WCRG Global Research
WORLD COMET RESEARCH GROUP

International Journal of Applied Data Science & Modern Computing
Vol. 1, No. 1, 2026
Doi: 10.XXXX/XXXXXXXX/IJADSMC-V1I1P101
PP. 01-13

Original Article

A Study on Federated Learning Techniques for Privacy Preservation

Dr. Aarav Mehta

Department of Computer Science and Engineering, Orion Institute of Management and Technology Vadodara, India.

Received: 30-11-2025 Revised: 25-12-2025 Accepted: 28-12-2025 Published: 02-01-2026

ABSTRACT

The fast proliferation of data-driven applications and intelligent systems has raised problem awareness as far as the concerns regarding data privacy, data security, and regulatory compliance are concerned to a very high level. The conventional centralized machine learning models demand the coalescence of raw data situated in disseminated sources, which is a grave threat of information leaks, unauthorized data accessibility, and breaking a privacy policy like GDPR and HIPAA. Federated Learning (FL) has become a hopeful decentralized learning framework so as to facilitate joint model training among various customers without relocating crude data to a central point. Rather, model updates of the local models are shared and aggregated, hence retaining locality of data and improving privacy. This paper will provide an in-depth analysis of federated learning methods and pay special attention to the issue of privacy. Its research paper investigates the very principles of federated learning, architecture designs, communication scheme, and ways of aggregation. The diverse privacy-enhancing schemes are secure aggregation, differential privacy, homomorphic encryption, trusted execution environments, and are critically analyzed. Moreover, this article examines new developments, concerns and tradeoffs connected with privacy, effectiveness of communication, model noise, and scalability of systems. There is a comparative analysis of federated learning methods presented in organized tabular form and mathematical equations. The existing studies are discussed with their experimental results in order to outline the effectiveness of federated learning to maintain the privacy and achieve the acceptable model accuracy. Lastly, the research issues and future paths are outlined in order to direct the further development of the privacy-preserving federated learning systems.

KEYWORDS

Federated Learning, Privacy Preservation, Distributed Machine Learning; Secure Aggregation, Differential Privacy, and Data Security.

ISSN: XXXX-XXXX
PP: 01-13.

1. INTRODUCTION

1.1. Background

The previously mentioned rapid development of modern machine learning (ML) and deep learning methods has largely been fueled by the existence of high quality volumes of high quality data. Conventional centralised learning models are based on the aggregate of data of various sources to centralised cloud servers, where the model training and model inference is carried out through computationally intensive methods. Though such strategy has realised state-of-the-art performance in a broad area of applications, serious privacy and security issues are introduced. This data storage centralization is the risk of the data breach and unauthorized access, as well as misuse, especially in the case of working with sensitive information, e.g., medical records, financial transactions, location tracks, user data. The rising trend of imposing stricter data protection laws and the rising awareness among users of the privacy of their data in recent years has shed light on the shortcomings of centralized machine learning models. Such regulations as the requirement to minimize data and the consented of users so that only necessary data is obtained and directly shared make traditional methods inapplicable in a variety of the practical setting. Consequently, there is a high demand in the learning frameworks capable of deriving applicable insights in the distributed data sources without violating individual privacy. This has inspired the development of decentralized and privacy conscious learning models that will minimize data disclosures without compromising model utility. A combination of localization of data and collaboration in shared model updates, these techniques promise a good balance between performance and privacy. These paradigms are not only compliant with the regulatory and ethical mandates and enable development of trust between the data owners but also initiate the development of scalable and secure intelligent systems in privacy-sensitive domains.

1.2. Importance of Study on Federated Learning Techniques

Fig 1 - Importance of Study on Federated Learning Techniques

1.2.1. Privacy Preservation in Distributed Data Environments

The key idea about the federated learning techniques is that they help to maintain data privacy in the distributed environment, which is one of the main motives to study this technique. As opposed to the conventional centralized machine learning, federated learning allows to train models without having to transfer raw data to local devices or organizations. It is especially essential in the sphere of healthcare, finance, and smart devices, where legally, ethically, and security considerations do not allow sharing sensitive information freely. A closer examination of the techniques of federated

ISSN: XXXX-XXXX
PP: 01-13.

learning assists in the discovery of efficient mechanisms of reducing the leakage of privacy with at the same time supporting collaborative intelligence.

1.2.2. Compliance with Data Protection Regulations

The increased application of data protection laws has resulted in privacy-conscious learning systems becoming an essential and no longer a choice. The concept of federated learning complies with the regulatory principles of data minimalization, user consent, and locality of data processing. The analysis of the federated learning techniques would be significant in order to know how such systems can be configured to correspond to the compliance criteria and how these systems would be configured to ensure the performance and scalability as acceptable.

1.2.3. Balancing Privacy, Accuracy, and Efficiency

Federated learning presents new trade-offs concerned with model accuracy, privacy guarantees, and system efficiency. Secure aggregation technique, the concept of differential privacy, and encryption are methods that can improve privacy but might create a significant burden on computation and communications. These techniques need to be studied to observe their implication on learning performance and to come up with optimism in designing solutions that do not compromise on protection of privacy and practical deployment limitations.

1.2.4. Scalability and Real-World Applicability

With the growing use of federated learning in practical settings with a significant number of millions of devices, it is important to learn how it can be scaled. Such factors like heterogeneity of clients, untrustworthy connection, and uneven distribution of data can have serious impact on systems work. The studies of federated learning methods contribute to the solution of these problems and contribute to the creation of efficient and scalable systems that can be implemented on large scale.

1.2.5. Foundation for Future Privacy-Aware AI Systems

Lastly, the research on federated learning methods gives a robust basis of the future of privacy-conscious artificial intelligence. With the increasing interest in the issue of data ownership and ethical AI, federated learning will become a key to collaborative data analytics and intelligent decision-making. Extensive knowledge of its methods makes it possible to build secure, trust-magnifying, and sustainable AI systems to be used in the next generation.

1.3. Limitations of Centralized Learning

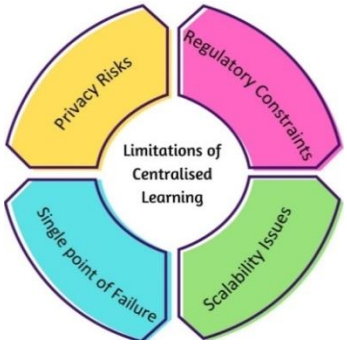

Fig 2 - Limitations of Centralized Learning

ISSN: XXXX-XXXX
PP: 01-13.

1.3.1. Privacy Risks

The centralized machine learning solutions demand that all raw data of various sources should be contained in one repository that poses high chances of privacy violations. Personal identifiers, medical history or financial data among other sensitive information may suffer unauthorized hacking, cybercrimes, and insider attacks. Despite a high level of security, gathering large quantities of data in a single location becomes an easy target of the attacker and leads to the problem of misuse of data and loss of user confidence.

1.3.2. Regulatory Constraints

Data processing and collection that is centralized with privacy regulations further complicated the process of adhering to them. These data protection standards usually have very high standards on the way data is stored, processed and cross-border transferred which are not easy to apply to centralized systems. Companies have to cope with consent, data retention, and access control on a large scale that adds legal and administrative workload. Aviation can cause heavy fines and loss of reputation in case of failure to adhere to such rules.

1.3.3. Scalability Issues

With the increasing amount of the data and the number of users, centralized servers are highly scaled down. The computational requirements and infrastructure costs are enormous to store the large datasets and also in training models on a large scale. Network bandwidth and higher latency also undermine the performance of the system, and extending centralized learning systems to real-time or large-scale the system is challenging.

1.3.4. Single Point of Failure

With centralized architectures, a server or a data center play a massive role in centralized learning making it a highly important single point of failure. The whole process of learning may be stopped by hardware malfunctions, bugs or computer network outages, which threaten the availability of the entire system. Moreover, in case of successful attacks on the central server, complete breakdowns of the system may occur. It is not redundant, which leads to a decrease in the system reliability and resilience.

2. LITERATURE SURVEY

2.1. Early Developments in Federated Learning

Decentralized machine learning Federated Learning (FL) is proposed as a decentralized machine learning paradigm to resolve increasing privacy concerns on the growing numbers of distributed data setting, especially on mobile and edge computing. FL also allows simultaneous training of a shared, global model by multiple clients instead of raw-data collection on a single central server but carrying out local training and only sending model updates. Federated Averaging (FedAvg), one of the longest running and most widely used algorithms in this field, is an algorithm that compresses model parameters of the clients with weights based on the size of the local data set. This was shown to be the case with FedAvg because the competitive model performance could be reached without direct data sharing. Nevertheless, initial FL models relied predominantly on communication ability and convergence behavior as early FL frameworks based their assumptions largely on honest and non-adversarial participants with a lack of attention to privacy concerns.

ISSN: XXXX-XXXX
PP: 01-13.

2.2. Privacy Threat Models

The latter study showed that federated learning which does not share raw data is not immune to various privacy attacks. Model inversion attacks seek to reconstruct subtle information of training data using shared model parameters. Membership inference attacks attempt to assess the use of a certain data sample by a model in training, potentially uncovering sensitive information of participation. Also, gradient leakage attacks have demonstrated that gradient or model updates are reversible to reconstruct original training samples, especially in deep neural networks. These attack models indicate that the deployment of models can inadvertently embed personal information, which shows that federated learning is not a system that offers good privacy protection without further security measures.

2.3. Secure Aggregation Techniques

To reduce instances of information leakage by individual client updates, the secure aggregation protocols have been suggested as a core privacy-enhancing tool of federated learning. These methods guarantee that the central server is able to retrieve only the total amount of updates about the clients and not about each one of them separately. Cryptographic techniques like secret sharing and secure multi-party computation (SMPC) are usually used to attain this objective. Secure aggregation is achievable by sharing encrypted, or masked updates with many parties, so that the server does not get to know any client-specific information in the case it is curious or partially compromised. Although these techniques go a long way to protect privacy, they also add more computational overhead, communication overhead and system latency which may be limiting in large-scale federated systems.

2.4. Differential Privacy in Federated Learning

Differential privacy (DP) offers a mathematically sound framework of calculating and managing the privacy leaks in federated learning. Distributing gradients or model updates with randomly added noise carefully calibrated ensures that each data point can have an impact on the resulting model. The privacy guarantee is conceptualized as:

$$\mathbf{Pr[M(D)\in S]\leq e^{\epsilon} Pr[M(D')\in S]}$$

ϵ means that there will be enhanced privacy at the expense of decreased accuracy of models. In federated, both local DP and central DP can be used in the client level or in the server level after aggregation. Throughout, differential privacy is effective in protecting against inference attacks; however, when applicants produce high noise, this may jeopardize the learning outcomes, and therefore privacy-utility balancing is a major research problem.

2.5. Comparative Summary of Literature

As has been proven in the existing literature, various privacy-preserving methods employed in federated learning have varying trade-offs of privacy strength, model accuracy, and cost of communication. Although FedAvg is communication efficient and accurate, it has low protection of privacy. Secure aggregation maximizes privacy by concealing single updates at the expense of communication and computation. Differential privacy provides powerful theoretical guarantees at relatively low cost of communication, though it can decrease the accuracy asa result of noise

ISSN: XXXX-XXXX
PP: 01-13.

injection. Homomorphic encryption is encryption, but encodes data in such a way that it is possible to compute on the encrypted data; this gives privacy that is very high but the cost of computation and communication is very high. These trade-offs suggest that there is no universal single best solution and that hybrid solutions involving a combination of techniques may be needed in case of practical and privacy sensitive federated learning deployments.

3. METHODOLOGY

3.1. System Architecture

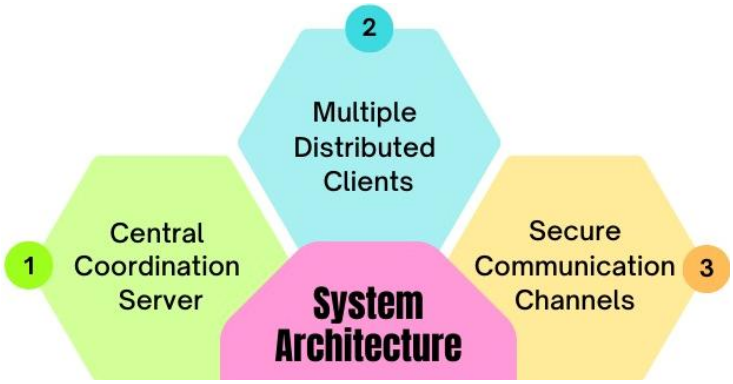

Fig 3 - System Architecture

3.1.1. Central Coordination Server

The federated learning of the process is orchestrated by the central coordination server. It is in charge of starting the global model, choosing participating clients in every training round and combining the received model updates by the clients. The server does not need to be exposed to raw client data as it will selectively act based on the model parameters or gradients, and in general fewer direct exposures to sensitive information occurs. The server further assumes the training schedules, convergence monitoring, and global model distribution, but is not aware of specific datasets of clients.

3.1.2. Multiple Distributed Clients

Federated learning model has various distributed clients, which can include mobile gadgets, edge nodes, or organizational servers with local data and possibly sensitive information. These customers undergo on-their-own model training using their own data and calculate local updates according to the global model parameters offered by the server. This decentralized paradigm of training is more confidential because the data does not move out of the client environment since it does not affect data locality. Nevertheless, multiple heterogeneities of clients in availability of data, computer strength, and network access provide difficulties to effective and equitable model training.

3.1.3. Secure Communication Channels

It must provide a secure channel of communication between the server and the clients in case of exchanging model updates and control messages. To discourage the occurrence of eavesdropping, tampering and replay attacks, there is the use of encryption, authentication procedures, and safe transmission criteria in the communication process. Such channels provide confidentiality and integrity of model updates along transit, and thus complement other privacy-preserving measures

ISSN: XXXX-XXXX
PP: 01-13.

like secure aggregation and differential privacy. Strong communication protection is essential to the preservation of trusts and reliability in federated learning systems that use untrusted networks.

3.2. Federated Training Workflow

3.2.1. Server Initializes the Global Model

Federated learning process entails the initiation of a global model by the central server, whether randomly or with a pre-trained baseline. The first model outlines the common learning goal and gives all the clients a point to start. Other training parameters that the server triggers include learning rate, the number of communication rounds, or participation of the clients to guarantee regular and controlled model update.

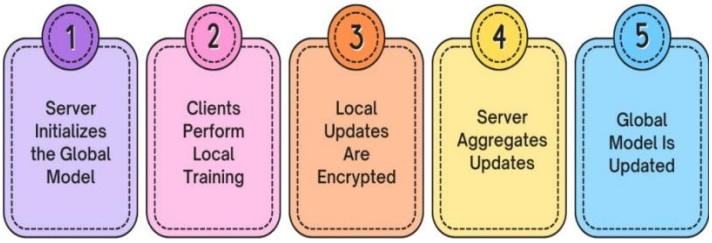

Fig 4 - Privacy-Preserving Mechanisms

3.2.2. Model Is Broadcast to Selected Clients

The server in each round of training picks a subset of potential clients using a specified criteria which could be availability, computational ability or random choice. The available global model parameters are then safely disseminated to the targeted clients. This selective participation process enhances scalability of it and reduces overhead in communication as well as preserving representative learning even in scattered data sources.

3.2.3. Clients Perform Local Training

On being provided with the global model, every client trains a local training on its own dataset. This generally requires running several epochs of stochastic gradient descent or other optimization methods. Local training enables the clients to localise the global model in their data distribution whilst ensuring that the raw data is confined within the client environment as, therefore, maintaining data privacy.

3.2.4. Local Updates Are Encrypted

Clients can produce model updates (e.g. gradients or new weights) after local training and transmit this encrypted. Encryption ensures that no other entity such as potentially interested servers or network enemies can access specific client updates. This is important to safeguard sensitive data that is stored in model updates and facilitates privacy enhancing aggregation protocols.

3.2.5. Server Aggregates Updates

The central server gathers the encrypted changes reported by the participating clients and it does the aggregation but it does not access any contribution. The method of aggregation is by weighted averaging of updates, or secure aggregation protocols, which are the techniques used to

ISSN: XXXX-XXXX
PP: 01-13.

combine updates into one global update. This procedure assures proper sensitivity of the impact of every client and the privacy of the local model modifications.

3.2.6. Global Model Is Updated

Lastly, the server changes the global model with the aggregated result and prepares it to be used by the subsequent round of training. The revised model reflects the overall experience acquired among the distributed clients and is refined on several rounds, after which convergence is reached. This is a continuous cyclical process until the achievement of defined performance requirements or training boundaries.

3.3. Privacy-Preserving Mechanisms

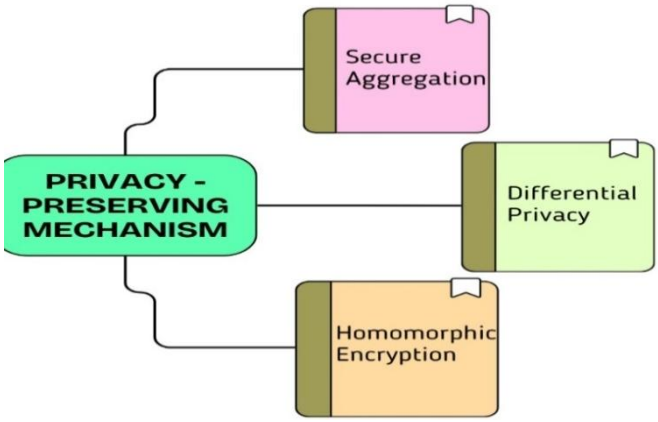

Fig 5 - Privacy-Preserving Mechanisms

3.3.1. Secure Aggregation

Secure aggregation will make sure that individual verdicts of clients do not reach the central server due to the encryption or masking of model updates before transmission. The client is coded such that the local update of each client will be encoded in a manner that only the aggregated result can be recovered by the server upon aggregating the updates sent by more than one client. This does not allow the server to monitor or assume any information based on the contribution of any one of the clients, even when the server is an honest-but-curious. Privacy This method can be used to securely aggregate the data obtained in federated learning and thus protect it. The cryptographic operation used in the secure aggregation contributes to more computational and communication overhead.

3.3.2. Differential Privacy

Differential privacy gives explicit guarantees against information leakage by it introducing random noise very carefully tuned to model updates or gradients. The set of hotspot generation mechanisms typically involves the use of noise created by the Gaussian or Laplace mechanisms that prevents the impact of a single data point on the final model. Diffidential privacy allows a privacy-accuracy trade-off to be tuned by fixing the privacy budget parameter. This method is useful to overcome inference and reconstruction attacks, but when there is excess noise it may impair the results of the learning.

ISSN: XXXX-XXXX
PP: 01-13.

3.3.3. Homomorphic Encryption

In homomorphic encryption, mathematical operations can be executed without the need to decrypt the encrypted data directly on the encrypted data. In federated learning, it allows the server to sum encrypted update in the model and maintain full confidentiality of an individual contribution. Consequently, sensitive information is guarded all through the training process including in the case of the aggregation. Although it has even the most powerful privacy guarantees, homomorphic encryption is computationally intense and may increase training time and communication expenses considerably, which restricts its use in practice in large-scale deployments.

3.4. Mathematical Formulation of Aggregation

The mathematical technique of combining models learnt locally into a global model is a fundamental operation of federated learning that facilitates collaborative learning without central data collection. At each training round is calculated by adding the contributions of all the participating clients, in which contribution of each client is weighted by the size of its local dataset. The weighting is that the more the data one client has, the larger a role in the global model it will have, which enhances the efficiency of the statistics and stability of convergence.More precisely, aggregation process entails scaling all the client model parameters by the fraction of local data in proportion to overall data size amongst all the involved clients. These weighted parameters are then added by the server to constitute the new global model. The formulation makes the assumption that every client is learning the same architecture and updates at the location are computed using the same optimization strategies. The aggregation mechanism promotes frequency of communication reduction and scalability in large distributed systems by using the parameter averaging to share gradient as opposed to gradient sharing. Privacy From an privacy point of view, this mathematical formulation is usually paired with secure aggregation or encryption methods where the server only gets access to the aggregated output and not at individual model parameters. Besides, the formulation can withstand non-identically distributed data among clients, which is a typical feature in federated learning settings. Altogether, weighted model aggregation is an efficient solution balancing between the accuracy of learning, the efficiency of communication, and the preservation of privacy and is an essential part of contemporary federated learning systems.

4. RESULTS AND DISCUSSION

4.1. Performance Evaluation Metrics

The measurement of the effectiveness of a federated learning system must have a well-defined set of performance measures that reflect not only the predictive performance, but also the privacy and the efficiency of the system. One of the measures used to determine the predictive power of the trained global model on unseen test data concerns the accuracy of the model. It is a measure of the continuity of learning quality of the federated learning method with centralized training. A high degree of accuracy implies that knowledge is properly aggregated between the distributed clients even when there are non-uniform distributions of data and minimal communication. The privacy - budget is another important measure and it is represented by. The value of ϵ that is used to measure the level of privacy protection under the effect of the usage of differential privacy mechanisms. The closer the value of ϵ, the more privacy, because epsilon constrained the differentiation between neighbouring datasets. Usually, it is the case that to balance ϵ, more noise in model updates is needed, and this will adversely affect model accuracy. As such, it is crucial to have a balance between

ISSN: XXXX-XXXX
PP: 01-13.

the privacy budget and accuracy to know about the privacy utility trade-off of the privacy-preserving federation-based learning system. Overhead Communication overhead is used to measure the interactions of total data that are exchanged between the central server and the involved clients during training. This extends to a relaying of world-models, international updates of encrypted local-updates and control messages of several communication rounds. Federation Communication performance in a federated environment is especially relevant when using bandwidth-limited or resource-constrained nodes, e.g. mobile and edge nodes. The reduced communication overhead enhances the scalability and training latency thereby making the system more realistic with real life applications. Lastly, the convergence rate of the training measures the speed at which the global model achieves some stable or optimal level of performance using repeated training. The quicker convergence means the efficient aggregation and optimization, and the minimization of the computational and communication costs. Collectively, these metrics offer a comprehensive system of evaluation of the effectiveness, efficiency and privacy strength of federated learning systems.

4.2. Comparative Results

Table 1: Comparative Results

Method	Accuracy (%)	Privacy Guarantee (%)	Overhead (%)
Centralized ML	95.20%	0%	20%
FL (FedAvg)	93.80%	40%	50%
FL + Differential Privacy	91.50%	80%	55%
FL + Secure Aggregation	93.10%	85%	75%

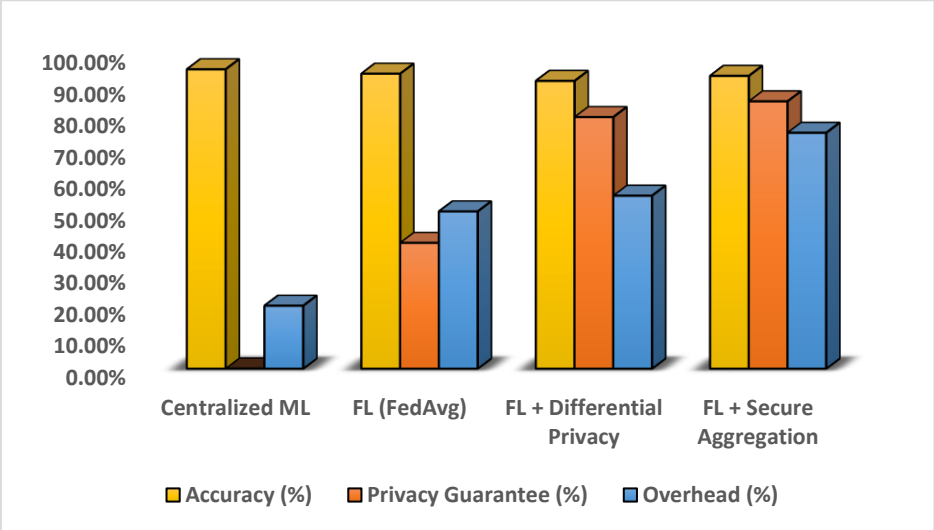

Fig 6 - Graph representing Comparative Results

4.2.1. Centralized Machine Learning

The highest accuracy of 95.2 is realized under centralized machine learning because the entire dataset gets to be accessed at the same location. It does not protect privacy, however, because the entire raw data have to be gathered and stored in the central place and this also enhances the chances of data breach and misuse. System overhead is also low at 20% since training and communication are

ISSN: XXXX-XXXX
PP: 01-13.

done within the same infrastructure without the necessity of encryption and distributed coordination.

4.2.2. Federated Learning (FedAvg)

The FedAvg algorithm of federated learning attains an accuracy of 93.8, which is competitive with centralized learning and at the same time the data remains decentralized. It offers a mid-range level of privacy of 40% since it does not allow sharing of raw data but still updates to a model may expose information. It is 50 percent since overhead costs consist of the repetitive communication between clients andaiter and distributed coordination during training.

4.2.3. Federated Learning with Differential Privacy

Federated learning offers better privacy protection with incorporation of the concept of differential privacy that guarantees protection of privacy at 80 percent. This method restricts the information leakage by individual clients by introducing noise into any model update. Nonetheless, the noise gives a reduced model accuracy of 91.5% which is the privacy utility trade-off. The overhead is also quite moderate with 55 per cent, since the difference in privacy does not boost the extra communication expenses significantly, and on the contrary, the local computation cost is increased.

4.2.4. Federated Learning with Secure Aggregation

Secure aggregation in federated learning offers the best privacy protection of all the considered approaches, and the privacy is ensured at 85 percent. Client updates are encrypted individually making the server unable to retrieve individual updates. This approach is relatively accurate with a high accuracy of 93.1, albeit with high overhead of 75 per cent of cryptographic keys and complex communication. Although costly, secure aggregation would be effective in the context of highly privacy sensitive applications.

4.3. Discussion

The relative outcomes clearly show that federated learning provides a persuasive trade-off in terms of data privacy as well as model performance in comparison to the conventional centralized machine learning methods. The federated learning method inherently lowers the risk of exposing data as the sensitive data is kept locally on client machines by keeping the data localized and yet being able to compete with the same level of accuracy. Though the performance would slightly be worse than with centralized learning, the loss is minimal and in many cases can be tolerated in privacy-sensitive contexts like healthcare, financial and mobile analytics. This brings into an emphasis the concept of federated learning as an effective feasible replacement to the collaborative learning model when data sharing is limited. Privacy-enhancing techniques are integrated with federated learning, which can entail a set of significant trade-offs to be effectively handled. Differential privacy offers considered and rigorous protection against privacy by capping the impact of individual items of data on the worldwide model. Nevertheless, the noise that is introduced to gradients or model updates is always bound to have an effect on the quality of learning which is clearly being decreased. This trade-off is even more exaggerated as tightening privacy budgets are produced that system designers have to trade regulatory needs against performance expectations. Secure aggregation, conversely, is concentrated on the safety of client updates on transmission and

ISSN: XXXX-XXXX
PP: 01-13.

aggregation without modifying the learning signal. Consequently, it maintains the accuracy of models better than the case of differential privacy. Nonetheless, such an asset is associated with the expenses of higher computational and communication costs in terms of cryptographic functions and communication among clients. This overhead has the potential to cause scalability, especially with high-scale federated networks over limited bandwidth or resource-constrained devices. In general, the discussion shows that there is no privacy mechanism that would be the best across the board. These two issues are differential privacy and secure aggregation, which is determined by the needs of an application, the threat, and the constraints of the system. Hybrid methods which will encompass several techniques could provide a path promising high privacy protection and at the same time a positive result of federated learning performance.

5. CONCLUSION

This work has provided an in-depth analysis of federated learning as a potentially valuable paradigm of privacy-preserving machine learning in the distributed world. Federated learning meets basic privacy requirements related to the centralized collection of data by allowing joint training of models without necessarily sharing raw data. This decentralized solution is one of the most important elements that will minimize the threat of information breaches and unauthorized access and, therefore, will be especially applicable to sensitive sectors including healthcare, finance, and the Internet-of-Things ecosystems. Moreover, the incorporation of cryptographic schemes, including secure aggregation and homomorphic encryption, and statistical privacy schemes, including differential privacy, offers a series of layers of protection against inference assaults, model inversion, and information reconstruction attacks. Although with such strengths, there are limitations with federated learning. The ability to deal with system heterogeneity is one of the main issues as the clients, involved into such processes, tend to vary in terms of computational power, storage, data distribution, and network connection. This heterogeneity may cause biased updates on the model, slower convergence, and overall poor performance. Communication efficiency is another bottleneck issue as the need to re-communicate model parameters to many training rounds could be very expensive, both in terms of bandwidth and latency, particularly in big-scale or resource-limited environments. Between, even though federated learning presupposes honesty among the participants of a particular system under most conditions, the existence of malicious or compromised clients poses a threat, including poisoning attacks and model manipulation, thereby deteriorating the model accuracy and privacy guarantees. In the perspective, future work ought to be based on creating adaptive and context-sensitive privacy controls that adjust privacy protection and model utility in accordance with the needs of application and the level of threat. Cryptographic protocols which are lightweight, and minimize computational and communication costs will be necessary to enhance scalability and applications in real time. Additionally, the research into strong aggregation strategies and anomaly detection is needed in order to protect against evil clients and integrity of the system. However, in the real world, the scale of federated learning systems can only be deployed and evaluated with a joint effort of the academia and the industry. Federated learning will have a crucial role in the future of privacy-friendly artificial intelligence and joint data analytics as data privacy regulations increase.

ISSN: XXXX-XXXX
PP: 01-13.

REFERENCES

[1] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). *Communication-efficient learning of deep networks from decentralized data*. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS).

[2] Kairouz, P., et al. (2021). *Advances and open problems in federated learning*. Foundations and Trends® in Machine Learning, 14(1–2), 1–210.

[3] Bonawitz, K., et al. (2017). *Practical secure aggregation for privacy-preserving machine learning*. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS).

[4] Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). *Calibrating noise to sensitivity in private data analysis*. Theory of Cryptography Conference (TCC).

[5] Dwork, C., & Roth, A. (2014). *The algorithmic foundations of differential privacy*. Foundations and Trends® in Theoretical Computer Science.

[6] Fredrikson, M., Jha, S., & Ristenpart, T. (2015). *Model inversion attacks that exploit confidence information and basic countermeasures*. Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security.

[7] Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). *Membership inference attacks against machine learning models*. IEEE Symposium on Security and Privacy.

[8] Zhu, L., Liu, Z., & Han, S. (2019). *Deep leakage from gradients*. Advances in Neural Information Processing Systems (NeurIPS).

[9] Geiping, J., Bauermeister, H., Drozdzal, M., & Moeller, M. (2020). *Inverting gradients – how easy is it to break privacy in federated learning?* Advances in Neural Information Processing Systems (NeurIPS).

[10] Geyer, R. C., Klein, T., & Nabi, M. (2017). *Differentially private federated learning: A client level perspective*. NIPS Workshop on Machine Learning on the Phone and other Consumer Devices.

[11] Abadi, M., et al. (2016). *Deep learning with differential privacy*. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security.

[12] Truex, S., et al. (2019). *A hybrid approach to privacy-preserving federated learning*. Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security.

[13] Acar, A., et al. (2018). *A survey on homomorphic encryption schemes: Theory and implementation*. ACM Computing Surveys, 51(4).

[14] Hardy, S., et al. (2017). *Private federated learning on vertically partitioned data via entity resolution and additive secret sharing*. Proceedings of the VLDB Endowment.

[15] 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.

[16] Tirumalasetty, P. (2022). Coded Machine Unlearning using Machine Learning.