Artificial Intelligence Review
https://doi.org/10.1007/s10462-023-10563-8

A systematic review of federated learning from clients' perspective: challenges and solutions

Yashothara Shanmugarasa[1] · Hye-young Paik[1] · Salil S. Kanhere[1] · Liming Zhu[2]

Abstract

Federated learning (FL) is a machine learning approach that decentralizes data and its processing by allowing clients to train intermediate models on their devices with locally stored data. It aims to preserve privacy as only model updates are shared with a central server rather than raw data. In recent years, many reviews have evaluated FL from the system (general challenges) or server's perspectives, ignoring the importance of clients' perspectives. Although FL helps users have control over their data, there are many challenges arising from decentralization, specifically from the perspectives of clients who are the main contributors to FL. Therefore, in response to the gap in the literature, this study intends to explore client-side challenges and available solutions by conducting a systematic literature review on 238 primary studies. Further, we analyze if a solution identified for one type of challenge is also applicable to other challenges and if there are impacts to consider. The conclusion of this survey reveals that servers and platforms have to work with clients to address client-side challenges.

Keywords Federated learning · Client-side challenges · Client-side solutions · Systematic literature review

✉ Yashothara Shanmugarasa
y.shanmugarasa@unsw.edu.au

Hye-young Paik
h.paik@unsw.edu.au

Salil S. Kanhere
salil.kanhere@unsw.edu.au

Liming Zhu
liming.zhu@data61.csiro.au

[1] School of Computer Science and Engineering, UNSW Sydney, Kensington, NSW, Australia

[2] Data61/CSIRO, Sydney, Australia

Published online: 07 August 2023

1 Introduction

Despite the heightened public awareness and technological and regulatory efforts, frequent data breaches and privacy violations are still being reported.[1] The problem is that large amounts of data are centralized to a single service provider with current data storage and processing architectures, leading to a single point of privacy failure. The recently emerged Machine Learning (ML) architecture, named Federated Learning (FL), decentralizes data and its processing pipelines by allowing users[2] to train intermediate models in their devices, effectively collaborating with the central service to build a global model for all clients without having to surrender the raw data to the central service.

Figure 1 describes the system architecture of FL and the training procedure. Steps 2, 3, 4, and 5 are repeated over time to keep the global model up to date across clients.

Formally, FL can be considered an optimization problem where the goal is to minimize a global objective function that aggregates local models while respecting the constraints imposed by the distributed nature of the contributing clients and data (Wang et al. 2021a). Let $\mathcal{C}$ represent the N number of clients participating in FL. Each client $i \in \mathcal{C}$ has a local dataset $\mathcal{D}_i$. The objective is to train a global model M by aggregating the local models of the clients. In FL, the learning process involves minimizing a loss function that is calculated on each client. This is achieved through a weighted aggregation method. The objective of FL is to minimize the following objective function as in Eq. (1):

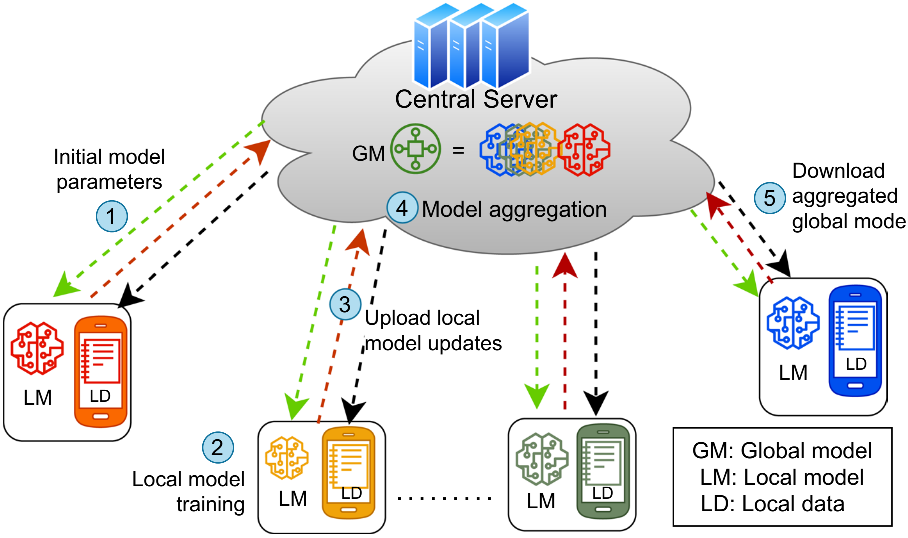


Fig. 1 The system architecture of FL

[1] https://tinyurl.com/f4z6bbr9.

[2] We use the terms “users” and “clients” in this paper. Users are individuals with devices. Clients refer to users’ devices that train the data in the FL process. As a user participates in the FL process through the client (devices), thus client can be considered the endpoint for the user. Hence, we refer “clients” to both devices and users. When required to distinguish the users (individuals) from devices, we will explicitly refer to them as “users”.

$$min_w \mathcal{F}(w) = \frac{1}{N} \sum_{i=1}^{N} w_i \cdot f_i(w_i) \quad (1)$$

Where w represents the model parameters, N is the total number of clients, w_i is the weight assigned to client i, and $f_i(w)$ is the local loss function computed on client i. The objective function is minimized by iteratively updating the model parameters based on the aggregated contributions from each client. The weights assigned to each client can be influenced by factors such as client performance, available resources, or fairness considerations. To ensure privacy, FL employs techniques like federated averaging or secure aggregation.

Although this paradigm shift helps users have greater control and transparency over their data, many challenges arise from decentralization. For example, clients may "drop out" during the training phase due to poor network connectivity; or maintaining model performance is challenging due to unreliable model updates from clients; or users may contribute heterogeneous data/devices, causing model divergence and straggling (Lyu et al. 2020b; Lo et al. 2021b; Kulkarni et al. 2020). Many researchers have extensively studied and surveyed many of these general FL challenges recently. For instance, previous papers have focused on the system (design aspects (Rahman et al. 2021) and general challenges such as communication efficiency (Shahid et al. 2021), model performance (Wang et al. 2021b), and security (Lyu et al. 2020b) or the server's perspective (statistical heterogeneity, client motivatability, and scalability) (Imteaj et al. 2021; Rahman et al. 2021).

However, the challenges from the clients' perspectives are still under-explored. We refer to "client-side challenges" as the challenges clients face during the FL training procedures. The challenges may arise from security and privacy viewpoints (e.g., malicious servers or dishonest "peers") and the complexity of FL processes as the computational burden is now placed on the clients.

These client-side challenges can affect a few or all clients on the network. For example, being able to personalize (fine-tune) a global model to a particular client would only affect those who want the capability. Privacy management challenges are relevant to every client.

We choose client-side challenges as our study focus because (i) client participation plays an important role in FL as they contribute resources. Shifting the data processing to clients may cause unintended mishaps and privacy risks as the wider population has limited technical knowledge (Kairouz et al. 2021), (ii) as clients contribute data and resources for FL, they should certainly be entitled to receive some benefits for their contribution. However, most clients are not technically savvy to define their requirements or understand the internal black-box mechanisms. For example, many mobile phone users are unaware that the predictions on Google keyboards are built using their data and resources,[3] and (iii) most surveys focused on the FL challenges in general. To the best of our knowledge, the review of client-side challenges and solutions is not yet well documented in the literature.

We conducted a comprehensive literature review on the selected research papers, tutorials, dissertations, and magazines in the FL domain to lay out the challenges from the clients' perspectives. We categorized and grouped the articles according to their primary focus areas. We combine some focus areas to illustrate the challenges better. For example, data management, computation cost management, and communication cost management are combined as resource management. Further, we study the survey papers on the general FL challenges such as lack of motivation of clients, computational/communication cost,

[3] https://tinyurl.com/2p98x899.

and privacy/security attacks (Lyu et al. 2020b; Rahman et al. 2021; Blanco-Justicia et al. 2021) and define the effect of these challenges from the client's perspective. For example, incentive mechanism is a widely discussed issue for motivating clients to participate in FL. Previous papers discuss this issue in terms of incentive mechanism processes, algorithms, and client motivation. But, clients are more interested in the benefits and transparency of incentives to compare the benefits among themselves. In another example, despite extensive discussions on privacy challenges in the literature, a thorough analysis of these challenges from the clients' perspective is lacking. Our study specifically focuses on examining five client-specific challenges: auditability, data granularity, re-identification, and consented data sharing under the privacy management challenge. By delving into these aspects, we aim to shed light on clients' unique issues. Correspondingly, our work concluded with six main categories of client-side challenges.

The overall objectives of this survey are (i) outlining FL challenges from clients' perspectives, (ii) providing an overview of the current research activities for the client-side challenges, (iii) summarising the challenges with existing approaches, and (iv) helping researchers to understand the open problems and future trends. We derived research questions (RQ) to achieve all the objectives and discuss them in the following section.

1.1 Research questions

1. **RQ 1** What are the client-side challenges in FL? (Sect. 3): To enhance the usability of FL to clients, we first need to understand the challenges from the clients' perspectives. Therefore the first research question focuses on client-side challenges.
2. **RQ 2** What are the state-of-the-art solutions that address these challenges? (Sect. 4): As we analyze the challenges in RQ 1.0, we examine state-of-the-art solutions to the given challenges in this question.
3. **RQ 3** Can a solution identified for one type of challenge be applied to other types of challenges? Are there any impacts to consider? (Sect. 5): Drawing on previous research and our understanding of existing solutions for the identified challenges, we assess a particular solution's impact on the other challenges. Specifically, we analyze whether the solution can be applied to solve or potentially exacerbate other challenges.
4. **RQ 4** What are the open challenges and possible future trends? (Sect. 6): This RQ focuses on open challenges and future trends in solving client-side challenges, which the literature does not cover fully.

The contributions of this paper are as follows: (i) identifying the challenges of FL from clients' perspectives, (ii) comprehensive analysis of the solutions given in state-of-the-art approaches with 238 studies, and (iii) analysis and discussion on the impacts of applying the solutions to the identified challenges.

1.2 Sources selection and strategy

We followed a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Page et al. 2021) approach to conduct the survey. We searched through Google Scholar search engine to identify research papers on different client-side challenges. We set the time frame from 01.01.2017 to 31.01.2022. The source collection statistics in search strings and the number of articles are tabulated in Table 1. Besides the initial search, we included some additional papers using snowballing process (Wohlin 2014) through the

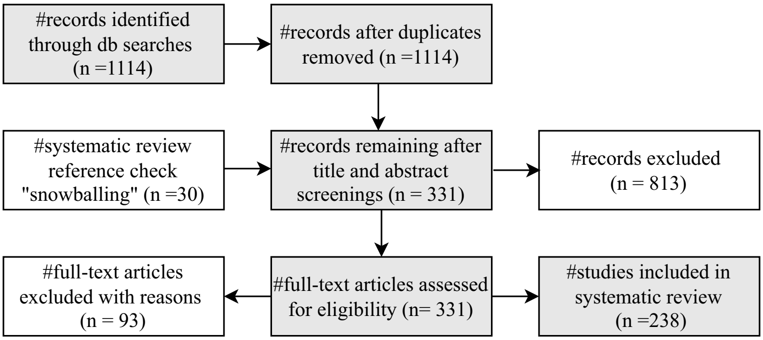

#records identified through db searches (n =1114)
#records after duplicates removed (n =1114)
#systematic review reference check "snowballing" (n =30)
#records remaining after title and abstract screenings (n = 331)
#records excluded (n = 813)
#full-text articles excluded with reasons (n = 93)
#full-text articles assessed for eligibility (n= 331)
#studies included in systematic review (n =238)

Fig. 2 PRISMA and snowballing approach for paper searching and selection (# implies number of)

bibliography of identified research papers. As our paper focuses on identifying the client-side challenges of FL, we included the search strings to retrieve the papers that discuss the challenges. Firstly, we screened the papers via title and abstract for including or excluding the papers. Figure 2 shows the PRISMA and snowball approach of the paper search and selection process.

After the initial screening through the title and abstract, we excluded 813 papers for the following reasons: long-short repetitive papers, non-relevance of the focused area, and a certain solution applied in multiple domains. We ended up with 331 papers for further full-text articles assessment with snowballing references. As we mainly focus on client-side challenges, we excluded 93 papers that were focused on the general challenges, survey papers (but we included them in our related works section), application papers (same technology applied in different domains), white papers, irrelevant to client-side issues, short versions of extended papers, and papers which adapt FL as privacy preserved approach (pure FL without any extensions). Finally, this study covered 238 studies.

The remainder of the article is organized as follows. Section 2 discusses related surveys in FL and how our study is unique from others. Section 3 introduces the client-side challenges in FL. Section 4 presents the state-of-art solutions to the client-side challenges. Section 5 discusses the linkage of challenges and applicability of current technologies. Section 6 opens the opportunities and trends for future work. Finally, we conclude the review in 7 concludes the survey.

2 Related works

This section provides an overview of the existing review papers on FL. Initially, we conducted a comprehensive search to identify and gather all the relevant surveys and reviews pertaining to FL. The majority of these surveys primarily focused on providing a general overview of FL, including its design aspects, application domains, and the overall challenges associated with FL.

However, to the best of our knowledge, no surveys have specifically delved into the FL challenges from the clients' perspectives. While certain literature surveys have addressed individual challenges clients face, they often lack a comprehensive analysis that considers the clients' viewpoints. Given this gap, our main focus was to thoroughly

Table 1 The statistics of source selection

Search strings	#Selected/retrieved	Focused challenge
allintitle: personalisation OR personalization OR personalised OR personalized "Federated learning"	50/96	Personalisation
allintitle: incentive OR "client motivation" "Federated learning"	37/53	Client Incentive
allintitle: Federated learning privacy OR "granular privacy" OR auditability	40/502	Privacy management
allintitle: fairness "Federated learning"	29/33	Fairness
allintitle: Federated learning communication OR "communication cost" OR allintitle: Federated learning "resource optimization" OR "computation cost" OR computation OR "energy efficient" OR "data update cost" OR "data cost" "Federated learning"	62/395	Resource management
allintitle: Federated learning "data security" OR "device security" OR defense	20/35	Data/device security
Total	238/1114	

# Implies number of

examine the client-side challenges in FL, considering the available state-of-the-art solutions from the clients' perspectives.

General surveys focused on the high-level view of the FL environments, such as definitions, components, algorithms, optimization, importance, design aspects, application domains, trends, general challenges, and evaluation approaches. Previous surveys (Rahman et al. 2021; Li et al. 2020e; Aledhari et al. 2020) discussed the characteristics of FL, the general challenges, and the available solutions for FL with future trends. With these general views of FL, Yang et al. (2019) and Kairouz et al. (2021) included privacy and security aspects of FL also in their survey. Moreover, the studies (Li et al. 2021e; Zhang et al. 2021a) provided a comprehensive study of FL systems, incorporating model building, data partitioning, privacy, scalability, and communication architecture aspects of FL. On another note, Lo et al. (2021b) conducted a systematic review from a software engineering perspective. They covered FL lifecycles such as background understanding, requirements analysis, architecture design, implementation, and future trend evaluation.

Some other general surveys have investigated the algorithms and applications of FL. These surveys provide comprehensive overviews of the various algorithms and diverse application areas in which FL has been implemented. Li et al. (2020b) examines FL's evolution and prevailing applications in industrial engineering. It aims to guide future applications and optimization in FL by reviewing related studies, addressing challenges, and discussing realistic applications in IoT devices, industrial engineering, and healthcare. Likewise, Wang et al. (2021b) provides practical recommendations and guidelines for designing and evaluating federated optimization algorithms through concrete examples and practical implementation. They also address the lack of consensus on core concepts in FL and offer suggestions on problem formulation and algorithm design. Ding et al. (2022) provides an outlook on the challenges and opportunities in FL across five emerging directions: algorithm foundation, personalization, hardware and security constraints, lifelong learning, and nonstandard data. The paper also touches on the challenges of data incompleteness, polarity, and complex dependency in FL.

Moreover, various systematic reviews have extensively examined different domains in the context of FL, including resource-constrained Internet of Things (IoT) (Imteaj et al. 2021; Du et al. 2020), mobile edge networks (Lim et al. 2020), wireless communication (Niknam et al. 2020), and healthcare and informatics (Xu et al. 2021). These reviews have explored existing studies, assumptions, challenges, applications, and problems within each domain.

Several surveys, including (Lyu et al. 2020b; Alazab et al. 2021; Mothukuri et al. 2021; Blanco-Justicia et al. 2021; Ma et al. 2020; Kurupathi and Maass 2020; Enthoven and Al-Ars 2021; Briggs et al. 2021), specifically delve into the comprehensive analysis of privacy and security challenges, applications, key techniques, trends, and open problems in the overall system of FL. In addition, prior studies such as (Kulkarni et al. 2020; Tan et al. 2021) have specifically examined the challenge of personalization in FL, providing insights into motivation, taxonomies, strategies, and future opportunities. A comprehensive analysis of incentive mechanisms in FL has been conducted by exploring existing works and key techniques in studies (Zhan et al. 2021; Zeng et al. 2021). Communication challenges in FL have been addressed by Shahid et al. (2021), while fairness challenges have been surveyed by Shi et al. (2021), covering aspects such as basic assumptions, fairness notions, taxonomies, metric evaluation, and future directions. However, these surveys primarily focus on the challenges of FL from a general perspective, implying a lack of emphasis on the client or user viewpoint. Table 2 outlines the parallel surveys that focused on FL.

Our work distinguishes itself from others in the following ways: (i) We specifically address the client-side challenges in a federated environment, recognizing the significance of clients in contributing resources and data. We highlight the potential risks and privacy concerns associated with complex FL processes delegated to clients, (ii) We thoroughly analyze the client-side challenges and their interdependencies in a federated environment, (iii) We adopt the PRISMA approach, providing a clear methodology for our survey, which is lacking in many existing works, (iv) we extensively cover research papers and survey papers in our review, (v) We discuss the open client-side challenges and future trends, and (vi) our review is up-to-date, incorporating papers published until January 2022.

In contrast to existing surveys that primarily examine FL challenges from the system, server, technical, and taxonomy perspectives, our work takes a unique approach by analyzing the challenges and solutions specifically from the clients' perspectives. We provide insights into the impacts involved in addressing these challenges. To the best of our knowledge, this work represents the first comprehensive exploration of client-side challenges in FL, offering a fresh perspective on the existing literature in this domain.

3 Client-side challenges in federated learning

This section describes the client-side challenges in the FL environment. Through a comprehensive analysis of the literary works, we were able to identify six main categories of client-side challenges such as (i) personalization, (ii) privacy management, (iii) incentive management, (iv) resource management, (v) data and devices security management, and (vi) fairness management. We analyzed these challenges more granularly from the clients' perspectives and identified precise issues under the above categories as shown in Fig. 3.

Table 2 Summary of exciting surveys on federated learning

Scope	Survey papers
Federated learning challenges: clients' perspective	Our survey *(2016–2022 January)*
General overview of federated environment	Lo et al. (2021b) *(2016–2020)*, Rahman et al. (2021) *(2016–2020)*, Li et al. (2021e) *(2016–2019)*, Zhang et al. (2021a) *(2016–2019)*, Kairouz et al. (2021) *(2016–2019)* Li et al. (2020e) *(2016–2020)*, Aledhari et al. (2020) *(2016–2020)*, Yang et al. (2019) *(2016-2019)*
General overview federated environment	Application and algorithm of FL: Applications: Li et al. (2020b) *(2016–2020)*, Optimization algorithm: Wang et al. (2021b) (2016–2021), Algorithmic challenges: Ding et al. (2022) *(2016–2021)*
General overview federated environment: domain-wise	IoT: Imteaj et al. (2021) *(2016–2021)*, Du et al. (2020) (2016-2020), Mobile edge network: Lim et al. (2020) *(2016–2020)*, Wireless communication: Niknam et al. (2020) *(2016-2019)*, Healthcare and informatics: Xu et al. (2021) *(2016–2019)*
Certain challenge: privacy and security	Lyu et al. (2020b) *(2016–2020)*, Alazab et al. (2021) *(2016–2021)*, Mothukuri et al. (2021) *(2016–2020)*, Blanco-Justicia et al. (2021) *(2016–2020)*, Ma et al. (2020) *(2016-2020)*, Kurupathi and Maass (2020) *(2016–2020)*, Enthoven and Al-Ars (2021) *(2016–2020)*, Briggs et al. (2021) *(2016-2020)*
Certain challenge: personalization	Kulkarni et al. (2020) *(2016–2020)*, Tan et al. (2021) *(2016–2021)*
Certain challenge: incentive	Zhan et al. (2021) *(2016–2020)*, (Zeng et al. 2021) *(2016–2021)*
Certain challenge: communication	Shahid et al. (2021) *(2016-2021)*
Certain challenge: fairness	Shi et al. (2021) *(2016–2021)*

*The years in parentheses refer to the years of research publications that covered in the review article

3.1 Personalization

Generally, personalization means designing a product/service to meet a user's requirements.[4] Based on the literature (Tan et al. 2021; Kulkarni et al. 2020; Li et al. 2021f), we describe the personalization challenge from the clients' perspectives as follows.

In FL, clients with a common goal join the FL environment because they do not have enough data to attain their objective with high performance and generalization guarantees. In a traditional FL setting, every client receives the same global model. However, some clients may require personalized models catered for their preferences, much like the recommendation systems. For example, the next word prediction scenario in Gboard,[5] each client's data can have quite different distributions (e.g., due to different texting habits) Therefore, next-word predictions should be personalized to the client while handling generalized

[4] https://dictionary.cambridge.org/dictionary/english/personalization.

[5] https://en.wikipedia.org/wiki/Gboard.

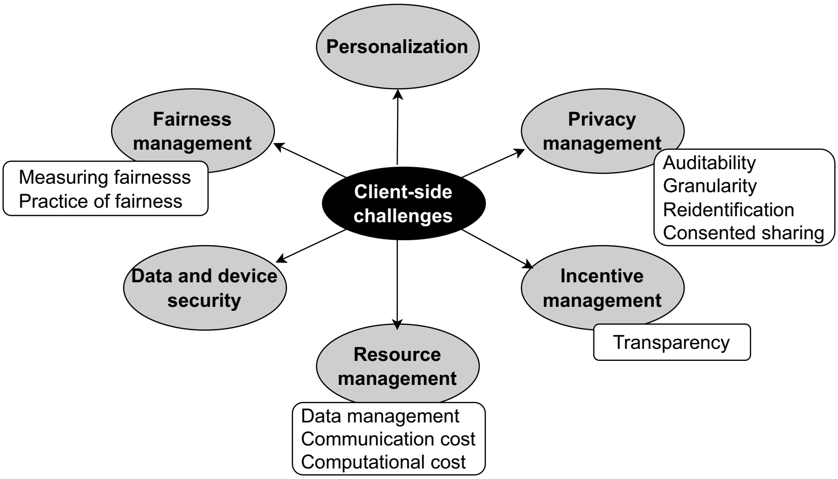


Fig. 3 The overview of the client challenges

situations (predictions of unknown sentences not in the client's data) with global predictions. Another case is that some clients' data are unique, which means the performance of the global model may not be satisfactory. The users may then may wish to fine-tune the model to get more personalized results.

Achieving an optimal balance between generalization and personalization poses a significant challenge for clients in the context of FL. It is crucial to strike the right balance to obtain the best results. The algorithm should be designed to handle the optimal balance by leveraging the available client data to address this.

3.2 Privacy management

Even in a federated environment, privacy is a concern, as sensitive information can be leaked through the built models (Zhu and Han 2020). It can occur due to the dishonest server or collusion attacks of other clients. The inference of raw data from the model poses a significant concern as it results in the loss of data control for the owners. This data leakage can potentially give rise to severe issues and vulnerabilities for the data owners.

We conducted a comprehensive analysis of the literature on privacy management (Mothukuri et al. 2021; Kurupathi and Maass 2020; Enthoven and Al-Ars 2021; Fang et al. 2022; Katevas et al. 2020), considering users' perspectives and categorizing precise issues under privacy management. Additionally, we examined general privacy threats such as linkability, identifiability, non-repudiation, detectability, information disclosure, content unawareness, and consent non-compliance (Deng et al. 2011), adapting the relevant challenges to the FL context from the users' viewpoints. Consequently, this survey addresses four key issues in privacy management: auditability, consent data sharing, data granularity, and reidentification.

3.2.1 Auditability

Auditing is an assessment mechanism for measuring the quality of a process's lifecycle to ensure accuracy and efficacy.[6] It allows users to make judgments based on all transactions that have occurred in the past. In the context of FL, auditing becomes more complex due to the involvement of multiple parties. It involves recording all FL activities, including model parameter transactions, participation records, model attributes, data or resource contributions, and accuracy. By providing visibility into the FL environment, auditing enables users to understand ongoing activities and make informed decisions about their continued participation.

3.2.2 Consented data sharing

Consented data sharing involves participants willingly agreeing to share their data, fully aware of the associated risks, benefits, and purpose(Deng et al. 2011). In the context of FL, individuals may unknowingly provide excessive information,[7] ultimately relinquishing control over their data. The more information a client discloses, the greater the potential risk of privacy breaches.

3.2.3 Data granularity

Data granularity refers to the level of detail present in a data.[8] In the context of FL, where a large amount of client data is involved, users should be able to determine the granularity level when building models. Each user may have different privacy preferences, with some preferring to share more data and others opting to limit the sharing of specific data items, such as location, medical history, and ethnicity, or reducing the granularity of address information to city or zip code level. However, manually managing these settings can be cumbersome and inconvenient for users.

3.2.4 Reidentification

While FL is appreciated by many users for its ability to protect their actual data, recent studies (Orekondy et al. 2018) have demonstrated that data can still be reidentified from models built through FL. In order to prevent data reidentification and protect their valuable and current data, clients need mechanisms to safeguard against inference attacks. For instance, location trajectories can reveal sensitive information such as points of interest, social relationships, and user identities (Khalfoun et al. 2021), which poses a significant threat to clients, given the widespread collection of precise location data by IoT devices. Therefore, it is crucial to implement privacy protection mechanisms to safeguard users' privacy.

[6] https://www.egnyte.com/guides/governance/data-auditing.

[7] https://tinyurl.com/2p98x899.

[8] https://c3.ai/glossary/features/data-granularity/.

3.3 Incentive management

Clients play a crucial role in the FL process by contributing valuable resources such as computation capacity, battery, data, memory, and bandwidth. However, these resources are limited and costly for clients, who also face privacy and security risks when participating in FL. In order to incentivize clients and encourage their active involvement, the FL framework offers various incentives such as personalized models that perform well on their own test data, accurate global models trained on comprehensive datasets, monetary compensation, reputation benefits, user-defined incentives, and auxiliary information like bias and model fairness considerations (Tu et al. 2022). These incentives aim to offset the costs and provide clients with tangible benefits for their contributions in FL.

Incentive mechanisms are widely discussed in the literature (Zhan et al. 2021; Tu et al. 2022; Ding et al. 2020; Kang et al. 2019a; Liu and Wei 2020), focusing on rewards types, motivation towards client participation, incentive calculation challenges, and incentive scheme design. They discussed the generic challenges of incentive mechanisms such as (i) complexity in determining an optimal incentive for the clients in a closed environment with no information on data structure, resource capability, and client infrastructure (Zeng et al. 2021), (ii) deriving an appropriate metric to quantify the contribution of clients. Generally, local model accuracy is used as the evaluation metric, which can be biased to some clients with unique values that do not contribute much to the global model (Zhan et al. 2021; Zeng et al. 2021), and (iii) consequently, motivating clients to actively participate in FL through incentives becomes a challenging task, as the design of an approximation scheme is often intricate (Kang et al. 2019b). It is important to note that these challenges are predominantly addressed by the servers, as they bear the responsibility of resolving these more generic issues in the context of incentive mechanisms.

However, the issue of transparency arises as a prominent concern in incentive management for clients, as it greatly influences clients' decisions regarding participation. Transparency entails being fully visible, open to scrutiny, and clear, with no hidden aspects.[9] However, due to confidentiality reasons, clients typically do not share their incentive decisions and compensation information with others, resulting in a lack of transparency in incentive scheme selection. This lack of transparency can lead to unfairness or ignorance of entitlements among clients. Furthermore, users with limited technical expertise may struggle to assert their rights and claim their rightful rewards. Balancing transparency with the need for confidentiality, fairness, security, and privacy presents a significant challenge in designing a robust and transparent incentive allocation mechanism.

3.4 Resource management

Given clients' devices' limited and costly resources, effective resource management becomes essential to ensure optimal performance within the given capacity limitations. The literature has extensively analyzed resource management in three key categories: data management (Moon et al. 2020; Jeong et al. 2018; Shin et al. 2020), computation management (Nour et al. 2021; Ren et al. 2019; Ji et al. 2021), and communication management (Shahid et al. 2021; Yue et al. 2022; Sattler et al. 2019b). Following a similar approach, we

[9] https://www.techtarget.com/whatis/definition/transparency.

address client-side challenges in resource management by comprehensively considering all three aspects in our survey.

3.4.1 Data management

The current digital age boosts the quantity and frequency of data generation from various resources (such as social networks, IoT devices, and health centres). The heterogeneous nature of devices results in varying amounts and quality of data generated by different clients. Consequently, the performance of FL global models is adversely affected when built using diverse datasets. Particularly, clients with limited, unique, or unbalanced data are significantly impacted by this challenge.

Poor performance is a common issue experienced by clients in FL, and one of the reasons behind this is data management challenges, such as data scarcity, data imbalance, and data representation. Our survey focuses on addressing data imbalance and representation challenges while acknowledging that data sparsity is commonly addressed in the literature through techniques like sampling. However, managing data in the FL environment is particularly challenging as the data remains decentralized. Consequently, the proposed solutions should consider client-side approaches (considering clients' limited technical expertise) or platform-based mechanisms that respect user privacy.

3.4.2 Computation cost management

Many researchers (Imteaj et al. 2021; Nour et al. 2021; Ren et al. 2019; Ji et al. 2021) widely analyzed the computation cost management challenge as it directly affects clients' participation. Typically, servers run these complex algorithms in a centralized architecture with hundreds of GPU machines to produce their model. But, resource-constrained devices run the FL algorithms in the background while managing their main tasks. Given the enormous resource requirements of FL algorithms, clients' devices are often limited and energy-consuming. Consequently, many clients are hesitant to dedicate all their resources solely to the FL process. However, clients often lack control over their devices, limiting their ability to decide when to participate. This lack of authority over participation poses challenges for clients in achieving better local models, efficiently utilizing resources, and minimizing computational costs.

3.4.3 Communication cost management

Besides preserving privacy, FL divides the computational power among clients and reduces the communication burden by transferring models instead of raw data. However, despite these advantages over traditional centralized architectures, clients still encounter communication challenges. These challenges arise from the resource constraints of devices, unreliable network connections, communication frequency, the transmission of large gradient vectors in complex deep neural networks, and the distance between clients and servers (Shahid et al. 2021).

3.5 Data and devices security management

The multi-party closed nature of the FL environment, where client and server information is not exchanged, introduces vulnerabilities and challenges in monitoring trustworthiness. Additionally, the FL environment is dynamic, with new clients and models constantly being introduced, requiring a continuous verification process. This ongoing verification process adds complexity to ensuring the trustworthiness of participants in the FL ecosystem.

The literature (Alazab et al. 2021; Ma et al. 2020; Fang et al. 2020; Bagdasaryan et al. 2020; Zhang et al. 2019; Lin et al. 2019) discussed several security breaches such as model invalidation, data/model poisoning, model inference, backdoor attacks, malicious clients, and malicious server. Adversaries use various vulnerable ways such as communication medium, client data manipulation, dis-honest server, and aggregation algorithm to attack the environment (Mothukuri et al. 2021). Hence, clients must employ defense mechanisms to protect themselves from adversarial attacks that can be initiated by other clients, servers, or external attackers. These attacks can compromise the shared goal of FL and put data at risk.

3.6 Fairness management

Fairness in FL entails treating every client impartially, without any bias or discrimination (Ezzeldin et al. 2021). Consider a face recognition scenario where the FL server has access to many mobile devices used by white users but only a few used by black users. Consequently, the model may exhibit better performance in recognizing the faces of white individuals compared to black individuals.[10] However, achieving fairness among clients is challenging due to statistical and system heterogeneity. Defining fairness itself lacks consensus, with different notions representing specific interests and aspects of participant groups. Therefore, attaining acceptable fairness in a multiparty collaboration environment is complex.

In the context of FL, fairness is a multifaceted concept, and different mathematical criteria have been proposed to capture fairness. One widely used criterion is equalized odds, which aims to ensure that the probability of a positive prediction is the same across different groups, irrespective of their protected attributes.

A predicting algorithm satisfies equalized odds if it ensures that both the true positive rate (TPR) and the false positive rate (FPR) are equal across different groups (Garg et al. 2020). More formally, equalized odds requires that the group-specific TPR satisfy Eq. (2) and FPR satisfies Eq. (3).

$$P(y' = 1 \mid y = 1, G = 0) = P(y' = 1 \mid y = 1, G = 1) \tag{2}$$

$$P(y' = 1 \mid y = 0, G = 0) = P(y' = 1 \mid y = 0, G = 1) \tag{3}$$

In both cases, the equation is comparing the conditional probabilities of the predicted label y' being 1, under different scenarios based on the values of the ground truth label y and the protected attribute G, which indicates different groups.

[10] https://tinyurl.com/yc4c2wb7/.

Other notions of fairness discussed in the literature are accuracy parity (uniformity in performance across clients), good-intent fairness (minimizing loss for underlying protected client classes), group fairness (minimizing disparities in algorithmic decision-making across groups), selection fairness (reducing FL model bias by increasing the participation of under or never-represented clients), contribution fairness (rewarding clients in proportion to the client's contribution), regret distribution fairness (minimizing the regret difference among clients. Regret indicates the difference between what the client has received so far and what they deserve), and expectation fairness (minimizing inequality between clients over a period until receiving rewards) (Shi et al. 2021). A considerable amount of literature (Li et al. 2021f; Garg et al. 2020; Divi et al. 2021b; Yu et al. 2020a) has been published on the fairness concept in the FL domain. We comprehensively analyzed the challenges and approaches focused on those studies from clients' perspectives. We derived two critical issues under fairness management: measuring fairness and practicing fairness in different disciplines.

3.6.1 Measuring fairness

As fairness is a variable and complex concept in FL, it is much more difficult for users to understand what is happening and whether they are treated fairly. Existing FL models primarily rely on performance evaluation metrics like accuracy and efficiency, which may not adequately capture fairness considerations. Adapting fairness metrics to evaluate model quality in a collaborative environment is crucial, as global model accuracy may vary among clients and may not align with individual client contributions. However, measuring fairness using all clients' data in a closed environment is often infeasible, and measuring fairness locally using only client data is inadequate due to limited data and the unknown distribution of other clients' data. A user-friendly and explainable fairness framework that accounts for fairness among clients would be an ideal approach to alleviate the challenges associated with fairness management.

3.6.2 Practice of fairness in different disciplines

Fairness is a widely studied concept in various disciplines, encompassing incentives, resource allocation, performance evaluation, privacy, client reputation, and addressing group bias. Incorporating fairness into FL requires algorithmic modifications, which primarily rely on the support and intervention of platforms and servers. Although clients may not directly influence algorithmic changes, the issue of ensuring fairness poses a significant challenge for them in the context of FL. The lack of control over the implementation of fairness becomes a noteworthy obstacle that affects their participation.

4 State-of-art-solutions on the client-side challenges

In this section, we will explore the current state-of-the-art solutions for the challenges faced by clients in FL. The research focus on these challenges has significantly grown in recent years, as evidenced by the increased number of papers published between 2020 and 2021.

Addressing client-side challenges in FL is typically expected to rely on users' involvement, as the architecture empowers them with greater control over their data. However, this approach can lead to unintended consequences when complex challenges are delegated to users with limited technical expertise. Many client-side challenges, such as resource management, fairness, security, and incentive management, require the involvement of platforms and servers for effective solutions. These challenges are inherently tied to the tasks and algorithms of the system.

For instance, achieving fairness in FL requires the integration of fairness considerations within server-side algorithms to ensure equal treatment of all clients. Similarly, platform-level changes are necessary to enhance transparency in incentive mechanisms, such as incorporating blockchain technology. Therefore, this section will explore solutions that involve collaboration between clients, servers, and platforms to address these challenges.

4.1 Solutions for personalization challenges

This section reviews the existing approaches proposed in the literature to address the personalization challenge in the FL. These approaches can be categorized into two main types: single-client-based personalization and cluster-based personalization. The single client-based personalization approaches focus on enhancing personalization by employing additional algorithms directly on individual clients. These algorithms aim to adapt the model to better suit each client's specific characteristics and preferences, resulting in a more personalized model.

On the other hand, the cluster-based personalization approaches involve grouping together clients with similar data distributions and objectives. By forming clusters of similar clients, the FL process can generate a more tailored and personalized model that aligns with the common characteristics and goals of the cluster members.

By exploring these two categories of approaches, we gain insights into the diverse strategies employed to address the personalization challenge in FL.

4.1.1 Single client-based personalization

Single client-based personalization algorithms aim to enhance personalization by involving clients directly in the process. These algorithms introduce additional calculations or modifications on the client side, allowing clients to actively participate in improving their own personalization. They tune the global model based on the client's data to improve the model's accuracy. We classify these approaches as fine-tuning methods and local model-global model closeness methods.

Fine-tuning approaches try to minimize the individual loss of each client by making small adjustments in the global model using a few gradient steps on gradient values based on the client's data (Deng et al. 2020). Adjusted models give accurate and personalized results for the clients. It can be considered a post-processing method. Local-model-global-model closeness approaches find the closeness between local and global models to achieve optimal personalization points for the client (Li et al. 2021f).

4.1.1.1 Fine-tuning approaches Fine-tuning approaches assume a similarity in the task across all clients but adjust the loss function based on each client's data distribution. These methods involve applying fine-tuning algorithms to the global model on the clients' edge, enabling personalization based on their individual data.

Meta-learning is a fine-tuning approach that involves training an initial model on multiple tasks, enabling it to quickly adapt and learn new tasks with limited training data on the client's end (Kulkarni et al. 2020; Jiang et al. 2019a). One notable concept in meta-learning is Model-Agnostic Meta-Learning (MAML), introduced by Finn et al. (2017), which is compatible with various deep learning models trained using gradient descent. The MAML framework consists of two main steps: meta-learning, where the model is trained on multiple tasks, and meta-testing, where the model adapts to a new task. Researchers Jiang et al. (2019a) and Fallah et al. (2020) have applied the MAML concept in the context of FL. Another meta-learning approach is parameterized algorithms (Chen et al. 2018), where clients receive algorithm parameters instead of a global model. This allows clients to fine-tune the algorithm based on their specific data for personalized learning. Additionally, studies by Khodak et al. (2019) and Balakrishnan et al. (2021) have extended meta-learning techniques to address dynamic environments and efficient resource allocation, respectively.

Base + Personalization layers with local parameters is a fine-tuning approach that involves clients sharing a common set of base layers with consistent weights while each client maintains individual personalization layers tailored to their specific data (Arivazhagan et al. 2019). This approach allows clients to incorporate their unique data characteristics while benefiting from the shared knowledge in the base layers. The model parameters derived from base layers are shared with the server, while calculation from personalized layers is retained in the client. Cheng et al. (2021) and Jourdan et al. (2021) further improved personalized layers with a stylized regression model and local adaptation.

Collins et al. (2021) introduced an approach that combines a low-dimensional local model with a learned global model to address the personalization challenge. The algorithm utilizes gradient updates to learn a global representation, enabling clients to compute personalized low-dimensional classifiers for individual labeling (Liang et al. 2020). Similarly, the approach proposed by Liang et al. (2020) also adopts a similar strategy of learning features locally and globally.

Transfer learning is learning a new task by transferring the knowledge gained from the other tasks (Torrey and Shavlik 2010). Wang et al. (2017) adapt this technology in FL to tackle the personalization challenge. FedHealth (Chen et al. 2020c) applied transfer learning in the healthcare FL domain for personalization.

Knowledge distillation is another technique where a smaller model (student) learns from a larger network (teacher) by mimicking its behavior (Li and Wang 2019). The studies (Li and Wang 2019; Ozkara et al. 2021; Divi et al. 2021a; Yu et al. 2020d) adapted the knowledge distillation technique into FL to improve personalization and communication efficiency.

Nadiger et al. (2019) adapted the reinforcement learning technique as the fine-tuning approach to make decisions sequentially. It employs trial-and-error procedures until a solution is found for a task in the client (Kaelbling et al. 1996). Hard et al. (2018) used contextual information such as logs and caches to fine-tune the character recognition task. They showed that adding contextual information boosted personalized performance.

Recent works (Yurochkin et al. 2019; Achituve et al. 2021; Yue and Kontar 2021; Kontoudis and Stilwell 2022) have integrated Bayesian and Gaussian Process (GP) techniques in FL to achieve personalized global models. More specifically, by incorporating prior information, the local data on each client can be leveraged as a personalization role in training FL algorithms. Yurochkin et al. (2019) proposed Bayesian nonparametric FL of neural networks, synthesizing a more expressive global network without additional supervision.

Achituve et al. (2021) shared a kernel function across all clients, employing a personal GP classifier for each client. Similarly, Yue and Kontar (2021) utilized GP in their regression framework (FGPR), resulting in personalized global models by jointly learning a global GP prior across all clients. Kontoudis and Stilwell (2022) incorporated GP in training and optimization using alternating direction method of multipliers, employing decentralized aggregation techniques for GP prediction through iterative and consensus methods.

Apart from discussed techniques, Li et al. (2021a) proposed a heterogeneous masking technology for fine-tuning, where clients learn a personalized and structured sparse model without changing local model parameters. Dinh et al. (2020) regularises the loss algorithm using the Moreau envelope to improve the personalized results. The works (Hu et al. 2020b) and (Yang et al. 2021b) applied differential privacy (DP) to achieve personalization. Zhang et al. (2021d) achieved personalization by allowing users to transfer personalized knowledge (update prediction) to the server. The global is getting updated based on the clients' predictions updates.

The fine-tuning process is efficient and rapid due to the internal representation of multiple models, allowing for excellent performance on new tasks with minimal data points and training iterations.

4.1.1.2 Local model-global model closeness approaches

Smith et al. (2017) adapted multitask learning to measure the closeness between the local model and global model. Multitask learning is the process of modeling naturally related tasks at a time and measuring the relationship among them. The studies (Mills et al. 2020; Mahara et al. 2021) extended Smith et al. (2017)'s research in different domains. Yu et al. (2020b) combined reinforcement learning with multitasking to achieve better results. Recently, Li et al. (2021f) proved that multitask learning can improve fairness and robustness along with personalization.

Model interpolation is defined as training a separate local model based on the local and global data and combining them for better performance in FL (Mansour et al. 2020). The studies (Peterson et al. 2019; Hanzely and Richtárik 2020) adapted this technology with experts' opinions in the domain to build personalized models for clients. Mansour et al. (2020) integrated model interpolation with clustering and data interpolation (training a model on combined local and global data) for better results. The studies (Deng et al. 2020; Zhang et al. 2020c; Luo and Wu 2021) achieved personalization by allowing the clients to build their local models simultaneously with global model building. They used the optimal mixing parameter to mix global and local models. Wu et al. (2021b) presented a hierarchical personalized FL framework in which clients initially define hierarchical information about their data (public and private). Only the public component will be uploaded to the server.

While the techniques mentioned above are commonly employed in the literature to achieve personalization, they rely on algorithms that clients have no control over. Moreover, the resource-constrained client environment makes it computationally challenging to perform the required additional calculations. Clients must allocate their limited computational power to accommodate these algorithms in order to obtain personalized models.

4.1.2 Cluster-based personalization approaches

Single client-based personalization approaches conform when clients' data distributions are similar. But, when the data distribution is naturally clustered among clients, finding an optimal personalized solution for all clients is difficult. Clustering has been proposed as a personalization solution in the literature. By grouping similar clients together, the model can mitigate the impact of heterogeneous data distribution.

Various clustering techniques have been proposed in the literature to address personalization in FL. These techniques include hard clustering, soft clustering, hypothesis-based clustering, attribute-based clustering, hierarchical clustering, and user-centric clustering. These approaches are typically implemented on the server side, as they involve collecting and aggregating local models. The clustering of clients is often determined based on the values of model parameters (Table 3).

Hard clustering means assigning a client to only one cluster; A client cannot belong to two clusters. The studies (Ghosh et al. 2020; Vahidian et al. 2021; Huang et al. 2019; Duan et al. 2021; Xie et al. 2021; Li et al. 2020f) adapted hard clustering technique in a federated environment to iteratively assign the clients in clusters. The cluster that provides the least loss updates was selected as the appropriate cluster for that client. Sattler et al. (2020) incorporated the hard clustering technique with multi-task learning based on cosine similarity between the gradient updates.

However, hard clustering in FL faces certain challenges, including unstable training, sub-optimal user assignment, and inefficiency when dealing with a large mix of data distributions. Researchers have introduced a solution known as soft clustering to address these issues. Unlike hard clustering, soft clustering allows clients to be partially assigned to multiple clusters, creating overlapping clusters. The approach presented by Li et al. (2021b) provides enhanced flexibility and robustness in addressing the challenges arising from client heterogeneity in FL.

As another approach, Mansour et al. (2020) applied hypothesis-based clustering, where clients are partitioned according to the best hypothesis based on a stochastic expectation maximization algorithm. Further, a hierarchical clustering approach was adapted in (Briggs et al. 2020; Yoo et al. 2021) to group clients using the similarity between local updates and the global server. The hierarchical clustering algorithm iteratively merges the most similar clients in each round until a given threshold.

In addition to the ones previously discussed, a user-centric federated clustering approach was proposed by Mestoukirdi et al. (2021), aiming to minimize communication overhead in FL. Instead of relying on the generic federated averaging algorithm, they introduced multiple user-centric aggregation rules in the server to obtain clustering results. Unlike traditional approaches that use model parameter values for clustering, they focused on client characteristics such as data size and distribution from the server's perspective. Another recent study by Kim et al. (2021) explored dynamic clustering, which adapts the clusters based on the changing environments.

However, clustering approaches in FL have limitations such as limited client control, privacy risks during data transfer, and a fixed number of clusters, except in the case of dynamic clustering (Kim et al. 2021).

4.2 Solutions for privacy management challenges

This section delves into state-of-the-art solutions that target different client privacy challenges. These challenges are classified based on the precise issues we discussed in subsection 3.2 regarding privacy management.

Table 3 Summary of personalization approaches

Approaches	Technology	References	Solution end
Fine tuning approaches	Meta-learning	Jiang et al. (2019a), Fallah et al. (2020), Chen et al. (2018), Khodak et al. (2019), and Balakrishnan et al. (2021)	Clients
	Base + Personalization layers	Arivazhagan et al. (2019), Cheng et al. (2021), and Jourdan et al. (2021)	Clients
	Combining a low dimensional local model with a learned global model	Collins et al. (2021) and Liang et al. (2020)	Clients
	Transfer Learning	Chen et al. (2020c) and Wang et al. (2017)	Clients
	Knowledge Distillation	Li and Wang (2019), Ozkara et al. (2021), Divi et al. (2021a) and Yu et al. (2020d)	Clients
	Reinforcement learning	Nadiger et al. (2019)	Clients
	Contextualization	Hard et al. (2018)	Clients
	Bayesian and Gaussian FL	Yurochkin et al. (2019), Achituve et al. (2021), Yue and Kontar (2021) and Kontoudis and Stilwell (2022)	Clients
	Other	Li et al. (2021a), Dinh et al. (2020), Hu et al. (2020b), Yang et al. (2021b) and Zhang et al. (2021d)	Clients
Local–global model closeness approaches	Multi-task learning	Li et al. (2021f), Smith et al. (2017), Mills et al. (2020), Mahara et al. (2021) and Yu et al. (2020b)	Clients
	Model interpolation	Deng et al. (2020), Mansour et al. (2020), Peterson et al. (2019), Hanzely and Richtárik (2020), Zhang et al. (2020c), Luo and Wu (2021), Wu et al. (2021b), Chou et al. (2021) and Li et al. (2020a)	Clients
Technologies to cluster similar clients			
Clustered model approaches	Hard clustering	Ghosh et al. (2020), Sattler et al. (2020), Vahidian et al. (2021), Ma et al. (2021), Cho et al. (2021), Huang et al. (2019), Duan et al. (2021) and Xie et al. (2021)	Server, Clients
	Soft clustering	Li et al. (2021b)	Server, Clients
	Hypothesis-based clustering	Mansour et al. (2020)	Server, Clients
	Hierarchical clustering	Briggs et al. (2020) and Yoo et al. (2021)	Server, Clients, Platform
	User-centric clustering	Mestoukirdi et al. (2021)	Server, Clients
	Dynamic clustering	Kim et al. (2021)	Server, Clients

4.2.1 Auditability in privacy management

The current state-of-the-art is on the basics of two main techniques for auditing: blockchain and visual analytics.

4.2.1.1 Blockchain technology for auditability Researchers utilize Blockchain as a popular technology for auditing FL transactions due to its robustness, immutability, and auditability (Swan 2015). These platform-based solutions involve integrating blockchain with FL's architecture to address privacy challenges.

A study by Lu et al. (2019) proposed a privacy-preserving FL data-sharing architecture that leverages blockchain. The blockchain is utilized to store information related to client selection, data statistics, encrypted retrieval transactions, data sharing requests, and transactions while ensuring the privacy of raw data. Similarly, FLchain (Majeed and Hong 2019) utilizes blockchain to store local model parameters as blocks, enabling trackability throughout global iterations.

The studies (Lu et al. 2019; Zhao et al. 2020a) also adapted blockchain in the IoT domain. Zhao et al. (2020a) replaced the central server with a blockchain-based system to store and aggregate local models, enabling tracking of malicious activities. BlockFlow (Mugunthan et al. 2020b) is another decentralized FL system that leverages blockchain to provide auditability, reward clients for their contributions, and offer protection against malicious adversaries.

Another study, VFChain (Peng et al. 2021), is a verifiable and auditable blockchain-based framework. A selected committee aggregates and records the local models after verifying them. Lo et al. (2021a) enhance FL architecture with reliability, accountability and fairness by integrating blockchain. Accountability is achieved by designing a smart contract-based data-model provenance registry.

To address the auditability challenge and comply with data privacy regulations, servers have implemented blockchain-based architectural changes. These changes aim to increase participant transparency and trust, attracting more client participation. Although users may not possess the technical knowledge to fully understand blockchain architecture, they can rely on its immutable and transparent nature. By recording all FL activities, the blockchain allows users to verify their transactions with the support of legal and technical expertise when necessary.

4.2.1.2 Visual analytics for auditability Visual analytics approaches empower users by involving them directly in the FL process, providing visibility into various aspects of FL activities. These approaches enable users to monitor and analyze data usage, client information, model aggregation, and accuracy distribution. By offering this level of transparency, users can fulfill their audit objectives and gain insights into the inner workings of FL.

Turbo Tucoon (Mike 2018) and FATE-Board (Fan 2018) are shallow-level analytical tools for FL. Turbo Tucoon summarises process logs and model performance for users to monitor and visualize the system. FATE-Board visualizes real-time log metrics, dataset information, task workflow, model output, and evaluation metrics. Wei et al. (2019) proposed a multi-agent visualization system demonstrating FL, multi-client coordination, input, and output through a game. However, the domain of this approach is for car racing games, and it is difficult to generalize to all the domains.

LEAF (Caldas et al. 2018) is a benchmark visual analysis tool with statistical and system metrics. LEAF can be applied in FL, meta-learning, multitask learning, and on-device learning. But, LEAF is mainly designed for tech-savvy users, mainly software engineers; It can be complicated for general users. PrivacyFL (Mugunthan et al. 2020a) helps users ensure collaboration is feasible and improve their model accuracy.

FedEval is a comprehensive, easy-to-use benchmarking framework that comprises accuracy, communication, time efficiency, privacy, and robustness. Although the system is primarily built for researchers to perform evaluations, users can visualize their performances.

HFLens (Li et al. 2021d) is a comparative visual interpretation system for fine-grained analysis of communication rounds and client instance levels. It analyses the overall client processes, correlation of clients' information with communication rounds, potential anomalies, data quality, and client contribution.

4.2.2 Consented data sharing in privacy management

Limited research focuses on addressing the challenge of consented data sharing in the FL environment. These approaches typically require collaboration between clients and servers. The server is responsible for obtaining user consent before collecting and sharing their data, while users have the autonomy to decide whether or not to share their data with the server.

The policy-based privacy setting framework "PoliFL" (Katevas et al. 2020) offers users a feature to choose which data to share in FL. Users can opt out of certain data based on their privacy preferences. DS2PM (Chen et al. 2021) protects privacy, integrity, and data ownership using blockchain. Data sharing occurs using an on-chain data retrieval mechanism with owner permission. The framework ensures the auditing and verification of transactions too.

4.2.3 Granularity in privacy management

Granularity solutions primarily rely on user-driven privacy settings mechanisms, allowing users to control the level of granularity in data sharing. Users have the flexibility to define the extent and specifics of data they are willing to share.

PoliFL (Katevas et al. 2020) offers heterogeneous privacy policies as users may have different privacy requirements. The server is responsible for aggregating locally processed models with different datasets. PoliFL considered three policies: a policy that permits all FL activities, a policy that permits FL with DP, and a policy that restricts specific data sources (varied among users) when training the FL model. The results showed that PoliFL performs well with heterogeneous policies within reasonable resource and time budgets.

Similarly, using an opt-out DP algorithm, the FeO2 framework (Aldaghri et al. 2021) protects clients' privacy. Users may opt out of certain features of their data or additional privacy-enhancing mechanisms based on their privacy needs.

4.2.4 Re-identification in privacy management

Privacy-preserving techniques in the literature for the re-identification challenge can be categorized into three main approaches: perturbation, Secure Multiparty Computation (SMC), and Homomorphic Encryption (HE). These approaches typically involve clients and platforms in order to safeguard against re-identification and uphold privacy. They achieve this by introducing an additional layer of protection in FL through techniques such as noise injection, secret sharing, or encryption of the model parameters.

4.2.4.1 Perturbation approaches Perturbation techniques, such as DP, involve adding noise to local parameters to ensure privacy in FL. DP creates anonymous data by introducing noise, allowing for statistical analysis without revealing sensitive personal information or individual client identities. Client-side perturbation, known as local differential privacy (LDP), involves data owners adding randomization or noise to their data before sharing it with a third party, addressing privacy concerns in FL (Dwork 2009; Tyagi 2022).

If a randomized algorithm $\mathcal{A}$ satisfies Eq. (4), then it provides ε-LDP.

Definition 1 An algorithm $\mathcal{A}$ satisfies ε-LDP if, for any two data values v_1 and v_2, and for any output Q within the range of outputs of $\mathcal{A}$, the following holds Eq. (4):

$$P\left[\mathcal{A}(v_1) \in Q\right] \leq exp(\varepsilon)P[\mathcal{A}(v_2 \in Q)] \tag{4}$$

Here, ε represents the privacy budget, quantifying the level of privacy. This definition ensures that the probability (P) of obtaining an output Q from $\mathcal{A}$ on v_1 is at most e^{ε} times the probability of obtaining the same output Q on v_2.

Geyer et al. (2017) tackled the re-identification issue from the client's standpoint. They used client-side DP to preserve complete data privacy and optimize performance. The studies (Wei et al. 2020a; Choudhury et al. 2019) also adapted DP to prevent information leakage by adding noise before aggregation. They showed that different variations of artificial noise lead to different levels of protection.

Another client-side approach, LDP, was adapted in (Zhao et al. 2020b; Seif et al. 2020; Truex et al. 2020), where clients locally perturb their data before sharing. The local privacy approach reduces communication costs and privacy threats. Wei et al. (2021) derived user-level DP algorithm extending the local privacy. Rather than guaranteeing only the privacy of individual samples, user-level DP protects a client's entire contribution. As the extension of DP, Triastcyn and Faltings (2019) adapted Bayesian DP, which adjusts noise according to the data distribution instead of the random adjustment.

In their work, Marathe and Kanani (2022) focused on subject-level privacy in FL, where a subject's private information is represented by multiple data items within or across federation clients. They achieved subject-level privacy by introducing noise to the data and training the noisy data in mini-batches. This approach aimed to protect the privacy of individual subjects while enabling effective collaborative learning in FL.

Sherpa.ai framework (Rodríguez-Barroso et al. 2020) was built based on FL and DP. This framework helps to build FL with DP without developing from scratch using the offered functionalities. So, clients with limited technical knowledge can use this framework to integrate DP technology into their data.

4.2.4.2 Secure Multiparty Computation approaches SMC is a platform-based solution that aims to enable the secure computation of a consensual function among clients without relying on any trusted third parties (Goldreich 1998). In SMC, input data is either masked or secret shared, and the computed result is typically disclosed to all parties involved. SMC offers the advantage of relatively low computational overhead, but it necessitates multiple rounds of interaction among the participating parties to achieve secure computation.

Figure 4 provides an overview of the SMC approach, depicting the secure sharing of private inputs among parties, the use of secure protocols for computations (function *F*), iterative interaction for communication and secure computations, and result in reconstruction for revealing the final result while preserving privacy.

Within the domain of FL, Truex et al. (2019) presented a privacy-preserving framework that combines DP and SMC. Their approach aimed to strike a balance between these two techniques, minimizing noise injection while preserving privacy. The framework incorporated a tunable trust parameter to accommodate various trust scenarios, ensuring both accuracy and privacy assurance in the FL process.

Another approach, HybridAlpha (Xu et al. 2019a) uses DP and functional encryption to employ SMC protocol. Functional encryption protocol supports mitigating inference attacks from curious aggregators and colluded clients.

4.2.4.3 Homomorphic Encryption approaches Homomorphic Encryption (HE) allows computations to be performed on encrypted data. In HE, as shown in Fig. 5, clients send encrypted data to a server and request the evaluation of a function on this encrypted data. The computation operates solely on encrypted data, with the inputs and outputs encrypted using the client's secret/public key, ensuring the privacy and security of the data.

In the FL domain, local models and public and private keys are encrypted before being sent to the server. HE allows performing operations over encrypted models. So, clients are mainly involved in these solutions to protect their data.

Hao et al. (2019) adapted HE in industrial FL, which prevents data reidentification even though many clients collude with each other to attack the system. Pivot (Wu et al. 2020b) protects clients' data against semi-honest adversaries. It is a hybrid framework with Threshold Partially Homogeneous Encryption (TPHE) and Multipartite Computation (MPC). Another framework, PFMLP (Fang and Qian 2021), transfers local models' encrypted gradients instead of raw gradients. However, they showed that the accuracy in homomorphic operation after decryption did not change much compared to plain text data.

Rather than adapting technology to reduce reidentification, Wei et al. (2020b) presented a framework for evaluating and comparing different forms of client privacy leakage attacks and methods to solve adversaries. The framework first provides experimental evidence of data reconstruction from model parameters. They then investigated how different hyperparameter configurations, serial compression ratios, and different settings of attack algorithms influence attack effectiveness and cost.

Despite various attempts to address re-identification challenges, recent research (Naseri et al. 2022) has demonstrated that privacy attacks can still succeed even with privacy-preserving mechanisms.

In response, blockchain techniques have emerged as a potential solution, offering benefits in terms of auditability and consented data sharing. However, the computational complexity and user-friendliness of blockchain approaches remain significant drawbacks. Nonetheless, adopting a single blockchain approach can simplify the environment by eliminating the need for multiple techniques.

In Table 4, a summary of privacy management techniques is provided, focusing on auditability, consented data sharing, data granularity, and re-identification. The table highlights the technologies utilized in state-of-the-art approaches for addressing these privacy concerns.

4.3 Solutions for incentive management challenges

The literature used different techniques for incentive management in FL such as shapely value (Yu et al. 2020a; Song et al. 2019; Wang et al. 2019a; Lim et al. 2021), contract theory (Kang et al. 2019a, b; Saputra et al. 2020), auction theory (Zhang et al. 2021e; Le et al. 2020; Zeng et al. 2020), game theory (Tu et al. 2022; Sarikaya and Ercetin 2019; Ng et al. 2021), blockchain (Weng et al. 2019; Zhang et al. 2021f), and reinforcement learning (Zhan et al. 2020; Jiao et al. 2020). In this survey, we will not delve into these techniques as our primary focus in incentive management is addressing the client-side challenge of "Transparency". Blockchain technologies and visual analytics tools are leveraged in conjunction with the aforementioned technologies to achieve transparency in incentive calculation.

4.3.1 Blockchain based solutions

Blockchain, a decentralized peer-to-peer digital ledger, offers robustness. Integrating blockchain into FL requires platform-based modifications to enable transparency. Additionally, blockchain addresses server-side challenges such as identifying malicious clients, task publication, client selection, incentive calculation or allocation, and regulatory compliance, making it an attractive solution for FL.

FLchain (Bao et al. 2019), DeepChain (Weng et al. 2019), and FIFL (Gao et al. 2021) are reputation-based incentive approaches that adapt blockchain to prevent malicious transactions by storing and monitoring all transactions. The probability of receiving rewards on the blockchain nodes is determined based on the client's confidential, transparent, and auditable previous rewards.

The studies (Zhang et al. 2021e; Toyoda and Zhang 2019) are incentive approaches based on auction theory where auxiliary functions such as task request, client selection, incentive allocation, and logging are made transparent by blockchain. Based on the data in the blockchain, rewards are transparently distributed among clients. FedCoin (Liu et al. 2020a) immutably records the incentive allocations based on proof of Shapley protocol in the blockchain. Fedcoin does not rely on a central server to distribute payments between clients with non-repudiation and tamper-resistant properties.

Refiner (Zhang et al. 2021f) handles malicious participants and incentives by auditing records on the blockchain using trusted validators. Participants randomly select validators to test local model updates with the validation data set. Incentives are distributed based on model quality assessed by validators.

An incentive mechanism based on Bayesian game theory, the Fedserving framework (Weng et al. 2021), adapted the blockchain to regulate transparent transactions between participants. They incorporated a "truth-finding" algorithm to learn accurate predictions and made them transparent using the blockchain.

While blockchain technology has successfully addressed the transparency and confidentiality issues in incentive mechanisms, it does come with certain drawbacks. The implementation of blockchain can be resource-intensive and costly. Moreover, a portion of the FL profits needs to be shared with blockchain miners who validate transactions on the network.

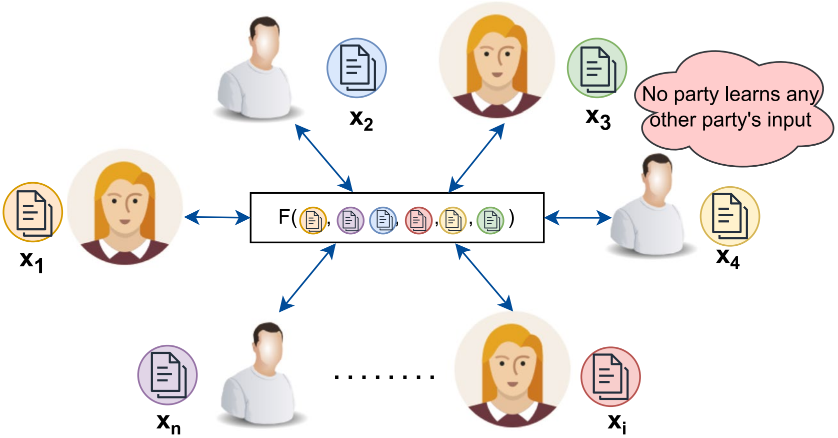
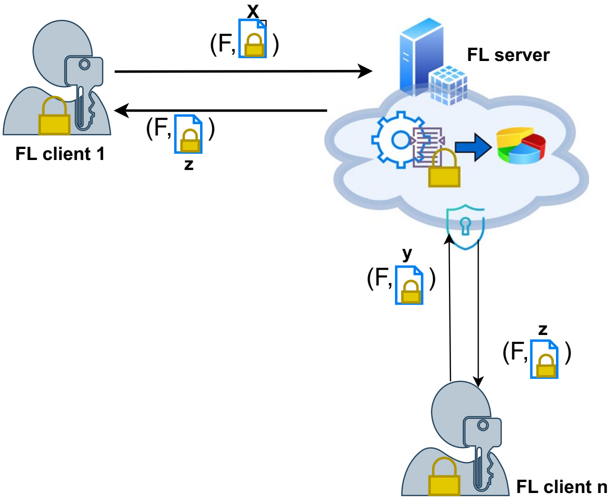


Fig. 4 The overview of secure multiparty computation approach

Fig. 5 The overview of homomorphic encryption approach



4.3.2 Visual analytics-based solutions

In addition to blockchain-based solutions, visual analytics tools play a crucial role in providing users with insights and transparency in the FL environment. These tools use graphical representations such as graphs, charts, and maps to visually present data and enable users to identify patterns and processes. For example, a study by Ng et al. (2021) implemented a visual analytics tool inspired by multiplayer games to enhance transparency in client incentive schemes. This tool offers clients an overview of the FL system, federation information, client statistics, data quality and quantity, market share changes, information about other clients, profit/loss details, and summary information. By leveraging these visualizations, clients can effectively assess their incentive scheme and make informed decisions.

Table 5 summarises the research based on all technologies in terms of the technology used and the factors driving the incentive scheme.

4.4 Solutions for resource management challenges

This section describes solutions to efficiently manage client resources such as data, computation, and communication costs.

4.4.1 Data management

Only a few studies focused on data management issues such as data imbalance and data representation. The literature used augmentation and reinforcement learning techniques for data imbalance and scarcity challenges. The majority of these approaches involve server or platform-based solutions, wherein the server simulates the clients' environments by requesting certain data from clients and generating solutions to be implemented on the client side. It reduces the burden on users while trading off privacy.

4.4.1.1 Solutions for data imbalance Data augmentation is an approach to address the data imbalance challenge by expanding the dataset through techniques such as generating slightly modified copies of existing data or synthesizing new data (Van Dyk and Meng 2001).

A server-side approach to address the data imbalance challenge was proposed in a study by Jeong et al. (2018), where the server simulated a client environment using a few data samples from clients and augmented the data to build a generative model. Shin et al. (2020) introduced a privacy-preserving XOR-based mixup data augmentation technique that involved adding encoded data from other clients to balance the training data. Duan et al. (2020) employed Z-score-based data augmentation and used Kullback Leibler divergence-based rescheduling to handle data imbalance. Wu et al. (2020a) utilized the synthetic minority oversampling technique (SMOTE) among trustworthy clients in a smart home setting, requiring platform-based changes. Apart from that, using reinforcement learning, Zhang et al. (2021b) addressed data imbalance without sending data to servers by optimizing client selection and global update frequency.

4.4.1.2 Solutions for data representation Blockchain technologies have been widely utilized to address data storage management challenges. Martinez et al. (2019) leveraged blockchain to securely store client data, enabling secure uploading, recording, and tracking. In a similar approach, Moon et al. (2020) proposed an AI-based data management system that effectively manages data between servers and clients. This system stores important client data characteristics, such as size and distribution, to assist clients in organizing their data efficiently.

4.4.2 Computation cost management

Researchers have devoted significant attention to managing computational and communication costs, as these directly impact users' participation, especially considering the resource-constrained nature of clients' devices. Various techniques have been explored to

Table 4 Summary of privacy management approaches

Precise issues	Techniques and references	Solution end
Auditability	Blockchain (Lu et al. 2019; Majeed and Hong 2019; Zhao et al. 2020a; Mugunthan et al. 2020b; Peng et al. 2021; Lo et al. 2021a; Chen et al. 2021; Fan et al. 2020; Bao et al. 2019)	Platform
	Visual analytics (Mike 2018; Fan 2018; Wei et al. 2019; Caldas et al. 2018; Mugunthan et al. 2020a; Li et al. 2021d; Han et al. 2019; Chai et al. 2020a)	Clients
Consented data sharing	Policy based (Katevas et al. 2020), Blockchain (Chen et al. 2021)	Clients, Platform
Data granularity	Policy based (Katevas et al. 2020), opt-out algorithm (Aldaghri et al. 2021)	Clients
Data re-identification	DP (Lu et al. 2019; Zhao et al. 2020a; Mugunthan et al. 2020b; Aldaghri et al. 2021; Geyer et al. 2017; Wei et al. 2020a; Choudhury et al. 2019; Rodríguez-Barroso et al. 2020), Variations of DP (Mugunthan et al. 2020a; Zhao et al. 2020b; Seif et al. 2020; Truex et al. 2020; Wei et al. 2021; Triastcyn and Faltings 2019; Marathe and Kanani 2022), Combination of DP and Secure multiparty computation (Truex et al. 2019; Xu et al. 2019a), Homomorphic Encryption (Chen et al. 2021; Hao et al. 2019; Fang and Qian 2021), Combination of Homomorphic encryption and multiparty computation (Wu et al. 2020b), Other techniques (Rodríguez-Barroso et al. 2020; Wei et al. 2020b; Yuan et al. 2021)	Clients, server, Platform

address computation-related challenges, including computation reuse, edge/fog computing, algorithm optimization, blockchain, reinforcement learning, and clustering.

Nour et al. (2021) proposed a computation reuse approach to store model parameters of previously executed tasks with a high probability of repetition, aiming to eliminate redundant computations. Edge-assisted FL techniques, as explored in studies by Ren et al. (2019), Ji et al. (2021), and Al-Abiad et al. (2021), leverage edge computing to reduce the computational burden on clients. These approaches allow clients to compute model parameters locally or offload computations to edge devices based on resource availability. Wang et al. (2021c) introduced the use of high-altitude balloons (HABs) as flying wireless base stations to offload clients' computational burden. These HABs dynamically adjust user association, service sequence, and task partition schemes to cater to clients' needs over time.

The studies (Xu et al. 2019b; Chen et al. 2020a; Prathiba et al. 2021; Do et al. 2021) optimized the FL algorithm on client devices or the global model in each training cycle to achieve better computational allocation. The server handles most optimization techniques, although they need to be executed on the client devices. The intelligent UDEC (I-UDEC) framework (Yu et al. 2020c) combines reinforcement learning and blockchain to obtain computation offload decisions in real time with low overhead and resource allocation strategies.

A resource management scheme based on clustering was proposed by Balakrishnan et al. (2021). The authors clustered clients according to their data and learned a federated meta-model from a subset of clients within each cluster. This approach allowed for efficient model building by organizing the process based on client clusters, resulting in personalized results while reducing communication and computation time.

4.4.3 Communication cost management

Researchers have explored several approaches to effectively manage communication costs in FL. These include compression techniques, reducing communication rounds, minimizing communication distance, and optimizing client selection.

4.4.3.1 Data compression Model compression schemes such as sparsification and quantization are widely used in the literature to reduce the size of local and global models during transfer (Sattler et al. 2019a). Sparsification methods limit the changes to only a small subset of the model parameters to reduce the entropy of the updates. Several studies adapted sparsification techniques such as transferring only model gradients greater than a predefined threshold (Strom 2015), updating only significant gradients (Li et al. 2020d), uploading the model after gradient sparsification (Li et al. 2020c, 2021c; Asad et al. 2020), optimally compressing parameter matrix of model convolutional layers (Zhou et al. 2020), downstream and upstream model compression with an encoding of the weight updates (Sattler et al. 2019a), compressing model gradient to Count Sketch data structure (Rothchild et al. 2020), gradient perturbation (Hu et al. 2020a), using the sparse binary mask technique (Li et al. 2021a), and dynamically adjusting sparsity budgets of the gradient compression variables (Nori et al. 2021). However, sparsification compression methods may not be suitable for many clients.

Quantization methods reduce the entropy of the model updates by restricting all updates to a reduced set of values (Sattler et al. 2019a). The quantization compression scheme is adapted in many studies, such as quantizing each gradient update to its binary sign

Table 5 Summary of incentive schemes in terms of technology and incentive factor

Technology	Incentive schemes and references	Solution end
Blockchain	Clients reputation (Bao et al. 2019; Weng et al. 2019; Gao et al. 2021)	Platform, Clients
	Auction theory (Zhang et al. 2021e; Toyoda and Zhang 2019)	Platform, Clients
	Shapley protocol (Liu et al. 2020a)	Platform, Clients
	Model quality (Zhang et al. 2021f)	Platform, Clients
	Bayesian game theory (Weng et al. 2021)	Platform, Clients
Visual tool	Data contribution (Data quality and quantity) (Ng et al. 2021)	Clients

(Bernstein et al. 2018), stochastically quantize the gradients during upload in an unbiased way (Wen et al. 2017; Chang and Tandon 2020), using vector quantization technique per iteration (Dai et al. 2019; Shlezinger et al. 2020), applying encoding-based compression (Chai et al. 2020b; Malekijoo et al. 2021), and using lossy compression (Amiri et al. 2020). Although the existing approaches for managing communication costs in FL are theoretically sound and exhibit convergence properties, their empirical performance falls short compared to the sparsification method (Sattler et al. 2019a).

Konečný et al. (2016) introduced a combination of sparsification and probabilistic quantization to effectively reduce communication delays. Their method significantly decreased uplink and downlink communication time, making it suitable for large-scale deployments involving numerous clients.

One drawback of data compression methods is the unavoidable loss of information during the compression process.

4.4.3.2 Reducing the Communication Rounds Another approach to enhance communication efficiency is reducing communication rounds. This can be achieved by allowing clients to perform multiple local epochs before sending their results to the server. Instead of updating the server for every small model change, clients aggregate their updates and communicate less frequently. This approach helps reduce communication overhead and improves overall efficiency in FL. Federated Averaging (FedAvg) (McMahan et al. 2017) algorithm is commonly used to reduce the communication rounds of FL through periodic connections. FedMMD (Yao et al. 2018) adapted a two-stream model with maximum mean discrepancy to integrate more knowledge from the local and global models. However, this method increased the computational cost for the clients to reduce communication rounds.

The CMFL approach proposed by Wang et al. (2019b) effectively reduces communication overhead by controlling irrelevant client updates. Clients evaluate whether their updates align with the server's feedback and contribute to model improvement before uploading them to the communication network. In contrast, Guha et al. (2019) introduced one-shot federated learning, which aims to learn a global model in a single round of communication across a set of clients. They leverage ensemble learning and knowledge aggregation to capture global information using client-specific models. Another strategy, presented by Avdiukhin and Kasiviswanathan (2021), is the adaptation of local Stochastic Gradient Descent (SGD). This approach allows clients to evolve locally on their own asynchronously and then average the sequences in a global server after multiple iterations.

However, these methods have limitations such as increased local computation cost, potential bias in the global model due to sampling, and the absence of consideration for data heterogeneity among different clients.

4.4.3.3 Reducing the communication distance Edge computing is a notable approach mentioned in the literature to reduce communication distance. It involves deploying edge servers in close proximity to clients, facilitating communication within the edge computing infrastructure. Some studies, such as Wang et al. (2019c), Lu et al. (2020), and Liu et al. (2021a), have incorporated edge computing into their FL systems. Partial aggregation or partial training may also be performed on edge servers to further optimize communication.

Another technique to reduce communication distance is peer-to-peer learning, where clients can leverage the knowledge and expertise of other clients in the network. BrainTorrent (Roy et al. 2019) and Online Push-Sum (He et al. 2019) are examples of central server-free algorithms enabling clients to communicate exclusively with trusted neighboring clients. In the LotteryFL framework (Li et al. 2020a), a subnetwork is formed based on the lottery ticket hypothesis, allowing clients to learn personalized models instead of a single global model. Similarly, RingFed (Yang et al. 2021a) employs a ring topology instead of a star topology, enabling clients to communicate with each other before transmitting the final model to the server.

4.4.3.4 Client selection Client selection serves as another mechanism to minimize the amount of data transmitted between clients and the server to reduce communication costs. These algorithms restrict the number of participating clients in a round (only a fraction of clients in a round). The client selection algorithms are mainly implemented on the server. Clients' influence on these algorithms is limited.

The studies (Nguyen et al. 2020; AbdulRahman et al. 2020; Liu et al. 2021b; Nishio and Yonetani 2019) chose only a subset of the clients in each round to decrease the number of uploading clients. They performed sampling based on device capabilities regardless of the heterogeneity of the data. Cho et al. (2020) reduced client selection bias and addressed data heterogeneity by selecting the highest-loss clients. Instead of static sampling, Ji et al. (2020) and Zhuang et al. (2020) adapted dynamic sampling to choose the fraction of available client models and model parameters. Ribero and Vikalo (2020) used an optimal sampling strategy to select a subset of clients with significant weight updates. FedPaq (Reisizadeh et al. 2020) uses periodic global updates, partial participation of devices, and compression techniques for efficient communication.

Clients with limited resources and data often face bias in most of these approaches. However, recent studies have aimed to address the bias against clients with fewer resources and data by investigating the differentiation in local models and considering the availability of client resources.

Besides these main approaches, there are some other approaches such as ensemble (Hamer et al. 2020), model minimization (Bouacida et al. 2020; Kang and Ahn 2021), pruning (Jiang et al. 2019b), overlapping training and communication (Zhou et al. 2021), feature fusion (Yao et al. 2019b, a), and knowledge distillation techniques (Wu et al. 2021a) focused on communication efficiency (Table 6).

4.5 Solutions for data and device security challenges

This section examines the existing defense techniques for FL, focusing on three main categories: defense mechanisms against malicious clients or external attackers, defense mechanisms against malicious servers, and approaches for verifying participants and models. These techniques aim to protect clients from adversarial attacks and mitigate potential risks in the FL environment.

4.5.1 Defence mechanism against malicious clients or external attackers

Data/model poisoning is a common attack in FL where malicious attackers (can be clients or malicious agents who take control over clients) incorporate malicious data in the training phase or manipulate the global model using fake data. This attack greatly affects the client as the global model predictions will be incorrect. The literature proposed defence mechanisms such as similarity-based approach (Cao et al. 2019), generative adversarial network approach (Zhao et al. 2019), validation test set-based approach (Wang et al. 2020b; Vy et al. 2021), notions of stealth approaches (Bhagoji et al. 2019), model-agnostic defence technique (Manna et al. 2021), and anomaly detection techniques (Shen et al. 2016; Wan et al. 2021; Li et al. 2021g).

A backdoor attack is a method of injecting a malicious task into an existing model without compromising the accuracy of the original task. This attack aims to introduce a hidden trigger or pattern that can be exploited by an adversary to manipulate the model's behavior. Backdoor attacks were defended using fine-tuning (Liu et al. 2018), model pruning (Jiang et al. 2022), clients' contribution similarity (Fung et al. 2018), reverse engineering (Zhao et al. 2021), additive feature importance strategy (Manna et al. 2021), and testing the clients' data accuracy on the high-quality test set belongs to the central server (Su et al. 2022).

In addition to the defense-oriented methods discussed earlier, there are other techniques such as knowledge distillation (Li and Wang 2019) (sharing only the knowledge instead of model parameters for security), pruning (Jiang et al. 2019b) (reducing the model size without affecting the accuracy), multi-task learning (Smith et al. 2017; Sattler et al. 2020; Li et al. 2019a; Yu et al. 2020e) (personalized model to reduce the impact of affected global models) that contribute to the defense mechanism in FL. While these techniques have different primary objectives, they also contribute to improving the overall defense mechanism in FL.

4.5.2 Defence mechanism against malicious servers

In the FL environment, it is important to consider the possibility of malicious behavior from the central server. If the server is compromised, the impact of the attack can be severe, as the server has access to clients' sensitive data through model updates. However, the literature on attacks originating from malicious servers is limited, with only a few studies specifically addressing this issue.

The studies (Mo and Haddadi 2019; Chen et al. 2020b) used Trusted Execution Environment (TEE) approach to defend against the malicious server. TEE allocates private memory regions to compute with hardware and software isolation. The server's memory usage patterns are monitored every time to defend against the server's malicious attacks. Each participant is compelled to execute secure and privacy algorithms in this environment. However, hardware changes are needed to adapt these approaches on the clients' end.

Security consortiums within trusted clients and peer-to-peer learning techniques can also protect against a malicious server, as clients do not need to communicate with the central server. The studies (Roy et al. 2019; He et al. 2019; Yang et al. 2021a) built a peer-to-peer network with only trusted clients. Clients only need to know about their neighbors rather than the global network. Because these approaches bypass the central server or minimize central server communication, the impact of malicious server attacks is less than in the traditional FL network. But the client's responsible for identifying other trusted clients within a huge network.

During the FL process, inference attacks pose a threat by attempting to extract sensitive information from clients. These attacks can be initiated by either the clients themselves or a potentially malicious centralized server involved in the FL system (Hu et al. 2021). Inference attacks are defended using techniques such as DP (Liu et al. 2020), knowledge distillation (Li and Wang 2019), secret sharing or secure boost protocol (Wang et al. 2020c), Generative Adversarial Networks (GAN) based algorithm (Zhang and Luo 2020), and fake data generation at client node (Triastcyn and Faltings 2020).

4.5.3 Verification of participants and models approaches

The verification process validates whether the model, clients, and server are trustworthy or attack-free. Wainakh et al. (2020) adapted the hierarchical FL to verify participants and models. Unlike FL, hierarchical FL is not controlled by a central server; It connects to multiple servers in a tree structure, leading to granular monitoring of clients.

Blockchain technology is used in many studies (Fang et al. 2022; Liu et al. 2020; Rahman et al. 2020; Yi Ming et al. 2021; Jiang et al. 2021) to verify the models and participants. Blockchain handles verification and stores the proofs of clients in the blockchain. Fang et al. (2022) secure the confidential properties of model gradients using a secure aggregation protocol. They verified the global model gradients using blockchain to avoid a possible tampering attack.

Table 7 provides a comprehensive summary of research works, categorizing them based on the malicious actor, attack types, solution techniques, and solution end. Various approaches have been developed to detect attacks, including checking accuracy, model similarity, client contribution, and client similarity. However, it is important to note that these approaches may impact resource-constrained clients, as they could be mistakenly identified as malicious. In recent studies, blockchain-based techniques have incorporated additional factors such as client model updates, user traces, and model participation to enhance the detection of malicious activities.

4.6 Solutions for fairness management challenges

This section is divided into two parts to examine state-of-the-art solutions: fairness measurement and the application of fairness in various disciplines of FL.

Table 6 Summary of resource management approaches

Precise issues	Techniques and references	Solution end
Data management—data imbalance/scarcity	*Augmentation techniques* (over sampling/ down sampling) (Jeong et al. 2018; Shin et al. 2020; Duan et al. 2020), *SMOTE over sampling* (Wu et al. 2020a), *Reinforcement learning* (Zhang et al. 2021b)	Clients, Server
Data management—data representation	*AI based data management system* (Moon et al. 2020), *Blockchain*(Martinez et al. 2019)	Platform, Clients
Computation cost management	*Computation reuse* (Nour et al. 2021)	Clients
	Edge computing (Ren et al. 2019; Ji et al. 2021; Al-Abiad et al. 2021; Wang et al. 2021c)	Platform, Clients
	Algorithm optimization (Xu et al. 2019b; Chen et al. 2020a; Prathiba et al. 2021; Do et al. 2021)	Server
	Blockchain, Reinforcement learning (Yu et al. 2020c)	Platform
	Client clustering (Balakrishnan et al. 2021)	Clients, Server, Platform
	Data compression: Sparsification compression techniques (Li et al. 2021a; Sattler et al. 2019a; Strom 2015; Li et al. 2020d, c, 2021c; Asad et al. 2020; Zhou et al. 2020; Rothchild et al. 2020; Hu et al. 2020a; Nori et al. 2021) Quantisation compression techniques(Bernstein et al. 2018; Wen et al. 2017; Chang and Tandon 2020; Dai et al. 2019; Shlezinger et al. 2020; Chai et al. 2020b; Malekijoo et al. 2021; Amiri et al. 2020), Combined both techniques (Konečný et al. 2016)	Platform, Clients
Communication cost management	*Reducing the communication rounds/frequency of client updates* (McMahan et al. 2017; Yao et al. 2018; Wang et al. 2019b; Guha et al. 2019; Avdiukhin and Kasiviswanathan 2021)	Clients, server
	Reducing the communication distance: Edge computing (Wang et al. 2019c; Lu et al. 2020; Liu et al. 2021a), Peer-to-peer learning (Li et al. 2020a; Roy et al. 2019; He et al. 2019; Yang et al. 2021a)	Platfrom, clients
	Client selection (Nguyen et al. 2020; AbdulRahman et al. 2020; Liu et al. 2021b; Nishio and Yonetani 2019; Cho et al. 2020; Ji et al. 2020; Zhuang et al. 2020; Ribero and Vikalo 2020; Reisizadeh et al. 2020)	Server
	Other: Ensemble (Hamer et al. 2020), Model minimisation (Bouacida et al. 2020; Kang and Ahn 2021), Pruning (Jiang et al. 2019b), Overlapping the training and communication (Zhou et al. 2021), Feature fusion (Yao et al. 2019b, a), Knowledge distillation (Wu et al. 2021a)	Clients, server, platform

4.6.1 Approaches to measuring fairness in FL

With regard to measuring fairness, researchers formulated measurement metrics such as average variance, distance metric (such as cosine distance, euclidean distance, maximum difference), Pearson correlation coefficient, and Jain's fairness index (Shi et al. 2021). However, the lack of standardization among these values poses a challenge in selecting a single metric. Factors such as the metric's definition, trade-offs, and compatibility with other metrics need to be considered. Additionally, non-technical users should easily understand the chosen metric and encompass various aspects of fairness.

In the pursuit of standardization, Garg et al. (2020) introduced a mathematical framework that outlines the commonly used fairness metrics and their interrelationships. The relational representation helps users find the most suitable metric. Chu et al. (2021) proposed a new FL framework called FedFair to train models with high performance and fairness without violating client privacy. They propose an estimation method to estimate model fairness in a privacy-constrained environment that is more efficient than estimating fairness locally. The framework includes the fairness estimation function of the loss function as a constraint.

Rather than assessing the accuracy of the global model across all clients, Divi et al. (2021b) focused on evaluating the effectiveness of individualized models for each client. They examined whether the accuracy of personalized models improved for each user and observed a fair perception overall. To evaluate the quality of the personalized models, they introduced five performance metrics and four fairness metrics, which assessed whether the personalized models provided equal improvements for all users.

These approaches encompass platform-based or client-based algorithmic solutions that enable users to visualize and assess their fair treatment within the system using a range of metrics.

4.6.2 Practice of fairness in different disciplines

The concept of fairness is practiced in various FL disciplines, including contribution evaluation, client selection, model optimization, incentive mechanism, and social good. This section discusses the existing works practicing fairness in these disciplines, techniques, and adapted notions of fairness.

4.6.2.1 Client selection Unfair treatment can start during the client selection process. However, many existing client selection approaches prioritize server interests, such as accuracy improvement and convergence rate, while disregarding the interests of individual clients. These approaches often prioritize clients based on factors like bandwidth, data quality, transmission speed, and computing power. Consequently, client selection can be unfair due to over-representation, under-representation, and the exclusion of certain clients (Shi et al. 2021).

Recent studies focus on reducing bias against under-represented clients (lower computational capabilities and smaller datasets). Huang et al. (2020a) modeled the client selection strategy as a Lyapunov optimization problem, where client participation rates were optimized through a dynamic queuing approach. The algorithm ensures that each client's average participation rate equals the expected guarantee rate. Similarly, Yang et al. (2021c) proposed a multi-arm bandit-based algorithm to encourage the selection of under-represented clients. The choice depends on the class distribution of the data. Clients with minimal class imbalance will receive the highest rewards, while the system allows clients with a maximal class imbalance to participate in at least a specified number of rounds. Another approach Kang et al. (2020) used reputation measurement in terms of honesty and contribution to choose clients. Highly reputed clients get more opportunities to be selected than low-reputed clients.

4.6.2.2 Contribution evaluation Contribution evaluation assesses the individual contributions of clients within the FL system without requiring access to their data. Various methods have been proposed to evaluate client contributions, including self-reported information, individual assessment, utility game, Shapley value, and empirical approaches (Shi et al. 2021).

The studies (Kang et al. 2019b; Sarikaya and Ercetin 2019; Zhang et al. 2020b; Le et al. 2021) use clients' self-reported information to evaluate client contributions. Self-reported information can contain data quality & quantity, data collection costs, and computational & communication capabilities they can commit to FL. The server uses this information to assign ratings for the clients. This approach assumes that clients are trustworthy and capable of assessing their data environment. Clients and server have to be involved to achieve this approach's fairness goal.

Individual reputation evaluation is based on the performance of clients on specific tasks. Reputation mechanisms are designed to track the clients' reliability and contribution. Client reputation is calculated based on client validation accuracy (Lyu et al. 2020a), the similarity between local model-global model (Xu and Lyu 2020), loss function values (Song et al. 2021), and direct or indirect reputation of clients from history with task publishers (Kang et al. 2019a; Zhang et al. 2021e; Kang et al. 2020). A task publisher is responsible for assigning tasks to clients.

Along with reputation, Zeng et al. (2020) incorporates resource quality information to evaluate the individual contributions of clients. Another approach proposed by Lyu et al. (2020c) involves a mutual evaluation process among FL clients to assess their potential value. These approaches take into account the involvement of clients, server, and platform to achieve fairness in the system.

Utility games are employed to translate clients' utility into rewards, offering another avenue for fairness adoption. Wang et al. (2019a) and Nishio et al. (2020) adopted the marginal loss approach to evaluate clients' contributions. The concept of marginal loss suggests that a client's gain is equivalent to the utility lost when the client departs from the system (Shi et al. 2021). Primarily, the server plays a significant role in this approach.

Shapley value evaluates contribution by calculating the weighted average of the marginal contribution from the utility perspective and the clients' impact. The studies (Song et al. 2019; Wang et al. 2020a) evaluated the impact by calculating Shapley value in the entire training session. Wang et al. (2019a) used the Shapley value to calculate the feature importance in VFL instead of considering all client data. If a client has important features that greatly influence the model, then the client receives high Shapley values.

To reduce the computational cost of the aforementioned theoretical methods, Shyn et al. (2021) proposed FedCCEA, which approximates the client contribution using the sample data size weights in the model. The server is mainly responsible for this approach by getting sample sizes from clients.

Table 7 Summary of data/device security management approaches

Malicious actor	Threats	Techniques and references	Solution end
Clients/ external attackers	Data/model poisoning	Similarity-based approach (Cao et al. 2019), GAN approach (Zhao et al. 2019), Validation test set-based approach (Wang et al. 2020b; Vy et al. 2021), Notions of stealth approaches (Bhagoji et al. 2019), Model-agnostic defence technique (Manna et al. 2021), Anomaly detection techniques (Shen et al. 2016; Wan et al. 2021; Li et al. 2021g)	Server
Client/ external attacker	Backdoor attacks	Fine tuning (Liu et al. 2018), Model pruning (Jiang et al. 2022), Clients' contribution similarity (Fung et al. 2018), Reverse engineering (Zhao et al. 2021), Additive feature importance strategy (Manna et al. 2021), Clients' data accuracy on a test set (Su et al. 2022)	Server
Client/ external attacker	Common attacks	Knowledge distillation (Li and Wang 2019), Pruning (Jiang et al. 2019b), Multi-task learning (Smith et al. 2017; Sattler et al. 2020; Li et al. 2019a; Yu et al. 2020e)	Server
Server	Malicious server attacks	TEE (Mo and Haddadi 2019; Chen et al. 2020b), Peer-to-peer learning or security consortium (Roy et al. 2019; He et al. 2019; Yang et al. 2021a)	Platform, Clients
Clients/ Server	Inference attacks	DP (Liu et al. 2020), Knowledge distillation (Li and Wang 2019), Secret sharing/ secure boost protocol (Wang et al. 2020c), GAN based algorithm (Zhang and Luo 2020), Fake data generation at client node (Triastcyn and Faltings 2020)	
Clients/ server	Verification of participants/models	Hierarchical FL (Wainakh et al. 2020), Blockchain (Fang et al. 2022; Liu et al. 2020; Rahman et al. 2020; Yi Ming et al. 2021; Jiang et al. 2021)	Platform, Clients, Server

4.6.2.3 Model optimization The process of model optimization can also introduce biases in global models, favoring certain groups or relying heavily on a small subset of clients. As a result, the performance of the global model may excel for some clients while neglecting others. Fairness in model optimization aims to achieve an even distribution of accuracy across all clients.

The agnostic FL framework (Mohri et al. 2019) and FedFa (Huang et al. 2020b) optimization algorithm were built to avoid bias towards clients while optimizing the FL model. Mohri et al. (2019) naturally yields fairness for any target distribution with a mixture of clients with a data-dependent Rademacher guarantee. The FedFa combines the double momentum gradient method and weighting strategy. The weights are calculated based on information quantity and training frequency.

Another approach, Ditto (Li et al. 2021f), focused on building personalized models for each client (by allowing clients to fine-tune) closer to the optimal global model. A regularization term is added to make the finely tuned model closer to the global model. It reduces the variation in accuracy between clients by approximately 10% and simultaneously improves fairness and robustness. These works are based on the accuracy parity fairness notion.

Fed-ZDAC (Hao et al. 2021) applied a zero-shot data augmentation technique to under-represented client data to achieve uniform accuracy across clients. The augmentation algorithm generates pseudo-exemplars of unseen classes to avoid under-representation of the client. Hao et al. (2021) considered good-intent fairness notion to minimize loss of underlying protected client classes. Michieli and Ozay (2021) proposed a fair aggregation algorithm, FairAvg, showing that the fair algorithm is useful in terms of accuracy and convergence rate.

Xu and Lyu (2020) proposed RFFL framework based on contribution fairness. They maintained a client reputation scheme based on clients' contributions via local model updates. The global model is weighted according to the client's reputation. Another approach is CFFL (Lyu et al. 2020a), where each client receives a different global model corresponding to their reputation. Alvi et al. (2021) also regulated global model quality according to the client's contributions and costs. The server adds noise to the global model based on the quality of the local model. They regulated utility fairness via adaptive calculations and transmission policies.

q-FFL (Li et al. 2019b) realized accuracy parity using fair resource distribution. They assigned more weight in aggregation to clients with higher losses. The degree of fairness can be adjusted by tuning q. It is a multi-objective algorithm to optimize the loss function of each client individually without sacrificing performance. Their approach to attain fairness in the optimization function, as defined in Eq. (5), involves reweighting the objective function of the traditional FL function (refer Eq. (1)). In this approach, they assign higher weights to devices with poor performance, thereby shifting the distribution of accuracies in the network towards greater uniformity. For given local non-negative cost functions f_k and parameter $q > 0$, they define the objective function as in Eq. (5).

$$min_w \mathcal{F}_q(w) = \sum_{k=1}^{m} \frac{p_k}{q+1} f_k^{q+1}(w) \tag{5}$$

The term f_k^{q+1} represents f_k raised to the power of $q + 1$. The parameter q controls the degree of fairness we aim to achieve. When $q = 0$, fairness remains at the level of the classical FL objective (refer Eq. (1)). A higher q places greater emphasis on devices with higher local empirical losses (w), leading to a more uniform training accuracy distribution and the potential induction of fairness. m mathematical m as represents the number of clients in the FL process.

FedFv (Wang et al. 2021d) was proposed to resolve conflicts between local models before averaging them when constructing a global model. The algorithm can handle two types of conflicts: internal conflicts (between selected clients) and external conflicts (between selected and unselected clients).

However, it is important to note that these algorithms assume identical data distributions in all scenarios. In reality, data distributions are dynamic; therefore, it is crucial to consider the applicability of these algorithms in dynamic situations.

4.6.2.4 Incentive mechanism FLI (Yu et al. 2020a) dynamically and fairly allocates incentives to clients in a context-aware way. A given budget is equitably divided among clients to maximize utility and minimize inequality among clients. The algorithm satisfies contribution fairness, regret distribution fairness, and expectation fairness.

Several studies have focused on rewarding clients based on their contribution rate (Zhang et al. 2021e; Zeng et al. 2020; Cong et al. 2020). Likewise, previous studies (Kang et al. 2019b; Fan et al. 2021; Ye et al. 2020) distributed incentives based on data quality using Shapley value and contract theory methods.

In addition to monetary incentive schemes, the hierarchical fair FL framework proposed by Zhang et al. (2020b) focuses on rewarding clients based on their contribution rate. This approach classifies clients into different levels based on the quality or quantity of data and distributes models at the client level. Similarly, Lyu et al. (2020c) divided clients into clusters and trained one model for each cluster.

The achievement of fairness in incentive mechanisms requires collaboration among clients, servers, and platforms.

4.6.2.5 Social good The social good ensures that the model is not biased toward a specific individual or group. It is subject to the concept of group fairness. Ezzeldin et al. (2021) proposed a fairness-aware aggregation algorithm using debiasing strategies to provide a fair model across sensitive groups (such as race and gender). Likewise, Yue et al. (2021) obtained group and individual fairness by using a regularisation term to give more weight to low-performing individual clients or groups.

Rodríguez-Gálvez et al. (2021) introduced a modified method of differential multipliers to minimize empirical risks with fairness constraints, thereby enforcing group fairness in private FL. Padala et al. (2021) presented an ethical FL model to achieve demographic parity and equalized odds. Demographic parity indicates that the model's prediction must be independent of a sensitive attribute. Equalized odds focus on equating false positive and negative rates among different groups or individuals.

Zhang et al. (2020a) focused specifically on discriminatory bias against demographic groups. They addressed the challenges of fairness-performance trade-off, constrained coordination, and information limitation in privacy-sensitive FL settings by adapting a deep multi-agent reinforcement learning framework and a secure aggregation protocol. Another study, Zhang et al. (2021c), solved the unified group fairness problem through an optimization algorithm. They simultaneously investigated attribute level, client level, and agnostic fairness.

These solutions primarily operate at the platform level, with minimal involvement from clients. Table 8 provides an overview of fairness approaches, including fairness measurement, its application in various disciplines, and existing research and solutions.

5 Discussion

In this section, we examine prior studies by considering the impacts of a given solution in addressing a specific client-side challenge over other challenges. We classify the impacts between these challenges into three groups: those with a positive impact, those with a negative impact, and those that can have either a negative or positive impact on other challenges.

The positive impact category denotes that a solution targeting a specific challenge can also be effective for addressing multiple challenges simultaneously, resulting in time and effort savings. Conversely, solutions falling under the negative impact category may inadvertently exacerbate other challenges while attempting to resolve one. The third category encompasses solutions that can have either a positive or negative impact on other challenges, depending on the specific approach employed. While addressing one challenge may aid in resolving another, it can also inadvertently unveil or intensify other challenges in certain cases.

Table 9 illustrates the interrelationships among various challenges, based on the solutions available in the existing literature, along with the impact of these solutions on model performance. Each row represents the main challenge being addressed in the research, and each column represents a secondary challenge. The cells indicate the impact of a solution for the main challenge on the secondary challenge. As an example, consider the intersection of the "personalization" row and the "privacy" column. The corresponding solution emphasizes personalization and investigates its implications for privacy. However, it is observed that this approach may have a negative effect on privacy management.

5.1 Positive impact

Personalization solutions offer the opportunity to establish incentive mechanisms based on the best model, incorporating individual client contributions to incentivize clients. Clustering techniques within personalization can further assist in incentive mechanisms by grouping similar clients, aiding in the allocation of incentives effectively.

Personalization solutions can also contribute to fairness and robustness in model performance. A study by Li et al. (2021f) found that personalization solutions can improve performance in all three of these disciplines. Personalization helps to improve robustness by allowing the global model to be customized based on individual client data. This can help to protect against adversaries attacking the global model, as the impact on individual client performance is mitigated through personalization. Personalization solutions can also improve fairness by reducing the accuracy parity among clients through personalized models that are based on individual client data.

In addition, fairness solutions related to client selection, model optimization, and contribution evaluation can contribute to personalization as they consider the individual client's contribution to model building. This allows highly contributed clients to achieve high performance while also giving under-represented clients the opportunity to have their contributions recognized. These solutions can also help in resource and incentive management by distributing resources and incentives among clients in a fair manner.

The auditability problem under privacy management is usually solved through blockchain technology. Due to its auditing feature, blockchain technology can help with many other challenges, such as incentives, data, and security management.

Privacy and security mechanisms can work together to provide mutual benefits. Implementing privacy mechanisms can help to reduce information sharing and protect against adversarial attacks, while a secure environment can minimize the risk of privacy violations.

Incentive approaches based on client reputation can have a positive impact on both security and fairness. Client reputation is often determined by the performance and contribution of the client to the environment. These measurements can be used to identify honest (high contribution) and dishonest (low contribution) clients in the network, which can help security mechanisms to be more effective. Additionally, most incentive mechanisms are based on client contribution and data quality, which helps to ensure that fairness is maintained.

5.2 Negative impact

There is a trade-off between privacy and personalization challenges. For example, the clustering approach for personalization may compromise privacy as it requires additional client information, such as data distribution, data size, and client location. Similarly, incentive schemes may also negatively impact privacy by requiring auxiliary information about clients in order to distribute incentives fairly, which can reveal more client data.

Privacy approaches may negatively affect resource management as privacy algorithms (e.g. DP) require additional server computation and transmission power. Security approaches can negatively impact fairness as unique clients can be identified as malicious, leading to unfairness.

Except for privacy and resource management solutions, all other solutions tend to positively impact performance. Privacy approaches may reduce performance by adding noise to model parameters or data before building the model. On the other hand, performance-oriented algorithms that require large amounts of data and resources may improve performance but may also contribute to unfairness by eliminating low-performing and resource-constrained clients from the FL process. Resource management algorithms, on the other hand, may reduce resource usage, which can impact performance.

5.3 Negative or positive impact

One solution for the personalization challenge is clustering, which positively impacts resource management, while another personalization approach, fine-tuning, negatively affects it. Clients can be grouped by clustering based on location, performance, and resource availability. Due to clustering, clients do not need to communicate with the server frequently, thus reducing communication costs constantly. But fine-tuning approaches require additional client computing resources for tuning.

Efficient solutions addressing communication management challenges, such as data compression and reducing communication rounds, can contribute to mitigating security and privacy concerns. Data compression techniques can make it more difficult for adversaries to access client data while reducing the number of communication rounds minimizes the amount of data exchanged between clients and the server, reducing the risk of data being compromised. However, the use of edge-assisted FL technology, which is used to manage computation and communication costs, can increase the risk of privacy and security breaches. This is because it requires clients to send raw data to the edge server, which can increase the risk of privacy and security violations.

Table 8 Summary of fairness approaches

Approaches	References	Solution end
Measuring fairness	Visual framework (Garg et al. 2020; Chu et al. 2021), Metrics (Divi et al. 2021b)	Client, Platform
	Client selection (Huang et al. 2020a; Yang et al. 2021c; Kang et al. 2020)	Server
Practice of fairness in different disciplines	Contribution evaluation (Kang et al. 2019a, b; Song et al. 2019; Wang et al. 2019a; Zhang et al. 2021e; Zeng et al. 2020; Sarikaya and Ercetin 2019; Kang et al. 2020; Zhang et al. 2020b; Le et al. 2021; Lyu et al. 2020a; Xu and Lyu 2020; Song et al. 2021; Lyu et al. 2020c; Nishio et al. 2020; Wang et al. 2020a; Shyn et al. 2021)	Platform, Server, Clients
	Model optimization (Li et al. 2021f; Mohri et al. 2019; Hao et al. 2021; Michieli and Ozay 2021)	Server
	Incentive mechanism (Kang et al. 2019b; Yu et al. 2020a; Zhang et al. 2021e; Zeng et al. 2020; Zhang et al. 2020b; Lyu et al. 2020c; Cong et al. 2020; Fan et al. 2021; Ye et al. 2020)	Clients, Server, Platform
	Social good (Ezzeldin et al. 2021; Yue et al. 2021; Rodríguez-Gálvez et al. 2021; Padala et al. 2021; Zhang et al. 2020a, 2021c)	Platform

Table 9 The relationships between challenges and performance

Challenges	Person-alization	Privacy	Incentive	Resource	Security	Fairness	Performance
Personalization	–	×	✓	✓×	✓	✓	✓
Privacy	×	–	✓	×	✓	–	×
Incentive	–	×	–	–	✓	✓	✓
Resource	–	✓×	–	–	✓×	–	×
Security	–	✓	–	×	–	×	✓
Fairness	✓	✓×	✓	✓	✓×	–	✓
Performance	✓	×	✓	×	–	×	–

*Each row indicates the primary challenge of the research and each column is a secondary challenge. Each cell represents the impact of primary challenge solutions on another challenge. † ✓: Positive impact, ×: Negative impact

Fairness measurement, accuracy parity, and good-intent/group fairness approaches can all contribute to privacy management. Fairness measurement can help with privacy management by making the FL process transparent and visible to clients, addressing the issue of auditability. Accuracy parity and good-intent/group fairness approaches can help to mitigate the risk of re-identification by reducing differences between individuals or groups. However, the self-reported information solution used in contribution evaluation may increase the privacy risk, as it requires clients to report data quality and quantity, data collection costs, and computational and communication capabilities to the server for review. This can expose sensitive information about the client.

Most of the fairness approaches negatively impact security because fairness approaches reduce disparity among clients, reducing the chance of detecting malicious clients using anomaly detection. From the security perspective, the disparity can help distinguish between malicious and honest clients. It is applicable vice-versa, too. However, on the topic of fairness, an approach known as “client-reputation measurement regarding honesty and contribution” can be employed to identify dishonest clients.

5.4 A solution applicability for many challenges

This section emphasizes the importance of considering the applicability of solutions to multiple challenges and understanding the interrelationship between these challenges when designing solutions. By doing so, it is possible to reduce system complexity and avoid duplicative efforts in the federated environment.

Blockchain technology has been used in the literature to address privacy, data computation, incentive, and security management challenges. This is because blockchain has many features, such as robustness, immutability, transparency, append-only, and auditability, that make it suitable for addressing a wide range of challenges. Researchers can consider collaborating with blockchain technology to address various challenges, as it has the potential to simplify solutions by combining different methods in a single system.

Likewise, certain personalization approaches can also address fairness and security challenges. For instance, Ditto (Li et al. 2021f) is a personalization solution that offers both fairness and security benefits. It is important to analyze other personalization approaches to identify their potential benefits in different aspects.

6 Open challenges and trend of future works

In this section, we explore open challenges and future research trends by examining the reviewed research articles, surveys, and our own insights. As we delve into these discussions, one potential avenue for future investigation involves examining the impact of solutions for specific challenges on other related challenges.

- Personalization challenges:
  - Impact of personalization methods on other challenges: For example, Ditto (Li et al. 2021f) evaluated the fairness and robustness benefits of the proposed personalization method. Therefore, future research could focus on the impact of other personalization solutions on different challenges.
  - Context-aware personalization: Developing context-aware techniques in FL is a potential open problem. The consideration of sensitive contextual information in FL is an ongoing topic of interest. While FL does not involve data transfer to third-party applications, the question of whether context information can be leveraged to enhance personalization without compromising privacy requires further investigation.
- Incentive management challenges:
  - Incentive schemes based on other values (except monetary value): While we have discussed various incentives such as model performance, reputation, computational power, auxiliary information, and model fairness, there is limited research on other incentive schemes. Exploring and studying additional incentive approaches in the context of FL would be a valuable direction for future research.
  - Incentive schemes with multiple servers: Almost all literature focused on the one-to-many relationship where one server with multiple clients (monopoly market) (Shi et al. 2021). Clients have no option to choose another server if they are not convinced by the offer. This area needs further study to create a non-monopoly market with multiple server options.
- Privacy management challenges:
  - Privacy and performance trade-off: Current approaches (such as DP) forfeit performance and computation to enable privacy for clients. Though researchers are working on finding an optimal point to manage privacy and performance, the privacy-utility trade-off is still open to researchers.
  - Dynamic settings with context: Privacy approaches in the literature are static, consistently using the same noise level and settings. However, clients' preferences may vary depending on the context.
  - Explainable AI: Explainable AI refers to building tools or frameworks to describe ML models in a human-understandable format. Applying explainable AI concepts in FL is still an open problem.
  - Granular privacy management and consented data sharing: Very little literature focuses on granular privacy management to meet the diverse privacy needs of clients.
  - Sharing less sensitive model (Lo et al. 2021b): Since data can be inferred even after DP mechanisms are applied, it is useful to have mechanisms to understand the model's sensitivity before it is shared.

  - Compliance with regulatory (Lo et al. 2021b): The application of regulatory compliance for FL (model exposure, model retention) is still underexplored.

- Resource management challenges: While significant research has been conducted on computational and communication management in the context of FL, relatively less attention has been given to data management aspects. The focus has primarily been on improving performance through computational and communication strategies, leaving room for further exploration and investigation of data management techniques in FL.

  - Handling unlabelled data (Lo et al. 2021b): Labelled data may not always be available to clients, and labeling is also expensive. Some potential approaches, such as semi-supervised learning (Lo et al. 2021b) to label data based on other clients' data, can be expected in the future.

- Security: Current literature mainly focuses on attacks from malicious clients rather than from malicious servers.

  - Security approaches for malicious servers: In literature, only a few studies (Mo and Haddadi 2019; Chen et al. 2020b) focused on solutions for malicious servers. More theoretical and empirical studies are needed to address malicious server problems.
  - Consortium among clients: To avoid malicious attacks from clients or servers, clients can cluster within themselves and form a consortium among themselves without concern of malicious server. The grouping may be based on the reputation of clients.

- Fairness: The concept of fairness has recently received extensive attention in ML. However, applying these methods in FL is not straightforward due to the data distribution. Thus, new techniques should be introduced in FL.

  - Fairness approaches in FL life cycle: FL consists of different stages of data processing, such as pre-processing (collecting clients' data, feature selection/modification, data synthesis), in-process (building local models, adding global models), and post-processing (result prediction). Introducing fairness at each stage can enable fairness-aware FL.
  - User interactive fairness system: A framework for setting the boundaries of clients' expected fairness is appreciable. Clients can visualize and define their own fairness expectations in the framework.

7 Conclusion

To the best of our knowledge, this study is the first survey of client-side challenges in FL. We conducted this systematic survey by analyzing the literature and categorized the client-side challenges into six broad categories: personalization, privacy management, incentive management, resource management, data and device security, and fairness. We also presented the available state-of-art solutions for the identified challenges. In addition, we conducted an analysis of the relationships between challenges, trade-offs in addressing them, and the applicability of solutions. Based on this analysis, a potential future research direction would be to explore the impact of addressing one challenge on others. By applying a solution to multiple challenges, it is possible to reduce system complexity and eliminate redundant efforts.

Acknowledgements The work has been supported by the Cyber Security Research Centre Limited whose activities are partially funded by the Australian Government's Cooperative Research Centres Programme.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions YS: Conceptualisation, investigation, methodology, writing—original draft. HP: Conceptualisation, methodology, data curation, writing—original draft. SSK: Conceptualisation, supervision, writing—review and editing LZ: Supervision, writing—review and editing.

References

AbdulRahman S, Tout H, Mourad A et al (2020) FedMCCS: multicriteria client selection model for optimal IoT federated learning. IEEE Internet Things J 8(6):4723–4735

Achituve I, Shamsian A, Navon A et al (2021) Personalized federated learning with gaussian processes. Adv Neural Inf Process Syst 34:8392–8406

Al-Abiad MS, Hassan M, Hossain M, et al (2021) Energy efficient federated learning in integrated fog-cloud computing enabled internet-of-things networks. arXiv:2107.03520

Alazab M, RM SP, Parimala M, et al (2021) Federated learning for cybersecurity: concepts, challenges and future directions. IEEE Trans Ind Inf 18:3501–3509

Aldaghri N, Mahdavifar H, Beirami A (2021) FeO$_2$: federated learning with opt-out differential privacy. arXiv:2110.15252

Aledhari M, Razzak R, Parizi RM, et al (2020) Federated learning: a survey on enabling technologies, protocols, and applications. IEEE Access 8:140699–140725

Alvi SA, Hong Y, Durrani S (2021) Utility fairness for the differentially private federated learning. arXiv:2109.05267

Amiri MM, Gunduz D, Kulkarni SR, et al (2020) Federated learning with quantized global model updates. arXiv:2006.10672

Arivazhagan MG, Aggarwal V, Singh AK, et al (2019) Federated learning with personalization layers. arXiv:1912.00818

Asad M, Moustafa A, Ito T (2020) Fedopt: Towards communication efficiency and privacy preservation in federated learning. Appl Sci 10(8):2864

Avdiukhin D, Kasiviswanathan S (2021) Federated learning under arbitrary communication patterns. In: International conference on machine learning. PMLR, pp 425–435

Bagdasaryan E, Veit A, Hua Y, et al (2020) How to backdoor federated learning. In: International conference on artificial intelligence and statistics. PMLR, pp 2938–2948

Balakrishnan R, Akdeniz M, Dhakal S et al (2021) Resource management and model personalization for federated learning over wireless edge networks. J Sens Actuator Netw 10(1):17

Bao X, Su C, Xiong Y, et al (2019) Flchain: a blockchain for auditable federated learning with trust and incentive. In: 2019 5th international conference on big data computing and communications (BIGCOM). IEEE, pp 151–159

Bernstein J, Wang YX, Azizzadenesheli K, et al (2018) signSGD: compressed optimisation for non-convex problems. In: International conference on machine learning. PMLR, pp 560–569

Bhagoji AN, Chakraborty S, Mittal P, et al (2019) Analyzing federated learning through an adversarial lens. In: Chaudhuri K, Salakhutdinov R (eds) Proceedings of the 36th international conference on machine learning, proceedings of machine learning research, vol 97. PMLR, pp 634–643, https://proceedings.mlr.press/v97/bhagoji19a.html

Blanco-Justicia A, Domingo-Ferrer J, Martínez S et al (2021) Achieving security and privacy in federated learning systems: survey, research challenges and future directions. Eng Appl Artif Intell 106(104):468

Bouacida N, Hou J, Zang H, et al (2020) Adaptive federated dropout: improving communication efficiency and generalization for federated learning. arXiv:2011.04050

Briggs C, Fan Z, Andras P (2020) Federated learning with hierarchical clustering of local updates to improve training on non-iid data. In: 2020 international joint conference on neural networks (IJCNN). IEEE, pp 1–9

Briggs C, Fan Z, Andras P (2021) A review of privacy-preserving federated learning for the internet-of-things. Fed Learn Syst pp 21–50

Caldas S, Duddu SMK, Wu P, et al (2018) Leaf: a benchmark for federated settings. arXiv:1812.01097

Cao D, Chang S, Lin Z, et al (2019) Understanding distributed poisoning attack in federated learning. In: 2019 IEEE 25th international conference on parallel and distributed systems (ICPADS). IEEE, pp 233–239

Chai D, Wang L, Chen K, et al (2020a) Fedeval: a benchmark system with a comprehensive evaluation model for federated learning. arXiv:2011.09655

Chai Z, Chen Y, Zhao L, et al (2020b) Fedat: a communication-efficient federated learning method with asynchronous tiers under non-iid data. ArXivorg

Chang WT, Tandon R (2020) Communication efficient federated learning over multiple access channels. arXiv:2001.08737

Chen F, Luo M, Dong Z, et al (2018) Federated meta-learning with fast convergence and efficient communication. arXiv:1802.07876

Chen R, Li L, Xue K, et al (2020a) To talk or to work: energy efficient federated learning over mobile devices via the weight quantization and 5G transmission co-design. arXiv:2012.11070

Chen Y, Luo F, Li T et al (2020b) A training-integrity privacy-preserving federated learning scheme with trusted execution environment. Inf Sci 522:69–79

Chen Y, Qin X, Wang J et al (2020c) Fedhealth: a federated transfer learning framework for wearable healthcare. IEEE Intell Syst 35(4):83–93

Chen Y, Li J, Wang F et al (2021) DS2PM: a data sharing privacy protection model based on blockchain and federated learning. IEEE Internet of Things Journal

Cheng G, Chadha K, Duchi J (2021) Fine-tuning is fine in federated learning. arXiv:2108.07313

Cho YJ, Wang J, Joshi G (2020) Client selection in federated learning: convergence analysis and power-of-choice selection strategies. arXiv:2010.01243

Cho YJ, Wang J, Chiruvolu T, et al (2021) Personalized federated learning for heterogeneous clients with clustered knowledge transfer. arXiv:2109.08119

Chou YH, Hong S, Sun C, et al (2021) Grp-fed: Addressing client imbalance in federated learning via global-regularized personalization. arXiv:2108.13858

Choudhury O, Gkoulalas-Divanis A, Salonidis T, et al (2019) Differential privacy-enabled federated learning for sensitive health data. arXiv:1910.02578

Chu L, Wang L, Dong Y, et al (2021) Fedfair: Training fair models in cross-silo federated learning. arXiv:2109.05662

Collins L, Hassani H, Mokhtari A, et al (2021) Exploiting shared representations for personalized federated learning. arXiv:2102.07078

Cong M, Yu H, Weng X, et al (2020) A VCG-based fair incentive mechanism for federated learning. arXiv:2008.06680

Dai X, Yan X, Zhou K, et al (2019) Hyper-sphere quantization: communication-efficient sgd for federated learning. arXiv:1911.04655

Deng M, Wuyts K, Scandariato R et al (2011) A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. Requirements Eng 16(1):3–32

Deng Y, Kamani MM, Mahdavi M (2020) Adaptive personalized federated learning. arXiv:2003.13461

Ding N, Fang Z, Huang J (2020) Incentive mechanism design for federated learning with multi-dimensional private information. In: 2020 18th international symposium on modeling and optimization in mobile, ad hoc, and wireless networks (WiOPT), pp 1–8

Ding J, Tramel E, Sahu AK et al (2022) Federated learning challenges and opportunities: an outlook. In: ICASSP 2022–2022 IEEE international conference on acoustics. Speech and signal processing (ICASSP). IEEE, pp 8752–8756

Dinh CT, Tran NH, Nguyen TD (2020) Personalized federated learning with moreau envelopes. arXiv:2006.08848

Divi S, Farrukh H, Celik B (2021a) Unifying distillation with personalization in federated learning. arXiv:2105.15191

Divi S, Lin YS, Farrukh H, et al (2021b) New metrics to evaluate the performance and fairness of personalized federated learning. arXiv:2107.13173

Do QV, Pham QV, Hwang WJ (2021) Deep reinforcement learning for energy-efficient federated learning in uav-enabled wireless powered networks. IEEE Commun Lett 26:99–103

Du Z, Wu C, Yoshinaga T et al (2020) Federated learning for vehicular internet of things: Recent advances and open issues. IEEE Open J Comput Soc 1:45–61

Duan M, Liu D, Chen X et al (2020) Self-balancing federated learning with global imbalanced data in mobile systems. IEEE Trans Parallel Distrib Syst 32(1):59–71

Duan M, Liu D, Ji X, et al (2021) Fedgroup: efficient federated learning via decomposed similarity-based clustering. In: 2021 IEEE international conference on parallel and distributed processing with applications, big data and cloud computing, sustainable computing and communications, social computing and networking (ISPA/BDCloud/SocialCom/SustainCom). IEEE, pp 228–237

Dwork C (2009) The differential privacy frontier. In: Theory of cryptography conference. Springer, New York, pp 496–502

Enthoven D, Al-Ars Z (2021) An overview of federated deep learning privacy attacks and defensive strategies. Federated Learn Syst pp 173–196

Ezzeldin YH, Yan S, He C, et al (2021) Fairfed: enabling group fairness in federated learning. arXiv:2110.00857

Fallah A, Mokhtari A, Ozdaglar A (2020) Personalized federated learning: a meta-learning approach. arXiv:2002.07948

Fan T (2018) FATE-Board: FATE's visualization toolkit. https://github.com/FederatedAI/FATE-Board

Fan S, Zhang H, Zeng Y et al (2020) Hybrid blockchain-based resource trading system for federated learning in edge computing. IEEE Internet Things J 8(4):2252–2264

Fan Z, Fang H, Zhou Z, et al (2021) Improving fairness for data valuation in federated learning. arXiv:2109.09046

Fang H, Qian Q (2021) Privacy preserving machine learning with homomorphic encryption and federated learning. Future Internet 13(4):94

Fang M, Cao X, Jia J, et al (2020) Local model poisoning attacks to Byzantine-Robust federated learning. In: 29th USENIX security symposium (USENIX Security 20), pp 1605–1622

Fang C, Guo Y, Ma J, et al (2022) A privacy-preserving and verifiable federated learning method based on blockchain. Computer Commun 186:1–11

Finn C, Abbeel P, Levine S (2017) Model-agnostic meta-learning for fast adaptation of deep networks. In: International conference on machine learning. PMLR, pp 1126–1135

Fung C, Yoon CJ, Beschastnikh I (2018) Mitigating sybils in federated learning poisoning. arXiv:1808.04866

Gao L, Li L, Chen Y, et al (2021) FIFL: a fair incentive mechanism for federated learning. In: 50th international conference on parallel processing, pp 1–10

Garg P, Villasenor J, Foggo V (2020) Fairness metrics: a comparative analysis. In: 2020 IEEE international conference on big data (Big Data). IEEE, pp 3662–3666

Geyer RC, Klein T, Nabi M (2017) Differentially private federated learning: a client level perspective. arXiv:1712.07557

Ghosh A, Chung J, Yin D, et al (2020) An efficient framework for clustered federated learning. arXiv:2006.04088

Goldreich O (1998) Secure multi-party computation. Manuscript preliminary version 78:110

Guha N, Talwalkar A, Smith V (2019) One-shot federated learning. arXiv:1902.11175

Hamer J, Mohri M, Suresh AT (2020) Fedboost: a communication-efficient algorithm for federated learning. In: International conference on machine learning. PMLR, pp 3973–3983

Han X, Yu H, Gu H (2019) Visual inspection with federated learning. In: International conference on image analysis and recognition. Springer, New York, pp 52–64

Hanzely F, Richtárik P (2020) Federated learning of a mixture of global and local models. arXiv:2002.05516

Hao M, Li H, Luo X et al (2019) Efficient and privacy-enhanced federated learning for industrial artificial intelligence. IEEE Trans Ind Inf 16(10):6532–6542

Hao W, El-Khamy M, Lee J, et al (2021) Towards fair federated learning with zero-shot data augmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 3310–3319

Hard A, Rao K, Mathews R, et al (2018) Federated learning for mobile keyboard prediction. arXiv:1811.03604

He C, Tan C, Tang H, et al (2019) Central server free federated learning over single-sided trust social networks. arXiv:1910.04956

Hu R, Gong Y, Guo Y (2020a) Federated learning with sparsification-amplified privacy and adaptive optimization. arXiv:2008.01558

Hu R, Guo Y, Li H et al (2020b) Personalized federated learning with differential privacy. IEEE Internet Things J 7(10):9530–9539

Hu H, Salcic Z, Sun L, et al (2021) Source inference attacks in federated learning. https://doi.org/10.48550/ARXIV.2109.05659,https://arxiv.org/abs/2109.05659

Huang L, Shea AL, Qian H et al (2019) Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. J Biomed Inform 99(103):291

Huang T, Lin W, Wu W et al (2020a) An efficiency-boosting client selection scheme for federated learning with fairness guarantee. IEEE Trans Parallel Distrib Syst 32(7):1552–1564

Huang W, Li T, Wang D, et al (2020b) Fairness and accuracy in federated learning. arXiv:2012.10069

Imteaj A, Thakker U, Wang S et al (2021) A survey on federated learning for resource-constrained iot devices. IEEE Internet Things J 9(1):1–24

Jeong E, Oh S, Kim H, et al (2018) Communication-efficient on-device machine learning: federated distillation and augmentation under non-iid private data. arXiv:1811.11479

Ji S, Jiang W, Walid A, et al (2020) Dynamic sampling and selective masking for communication-efficient federated learning. arXiv:2003.09603

Ji Z, Chen L, Zhao N et al (2021) Computation offloading for edge-assisted federated learning. IEEE Trans Veh Technol 70(9):9330–9344

Jiang Y, Konečnỳ J, Rush K, et al (2019a) Improving federated learning personalization via model agnostic meta learning. arXiv:1909.12488

Jiang Y, Wang S, Valls V, et al (2019b) Model pruning enables efficient federated learning on edge devices. arXiv:1909.12326

Jiang C, Xu C, Zhang Y (2021) PFLM: privacy-preserving federated learning with membership proof. Inf Sci 576:288–311

Jiang Y, Wang S, Valls V, et al (2022) Model pruning enables efficient federated learning on edge devices. IEEE Trans Neural Netw Learn Syst. https://doi.org/10.1109/TNNLS.2022.3166101

Jiao Y, Wang P, Niyato D, et al (2020) Toward an automated auction framework for wireless federated learning services market. IEEE Trans Mob Comput 20:3034–3048

Jourdan T, Boutet A, Frindel C (2021) Privacy assessment of federated learning using private personalized layers. arXiv:2106.08060

Kaelbling LP, Littman ML, Moore AW (1996) Reinforcement learning: a survey. J Artif Intell Res 4:237–285

Kairouz P, McMahan HB, Avent B et al (2021) Advances and open problems in federated learning. Found Trends® Mach Learn 14(1–2):1–210

Kang D, Ahn CW (2021) Communication cost reduction with partial structure in federated learning. Electronics 10(17):2081

Kang J, Xiong Z, Niyato D et al (2019a) Incentive mechanism for reliable federated learning: a joint optimization approach to combining reputation and contract theory. IEEE Internet Things J 6(6):10,700-10,714

Kang J, Xiong Z, Niyato D, et al (2019b) Incentive design for efficient federated learning in mobile networks: a contract theory approach. In: 2019 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS). IEEE, pp 1–5

Kang J, Xiong Z, Niyato D et al (2020) Reliable federated learning for mobile networks. IEEE Wirel Commun 27(2):72–80

Katevas K, Bagdasaryan E, Waterman J, et al (2020) Policy-based federated learning. arXiv:2003.06612

Khalfoun B, Ben Mokhtar S, Bouchenak S, et al (2021) Eden: Enforcing location privacy through re-identification risk assessment: a federated learning approach. Proc ACM Interact Mob Wearable Ubiquitous Technol. https://doi.org/10.1145/3463502

Khodak M, Balcan MF, Talwalkar A (2019) Adaptive gradient-based meta-learning methods. arXiv:1906.02717

Kim Y, Al Hakim E, Haraldson J, et al (2021) Dynamic clustering in federated learning. In: ICC 2021-IEEE international conference on communications. IEEE, pp 1–6

Konečnỳ J, McMahan HB, Yu FX, et al (2016) Federated learning: strategies for improving communication efficiency. arXiv:1610.05492

Kontoudis GP, Stilwell DJ (2022) Fully decentralized, scalable gaussian processes for multi-agent federated learning. arXiv:2203.02865

Kulkarni V, Kulkarni M, Pant A (2020) Survey of personalization techniques for federated learning. In: 2020 fourth world conference on smart trends in systems, security and sustainability (WorldS4). IEEE, pp 794–797

Kurupathi SR, Maass W (2020) Survey on federated learning towards privacy preserving ai. In: Proceedings of computer science & information technology (CSIT), pp 1–19

Le THT, Tran NH, Tun YK, et al (2020) Auction based incentive design for efficient federated learning in cellular wireless networks. In: 2020 IEEE wireless communications and networking conference (WCNC), pp 1–6. https://doi.org/10.1109/WCNC45663.2020.9120773

Le THT, Tran NH, Tun YK, et al (2021) An incentive mechanism for federated learning in wireless cellular network: an auction approach. IEEE Trans Wirel Commun 20:4874–4887

Li D, Wang J (2019) FedMD: heterogenous federated learning via model distillation. arXiv:1910.03581

Li R, Ma F, Jiang W, et al (2019a) Online federated multitask learning. In: 2019 IEEE international conference on big data (Big Data). IEEE, pp 215–220

Li T, Sanjabi M, Beirami A, et al (2019b) Fair resource allocation in federated learning. arXiv:1905.10497

Li A, Sun J, Wang B, et al (2020a) Lotteryfl: personalized and communication-efficient federated learning with lottery ticket hypothesis on non-IID datasets. arXiv:2008.03371

Li L, Fan Y, Tse M et al (2020b) A review of applications in federated learning. Comput Ind Engi 149(106):854

Li M, Chen Y, Wang Y et al (2020c) Efficient asynchronous vertical federated learning via gradient prediction and double-end sparse compression. In: 2020 16th international conference on control, automation, robotics and vision (ICARCV). IEEE, pp 291–296

Li S, Qi Q, Wang J, et al (2020d) Ggs: general gradient sparsification for federated learning in edge computing. In: ICC 2020-2020 IEEE international conference on communications (ICC). IEEE, pp 1–7

Li T, Sahu AK, Talwalkar A et al (2020e) Federated learning: challenges, methods, and future directions. IEEE Signal Process Mag 37(3):50–60

Li T, Sahu AK, Zaheer M et al (2020f) Federated optimization in heterogeneous networks. Proce Mach Learn Syst 2:429–450

Li A, Sun J, Zeng X, et al (2021a) FedMASK: joint computation and communication-efficient personalized federated learning via heterogeneous masking. In: Proceedings of the 19th ACM conference on embedded networked sensor systems, pp 42–55

Li C, Li G, Varshney PK (2021b) Federated learning with soft clustering. IEEE Internet Things J

Li L, Shi D, Hou R, et al (2021c) To talk or to work: flexible communication compression for energy efficient federated learning over heterogeneous mobile edge devices. In: IEEE INFOCOM 2021-IEEE conference on computer communications. IEEE, pp 1–10

Li Q, Wei X, Lin H, et al (2021d) Inspecting the running process of horizontal federated learning via visual analytics. IEEE Trans Visual Comput Graph 28:4085–4100

Li Q, Wen Z, Wu Z, et al (2021e) A survey on federated learning systems: vision, hype and reality for data privacy and protection. IEEE Trans Knowl Data Eng. https://doi.org/10.1109/TKDE.2021.3124599

Li T, Hu S, Beirami A, et al (2021f) Ditto: fair and robust federated learning through personalization. In: International conference on machine learning. PMLR, pp 6357–6368

Li X, Qu Z, Zhao S, et al (2021g) Lomar: a local defense against poisoning attack on federated learning. IEEE Trans Depend Secure Comput

Liang PP, Liu T, Ziyin L, et al (2020) Think locally, act globally: federated learning with local and global representations. arXiv:2001.01523

Lim WYB, Luong NC, Hoang DT et al (2020) Federated learning in mobile edge networks: a comprehensive survey. IEEE Commun Surveys Tutor 22(3):2031–2063

Lim WYB, Huang J, Xiong Z, et al (2021) Towards federated learning in uav-enabled internet of vehicles: a multi-dimensional contract-matching approach. IEEE Trans Intell Transp Syst 22:5140–5154

Lin J, Du M, Liu J (2019) Free-riders in federated learning: attacks and defenses. arXiv:1911.12560

Liu Y, Wei J (2020) Incentives for federated learning: a hypothesis elicitation approach. arXiv:2007.10596

Liu K, Dolan-Gavitt B, Garg S (2018) Fine-pruning: Defending against backdooring attacks on deep neural networks. In: International symposium on research in attacks, intrusions, and defenses. Springer, New York, pp 273–294

Liu Y, Ai Z, Sun S, et al (2020a) Fedcoin: a peer-to-peer payment system for federated learning. In: Federated learning. Springer, New York, pp 125–138

Liu Y, Peng J, Kang J et al (2020b) A secure federated learning framework for 5G networks. IEEE Wirel Commun 27(4):24–31

Liu L, Zhang J, Song S, et al (2021a) Hierarchical quantized federated learning: convergence analysis and system design. arXiv:2103.14272

Liu S, Yu J, Deng X, et al (2021b) FedCPF: an efficient-communication federated learning approach for vehicular edge computing in 6G communication networks. IEEE Trans Intell Transp Syst

Lo SK, Liu Y, Lu Q, et al (2021a) Blockchain-based trustworthy federated learning architecture. arXiv:2108.06912

Lo SK, Lu Q, Wang C et al (2021b) A systematic literature review on federated machine learning: from a software engineering perspective. ACM Comput Surveys 54(5):1–39

Lu Y, Huang X, Dai Y et al (2019) Blockchain and federated learning for privacy-preserved data sharing in industrial iot. IEEE Trans Industr Inf 16(6):4177–4186

Lu Y, Huang X, Zhang K et al (2020) Communication-efficient federated learning for digital twin edge networks in industrial iot. IEEE Trans Industr Inf 17(8):5709–5718

Luo J, Wu S (2021) Adapt to adaptation: learning personalization for cross-silo federated learning. arXiv:2110.08394

Lyu L, Xu X, Wang Q, et al (2020a) Collaborative fairness in federated learning. In: Federated learning. Springer, p 189–204

Lyu L, Yu H, Yang Q (2020b) Threats to federated learning: a survey. arXiv:2003.02133

Lyu L, Yu J, Nandakumar K et al (2020c) Towards fair and privacy-preserving federated deep models. IEEE Trans Parallel Distrib Syst 31(11):2524–2541

Ma C, Li J, Ding M et al (2020) On safeguarding privacy and security in the framework of federated learning. IEEE Network 34(4):242–248

Ma Z, Lu Y, Li W, et al (2021) Pfedatt: attention-based personalized federated learning on heterogeneous clients. In: Asian conference on machine learning. PMLR, pp 1253–1268

Mahara SS, Bharath B, et al (2021) Multi-task federated edge learning (mtfeel) in wireless networks. arXiv:2108.02517

Majeed U, Hong CS (2019) FLCHIAN: lederated learning via MEC-enabled blockchain network. In: 2019 20th Asia-Pacific network operations and management symposium (APNOMS). IEEE, pp 1–4

Malekijoo A, Fadaeieslam MJ, Malekijou H, et al (2021) FEDZIP: a compression framework for communication-efficient federated learning. arXiv:2102.01593

Manna A, Kasyap H, Tripathy S (2021) Moat: Model agnostic defense against targeted poisoning attacks in federated learning. In: International conference on information and communications security. Springer, New York, pp 38–55

Mansour Y, Mohri M, Ro J, et al (2020) Three approaches for personalization with applications to federated learning. arXiv:2002.10619

Marathe VJ, Kanani P (2022) Subject granular differential privacy in federated learning. arXiv:2206.03617

Martinez I, Francis S, Hafid AS (2019) Record and reward federated learning contributions with blockchain. In: 2019 International conference on cyber-enabled distributed computing and knowledge discovery (CyberC). IEEE, pp 50–57

McMahan B, Moore E, Ramage D, et al (2017) Communication-efficient learning of deep networks from decentralized data. In: Artificial intelligence and statistics. PMLR, pp 1273–1282

Mestoukirdi M, Zecchin M, Gesbert D, et al (2021) User-centric federated learning. arXiv:2110.09869

Michieli U, Ozay M (2021) Are all users treated fairly in federated learning systems? In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 2318–2322

Mike (2018) Federated learning: distributed machine learning with data locality and privacy. https://blog.fastforwardlabs.com/2018/11/14/federated-learning.html

Mills J, Hu J, Min G (2020) Multi-task federated learning for personalised deep neural networks in edge computing. arXiv:2007.09236

Mo F, Haddadi H (2019) Efficient and private federated learning using TEE. In: Proceedings of EuroSys Conference, Dresden, Germany

Mohri M, Sivek G, Suresh AT (2019) Agnostic federated learning. In: International conference on machine learning. PMLR, pp 4615–4625

Moon J, Kum S, Kim Y, et al (2020) A decentralized ai data management system in federated learning. In: 2020 international conference on intelligent systems and computer vision (ISCV). IEEE, pp 1–4

Mothukuri V, Parizi RM, Pouriyeh S et al (2021) A survey on security and privacy of federated learning. Futur Gener Comput Syst 115:619–640

Mugunthan V, Peraire-Bueno A, Kagal L (2020a) PrivacyFL: a simulator for privacy-preserving and secure federated learning. In: Proceedings of the 29th ACM international conference on information and knowledge management, pp 3085–3092

Mugunthan V, Rahman R, Kagal L (2020b) Blockflow: an accountable and privacy-preserving solution for federated learning. arXiv:2007.03856

Nadiger C, Kumar A, Abdelhak S (2019) Federated reinforcement learning for fast personalization. In: 2019 IEEE second international conference on artificial intelligence and knowledge engineering (AIKE). IEEE, pp 123–127

Naseri AM, Lucia W, Youssef A (2022) Confidentiality attacks against encrypted control systems. Cyber-Physical Systems pp 1–20

Ng KL, Chen Z, Liu Z, et al (2021) A multi-player game for studying federated learning incentive schemes. In: Proceedings of the twenty-ninth international conference on international joint conferences on artificial intelligence, pp 5279–5281

Nguyen HT, Sehwag V, Hosseinalipour S et al (2020) Fast-convergent federated learning. IEEE J Sel Areas Commun 39(1):201–218

Niknam S, Dhillon HS, Reed JH (2020) Federated learning for wireless communications: motivation, opportunities, and challenges. IEEE Commun Mag 58(6):46–51

Nishio T, Yonetani R (2019) Client selection for federated learning with heterogeneous resources in mobile edge. In: ICC 2019-2019 IEEE international conference on communications (ICC). IEEE, pp 1–7

Nishio T, Shinkuma R, Mandayam NB (2020) Estimation of individual device contributions for incentivizing federated learning. In: 2020 IEEE Globecom workshops (GC Wkshps. IEEE, pp 1–6

Nori MK, Yun S, Kim IM (2021) Fast federated learning by balancing communication trade-offs. IEEE Trans Commun 69(8):5168–5182

Nour B, Cherkaoui S, Mlika Z (2021) Federated learning and proactive computation reuse at the edge of smart homes. IEEE Trans Netw Sci Eng 9:3045–3056

Orekondy T, Oh SJ, Zhang Y, et al (2018) Gradient-leaks: understanding and controlling deanonymization in federated learning. arXiv:1805.05838

Ozkara K, Singh N, Data D, et al (2021) QuPeD: Quantized personalization via distillation with applications to federated learning. Adv Neural Inf Process Syst 34

Padala M, Damle S, Gujar S (2021) Federated learning meets fairness and differential privacy. In: International conference on neural information processing. Springer, New York, pp 692–699

Page MJ, McKenzie JE, Bossuyt PM, et al (2021) The prisma 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372. https://doi.org/10.1136/bmj.n71,https://www.bmj.com/content/372/bmj.n71, https://arxiv.org/abs/https://www.bmj.com/content/372/bmj.n71.full.pdf

Peng Z, Xu J, Chu X et al (2021) Vfchain: enabling verifiable and auditable federated learning via blockchain systems. IEEE Trans Netw Sci Eng 9(1):173–186

Peterson D, Kanani P, Marathe VJ (2019) Private federated learning with domain adaptation. arXiv:1912.06733

Prathiba SB, Raja G, Anbalagan S, et al (2021) Federated learning empowered computation offloading and resource management in 6G-V2X. IEEE Trans Netw Sci Eng 9:3234–3243

Rahman MA, Hossain MS, Islam MS et al (2020) Secure and provenance enhanced internet of health things framework: a blockchain managed federated learning approach. IEEE Access 8:205,071-205,087

Rahman KJ, Ahmed F, Akhter N et al (2021) Challenges, applications and design aspects of federated learning: a survey. IEEE Access 9:124,682-124,700

Reisizadeh A, Mokhtari A, Hassani H, et al (2020) FedPAQ: a communication-efficient federated learning method with periodic averaging and quantization. In: International conference on artificial intelligence and statistics. PMLR, pp 2021–2031

Ren J, Wang H, Hou T et al (2019) Federated learning-based computation offloading optimization in edge computing-supported internet of things. IEEE Access 7:69,194-69,201

Ribero M, Vikalo H (2020) Communication-efficient federated learning via optimal client sampling. arXiv:2007.15197

Rodríguez-Barroso N, Stipcich G, Jiménez-López D et al (2020) Federated learning and differential privacy: software tools analysis, the sherpa. ai fl framework and methodological guidelines for preserving data privacy. Inf Fusion 64:270–292

Rodríguez-Gálvez B, Granqvist F, van Dalen R, et al (2021) Enforcing fairness in private federated learning via the modified method of differential multipliers. arXiv:2109.08604

Rothchild D, Panda A, Ullah E, et al (2020) Fetchsgd: communication-efficient federated learning with sketching. In: International conference on machine learning. PMLR, pp 8253–8265

Roy AG, Siddiqui S, Pölsterl S, et al (2019) Braintorrent: a peer-to-peer environment for decentralized federated learning. arXiv:1905.06731

Saputra YM, Nguyen DN, Hoang DT, et al (2020) Federated learning meets contract theory: energy-efficient framework for electric vehicle networks. arXiv:2004.01828

Sarikaya Y, Ercetin O (2019) Motivating workers in federated learning: a stackelberg game perspective. IEEE Netw Lett 2(1):23–27

Sattler F, Wiedemann S, Müller KR et al (2019a) Robust and communication-efficient federated learning from non-iid data. IEEE Trans Neural Netw Learn Syst 31(9):3400–3413

Sattler F, Wiedemann S, Müller KR, et al (2019b) Sparse binary compression: towards distributed deep learning with minimal communication. In: 2019 international joint conference on neural networks (IJCNN). IEEE, pp 1–8

Sattler F, Müller KR, Samek W (2020) Clustered federated learning: model-agnostic distributed multitask optimization under privacy constraints. IEEE Trans Neural Netw Learn Syst 32:3710—3722

Seif M, Tandon R, Li M (2020) Wireless federated learning with local differential privacy. In: 2020 IEEE international symposium on information theory (ISIT). IEEE, pp 2604–2609

Shahid O, Pouriyeh S, Parizi RM, et al (2021) Communication efficiency in federated learning: achievements and challenges. arXiv:2107.10996

Shen S, Tople S, Saxena P (2016) Auror: Defending against poisoning attacks in collaborative deep learning systems. In: Proceedings of the 32nd annual conference on computer security applications, pp 508–519

Shi Y, Yu H, Leung C (2021) A survey of fairness-aware federated learning. arXiv:2111.01872

Shin M, Hwang C, Kim J, et al (2020) Xor mixup: privacy-preserving data augmentation for one-shot federated learning. arXiv:2006.05148

Shlezinger N, Chen M, Eldar YC et al (2020) Federated learning with quantization constraints. In: ICASSP 2020–2020 IEEE international conference on acoustics. Speech and Signal Processing (ICASSP). IEEE, pp 8851–8855

Shyn SK, Kim D, Kim K (2021) Fedccea: A practical approach of client contribution evaluation for federated learning. arXiv:2106.02310

Smith V, Chiang CK, Sanjabi M, et al (2017) Federated multi-task learning. arXiv:1705.10467

Song T, Tong Y, Wei S (2019) Profit allocation for federated learning. In: 2019 IEEE international conference on big data (Big Data). IEEE, pp 2577–2586

Song Z, Sun H, Yang HH et al (2021) Reputation-based federated learning for secure wireless networks. IEEE Internet Things J 9(2):1212–1226

Strom N (2015) Scalable distributed dnn training using commodity gpu cloud computing. In: Sixteenth annual conference of the international speech communication association

Su L, Liu Z, Ye J (2022) Reputation-based defense scheme against backdoor attacks on federated learning. In: 2021 international conference on big data analytics for cyber-physical system in smart city. Springer, New York, pp 949–955

Swan M (2015) Blockchain: blueprint for a new economy. O'Reilly Media Inc, Sebastopol

Tan AZ, Yu H, Cui L, et al (2021) Towards personalized federated learning. arXiv:2103.00710

Torrey L, Shavlik J (2010) Transfer learning. In: Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. IGI global, p 242–264

Toyoda K, Zhang AN (2019) Mechanism design for an incentive-aware blockchain-enabled federated learning platform. In: 2019 IEEE international conference on big sata (Big Data). IEEE, pp 395–403

Triastcyn A, Faltings B (2019) Federated learning with bayesian differential privacy. In: 2019 IEEE international conference on big data (Big Data). https://doi.org/10.1109/bigdata47090.2019.9005465

Triastcyn A, Faltings B (2020) Federated generative privacy. IEEE Intell Syst 35(4):50–57

Truex S, Baracaldo N, Anwar A, et al (2019) A hybrid approach to privacy-preserving federated learning. In: Proceedings of the 12th ACM workshop on artificial intelligence and security, pp 1–11

Truex S, Liu L, Chow KH, et al (2020) LDP-Fed: federated learning with local differential privacy. In: Proceedings of the third ACM international workshop on edge systems, analytics and networking, pp 61–66

Tu X, Zhu K, Luong NC, et al (2022) Incentive mechanisms for federated learning: from economic and game theoretic perspective. IEEE Trans Cognit Commun Netw pp 1–1. https://doi.org/10.1109/TCCN.2022.3177522

Tyagi N (2022) What is differential privacy and how does it work? Analytics steps. https://www.analyticssteps.com/blogs/what-differential-privacy-and-how-does-it-work

Vahidian S, Morafah M, Lin B (2021) Personalized federated learning by structured and unstructured pruning under data heterogeneity. arXiv:2105.00562

Van Dyk DA, Meng XL (2001) The art of data augmentation. J Comput Graph Stat 10(1):1–50

Vy NC, Quyen NH, Pham VH, et al (2021) Federated learning-based intrusion detection in the context of iiot networks: poisoning attack and defense. In: International conference on network and system security. Springer, New York, pp 131–147

Wainakh A, Guinea AS, Grube T, et al (2020) Enhancing privacy via hierarchical federated learning. In: 2020 IEEE European symposium on security and privacy workshops (EuroS &PW). IEEE, pp 344–347

Wan W, Lu J, Hu S, et al (2021) Shielding federated learning: a new attack approach and its defense. In: 2021 IEEE wireless communications and networking conference (WCNC). IEEE, pp 1–7

Wang J, Chen Y, Hao S, et al (2017) Balanced distribution adaptation for transfer learning. In: 2017 IEEE international conference on data mining (ICDM). IEEE, pp 1129–1134

Wang G, Dang CX, Zhou Z (2019a) Measure contribution of participants in federated learning. In: 2019 IEEE international conference on big data (Big Data), pp 2597–2604, https://doi.org/10.1109/BigData47090.2019.9006179

Wang L, Wang W, Li B (2019b) CMFL: mitigating communication overhead for federated learning. In: 2019 IEEE 39th international conference on distributed computing systems (ICDCS), pp 954–964

Wang X, Han Y, Wang C et al (2019c) In-edge ai: intelligentizing mobile edge computing, caching and communication by federated learning. IEEE Network 33(5):156–165

Wang T, Rausch J, Zhang C, et al (2020a) A principled approach to data valuation for federated learning. In: Federated learning. Springer, p 153–167

Wang Y, Zhu T, Chang W, et al (2020b) Model poisoning defense on federated learning: a validation based approach. In: International conference on network and system security. Springer, New York, pp 207–223

Wang Z, Yang Y, Liu Y, et al (2020c) Cloud-based federated boosting for mobile crowdsensing. arXiv:2005.05304

Wang C, Liu Z, Wei H et al (2021a) Hybrid deep learning model for short-term wind speed forecasting based on time series decomposition and gated recurrent unit. Complex Syst Model Simul 1(4):308–321

Wang J, Charles Z, Xu Z, et al (2021b) A field guide to federated optimization. arXiv:2107.06917

Wang S, Chen M, Yin C et al (2021c) Federated learning for task and resource allocation in wireless high-altitude balloon networks. IEEE Internet Things J 8(24):17,460-17,475

Wang Z, Fan X, Qi J, et al (2021d) Federated learning with fair averaging. arXiv:2104.14937

Wei X, Li Q, Liu Y, et al (2019) Multi-agent visualization for explaining federated learning. In: IJCAI, pp 6572–6574

Wei K, Li J, Ding M et al (2020a) Federated learning with differential privacy: algorithms and performance analysis. IEEE Trans Inf Forensics Secur 15:3454–3469

Wei W, Liu L, Loper M, et al (2020b) A framework for evaluating client privacy leakages in federated learning. In: European symposium on research in computer security. Springer, New York, pp 545–566

Wei K, Li J, Ding M, et al (2021) User-level privacy-preserving federated learning: analysis and performance optimization. IEEE Trans Mob Comput

Wen W, Xu C, Yan F, et al (2017) Terngrad: Ternary gradients to reduce communication in distributed deep learning. In: Advances in neural information processing systems, p 30

Weng J, Weng J, Zhang J, et al (2019) Deepchain: auditable and privacy-preserving deep learning with blockchain-based incentive. IEEE Trans Depend Secure Comput

Weng J, Weng J, Huang H, et al (2021) Fedserving: a federated prediction serving framework based on incentive mechanism. In: IEEE INFOCOM 2021-IEEE conference on computer communications. IEEE, pp 1–10

Wohlin C (2014) Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: Proceedings of the 18th international conference on evaluation and assessment in software engineering. Association for Computing Machinery, EASE '14, https://doi.org/10.1145/2601248.2601268

Wu Q, Chen X, Zhou Z, et al (2020a) Fedhome: Cloud-edge based personalized federated learning for in-home health monitoring. IEEE Trans Mobile Comput

Wu Y, Cai S, Xiao X, et al (2020b) Privacy preserving vertical federated learning for tree-based models. arXiv:2008.06170

Wu C, Wu F, Liu R, et al (2021a) Fedkd: Communication efficient federated learning via knowledge distillation. arXiv:2108.13323

Wu J, Liu Q, Huang Z et al (2021b) Hierarchical personalized federated learning for user modeling. Proc Web Conf 2021:957–968

Xie M, Long G, Shen T, et al (2021) Multi-center federated learning. arXiv:2108.08647

Xu X, Lyu L (2020) Towards building a robust and fair federated learning system. arXiv e-prints pp arXiv–2011

Xu R, Baracaldo N, Zhou Y, et al (2019a) Hybridalpha: An efficient approach for privacy-preserving federated learning. In: Proceedings of the 12th ACM workshop on artificial intelligence and security. pp 13–23

Xu Z, Yang Z, Xiong J et al (2019b) Elfish: Resource-aware federated learning on heterogeneous edge devices. Ratio 2(r1):r2

Xu J, Glicksberg BS, Su C et al (2021) Federated learning for healthcare informatics. J Healthc Inf Res 5(1):1–19

Yang Q, Liu Y, Chen T, et al (2019) Federated machine learning: concept and applications. ACM Trans Intell Syst Technol 10(2). https://doi.org/10.1145/3298981

Yang G, Mu K, Song C, et al (2021a) Ringfed: Reducing communication costs in federated learning on non-iid data. arXiv:2107.08873

Yang G, Wang S, Wang H (2021b) Federated learning with personalized local differential privacy. In: 2021 IEEE 6th international conference on computer and communication systems (ICCCS). IEEE, pp 484–489

Yang M, Wang X, Zhu H, et al (2021c) Federated learning with class imbalance reduction. In: 2021 29th European signal processing conference (EUSIPCO). IEEE, pp 2174–2178

Yao X, Huang C, Sun L (2018) Two-stream federated learning: reduce the communication costs. In: 2018 IEEE visual communications and image processing (VCIP). IEEE, pp 1–4

Yao X, Huang T, Wu C, et al (2019a) Towards faster and better federated learning: a feature fusion approach. In: 2019 IEEE international conference on image processing (ICIP). IEEE, pp 175–179

Yao X, Huang T, Wu C, et al (2019b) Federated learning with additional mechanisms on clients to reduce communication costs. arXiv:1908.05891

Ye D, Yu R, Pan M et al (2020) Federated learning in vehicular edge computing: a selective model aggregation approach. IEEE Access 8:23,920-23,935

Yi Ming W, Ge Hao L, Li Yu F, et al (2021) Research on block chain defense against malicious attack in federated learning. In: 2021 the 3rd international conference on blockchain technology, pp 67–72

Yoo JH, Son HM, Jeong H, et al (2021) Personalized federated learning with clustering: non-iid heart rate variability data application. In: 2021 International conference on information and communication technology convergence (ICTC). IEEE, pp 1046–1051

Yu H, Liu Z, Liu Y, et al (2020a) A fairness-aware incentive scheme for federated learning. In: Proceedings of the AAAI/ACM conference on AI, ethics, and society. pp 393–399

Yu P, Kundu A, Wynter L, et al (2020b) Fed+: a unified approach to robust personalized federated learning. arXiv:2009.06303

Yu S, Chen X, Zhou Z, et al (2020c) When deep reinforcement learning meets federated learning: intelligent multi-timescale resource management for multi-access edge computing in 5G ultra dense network. arXiv:2009.10601

Yu T, Bagdasaryan E, Shmatikov V (2020d) Salvaging federated learning by local adaptation. arXiv:2002.04758

Yu T, Li T, Sun Y, et al (2020e) Learning context-aware policies from multiple smart homes via federated multi-task learning. In: 2020 IEEE/ACM fifth international conference on internet-of-things design and implementation (IoTDI). IEEE, pp 104–115

Yuan X, Ma X, Zhang L, et al (2021) Beyond class-level privacy leakage: breaking record-level privacy in federated learning. IEEE Internet Things J

Yue X, Kontar RA (2021) Federated gaussian process: convergence, automatic personalization and multi-fidelity modeling. arXiv:2111.14008

Yue X, Nouiehed M, Kontar RA (2021) Gifair-fl: An approach for group and individual fairness in federated learning. arXiv:2108.02741

Yue K, Jin R, Wong CW, et al (2022) Communication-efficient federated learning via predictive coding. IEEE J Select Top Signal Process 16:369–380

Yurochkin M, Agarwal M, Ghosh S, et al (2019) Bayesian nonparametric federated learning of neural networks. In: International conference on machine learning. PMLR, pp 7252–7261

Zeng R, Zhang S, Wang J, et al (2020) Fmore: An incentive scheme of multi-dimensional auction for federated learning in mec. In: 2020 IEEE 40th international conference on distributed computing systems (ICDCS). IEEE, pp 278–288

Zeng R, Zeng C, Wang X, et al (2021) A comprehensive survey of incentive mechanism for federated learning. arXiv:2106.15406

Zhan Y, Li P, Qu Z et al (2020) A learning-based incentive mechanism for federated learning. IEEE Internet Things J 7(7):6360–6368

Zhan Y, Zhang J, Hong Z, et al (2021) A survey of incentive mechanism design for federated learning. IEEE Trans Emerg Top Comput 10:1035–1044

Zhang X, Luo X (2020) Exploiting defenses against gan-based feature inference attacks in federated learning. arXiv:2004.12571

Zhang J, Chen J, Wu D, et al (2019) Poisoning attack in federated learning using generative adversarial nets. In: 2019 18th IEEE international conference on trust, security and privacy in computing and communications/13th IEEE international conference on big data science and engineering (TrustCom/BigDataSE). IEEE, pp 374–380

Zhang DY, Kou Z, Wang D (2020a) FairFL: a fair federated learning approach to reducing demographic bias in privacy-sensitive classification models. In: 2020 IEEE international conference on big data (Big Data). IEEE, pp 1051–1060

Zhang J, Li C, Robles-Kelly A, et al (2020b) Hierarchically fair federated learning. arXiv:2004.10386

Zhang M, Sapra K, Fidler S, et al (2020c) Personalized federated learning with first order model optimization. arXiv:2012.08565

Zhang C, Xie Y, Bai H et al (2021a) A survey on federated learning. Knowl-Based Syst 216(106):775

Zhang DY, Kou Z, Wang D (2021b) Fedsens: a federated learning approach for smart health sensing with class imbalance in resource constrained edge computing. In: IEEE INFOCOM 2021-IEEE conference on computer communications. IEEE, pp 1–10

Zhang F, Kuang K, Liu Y, et al (2021c) Unified group fairness on federated learning. arXiv:2111.04986

Zhang J, Guo S, Ma X, et al (2021d) Parameterized knowledge transfer for personalized federated learning. Adv Neural Inf Process Syst 34:10092–10104

Zhang J, Wu Y, Pan R (2021e) Incentive mechanism for horizontal federated learning based on reputation and reverse auction. Proc Web Conf 2021:947–956

Zhang Z, Dong D, Ma Y et al (2021f) Refiner: a reliable incentive-driven federated learning system powered by blockchain. Proc VLDB Endow 14(12):2659–2662

Zhao Y, Chen J, Zhang J, et al (2019) PDGAN: a novel poisoning defense method in federated learning using generative adversarial network. In: International conference on algorithms and architectures for parallel processing. Springer, New York, pp 595–609

Zhao Y, Zhao J, Jiang L et al (2020a) Privacy-preserving blockchain-based federated learning for iot devices. IEEE Internet Things J 8(3):1817–1829

Zhao Y, Zhao J, Yang M et al (2020b) Local differential privacy-based federated learning for internet of things. IEEE Internet Things J 8(11):8836–8853

Zhao C, Wen Y, Li S, et al (2021) Federatedreverse: a detection and defense method against backdoor attacks in federated learning. In: Proceedings of the 2021 ACM workshop on information hiding and multimedia security, pp 51–62

Zhou H, Cheng J, Wang X, et al (2020) Low rank communication for federated learning. In: International conference on database systems for advanced applications. Springer, New York, pp 1–16

Zhou Y, Ye Q, Lv J (2021) Communication-efficient federated learning with compensated overlap-fedavg. IEEE Trans Parallel Distrib Syst 33(1):192–205

Zhu L, Han S (2020) Deep leakage from gradients. In: Federated learning. Springer, New York, pp 17–31

Zhuang W, Wen Y, Zhang X, et al (2020) Performance optimization of federated person re-identification via benchmark analysis. In: Proceedings of the 28th ACM international conference on multimedia, pp 955–963