Journal of Information Technology Management
ISSN-OnLine: 2980-7972
University of Tehran

Enhancing Privacy and Efficiency Techniques in Federated Learning Systems: Applications in Healthcare, Finance, and Smart Devices

Ravi Shankar Shukla *

*Corresponding author, Department of Computer Science, College of Computing and Informatics, Saudi Electronic University, Saudi Arabia. E-mail: ravipraful@gmail.com

Journal of Information Technology Management, 2025, Vol. 17, Special Issue, pp.45-62.
Published by the **University** of Tehran, College of Management
doi: https://doi.org/10.22059/jitm.2025.102921
Article Type: Research Paper

Received: January 17, 2025
Received in revised form: March 03, 2025
Accepted: June 13, 2025
Published online: August 01, 2025

Abstract

Federated Learning (FL) has emerged as a revolutionary technique for distributed machine learning for training a model on shared data without sharing the data itself. Nevertheless, privacy-related concerns and scalability difficulties remain a problem. This paper discusses the state-of-the-art works to improve the privacy and convergence at FL frameworks for targeted healthcare and financial applications, as well as smart devices. It focuses on methodologies that preserve user privacy, such as differential privacy, homomorphic encryption, secure multi-party computation, and methods that enhance the model's efficiency, including model compression, communication optimization, and adaptive optimization algorithms. To overcome these challenges, this study helps in the future design of FL systems for vital domains with high scalability.

Keywords: Federated Learning (FL), Privacy Enhancement, Adaptive Federated Optimization, Heterogeneity, Scalability, Federated Averaging.

Introduction

Federated Learning (FL) is an innovative machine learning architecture that enables model training across multiple parties or devices, such as individual devices or organizations, without the need to transfer raw data. Indeed, this decentralized data handling and storage organization directly solves the emergent need for privacy protection, especially in emerging domains, including healthcare, finance, and smart devices, where data protection and compliance with privacy legislation like GDPR and HIPAA are paramount. Nevertheless, FL faces critical challenges, primarily centered on ensuring robust privacy guarantees and

achieving satisfactory computational performance. These challenges are further compounded in complex environments where the data distribution is irregular and different devices serve different distributions for clients and servers that need real-time interaction.

Privacy issues in FL are because gradients, model updates, or inference attacks can potentially leak the data. Several advanced techniques have been worked out to minimize these risks. Differential Privacy (DP) perturbs model updates directly with calibrated noise to offer the theoretical protection of hiding specific data details. However, it may not always address the problem of privacy while still trying to retain the accuracy of the model. Homomorphic Encryption (HE) enables computations to be performed directly on encrypted data, ensuring data confidentiality during processing. However, due to its high computational cost, it is not well-suited for real-time applications. Secure Multi-Party Computation (SMPC), on the other hand, allows multiple participants to jointly perform computations without exposing their private data to one another. Despite the high security that SMPC provides, there are important issues, such as communication overhead and synchronization, which have to be addressed in this protocol (Abadi et al., 2016). Another technique, which is also used, is known as pseudonymization, which anonymizes data so as not to be reverse-engineered. Although easy to apply, its effectiveness is contingent on the honesty of the aggregators and has the problem of re-identification. It is performed over various important segments like healthcare, finance, and IoT networks, where the difficulty and opportunities vary (Kairouz et al., 2021).

Flaws affecting efficiency in FL are also equally substantial and chiefly pertain to the computational and communication requirements of distributed learning (Bonawitz et al., 2019). The concept of model compression includes pruning, quantization, and distillation, which makes models compact for efficient training and deployment. However, these methods are at the cost of model accuracy and computational time as well (Li et al., 2020). As a way to reduce the amount of communicated data between individual participants and the central server, Communication-efficient algorithms like gradient sparsification and Federated Averaging or FedAvg have been proposed; however, they also have the problem of potential bottlenecks in large-scale systems. Instead, more complex approaches to adaptively optimize client updates include FedProx and learning rate control; these consider the diversity of the distribution of client data and client computation resources (Yang, Liu et al., 2019). The integration of Edge AI, which uses FL together with edge computing, enables on-device processing and inference, although resource restrictions at the edge may hinder the deployment of the technique (McMahan et al., 2017). All these techniques are designed to address the specific requirements of the healthcare and the financial system, as well as smart devices, to operate effectively under different conditions.

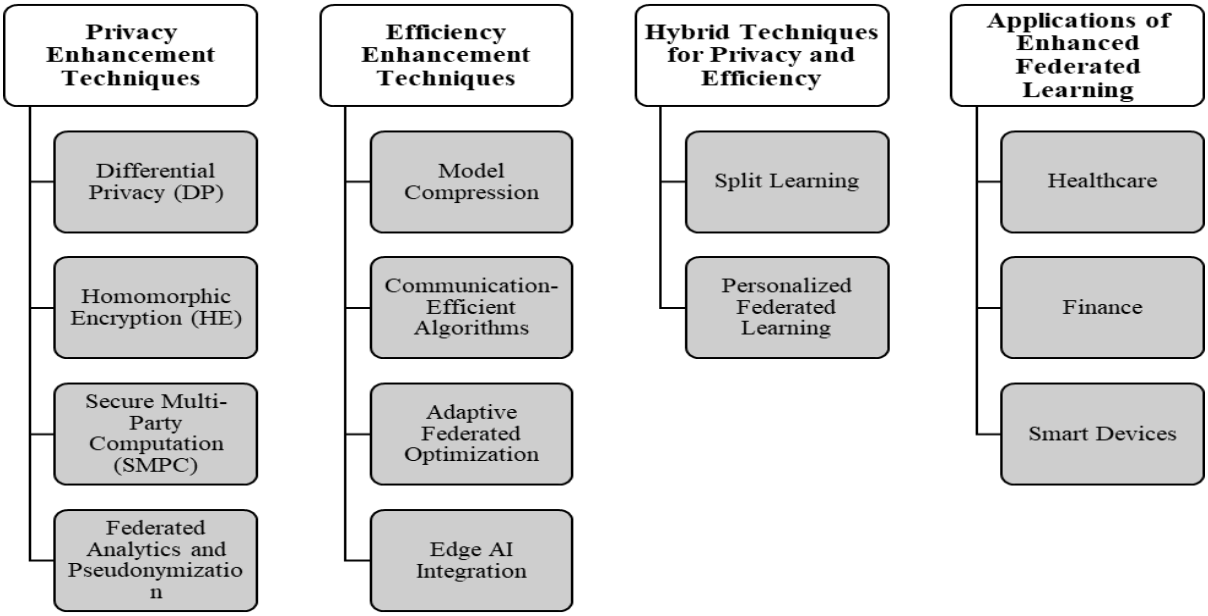

Figure 1. Overview of Federated Learning Techniques and Applications

In order to address the problem of the gap between privacy and efficiency, new approaches are being developed, which use a combination of both methods – split learning and personalized federated learning (Geyer, Klein, & Nabi, 2017). The proposed split learning approach performs computations on the client side and server side, where only the necessary data is exposed in the client (Truex et al., 2019). The personalized federated learning categorizes generic models from the World Wide Web as relevant to local clients while maintaining the confidentiality of outcomes (Zhao et al., 2018). Such approaches are particularly preferred to achieve both the goals of {preserving the privacy} and {developing efficient resource} management (Hardy, Berkovsky, Penschuck, Germanakos, & Bellotti, 2020).

The following tables summarize the privacy and efficiency techniques, their strengths, challenges, and applications:

Table 1. Privacy and efficiency techniques

Privacy Technique	Strengths	Challenges	Applications
Differential Privacy (Li et al., 2020)	Provides strong theoretical guarantees	May reduce model accuracy	Healthcare, finance, and smart devices
Homomorphic Encryption (Hynes et al., 2018)	Ensures secure computation on encrypted data	High computational costs	Medical data aggregation, IoT systems
Secure Multi-Party Computation (Smith et al. 2017)	Facilitates secure collaborative computation	Communication overhead	Collaborative diagnostics, financial risk models
Pseudonymization (Wang, et al. 2020)	Simplifies implementation through anonymity	Vulnerable to re-identification risks	Genomic analysis, trend forecasting

In conclusion, Federated Learning holds immense potential as a privacy-preserving, efficient framework for modern applications.

Table 2. Efficiency framework for modern applications

Efficiency Technique	Key Features	Challenges	Applications
Model Compression (Zhu et al. 2019)	Reduces model size through pruning, quantization, and distillation	Potential accuracy loss	Medical imaging, low-latency fraud detection
Communication Efficiency (Shokri & Shmatikov, 2015)	Minimizes data transfer through sparsification and aggregation techniques	Communication bottlenecks in large-scale systems	Real-time medical diagnostics, personalized recommendations
Adaptive Optimization (Wang et al. 2019)	Adapts to heterogeneous data and client conditions	Implementation complexity	Healthcare data diversity, IoT device training
Edge AI Integration (Caldas et al. 2018)	Enables localized processing and inference	Device resource limitations	On-device diagnostics, smart device automation

It is therefore possible to scale FL using state-of-the-art techniques and hybrid methods to address the requirements of various and complex domains, thus making it possible to develop reliable, more scalable, and secure decentralization of learning systems (Konečný et al., 2016). This paper provides a rich presentation of these techniques and their uses and importance in the next FL systems in healthcare, finance, and smart device networks.

Methodology

Privacy Enhancement Techniques

Some of the biggest issues in federated learning (FL) systems include privacy issues, which may be a big deal in business-oriented areas such as consideration in healthcare, finance, and smart devices (Yang et al., 2018). Several methods have been worked out to overcome all these problems, and all of these have been designed and developed using quite different approaches to support collaborative learning while preserving the privacy of the learners. This section gives a detailed understanding of the four IM/information protection methods, namely, Differential Privacy (DP), Homomorphic Encryption (HE), Secure Multi-Party Computation (SMPC), and Federated Analytics with Pseudonymization, uses, limitations, and comparison (Bonawitz et al., 2017).

2.1 Differential Privacy (DP)

Differential Privacy is a statistical procedure that prevents individual data contributions from being distinguished by adding calibrated noise into gradients or updates to models (Li et al., 2019). The first principle for dealing with data is to contribute due to any single item in a set so small that it can be regarded effectively as negligible. This mechanism is formally defined by the ϵ Differential Privacy metric, where the parameter ϵ determines the privacy level. A commonly used implementation is expressed as

$\Delta u = f(x) + \mathcal{N}(0, \sigma^2)$

Where:

$(f(x): The original function (e.g., model gradients or updates).$

$\mathcal{N}(0, \sigma^2)$: Gaussian noise with mean 0 and variance \(σ^2\).

Δu: The output after applying differential privacy.

In the field of health care, DP has been applied to preserve the privacy of patients' records during the shared training of diagnostic models (Mohri et al., 2019). For instance, healthcare facilities can make use of patients' information to train models with the restriction that no personal details of the patient will be used. In finance, it protects transactional data used in fraud detection and credit scoring and enables different institutions to improve their models based on the data while preserving customers' privacy (Nguyen et al., 2021). Likewise, in the context of smart devices, it hinders user activity data to promote privacy in the recommendation engines in smart devices and other personalization techniques (Chen et al., 2020). Nonetheless, the major constraint in deploying Differential Privacy is the way of optimally partitioning the privacy-sensitivity and performance-relevance spectrum. High noise levels reduce the performance of the model, and so, fine-tuning is crucial to return utility (Sattler et al., 2019).

2.2 Homomorphic Encryption (HE)

Homomorphic Encryption is a cryptographic technique that includes the capability to compute on encrypted data securely. While most approaches involve decrypting data before computation, HE means that the data does not have to be decrypted for computation to occur (Reisizadeh et al., 2020). Of particular importance in federated learning is this property because data that is normally sensitive has to be processed collectively (Zhang et al., 2021).

In healthcare, Homomorphic Encryption allows statistical data of patients' records in different institutions to be added together so results indicating prognosis of patient or disease pattern, etc, can be obtained without revealing individual patients' data (Luo et al., 2021). In finance, for instance, it allows encrypted training of machine learning models on the financial data while allowing institutions to exchange encrypted model updates rather than the data. For smart devices, HE allows computations in IoT systems and, at the same time, keeps data locally in each device while helping to create global models (Zhu, Zhou, & Xu, 2021). Although Homomorphic Encryption provides strong and flexible security, it comes with a high computational expense. The high requirement of computational power and relatively longer time for computation may curtail its usability in real-time or constrained environments (Mangal et al., 2023).

2.3 Secure Multi-Party Computation (SMPC)

SMPC or Secure Multi-Party Computation is a cryptographic method for carrying out computations where multiple parties possess data, but they all compute a joint function on it without each of the inputs being visible to the others. This approach is most suitable in situations where one party cannot be relied upon to manage full data.

In healthcare, SMPC assists multiple hospitals in employing cooperative diagnostic tools by enabling them to perform computations on the collective SMs model without sharing the patient-level data. For example, it is possible for different institutions to collaboratively develop an AI system for disease prediction without compromising on records. In the financial sector, it allows risk assessment of many separate institutions that combine advice and information without revealing individual data. For smart devices, SMPC keeps user-level data safe and allows devices to learn from each other's information while keeping it secret. However, high communication overhead and the necessity of coordinating all parties that join SMPC pose major issues for the further development of the protocol.

2.4 Federated Analytics and Pseudonymization

Federated analytics and pseudonymization both focus on obscuring and disguising data insights to prevent the reverse engineering of data for identification purposes. These methods help ensure the anonymity of sensitive information while still enabling data analysis by replacing personal identifiers with pseudonyms and aggregating numerical data. In healthcare, it improves genomic data through the centralization of de-identified data from diverse clinical care sites on patients. It ensures that multiple parties can have a look at genetic diseases without having access to other genomic sequences. In finance, Pseudonymization is used for maintaining correlation for trend analysis for market prediction that would involve later linking, while institutions can model finance without sharing customer data. For smart devices, it protects user-level data contribution to shared learning processes that preventing tracking of the individual device activity. However, the viability of Pseudonymization is based on trust in the aggregators as well as the efficacy of the anonymization processes. Insecure pseudonymization means that a greater number of data subjects can be easily identified from the data, which is an invasion of privacy.

Results

Comparative Analysis of Privacy Techniques

The table below provides a detailed comparison of the strengths, challenges, and applications of these privacy-enhancing techniques. A tick (✓) indicates a positive attribute, while a cross (×) highlights a limitation.

Table 3. Comparison Analysis of Privacy Techniques

Technique	Strengths	Challenges	Healthcare	Finance	Smart Devices
Differential Privacy	✓ Strong theoretical guarantees	× Potential accuracy loss	✓	✓	✓
Homomorphic Encryption	✓ Secure computation on encrypted data	× High computational demands	✓	✓	✓
Secure MPC	✓ High security for collaborative tasks	×Communication overhead	✓	✓	✓
Pseudonymization	✓ Simplicity and scalability	×Risk of re-identification	✓	✓	✓

The options presented and discussed in this analysis show the best and worst of specific approaches to privacy for federated learning systems. Each of the methods has its special characteristics and strengths; however, the choice of the method directly depends on the features of the application domain. For future work, it is essential to develop a blend of these techniques to eliminate respective deficits and provide a strong and elastic privacy solution for FL.

Efficiency Enhancement Techniques

One of the most important requirements of FL systems is efficiency since computation becomes limited, communication is expensive, and clients may have differing capabilities. Some of the advanced techniques that have been used in order to improve efficiency in FL include the following. They are model compression, communication-efficient algorithms, adaptive federated optimization, and integration of Edge AI. These are described in this section together with examples of their usage, and the possibility of redesigning FL systems.

Here is a detailed tabular comparison of efficiency enhancement techniques in federated learning using tick (✓) and cross (×) symbols to indicate their suitability for specific metrics and applications:

Table 4. Comparison of efficiency enhancement techniques

Technique	Communication Overhead	Computational Cost	Scalability	Deployment Ease
Model Compression	×	×	✓	✓
Communication-Efficiency	✓	~ (Moderate)	✓	✓
Adaptive Optimization	~ (Moderate)	~ (Moderate)	✓	~ (Moderate)
Edge AI Integration	✓	✓	×	✓

Model compression raises specific concerns, such as communication overhead and computational cost. Additionally, it often requires pre-processing, which may lead to some

loss of accuracy. Nevertheless, it offers significant advantages in terms of scalability and extensibility, and its ease of implementation makes it particularly suitable for organizations with limited resources. Communication efficiency, on the other hand, outperforms other PGAS metrics by significantly reducing communication overhead while maintaining good scalability and ease of deployment. Nevertheless, it has a relatively reasonable time complexity, and it can only be regarded as middle-range, which should be used and implemented methodically and with due measure. Adaptive optimization imagery bears reasonable scalabilities, and the system performs efficiently on various datasets and clients. However, it provides moderate computational cost, communication overhead, and ease of deployment, in which optimization is possible. Thus, the integration of edge AI is distinguished by reducing the communication overhead and improving the level of deployment ease due to the data's ability to be processed locally. It also has a low computational cost for its operation. However, its scalability is limited, as edge devices often struggle to manage large, distributed environments efficiently.

1. Model Compression

The majority of neural network-based models for machine intelligence must go through model compression methodologies to minimize their size for efficient training and deployment. These techniques work to realize compression by reducing the probability model space and any redundancy. Pruning eliminates useless parameters within the model, thus decreasing the number of calculations needed without considerable degradation of performance. This process decreases the model parameter precision, which in turn minimizes memory and computational usage. Knowledge distillation can be described as fine-tuning lightweight models using information derived from larger trained models. In the healthcare context, model compression makes tasks, including medical image analysis, less demanding computationally and thus deployable on less capable platforms. In finance, compressed models provide a means of optimizing fraud detection systems for high-speed, low-latency responses. In the case of optimizing for models of smart devices, the act of compression implies that they have to fit into the restricted memory and processing power of mobile phones and IoT systems.

2. Communication-Efficient Algorithms

Inter-client or inter-server coordination is another major FL issue: coordination of communicating model updates between clients and servers can put overwhelming pressure on bandwidth resources and slow down training. To overcome this, communication-efficient algorithms minimize the amount and rate of information exchanged. The Gradient sparsification means that during the training process, only the greatest gradients are transferred. Federated Averaging (FedAvg) only requires the server to aggregate the local model updates; thus, fewer communication rounds are needed for convergence. Thus,

asynchronous updates enable the clients to update independently with no regard to synchronization with the server, which advances the training progress. In healthcare, these algorithms accelerate the federated learning for medical diagnosis as the training delay is effectively shortened. They extend such applications in finance as real-time fraud detection models, by providing quick data consolidation and decision-making. For smart devices, communication-efficient algorithms support responsive learning, particularly by enabling rapid updates to recommendations.

3. Adaptive Federated Optimization

The inherent non-IID data and diverse client devices give rise to novel optimization strategies in FL. These practices make it possible to ensure that the global model will run properly on different client platforms. Other methods, like FedProx and adaptive learning rates, are able to alter contributions from different clients by optimizing the process. In healthcare, adaptive optimization means that the control is tuned to achieve stable prediction or diagnosis performance no matter how the patients' data changes in terms of distribution. In finance, it helps institutions with different characteristics of data to combine efforts in the system. In the case of smart devices, adaptive optimization uses individual data patterns to optimize the learning process for individual users effectively.

4. Edge AI Integration

FL is introduced together with edge AI that allows for the processing and inference on the device. Edge AI minimizes latency, improves privacy, and utilizes less amount of bandwidth since computations are done on the edge. In healthcare, edge AI makes real-time diagnostic evaluations through a federated model at the edges for wearable health monitoring. In finance, it facilitates local credit rating and evaluation, not requiring clearing at the center. Edge AI in smart devices allows for instantaneous personalization, as well as automation within an IoT setting for users. The heatmap provides a visual representation of the performance of various efficiency techniques used in federated learning across four key metrics, namely communication overhead, computation complexity or time consumption, scalability, and global deployment convenience. Each technique is represented by a grade from 1 to 5, where darker shades represent strength or excellent performance while light shades depict areas of weakness. Deployment Ease and Scalability are easy, and Model Compression gets an impressive 5 for the former and 4 for the latter, showing it can readily be applied where there is a limitation in the available resources. Yet it has limitations in communication overhead and computational complexity, rated 2 and 1, respectively, suggesting high computational demands and communication overhead when implemented. The class of Communication-efficient algorithms achieved the highest score of 5 in the ability to minimize the degree of communication used and 4 in scalability, as well as deployment ease. Their performance in computational cost is moderate, with a computed score of 3, which indicates that great caution

should be taken while optimizing. AOC is the best scheme, attaining a score of 5 in scalability, implying that the adaptive optimization technique can well cope with heterogeneous data distributions. Intermediate numbers in computational cost and communication overhead scores, as well as in relative ease of deployment, suggest moderate, but by no means outstanding, performance in these regards.

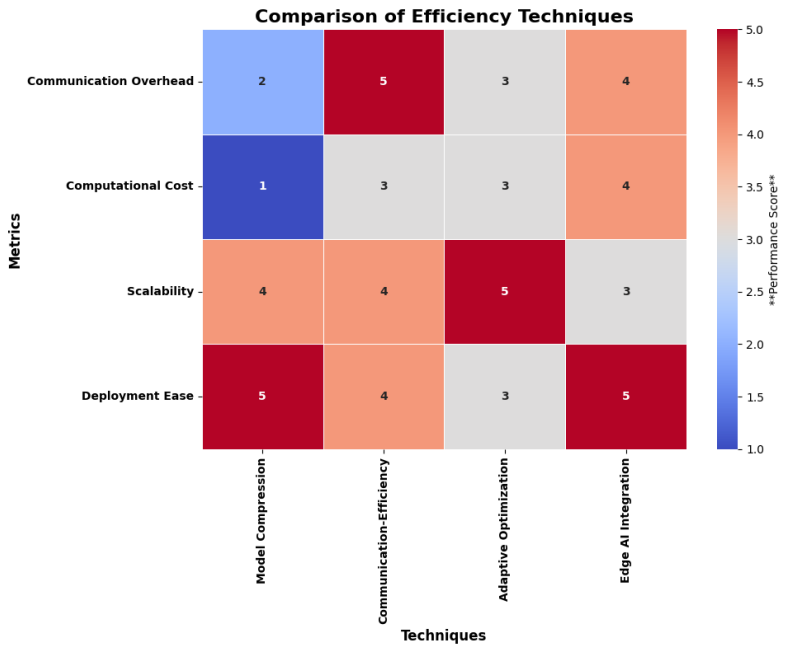

Figure 2. Comparison of Efficiency Techniques

Hybrid Techniques for Privacy and Efficiency

Since no single method of either privacy or efficiency can fully solve the problem while the others can, the best way to achieve both goals is to combine them into hybrid techniques to create sound and efficient federated learning (FL) systems. This section explores two significant hybrid approaches: Split Learning and Personalized Federated Learning.

1. Split Learning

Split Learning divides the model into two segments: client and server modules, such as the Server Side Include, or SSI. Some data manipulations are done on the client side; the other computations are done on the server side. This division decreases the pressure on the client's device and decreases data openness by restricting the aspects provided to the server. Some of the use cases of Split Learning include PHI protection during collaborative learning in healthcare, protection of financial data during collaborative learning in finance, etc. It also enhances low-latency services in smart devices by minimizing computations to other gadgets

and delay periods. For example, in the case of mobile wearables, only basic analysis might be carried out on the device while just sending selected parameters to the server.

2. Personalized Federated Learning

Global models to each client's needs is what Personalized Federated Learning does, which helps to overcome the problem of heterogeneity. This approach enhances model relevance, helps to exclude pointless computations, and strengthens generalization across differing sets of data. In healthcare, where everything is getting centralized, it helps in recommending an individually tailored treatment plan based on the patient's data collected on a global level. In finance, personalized FL promotes financial planning to meet the needs of different institutions and customers, as the model is developed to suit different audiences. For smart devices, user experience is thus to provide personalized recommendations while boosting the functionality of IoT applications such as smart home technology and wearables.

Applications of Enhanced Federated Learning

These improved techniques mentioned earlier are revolutionizing several fields through secure, efficient, and scalable FL systems.

1. Healthcare

In healthcare, the medical image analysis based on joint diagnostics enables training of models based on big data without disclosing patients' records between hospitals. Through federated learning models of drug discovery, the identification of possible drug candidates is enhanced, and the privacy of the patient's data is protected. Moreover, FL can contribute to privacy-preserving genomics and disease prediction to allow multi-institutional collaboration for genetic study data.

2. Finance

They work in finance to boost fraudulent detection through shared learning among institutions, since security against fraudulent people is well enhanced through the FL system. Personal credit scoring: A Large amount of credit data is analyzed locally, and individual credit scoring can be given without transmitting delicate financial details. FL also provides an effective means for distributing market trend prediction responsibility to different market areas in order that institutions can forecast economic trends.

3. Smart Devices

Smart devices also transform applications like smart home automation since user data is processed locally, thus improving privacy. Focused personalization in wearable technology

helps the targeted person adjust to personalized suggestions. Effective federated learning approaches enhance scalability in IoT networks and ensure secure D2D connections.

Table 5. Comparative Table for Hybrid Techniques

Technique	Strengths	Challenges	Applications
Split Learning	Reduces client computation and data exposure	Requires high coordination	Healthcare, Finance, Smart Devices
Personalized Federated Learning	Improves model relevance and adaptability	May increase complexity in updates	Healthcare, Finance, IoT

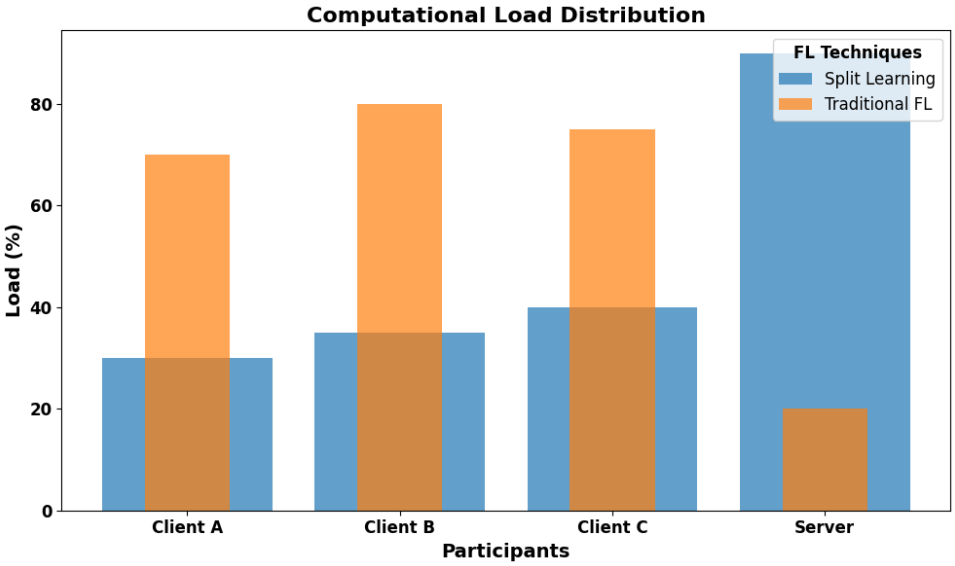

Figure 3. Computation load distribution

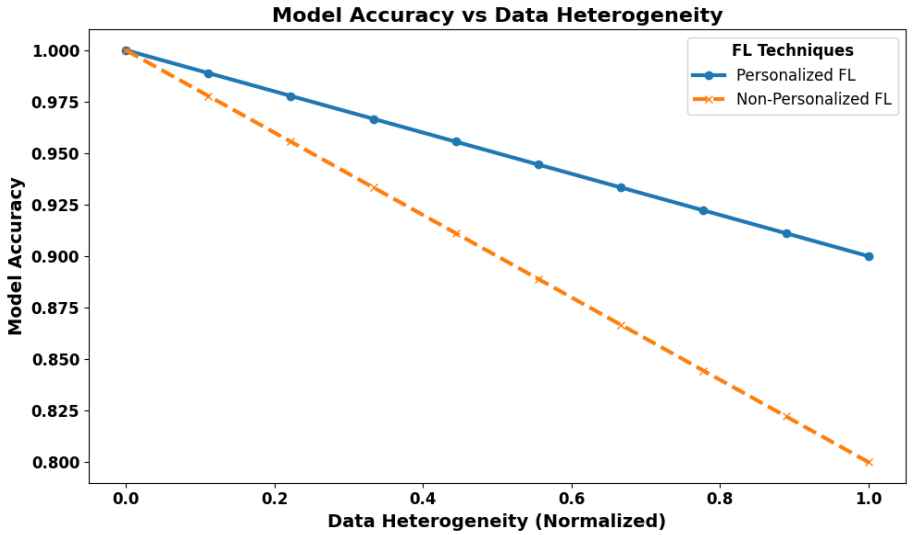

Figure 4. Model accuracy vs data heterogeneity

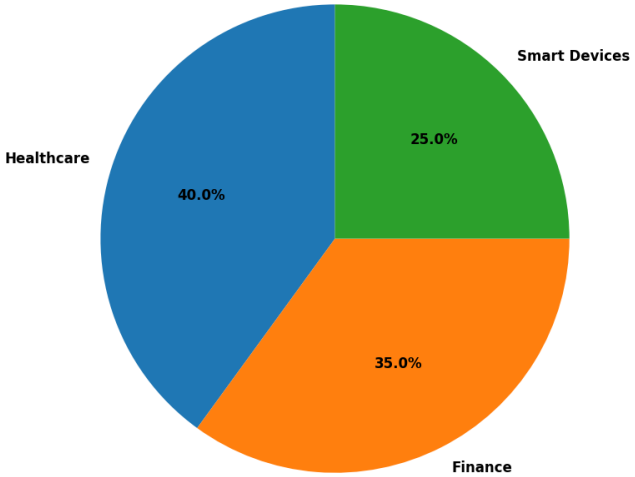

Figure 5. FL Application distribution

Graph 1: Split Learning Performance Overhead

It aims to show the share of the computation that is performed by clients and servers in Split Learning against to the traditional federated learning (FL) approach. By applying Split Learning, the end user has lesser amount of calculations to perform while the computational work is scaled up on the server.

Tabular Data: Computational Load Distribution

Graph 1. Personalized FL - Model Accuracy vs Data Heterogeneity

Participant	Split Learning (%)	Traditional FL (%)
Client A	30	70
Client B	35	80
Client C	40	75
Server	90	20

Tabular Data: Accuracy vs Heterogeneity

Graph 2. Applications Distribution

Heterogeneity Level	Personalized FL Accuracy (%)	Non-Personalized FL Accuracy (%)
0.0	100	100
0.2	98	96
0.4	96	92
0.6	94	88
0.8	92	84
1.0	90	80

In this pie chart, these are the percentages of federated learning applications where the three sectors: healthcare, finance, and smart devices. Healthcare has the greatest share, then finance, and afterwards smart devices.

Tabular Data: FL Applications

Application Domain	Percentage (%)
Healthcare	40
Finance	35
Smart Devices	25

Discussion

Challenges and Future Directions

Scalability

One of the biggest challenges of Federated Learning (FL) is the scalability. Contemporary FL systems have to coordinate millions of devices placed in decentralized settings with different computational capabilities and connection quality. Sustaining such a large ecosystem requires a tight coupling of communication protocols to reduce delay, learning algorithms to address resource use, and structures capable of handling heavy parallelism. Scalability must be a top priority in order to make FL systems remain efficient and coherent as more devices join the systems.

Heterogeneity

One major challenge in FL systems is heterogeneity, as the data at clients and the client devices themselves may differ greatly. The data collected from various devices is not identically distributed, thus it brings a pull of bias that challenges the development of global models. Moreover, devices span the spectrum of computing in terms of memory and network capabilities, from constrained devices to highly capable ones. This results in an unbalanced learning, which requires the use of adjustive measures to achieve a fair contribution amongst the clients as well as fair and productive distribution of work.

Regulation Compliance

Another important issue here associated with FL systems is the ability to follow data protection laws between different legal systems. Policies and laws like the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA) in the United States have constraints regarding data utilisation and storage. In their training processes, FL systems must be able to meet these laws' requirements in addition to being transparent and auditable. This is even more daunting, especially within industries such as healthcare and the financial industry, where the regulatory requirements are very stringent for any organization that fails to meet them and then faces severe consequences.

Future Research Directions

To address these issues, other future work should aim at creating new combined strategies which incorporate both privacy protection and improved efficiency. For instance, the conjoint use of differential privacy and model compression can produce high levels of security with low costs in terms of computation and communication. Self-assessment frameworks are also needed to perform similar comparisons across various applications and FL settings systematically to assess FL systems objectively. These frameworks can determine different areas that take a long time to perform and set fundamental criteria for FL.

Furthermore, incorporating FL with new generation technologies provides constructive solutions. The presented FL approach based on blockchain has the potential to more securely and transparently manage model updates and aggregation than using traditional methodologies. Quantum computing has the potential to improve computational capabilities and introduce enhanced security measures for data protection. With these innovations, FL systems should be able to address emerging challenges, making them more suitable for wide deployment across diverse domains requiring scalability, flexibility, and security.

Conclusion

Improving privacy and reducing computation overhead are critical to the achievement of FL's potential in healthcare, financial, and intelligent devices' applications. These regions require strong measures to help protect data from unauthorized access, meet legal requirements, and ensure optimal resource use, where privacy preservation and methodical performance enhancement remain crucial. Sophisticated approaches, including differential privacy, provide tremendous data security, and other techniques like model compression and adaptive optimization make computation and communication faster and efficient. Altogether, these approaches tackle the critical issues associated with FL so as to facilitate the functionality of large FL designs.

This study aims to present these techniques and evaluate their relevance and impact on enhancing FL systems. The analysis of the transitional approaches and the implementation of other applicable new technologies also contribute to future developments in FL, making it a foundation for effective protected collaboration in learning and other fields.

Acknowledgments

This study is a part of a PhD thesis. Acknowledgment is bestowed on honorable supervisors and examiners.

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 308–318. https://doi.org/10.1145/2976749.2978318

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & van Overveldt, T. (2019). Towards federated learning at scale: System design. *Proceedings of the 2nd SysML Conference*. Retrieved from https://arxiv.org/abs/1902.01046

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., ... & Seth, K. (2017). Practical secure aggregation for federated learning on user-held data. *Advances in Neural Information Processing Systems (NeurIPS), 30*, 1–12. Retrieved from https://arxiv.org/abs/1708.06689

Caldas, S., Konečný, J., McMahan, H. B., & Talwalkar, A. (2018). Expanding the reach of federated learning by reducing client resource requirements. *arXiv preprint*. Retrieved from https://arxiv.org/abs/1812.07210

Chen, Y., Xie, Y., Kairouz, P., & Song, L. (2020). Understanding model averaging in federated learning with structured optimization. *Advances in Neural Information Processing Systems (NeurIPS), 33*, 1–11. Retrieved from https://arxiv.org/abs/2003.10417

Geyer, R. C., Klein, T., & Nabi, M. (2017). Differentially private federated learning: A client level perspective. *Advances in Neural Information Processing Systems (NeurIPS) Workshop*. Retrieved from https://arxiv.org/abs/1712.07557

Hardy, S., Berkovsky, S., Penschuck, M., Germanakos, P., & Bellotti, F. (2020). Federated learning for user modeling. *Proceedings of the 28th ACM Conference on User Modeling, Adaptation, and Personalization (UMAP)*, 1–11. https://doi.org/10.1145/3340631.3394868

Hynes, N., Dao, D., Yan, C., Cheng, F., Song, D., & Popa, R. A. (2018). A demonstration of Sterling: A privacy-preserving data marketplace. *Proceedings of the VLDB Endowment, 11*(12), 2086–2089. https://doi.org/10.14778/3229863.3236267

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning, 14*(1–2), 1–210. https://doi.org/10.1561/2200000083

Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. *arXiv preprint*. Retrieved from https://arxiv.org/abs/1610.05492

Li, Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., & He, B. (2019). A survey on federated learning systems: Vision, hype, and reality for data privacy and protection. *arXiv preprint*. Retrieved from https://arxiv.org/abs/1902.04885

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine, 37*(3), 50–60. https://doi.org/10.1109/MSP.2020.2975749

Li, X., Huang, K., Yang, W., Wang, S., & Zhang, Z. (2020). On the convergence of FedAvg on non-IID data. *International Conference on Learning Representations (ICLR)*. Retrieved from https://arxiv.org/abs/1907.02189

Luo, C., Zhang, F., Zhou, Y., & Huang, L. (2021). Privacy-preserving federated learning for speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 29*, 1495–1507. https://doi.org/10.1109/TASLP.2021.3050296

Mangal, A., Garg, H., & Bhatnagar, C. (2023). Assessing the performance of saliency detection method using various deep neural networks. *Journal of Information Technology Management, 15*(Special Issue), 23–34. https://doi.org/10.22059/jitm.2023.95243

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. Y. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 54*, 1273–1282. Retrieved from https://arxiv.org/abs/1602.05629

Mohri, M., Sivek, G., & Suresh, A. T. (2019). Agnostic federated learning. *International Conference on Machine Learning (ICML)*, 4661–4670. Retrieved from https://arxiv.org/abs/1902.00146

Nguyen, D. C., Ding, M., Pathirana, P. N., & Seneviratne, A. (2021). Federated learning for Internet of Things: A comprehensive survey. *IEEE Communications Surveys & Tutorials, 23*(3), 1622–1658. https://doi.org/10.1109/COMST.2021.3075439

Reisizadeh, A., Mokhtari, A., Hassani, H., Jadbabaie, A., & Pedarsani, R. (2020). FedPAQ: A communication-efficient federated learning method with periodic averaging and quantization. *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020, 2021–2029. Retrieved from https://arxiv.org/abs/1909.06335

Sattler, F., Wiedemann, S., Müller, K.-R., & Samek, W. (2019). Robust and communication-efficient federated learning from non-i.i.d. data. *IEEE Transactions on Neural Networks and Learning Systems, 31*(9), 3400–3413. https://doi.org/10.1109/TNNLS.2019.2944482

Shokri, R., & Shmatikov, V. (2015). Privacy-preserving deep learning. *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 1310–1321. https://doi.org/10.1145/2810103.2813687

Smith, V., Chiang, C.-K., Sanjabi, M., & Talwalkar, A. (2017). Federated multi-task learning. *Advances in Neural Information Processing Systems (NeurIPS), 30*, 4424–4434. Retrieved from https://arxiv.org/abs/1705.10467

Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Weber, B., & Pasupuleti, V. (2019). A hybrid approach to privacy-preserving federated learning. *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security (AISec)*, 1–11. https://doi.org/10.1145/3338501.3357370

Wang, H., Yurochkin, M., Sun, Y., Papailiopoulos, D., & Khazaeni, Y. (2020). Federated learning with matched averaging. *International Conference on Learning Representations (ICLR)*. Retrieved from https://arxiv.org/abs/2002.06440

Wang, J., Zhao, Y., Yao, Q., Kwok, J. T., & Ni, L. M. (2019). Federated learning with matched averaging. *International Joint Conferences on Artificial Intelligence (IJCAI)*, 5050–5056. https://doi.org/10.24963/ijcai.2019/703

Yang, Q., Liu, Y., Cheng, Y., Kang, Y., Chen, T., & Yu, H. (2019). Federated learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning, 13*(3), 1–207. https://doi.org/10.2200/S00960ED1V01Y201910AIM043

Yang, T., Andrew, G., Eichner, H., Sun, H., Li, W., Kong, N., ... & McMahan, H. B. (2018). Applied federated learning: Improving Google keyboard query suggestions. *arXiv preprint*. Retrieved from https://arxiv.org/abs/1812.02903

Zhang, Z., Liang, P. P., Zhang, C., & Gitter, A. (2021). Federated multi-modal learning with missing data. *Proceedings of the 38th International Conference on Machine Learning (ICML)*, 2021. Retrieved from https://arxiv.org/abs/2009.06186

Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D., & Chandra, V. (2018). Federated learning with non-iid data. *arXiv preprint*. Retrieved from https://arxiv.org/abs/1806.00582

Zhu, H., Zhou, J., & Xu, S. (2021). Data-free knowledge distillation for heterogeneous federated learning. *Proceedings of the 39th International Conference on Learning Representations (ICLR)*. Retrieved from https://arxiv.org/abs/2012.05387

Zhu, L., Liu, Z., & Han, S. (2019). Deep leakage from gradients. *Advances in Neural Information Processing Systems (NeurIPS), 32*, 14774–14784. Retrieved from https://arxiv.org/abs/1906.08935

Bibliographic information of this paper for citing:

Shukla, Ravi Shankar (2025). Enhancing Privacy and Efficiency Techniques in Federated Learning Systems: Applications in Healthcare, Finance, and Smart Devices. *Journal of Information Technology Management*, 17 (Special Issue), 45-62. https://doi.org/10.22059/jitm.2025.102921