yang Z, Fu L, Chen W. Federated machine learning for intelligent IoT via reconfigurable intelligent surface. IEEE Network 2020;34 (5):16–22.

[57] Ni W, Liu Y, Yang Z, Tian H, Shen X. Federated learning in multi-RIS aided systems. 2020. arXiv: 2010.13333.

[58] Huang C, Alexandropoulos GC, Yuen C, Debbah M. Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces. In: Proceedings of 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC); 2019 Jul 2–5; Cannes, France; 2019.

[59] Huang C, Yang Z, Alexandropoulos GC, Xiong K, Wei L, Yuen C, et al. Hybrid beamforming for RIS-empowered multi-hop terahertz communications: a DRL-based method. 2020. arXiv: 2009.09380.

[60] Xie H, Qin Z, Li GY, Juang BH. Deep learning enabled semantic communication systems. IEEE Trans Signal Process 2021;69:2663–75.

[61] Xie H, Qin Z. A lite distributed semantic communication system for Internet of Things. 2020. arXiv: 2007.11095.

[62] Chen M, Semiari O, Saad W, Liu X, Yin C. Federated echo state learning for minimizing breaks in presence in wireless virtual reality networks. IEEE Trans Wirel Commun 2020;19(1):177–91.

[63] Chen M, Challita U, Saad W, Yin C, Debbah M. Artificial neural networks-based machine learning for wireless networks: a tutorial. IEEE Commun Surv Tutor 2019;21(4):3039–71.

[64] Ding Z, Liu Y, Choi J, Sun Qi, Elkashlan M, Chih-Lin I, et al. Application of non-orthogonal multiple access in LTE and 5G networks. IEEE Commun Mag 2017;55(2):185–91.

[65] Liu Y, Qin Z, Elkashlan M, Ding Z, Nallanathan A, Hanzo L. Nonorthogonal multiple access for 5G and beyond. Proc IEEE 2017;105(12):2347–81.

[66] Liu Y, Xing H, Pan C, Nallanathan A, Elkashlan M, Hanzo L. Multiple-antenna-assisted non-orthogonal multiple access. IEEE Wirel Commun 2018;25 (2):17–23.

[67] Qin Z, Yue X, Liu Y, Ding Z, Nallanathan A. User association and resource allocation in unified NOMA enabled heterogeneous ultra dense networks. IEEE Commun Mag 2018;56(6):86–92.

[68] Yang Z, Xu W, Pan C, Pan Y, Chen M. On the optimality of power allocation for NOMA downlinks with individual QoS constraints. IEEE Commun Lett 2017;21 (7):1649–52.

[69] Yang Z, Pan C, Xu W, Pan Y, Chen M, Elkashlan M. Power control for multi-cell networks with non-orthogonal multiple access. IEEE Trans Wirel Commun 2018;17(2):927–42.

[70] Shin W, Vaezi M, Lee B, Love DJ, Lee J, Poor HV. Non-orthogonal multiple access in multi-cell networks: theory, performance, and practical challenges. IEEE Commun Mag 2017;55(10):176–83.

[71] Ni W, Liu X, Liu Y, Tian H, Chen Y. Resource allocation for multi-cell IRS-aided NOMA networks. IEEE Trans Wirel Commun 2021;20(7):4253–68.

[72] Andrieu C, de Freitas N, Doucet A, Jordan MI. An introduction to MCMC for machine learning. Mach Learn 2003;50(1–2):5–43.

[73] Freeman WT, Pasztor EC, Carmichael OT. Learning low-level vision. Int J Comput Vis 2000;40(1):25–47.

[74] Collobert R, Weston J. A unified architecture for natural language processing: deep neural networks with multitask learning. In: Proceedings of 25th International Conference on Machine Learning; 2008 Jul 5–9; Helsinki, Finland; 2008.

[75] Bishop CM. Pattern recognition and machine learning. Berlin: Springer; 2006.

[76] Lee H, Wicke M, Kusy B, Gnawali O, Guibas L. Predictive data delivery to mobile users through mobility learning in wireless sensor networks. IEEE Trans Veh Technol 2015;64(12):5831–49.

[77] Yao L, Chen A, Deng J, Wang J, Wu G. A cooperative caching scheme based on mobility prediction in vehicular content centric networks. IEEE Trans Veh Technol 2018;67(6):5435–44.

[78] Chen M, Mozaffari M, Saad W, Yin C, Debbah M, Hong CS. Caching in the sky: proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience. IEEE J Sel Areas Commun 2017;35 (5):1046–61.

[79] Yin J, Li L, Zhang H, Li X, Gao A, Han Z. A prediction-based coordination caching scheme for content centric networking. In: Proceedings of 27th Wireless and Optical Communication Conference; 2018 Apr 30–May 1; Hualien, China; 2018.

[80] Yang Z, Xu W, Xu H, Shi J, Chen M. Energy efficient non-orthogonal multiple access for machine-to-machine communications. IEEE Commun Lett 2017;21 (4):817–20.

[81] Yang Z, Xu W, Pan Y, Pan C, Chen M. Energy efficient resource allocation in machine-to-machine communications with multiple access and energy harvesting for IoT. IEEE Internet Things J 2018;5(1):229–45.

[82] Cui J, Ding Z, Fan P, Al-Dhahir N. Unsupervised machine learning-based user clustering in millimeter-wave-NOMA systems. IEEE Trans Wirel Commun 2018;17(11):7425–40.

[83] Li M, Zhou L, Yang Z, Li A, Xia F, Andersen DG, et al. Parameter server for distributed machine learning. In: Proceedings of Big Learning NIPS Workshop; 2013 Dec 5–10; Harrahs and Harveys, NV, USA, 2013.

[84] Bekkerman R, Bilenko M, Langford J. Scaling up machine learning: parallel and distributed approaches. Cambridge: Cambridge University Press; 2011.

[85] Kim M, Kim NI, Lee W, Cho DH. Deep learning-aided SCMA. IEEE Commun Lett 2018;22(4):720–3.

[86] Samarakoon S, Bennis M, Saad W, Debbah M. Federated learning for ultra-reliable low-latency V2V communications. In: Proceedings of IEEE Global Communications Conference; 2018 Dec 9–13; Abu Dhabi, United Arab Emirates; 2018.

[87] Ma C, Li J, Ding M, Yang HH, Shu F, Quek TQ, et al. On safeguarding privacy and security in the framework of federated learning. IEEE Netw 2020;34(4):242–8.

[88] Li X, Huang K, Yang W, Wang S, Zhang Z. On the convergence of FedAvg on non-IID data. 2020. arXiv: 1907.02189.

[89] Luo S, Chen X, Wu Q, Zhou Z, Yu S. HFEL: joint edge association and resource allocation for cost-efficient hierarchical federated edge learning. IEEE Trans Wirel Commun 2020;19(10):6535–48.

[90] Chen M, Poor HV, Saad W, Cui S. Convergence time optimization for federated learning over wireless networks. IEEE Trans Wirel Commun 2021;20 (4):2457–71.

[91] Yang HH, Liu Z, Quek TQS, Poor HV. Scheduling policies for federated learning in wireless networks. IEEE Trans Commun 2020;68(1):317–33.

[92] Khaled A, Mishchenko K, Richtárik P. Tighter theory for local SGD on identical and heterogeneous data. In: Proceedings of International Conference on Artificial Intelligence and Statistics; 2020 Aug 26–28; online; 2020.

[93] Wei K, Li J, Ding M, Ma C, Yang HH, Farokhi F, et al. Federated learning with differential privacy: algorithms and performance analysis. IEEE Trans Inf Forensics Secur 2020;15:3454–69.

[94] Chen M, Shlezinger N, Poor HV, Eldar YC, Cui S. Communication efficient federated learning. Proc Natl Acad Sci 2021; 118(17): e2017318118.

[95] Chen M, Gunduz D, Huang K, Saad W, Bennis M, Feljan AV, et al. Distributed learning in wireless networks: recent progress and future challenges. IEEE J Sel Areas Commun 2021;39(12):3579–605.

[96] Tian Y, Zhang Z, Yang Z, Yang Q. JMSNAS: joint model split and neural architecture search for learning over mobile edge networks. 2021. arXiv: 2111.08206.

[97] Tong X, Zhang Z, Wang J, Huang C, Debbah M. Joint multi-user communication and sensing exploiting both signal and environment sparsity. IEEE J Sel Topics Signal Process. In press.

[98] Yang Y, Zhang Z, Yang Q. Communication-efficient federated learning with binary neural networks. IEEE J Sel Areas Commun. In press.

[99] Qi Q, Chen X, Zhong C, Zhang Z. Integrated sensing, computation and communication in B5G cellular Internet of Things. IEEE Trans Wirel Commun 2021;20(1):332–44.