Received: 21 December 2023 | Revised: 12 January 2024 | Accepted: 16 January 2024

DOI: 10.1002/spy2.374

RESEARCH ARTICLE

WILEY

Federated learning with hybrid differential privacy for secure and reliable cross-IoT platform knowledge sharing

Oshamah Ibrahim Khalaf[1] | **Ashokkumar S.R**[2] | **Sameer Algburi**[3] | **Anupallavi S**[4] | **Dhanasekaran Selvaraj**[5] | **Mhd Saeed Sharif**[6] | **Wael Elmedany**[7]

[1]Department of Solar, Al-Nahrain Research Center for Renewable Energy, Al-Nahrain University, Jadriya, Baghdad, Iraq

[2]Department of Computer and Communication Engineering, Sri Eshwar College of Engineering, Coimbatore, India

[3]Al-Kitab University, Kirkuk, Iraq

[4]Artificial Intelligence & Machine Learning, Acharya Institute of Technology, Bengaluru, India

[5]Department of Electronics and Communication Engineering, Sri Eshwar College of Engineering, Coimbatore, India

[6]Intelligent Technologies Research Group, Computer Science and DT, ACE, UEL University, London, UK

[7]College of Information Technology, University of Bahrain, Zallaq, Bahrain

Correspondence
Mhd Saeed Sharif, Intelligent Technologies Research Group, Computer Science and DT, ACE, UEL University, London, UK.
Email: s.sharif@uel.ac.uk

Abstract

The federated learning has gained prominent attention as a collaborative machine learning method, allowing multiple users to jointly train a shared model without directly exchanging raw data. This research addresses the fundamental challenge of balancing data privacy and utility in distributed learning by introducing an innovative hybrid methodology fusing differential privacy with federated learning (HDP-FL). Through meticulous experimentation on EMNIST and CIFAR-10 data sets, this hybrid approach yields substantial advancements, showcasing a noteworthy 4.22% and up to 9.39% enhancement in model accuracy for EMNIST and CIFAR-10, respectively, compared to conventional federated learning methods. Our adjustments to parameters highlighted how noise impacts privacy, showcasing the effectiveness of our hybrid DP approach in striking a balance between privacy and accuracy. Assessments across diverse FL techniques and client counts emphasized this trade-off, particularly in non-IID data settings, where our hybrid method effectively countered accuracy declines. Comparative analyses against standard machine learning and state-of-the-art FL approaches consistently showcased the superiority of our proposed model, achieving impressive accuracies of 96.29% for EMNIST and 82.88% for CIFAR-10. These insights offer a strategic approach to securely collaborate and share knowledge among IoT devices without compromising data privacy, ensuring efficient and reliable learning mechanisms across decentralized networks.

KEYWORDS

differential privacy, dynamic allocation, federated learning, privacy budget

1 | INTRODUCTION

The exponential rise of data and the connection of items through networked sensors have driven recent advances in the Internet of Things (IoT). The interchange and use of enormous volumes of data become essential for optimizing intelligent services and goods as the IoT expands. The underlying problem, however, comes in protecting the privacy and

security of user data while capitalizing on the potential insights obtained from these linked devices.[1] An analysis of parameter variations between training and recording, including neural network training weights, has revealed that privacy can still be compromised to some extent. Furthermore, the need for repetitive uploading and compilation of parameters using a global model during federated learning poses a risk of exposing sensitive data, making it susceptible to model inversion attacks[2] that compromise personal information. In response to these formidable challenges, a new frontier has emerged—Big Data Security, where data security engineers specialize in safeguarding data against cyber threats, be they external or originating within an organization. In this paper, federated learning (FL) which enables numerous entities to collaboratively train a model using their individual data without the need to exchange raw data, while facilitating cooperative model training across many IoT platforms. This study focuses on FL, an approach that enables numerous entities to foster cooperative model training across various IoT platforms. Additionally, we delve into the concept of differential privacy, a privacy-protection technique that obscures query responses to conceal specific data points and enables safe knowledge sharing while maintaining data security and reliability across the cross-IoT infrastructure.

With the advancement in machine learning FL (FL),[3] is revolutionized since it allows numerous clients to collaborate on model training while maintaining data privacy through localized training. This approach efficiently lowers the expenses and privacy hazards associated with centralized machine learning systems. By enabling several organizations to work together on model improvement without exchanging sensitive data, FL radically changes the machine learning environment. FL protects data privacy and advances machine learning capabilities by concentrating on training algorithms on distinct local data sets owned by smart devices or platforms.

Differential privacy (DP) stands as a fundamental machine learning approach designed to safeguard data while protecting individual privacy. Its strength is in ensuring anonymity in outputs, protecting user-specific data, and enhancing privacy without relying on trust assumptions. FL and DP have gained traction in the recent years, especially in the IoT and healthcare industries. These combined strategies attempt to create a compromise between safeguarding privacy and training efficacy. Still, protecting model privacy without sacrificing accuracy is a recurring problem. Earlier attempts to apply DP to FL frequently involved adding noise to the training process and relying on a consistent degree of trust in the central server.

Here, two questions come to consideration: How may DP protections in FL systems be adjusted to account for the various data properties across participating clients?[4] Further, how can hybrid DP techniques that maximize privacy without sacrificing model performance be smoothly integrated into FL models across a variety of data sets?[5] To handle DP measures efficiently in FL, this research introduces detailed methods that are suited to the particular information properties across participating clients.

The examination of two fundamental data sets—EMNIST and CIFAR-10—provides a useful perspective for investigating the complexities of privacy in intricate IoT configurations. These data sets provide critical understanding of the issues and remedies necessary to protect privacy in real-world applications. The depiction of handwritten characters by EMNIST allows for a closer investigation of privacy approaches in scenarios demanding exact identification without revealing sensitive information. CIFAR-10, which covers image classification and object recognition, demonstrates the need of protecting privacy while working with visual data in networked IoT systems. These data sets are excellent resources that highlight the practical applications of our suggested methodology. These are situations in which maintaining data privacy is essential to maintaining the efficiency and accuracy of shared models in FL settings.

Our paper introduces a multi-level dynamic multi-participant privacy budget allocation mechanism that varies to the convergence of the global model and the heterogeneity of participant data. This mechanism ensures data privacy while minimizing noise injection and enhancing the model's utility. To maximize utility, our proposed method includes an active participant selection mechanism for each training iteration. This selection process relies on an appropriate participant assessment approach that uses the 1-Wasserstein distance to quantitatively evaluate the effect of participant data heterogeneity on the global accuracy. This method dynamically assigns the privacy budget across various entities and iterations to adaptively balance privacy and utility. Furthermore, we implement a noise scaling approach that reduces noise injection as the global model converges, leading to optimal accuracy at a local ideal point.

The paper makes noteworthy contributions in the following aspects:

1. We present a unique technique for secure and reliable cross-platform knowledge sharing by combining hybrid DP and FL.
2. We address the challenge of balancing privacy and utility in FL situations, where data privacy preservation is crucial while ensuring the trained model is accurate and useful.

3. By incorporating an active participant selection mechanism in each training iteration based on a suitable participant assessment technique that utilizes the 1-Wasserstein distance to quantitatively assess data heterogeneity's impact on the global model's correctness, we adaptively balance privacy and utility.
4. Our method employs a noise scaling adjustment approach that reduces noise injection as the global model converges, resulting in optimal accuracy at a local ideal point.

The paper is structured to offer a comprehensive exploration of proposed methodologies and their theoretical underpinnings. Section 2 meticulously reviews related work, encompassing DP, FL and the application of model compression techniques within this framework. In Section 3, fundamental concepts are introduced, elucidating differential Privacy in federated learning contexts. Section 4, outlines the proposed approaches detailing their implementations and functionalities. Section 5 undertakes a thorough extensive empirical evaluations derived from experiments, exploring generalization within the proposed methods. Following this Section 6 encapsulates the paper by outlining essential discoveries and proposing potential avenues for future research and advancements specifically within the domain of HDP-FL methodologies.

2 | RELATED WORK

With the massive growth of data collection and storage, ensuring the security of data emerges as a paramount priority for organizations. Although access restrictions and encryption are frequently employed to secure data, they have limitations such as scalability and susceptibility to attacks. To address these issues, novel strategies that hybridize DP and FL are used. The FL technique improves data privacy by decentralizing the data and minimizing the likelihood of data breaches or illegitimate access to confidential data. By utilizing the advantages of each strategy, HDP-FL has the efficacy to improvise big data security. The issues of data sharing and privacy can be addressed via FL, and hybrid DP can offer further security against data breaches.

Several research investigations evaluate and contrast the efficiency of differentially private FL methodologies. Liu et al.[6] find that a novel approach achieves higher privacy assurances while maintaining accuracy levels similar to non-differentially private FL. Another study shows that hybrid strategies are more accurate than client-side techniques when compared to both client-side and server-side DP methods. Kairouz et al.[7] compare DP in both client and server-side using FL, showing that hybrid techniques can perform better than each strategy by itself. The authors demonstrate that this method works well with big data sets and that great accuracy may be attained while maintaining anonymity.

To enhance the protection of large data, several research efforts look into ways to combine machine learning with DP. A recent work by Bonawitz et al.[8] proposes that machine learning on multipartite distributed data sets maintains data privacy. Li et al.[9] establish a novel framework for FL to make it more resistant to hostile assaults. Abdul Rahman et al.[10] offer a thorough analysis of FL, including its uses, difficulties, and privacy-preserving methods. For distributed data fusion, Liu et al.[11] provide a solid FL method to secure multiparty computing and combine DP.

With the rise in data collection and storage, organizations are increasingly concerned about data security and sharing. However, access restrictions[12] and encryption[13] limitations have led to the development of novel strategies like FL and deep learning. Singh and Shukla[14] review different methods for maintaining privacy in machine learning, utilizing homomorphic encryption, multiparty secure computation, and DP. Wang et al.[15] provide a survey of FL with DP, focusing on the challenges and recent developments in this field. Maeng et al.[16] propose an incremental learning approach that maintains privacy by exploiting local and global consistency, leveraging differential privacy to protect training data privacy. Some studies[17–19] demonstrate that a FL approach for healthcare outperforms other DP methods implementing deep learning approaches. Hassan et al.[20] suggest a system design for larger data set FL that uses a hybrid strategy to tackle privacy concerns in this study. The authors show that this method works well with big data sets and that great accuracy may be attained while maintaining anonymity. Nair et al.[21] devise and execute a privacy-focused FL approach for analyzing substantial IoT-generated data, employing edge computing. Their evaluation, performed on an authentic healthcare data set, demonstrates the framework's ability to maintain high accuracy levels while safeguarding data privacy.

The proposed work distinguishes itself from related studies through its innovative combination of DP and FL to address the critical balance between data privacy and model accuracy. HDP-FL demonstrates superiority in scenarios with non-I.I.D. This is a significant departure from many related works that often focus on scenarios with uniform data distribution. The ability of HDP-FL to maintain high accuracy in such non-I.I.D. settings set it apart from other methodologies. Unlike some existing methods that primarily focus on achieving high accuracy with minimal consideration for privacy,

HDP-FL excels in maintaining strong accuracy while simultaneously enhancing privacy protection. It showcases a meticulous balance between these two critical aspects in FL. The proposed work extensively delves into parameter tuning for DP, thoroughly analyzing the impact of noise on both privacy loss and accuracy over various training rounds. This depth of analysis contributes to a nuanced understanding of the trade-offs between privacy, noise levels, and model performance. HDP-FL goes beyond theoretical propositions by conducting detailed comparative assessments against state-of-the-art techniques and performing scenario-based evaluations. This empirical validation provides concrete evidence of its efficacy in real-world settings with diverse data distributions and privacy requirements.

3 | PRELIMINARIES

We present hybrid differential privacy and its collaboration with federated learning for trustworthy cross-platform knowledge exchange with a focus on noise scaling and dynamic privacy budget allocation mechanisms is figured out in Figure 1.

HDP-FL introduces a novel method ensuring data privacy in collaborative machine learning environments. In this system, users collaborate to train a shared machine-learning model through a main server. The unique approach involves users sending gradients based on their trust level: those trusting the server directly share their local data, while cautious users only send gradient direction information. This method, while preserving privacy, allows secure model updates without the server gaining individual user insights. Encrypted data from cautious users is aggregated securely, preventing any learning about specific users by the server. This iterative process of information sharing and model refinement continues until convergence, ultimately delivering a reliable, privacy-preserving shared model for all users involved.

3.1 | Hybrid differential privacy

One of the known methods for protecting in the privacy is DP which introduces random noise to data before release to shield the identities of people. This method is employed by FL to protect participant data privacy. The global model's accuracy, however, might be adversely impacted by the volume of added noise. Hence, a hybrid DP technique that dynamically

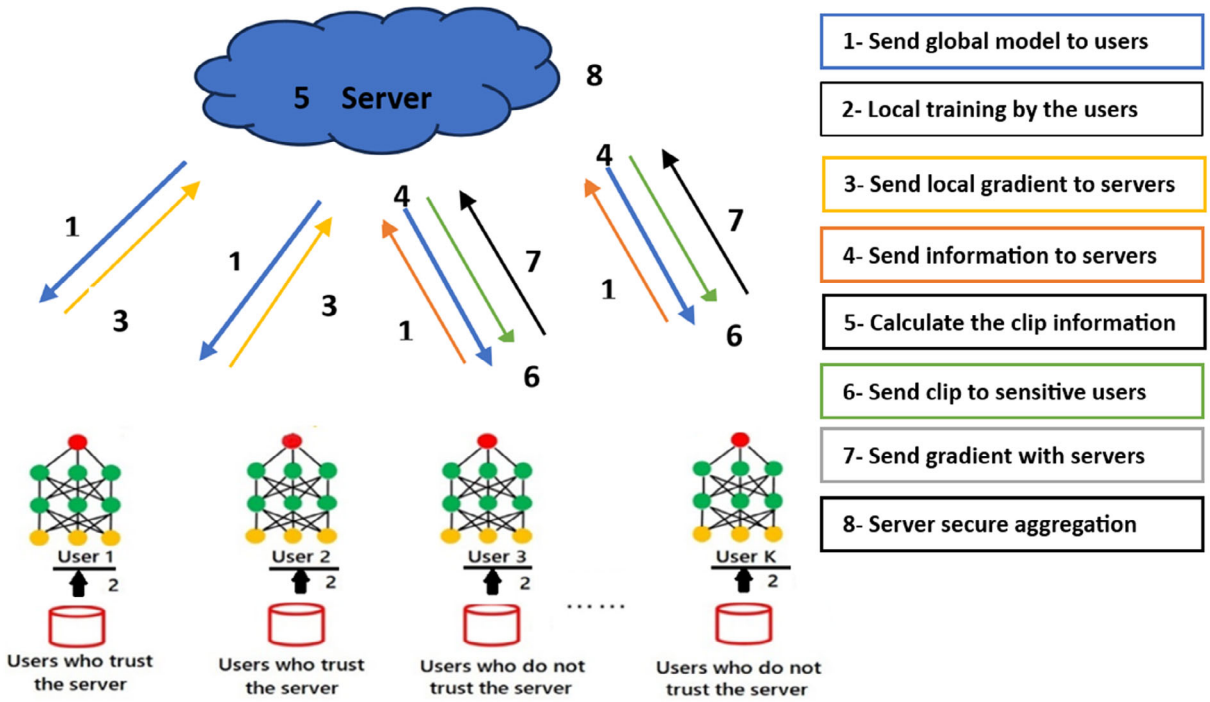

FIGURE 1 Architecture for HDP-FL.

distributes the privacy budget among a number of participants and iterations is provided in Equation (1)

$$\varepsilon = \varepsilon_0/((i/I)^{\alpha}), \tag{1}$$

where i is the present iteration, I is the number of iterations, α is the noise scaling factor, ε_0 is the initial privacy budget, and ε is the privacy budget. To establish a relationship between privacy and utility, use the noise scaling adjustment factor to modify the noise level at each iteration. Using this method, privacy is protected while the injected noise may be greatly reduced and the model's effectiveness increased. Moreover, it takes into account both global model convergence and heterogeneous participant data.

Let G_j be the local model's gradient for participant j. If that is done, the noise gradient G'_j prime may be computed using Equation (2)

$$G'_j = G_j + N\left(0, \varepsilon^2 * C_j\right), \tag{2}$$

where C_j is the gradient sensitivity for participant j and Gaussian noise $N\left(0, \varepsilon^2 * C_j\right)$ is introduced to the gradient with variance $\varepsilon^2 * C_j$.

The C_j sensitivity is equated from Equation (3) as

$$C_j = \max_{x,yind_j} ||\text{grad}(f(a,b))||, \tag{3}$$

where f(a, b) is the loss function and d_j is the participant j local data set. More privacy protection is provided by smaller sigma values, although updating the global model may be less accurate.

3.2 | Federated learning

Without the requirement for centralized data storage, FL enables numerous parties to cooperatively train a global model. After local training on individual devices, updates are consolidated to form a global model upon completion of the training process. Let $f_w(a)$ be a parameter in the neural network model L(W, a), where 'a' is the input data. With the huge data set the $d = a_1, a, \ldots, a_n$ objective is to learn the w* model that reduces the loss function. A data set d is divided into m distinct $d_1, d_2, \ldots, d_m$, non-overlapping subsets in FL so that $d = \bigcup_{j=1}^{m} d_j$. Each participant j's local model is first populated with the global model W_j, and each subset d_j is maintained by a distinct participant.

A randomly chosen subset of the p_t participants receive the most recent global model W_i during each training session. Each participant j at p_t initialises its local model $W_{j,i+1}$ using the current global model and trains it on the local subset d_j. The server receives the updated local model parameters and adds them to create the new global model W_{i+1}. Up to convergence, this process is performed several times. The global model update W_{i+1} in round i + 1 may be calculated from Equation (4)

$$W_{i+1} = \sum_{j \in p_t} \frac{n_i}{n} W_{j,i+1}, \tag{4}$$

where n_j is the size of the participant j's local data set, n is the total data set range, and $W_{j,i+1}$ is the participant j's updated local model at round i + 1. DP may be added to the FL process to assure privacy protection. This may be done by computing gradients locally, then before transferring them to the server, including the random noise to them. The privacy budget parameter can be used to regulate the level of extra noise and the local model j's noisy gradient update in round i + 1 is provided in Equation (5) as,

$$\tilde{W}_{j,i+1} = W_{j,i} - \eta_t \left(\nabla L\left(W_{j,i}, d_j\right) + \mathcal{N}\left(0, \sigma_t^2 K\right)\right), \tag{5}$$

where η_t is the learning rate at round t, $\sigma_t^2 K = [2\ \log(1.25/\delta)]\left[\varepsilon n_j\right]$ is the added Gaussian noise variance to the gradient variance, and $\mathcal{N}\left(0, \sigma_t^2 I\right)$ represents the Gaussian noise with mean 0 and variance σ_t^2.

Using a hybrid method to DP, noisy gradients are incorporated. This method assigns a privacy budget to each participant depending on the heterogeneity of the data and the contribution of the model. In this work, we have successfully

balanced privacy and usefulness by dynamically altering the allocation of the privacy budget, thereby enhancing the useability and precision of the FL process.

4 | PROPOSED METHOD

The HDP-FL methodology involves several key steps: First, gathering and preprocessing data from distributed sources to ensure uniformity. Then, initializing the global model parameters and sharing them across clients. Clients perform local training with their data sets while incorporating DP mechanisms to protect sensitive information. Next, the locally updated model parameters are aggregated to update the global model. Continuous evaluation and monitoring of metrics such as accuracy occur to assess the model's performance. This iterative process continues, improving the global model's accuracy while safeguarding data privacy, until predefined termination criteria are met. These steps emphasize collaborative learning among distributed entities, aiming to balance model enhancement and privacy preservation.

4.1 | 1-Wasserstein distance

1-Wasserstein distance is used to estimate the distinction amid global and local models in FL. A mapping between two probability distributions that reduces the overall transportation cost from one distribution to the other is known as an optimum transportation plan. In many cases, cost is determined based on the distance metric such as the Euclidean distance. For each participant, we determine out the 1-Wasserstein distance among the local and global models. The effect of data heterogeneity on the overall system is determined using the average distance.

The Wasserstein distance is determined by applying the Equation (6) below:

$$W(P, Q) = \inf_{\gamma \in \Pi(P,Q)} E_{(a,b)\sim\gamma}[|a - b|], \tag{6}$$

P and Q are considered to be the probability distributions for both participant data sets, and $\Pi(P, Q)$is the collection of joint distributions encompassing both P and Q as their marginal distributions.

Each j participant can add noise to their local model updates using the equation (7) to leverage hybrid DP in conjunction with 1-Wasserstein distance:

$$\tilde{\theta}_j = \theta_j + \mathcal{N}\left(0, \frac{2\sigma^2}{n_j \varepsilon_p}\right), \tag{7}$$

where n_j is the count of remaining data points, $\tilde{\theta}_j$ is signifies the noisy local model update for that participant, θ_j is the update of true local model for participant j. The noise provided for the updating of remaining data points by controlling the n_j, σ is the noise standard deviation, and ε_p is the privacy parameter. The noisy local model's updates may then be collected by the central server using the equation (8):

$$\tilde{\theta} = \frac{\sum_{j=1}^{N} n_j \tilde{\theta} j}{\sum_{j=1}^{N} n_j}, \tag{8}$$

where $\tilde{\theta}$ is the noisy global model update, N is the total participants, and n_j is the data points held by participant j.The procedure is continued until convergence, initiating subsequent rounds of training with a noisy global model update.

The algorithm for 1-Wasserstein distance.

---

1. Set the θ model parameter to zero
2. Do the following in each round t:
    a. Randomly choose the C participants.
    b. Every participant in j:

i. Based on the input's sensitivity, determine the privacy budget ϵ_j.
ii. Train a local model $M_j$with DP using ϵ_i.

c. Determine the local model's weighted average using the Wasserstein distance as the weight.
d. Include the weighted mean into the global model parameter theta.

3. Output of the final global model θ parameter.

Consider the CIFAR-10 data set, which contains thumbnail pictures of various objects, and the EMNIST data set, which comprises of handwritten numbers. The global model using FL with coupled DP and 1-Wasserstein distance that excels on both data sets despite data heterogeneity across users. Also, by incorporating noise to the local model updates, it is possible to guarantee the confidentiality of the data for each participant. A potent strategy for examining how data heterogeneity affects overall model performance in FL situations is the merge of DP methods with 1-Wasserstein distance.

4.2 | Dynamic allocation

In privacy budget allocation technique, the overall privacy budget among all participants is considered according to the degree of data heterogeneity. Based on the participant and data type, $\delta(h, p_a)$ is used to calculate the degree of heterogeneity. Let p_a be the participants count and H be the degrees of data heterogeneity. The overall privacy budget ε is then distributed according to the degree of heterogeneity $\alpha(h, p_a)$ to the level h, p_a: of each participant is provided in equation (9).

$$\varepsilon(h, p_a) = \alpha(h, p_a) / \sum \alpha(h, p_b) * \varepsilon, \tag{9}$$

where $\alpha\ (h, p_a)$ is the allocation ratio for each participant at level h and ε is the overall privacy budget.

The allotment for the privacy budget for each level and participant is then set up in an array. An iterative calculation is made for allocating the ratio, depending on the degree of heterogeneity for each level and participant. The privacy budget allocation is then determined for each level and participant, returning a list of allocations. We can confirm that privacy budgets are assigned depending on data and contributor heterogeneity by using this dynamic allocation mechanism. On the EMNIST and CIFAR-10 data sets, this can increase the overall accuracy of the HDP-FL models.

4.3 | A hybrid implementation of DP and FL

We integrate the proposed privacy-preserving method into the framework of FL on two different data sets. This hybrid approach is termed as HDP-FL. A multi-level and multi-participant dynamic allocation technique of privacy budgets are presented to decrease injection noise and effectively enhance model usability after examining the effect of data heterogeneity on the accuracy of global models.

Algorithm for HDP-FL for secure and reliable cross-platform knowledge sharing

---

1. Randomly initialize global model parameters: W(0)
2. Divide the data set into non-IID subsets: $D = D_1, D_2, \ldots, D_n$
3. For every round $i = 1, 2, 3, \ldots, R$ do:
   a. Randomly choose a subset of participants: p_{at}
   b. For each participant j in the subset, do the following:

      i. Evaluating a local model on a non-IID subset of data: $W_j(t) = \text{argmin}_w E_j(W)$
      ii. Calculate the gradients of the loss function and the model parameters for each local device from Equation (10).

$$G_j(t) = \nabla L_j\left(W_j(t)\right) \tag{10}$$

iii. Add differential privacy noise to the gradient using a hybrid DP mechanism from Equation (11).

$$\hat{G}_j(t) = G_j(t) + \mathcal{N}\left(0, \sigma_t^2 K\right) \tag{11}$$

iv. Give the server a noisy gradient.

c. Update the parameters of the global model gathering noisy gradients from all participants and using a learning technique like federated averaging.

$$W(i+1) = \mathrm{Agg}\left(p_{ai}, \hat{G}_1(t), \hat{G}_2(t), \ldots, \hat{G}_{|p_{ai}|}(t)\right) \tag{12}$$

d. Continue through steps b through c until convergence or a predetermined set of local repetitions.

4. Examine the global model's precision and privacy on the validation set.
5. Continue steps 3–4 until convergence or a predetermined number of rounds.

It also stratifies clients into tiers based on factors such as data sensitivity and privacy practices, allocating a more substantial budget to clients with greater privacy considerations. It dynamically adapts to changes in the client composition, data characteristics, and model progression, ensuring equitable and practical budget distribution.

From Equation (12) the global model parameter is W(i) at D is the entire data set, the subset of participants is chosen to take part in round t is called p_{ai}, the local model parameter for participant j in round t is $W_j(i)$, the empirical risk of the data subset for participant j is $E_j(W)$. Random sampling can also be employed with certain constraints to ensure non-IID characteristics. Constraints can be applied to the sampling process to prevent the formation of overly homogenous subsets, promoting diversity in data distribution.

The local gradient of participant j's data subset $G_j(t)$, the noise gradient of the data set for participant j with respect to the global model parameters at round t is represented by $\hat{G}_j(t)$. The noise scaling parameter at round t is represented by σ_t. $\mathcal{N}\left(0, \sigma_t^2 K\right)$ is Gaussian noise with mean 0 and covariance matrix sigma $\sigma_t^2 * I$. In order to update the global model parameters at round t + 1, the aggregation function $\mathrm{Agg}\left(p_{ai}, \hat{G}_1(t), \hat{G}_2(t), \ldots, \hat{G}_{|p_{ai}|}(t)\right)$ aggregates the noisy gradients of the participants in the subset p_{ai}. The loss function for the participant j's data subset with regard to model parameter W is $L_j(W)$.

For the EMNIST and CIFAR-10 data sets, deep learning models may be trained using a hybrid differential privacy-preserving and federated strategy. In this method, non-IID portions of the training data are separated, and a subset of participants is randomly chosen to take part in each training cycle. Using the validation set, the noise scale parameter sigma tis is dynamically modified to reduce accuracy loss while respecting privacy requirements. This ongoing iterative approach persists until it attains convergence, typically extending over a predetermined set of cycles or iterations. In order to confirm that the model is correct and private, the global model's accuracy and privacy are evaluated in the validation set.

5 | RESULTS

5.1 | Experimental setup

In this study, we developed the algorithm in Python using the TensorFlow federated learning package. For assessment purposes, two publicly available real-world data sets, namely EMNIST and CIFAR-10, were employed in this study. The 70 000 handwritten figures in grayscale found in the EMNIST data set are split into 60 000 training pictures and 10 000 test images. Each picture is 28 by 28 pixels and has the numbers 0 to 9 written on it. Ten categories of RGB color pictures in the CIFAR-10 data set are split into 50 000 training images and 10 000 test images. There are 10 picture categories, each with a 32×32-pixel image, including aircraft, vehicle, bird, cat, deer, dog, frog, horse, and truck.

The experiments are processed on a Linux server with Intel® Xeon® CPU E5-2690 v4 processors clocked at 2.60 GHz, 120 GB of RAM, and a V100 GPU. Due to hardware constraints, the FL clients trained sequentially. To ensure the server model's convergence without resorting to early-stop strategies, we meticulously trained all client models. In each iteration, each client model received one epoch of training using its raw data. The EMNIST and CIFAR-10 data sets were utilized to

TABLE 1 Parameter settings for FL with EMNIST and CIFAR-10 data sets.

Parameter	EMNIST	CIFAR-10	Parameter	EMNIST	CIFAR-10
σ	1.0	1.0	Hybrid DP noise multiplier (δ)	1.5	1.2
k	0.7	0.7	Hybrid DP clipping norm (C)	2.5	3.0
Batch size	20	20	Rounds for FL	100	200
Clipping threshold	4.0	7.0	Batch size for FL	10	10
Learning rate	0.05	0.05	Clients for FL	500	50
Number of participants	1000	100	Server learning rate for FL	0.1	0.01
Training round	50	50	Client learning rate for FL	0.01	0.001
Noise scale adjusting factor (α)	0.5	0.5	Federated averaging factor (β)	0.5	0.9

train neural network models using FL techniques, and their parameters are provided in Table 1. The Learning rate (0.05) gives the details how quickly the main server updates its global model based on contributions from all participants; it's higher for EMNIST (0.1) and lower for CIFAR-10 (0.01). The Client learning rate for FL determines how fast individual devices update their local models before distributing them with the server. EMNIST clients updated faster (0.01) than CIFAR-10 (0.001). EMNIST had 1000 potential contributors, with 500 actively participating per round. For CIFAR-10, out of 100 participants, 50 contributed per round. "Batch size for FL" determines how many clients updated their models each round, playing key roles in our model training strategy.

5.2 | Performance

In the HDP-FL approach's noise scale, which ranges from 0.1 to 1. Table 2 illustrates the privacy loss in DP with varying noise parameters (σ) over training rounds for the EMNIST data set. It shows how the privacy loss changes over time and with different noise scales. As the training rounds progress, we observe a decrease in privacy loss for each value of σ, demonstrating the impact of noise on privacy. It is important to note that while too much noise can compromise accuracy, our method aids in achieving an equilibrium between preserving privacy and maximizing utility. For instance, the noise scale at round 1 is 0.1 for $\varepsilon = 1.0$. This indicates that during training, the model introduces noise to the data that is 0.1 times the data's standard deviation. The noise scale grows along with the number of training rounds. The noise introduced during training is less at round 10 than it was at round 1 (the noise scale is 0.039 for $\varepsilon = 1.0$ at round 10). This demonstrates how privacy is significantly impacted by the amount of noise that is introduced to the data. But over-noising model might also make it less accurate. In a FL context, utilizing hybrid DP generally aids in striking a good compromise between accuracy and privacy. Local and centralized DP methods are combined to provide hybrid DP. According to the table, as the reduction in privacy loss is related to the increase in the number of trainings rounds and applies to each value of the noise scale.

Figure 2 depicts the correlation between increasing the number of training sessions and enhancing both the model's accuracy and privacy. This visual representation elucidates the intricate balance between privacy and accuracy, showcasing that as noise levels diminish, accuracy ascends and conversely declines. Table 3 is presented to compare the accuracy outcomes of three FL techniques—Client-based FL (CL-FL), Hybrid DP-FL(HDP-FL), and Differentiated Private FL (DP-FL)—on different client counts. As the privacy parameter value escalates, the degree of privacy safeguarding diminishes. The table also shows the accuracy comparison of these techniques on the EMNIST and CIFAR-10 data sets for various client counts and privacy budgets. With the exception of DP-FL 1000 clients, which have a better accuracy than DP-FL 100 customers, all FL techniques see a minor decline in accuracy as the number of clients rises. Figure 3 confirms this information by visualizing the accuracy comparison for EMNIST and CIFAR-10 data sets using these techniques. We observe that HDP-FL achieves the highest accuracy when no privacy protection is applied. However, as privacy parameters increase, accuracy decreases, emphasizing the trade-off.

Table 4 compares the EMNIST and CIFAR-10 data set's accuracy across 50 training cycles in Independent and Identically Distributed (IID) and non-IID environments. For EMNIST Starting at 0.968, it rises steadily until it reaches 0.979 at the halfway point. It highlights that differentiated privacy and FL can mitigate the detrimental impact of non-IID data distribution on model accuracy. It is indicated by the significantly higher accuracy in IID settings. The best accuracy attained for the IID tuning is 0.989, making it substantially more accurate than the non-IID tuning.

TABLE 2 Privacy loss in HDP-FL with varying noise parameter (σ) over training rounds.

Training Round	σ = 0.1	σ = 0.2	σ = 0.3	σ = 0.4	σ = 0.5	σ = 0.6	σ = 0.7	σ = 0.8	σ = 0.9	σ = 1.0
1	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
5	0.058	0.116	0.174	0.232	0.290	0.348	0.406	0.464	0.522	0.580
10	0.039	0.078	0.117	0.156	0.195	0.234	0.273	0.312	0.351	0.390
15	0.029	0.058	0.087	0.116	0.145	0.174	0.203	0.232	0.261	0.290
20	0.023	0.046	0.069	0.092	0.115	0.138	0.161	0.184	0.207	0.230
25	0.019	0.038	0.057	0.076	0.095	0.114	0.133	0.152	0.171	0.190
30	0.016	0.032	0.048	0.064	0.080	0.096	0.112	0.128	0.144	0.160
35	0.014	0.028	0.042	0.056	0.070	0.084	0.098	0.112	0.126	0.140
40	0.012	0.024	0.036	0.048	0.060	0.072	0.084	0.096	0.108	0.120
45	0.011	0.022	0.033	0.044	0.055	0.066	0.077	0.088	0.1	0.1
50	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.1	0.1	0.1

FIGURE 2 Effect of noise levels on training accuracy in HDP-FL.

TABLE 3 Comparison of FL Techniques on Different Numbers of Clients with Varying Privacy Parameters.

			Accuracy of HDP-FL		Accuracy of CL-FL		Accuracy of DP-FL	
	# of Clients	δ	EMNIST	CIFAR-10	EMNIST	CIFAR-10	EMNIST	CIFAR-10
Non-DP	10		0.991	0.891	0.925	0.882	0.872	0.854
DP	10	$\frac{1}{e^2}$	0.929	0.775	0.867	0.769	0.821	0.756
DP	100	$\frac{1}{e^3}$	0.938	0.846	0.878	0.797	0.856	0.779
DP	1000	$\frac{1}{e^4}$	0.989	0.884	0.917	0.862	0.869	0.788

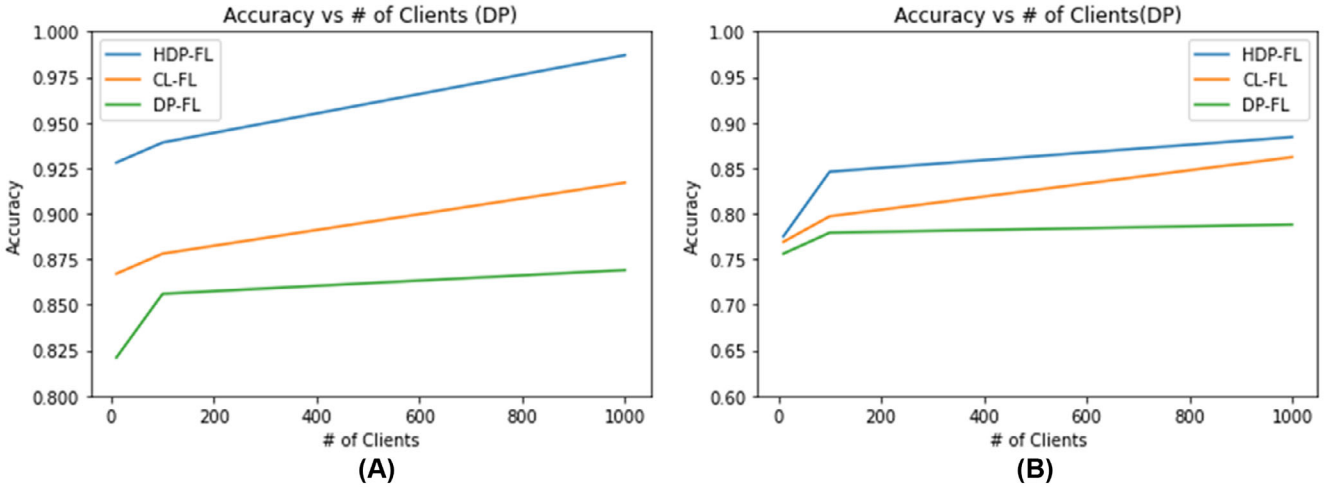

FIGURE 3 Comparative accuracy analysis of FL methods: (A) EMNIST, (B) CIFAR-10.

TABLE 4 Comparison of accuracy in different distributed scenarios.

Data sets	Training rounds	1	5	10	15	20	25	30	35	40	45	50
EMNIST	IID	0.942	0.956	0.960	0.963	0.966	0.970	0.973	0.981	0.983	0.985	0.989
	Non-IID	0.937	0.946	0.951	0.963	0.964	0.965	0.966	0.967	0.968	0.969	0.979
CIFAR-10	IID	0.537	0.696	0.764	0.805	0.831	0.846	0.857	0.866	0.873	0.878	0.884
	Non-IID	0.480	0.594	0.675	0.716	0.746	0.764	0.778	0.789	0.799	0.807	0.855

Figure 4A,B visualizes the effect of training rounds on model accuracy for CIFAR-10 and EMNIST respectively, emphasizing the potential of a hybrid strategy to handle privacy and accuracy challenges effectively. These findings demonstrate that, while retaining a respectable degree of privacy protection, differentiated privacy and FL may be employed to lessen the detrimental effect of data distribution for non-IID on model accuracy. The best accuracy for CIFAR-10 attained for the IID tuning was 0.884, making it much more accurate than the non-IID tuning. This study demonstrates that FL and DP when used together can achieve respectable accuracy in IID and non-IID settings, although non-IID conditions still have space for improvement.

Table 5 offers an extensive breakdown analysis of the performance of various traditional FL algorithms, including FedAvg,[22] FedSGD,[23] FedMA,[24] FedDyn,[25] and the proposed method, across two distinct data distribution scenarios: Non-IID and IID. Among them, "FedMA" consistently shows good training accuracy in Non-IID data sets. "FedSGD" also performs well in both settings. "FedAvg" is slightly less accurate, especially with CIFAR-10. The Proposed Method stands out in IID scenarios, achieving the highest overall accuracy of 96.29% in IID EMNIST as shown in Figure 5. This demonstrates its strength across various FL scenarios, especially when data is evenly distributed among clients. This hybrid method proves effective in addressing privacy concerns. Statistically, the differences in accuracy percentages among algorithms vary notably, with the proposed method showcasing significantly higher validation accuracy compared to FedAvg and FedSGD, particularly in IID data sets.

Furthermore, to benchmark our method against the latest State-Of-The-Art (SOTA) in FL privacy protection, we have incorporated referencing studies 26 and 27 as controls within our experimentation. Additionally, references [28–30] aim to ensure a rigorous comparison against established privacy protection mechanisms and highlight the advancements achieved by our proposed approach. We believe these enhancements significantly contribute to the robustness and relevance of our experimental section. Table 6 compares the precision of our innovative model compared to prevalent state-of-the-art analyses on EMNIST and CIFAR-10 data sets. FedSGD coupled with a deep neural network (DNN) achieves moderate accuracy, reaching 75% on EMNIST and 54% on CIFAR-10. Similarly, FL employing K-nearest neighbor (KNN) records 65% accuracy on EMNIST and 62% on CIFAR-10, presenting reasonably competitive results. However, the standout performer is the proposed model, demonstrating remarkable accuracy rates of 96.29% on

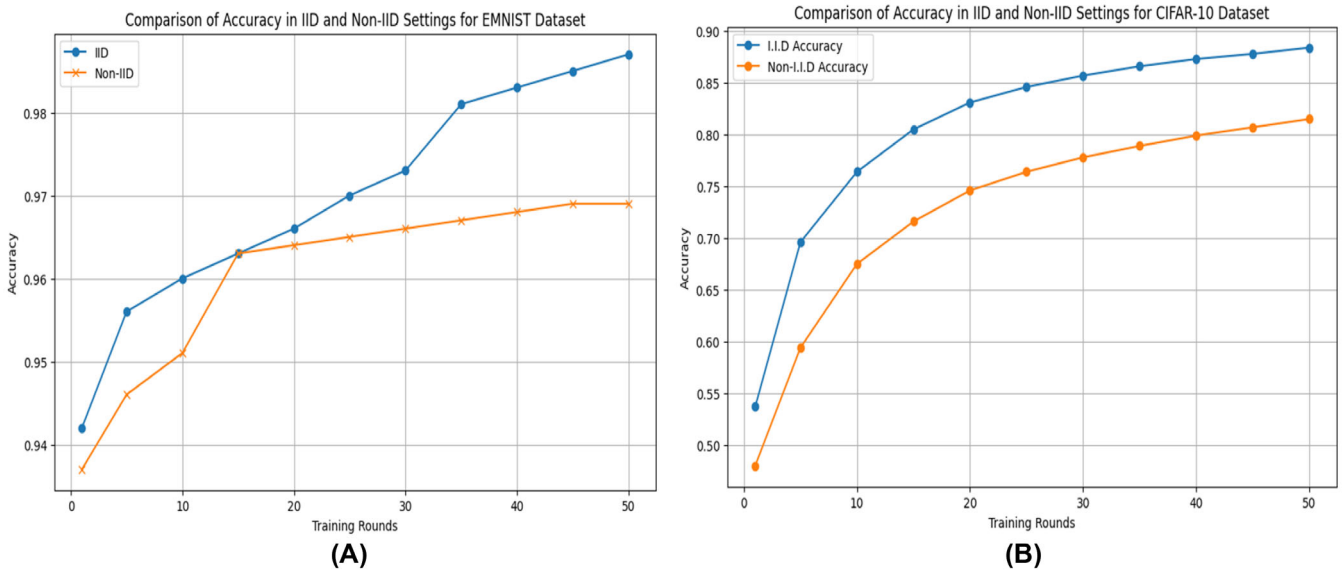

FIGURE 4 Model accuracy variation with training rounds (A) EMNIST data set(B) CIFAR-10.

TABLE 5 Performance metrics of various algorithms on EMNIST and CIFAR-10 data sets Under Non-IID and IID Settings.

		Non-IID				IID			
Data sets		**EMNIST**		**CIFAR-10**		**EMNIST**		**CIFAR-10**	
Algorithms	(%)	Train	Validation	Train	Validation	Train	Validation	Train	Validation
	FedAvg	96.78 ± 0.11	85.80 ± 0.60	79.87 ± 1.52	63.27 ± 0.67	96.12 ± 0.85	89.58 ± 0.25	80.70 ± 0.83	70.15 ± 0.47
	FedSGD	95.18 ± 0.64	83.32 ± 0.19	79.07 ± 0.45	59.81 ± 0.81	95.61 ± 0.94	68.87 ± 0.31	82.90 ± 0.72	67.56 ± 0.02
	FedMA	96.21 ± 0.02	84.18 ± 0.73	76.71 ± 0.01	63.52 ± 0.53	92.71 ± 0.61	77.75 ± 0.37	81.12 ± 0.83	69.56 ± 0.57
	FedDyn	96.28 ± 0.02	85.20 ± 0.35	77.55 ± 0.02	63.10 ± 0.18	93.66 ± 0.40	78.27 ± 0.19	80.90 ± 0.86	68.10 ± 0.22
	Proposed method	97.96 ± 0.04	93.41 ± 0.91	85.52 ± 0.02	76.75 ± 0.35	98.90 ± 0.40	94.90 ± 0.86	88.42 ± 0.52	80.86 ± 0.57

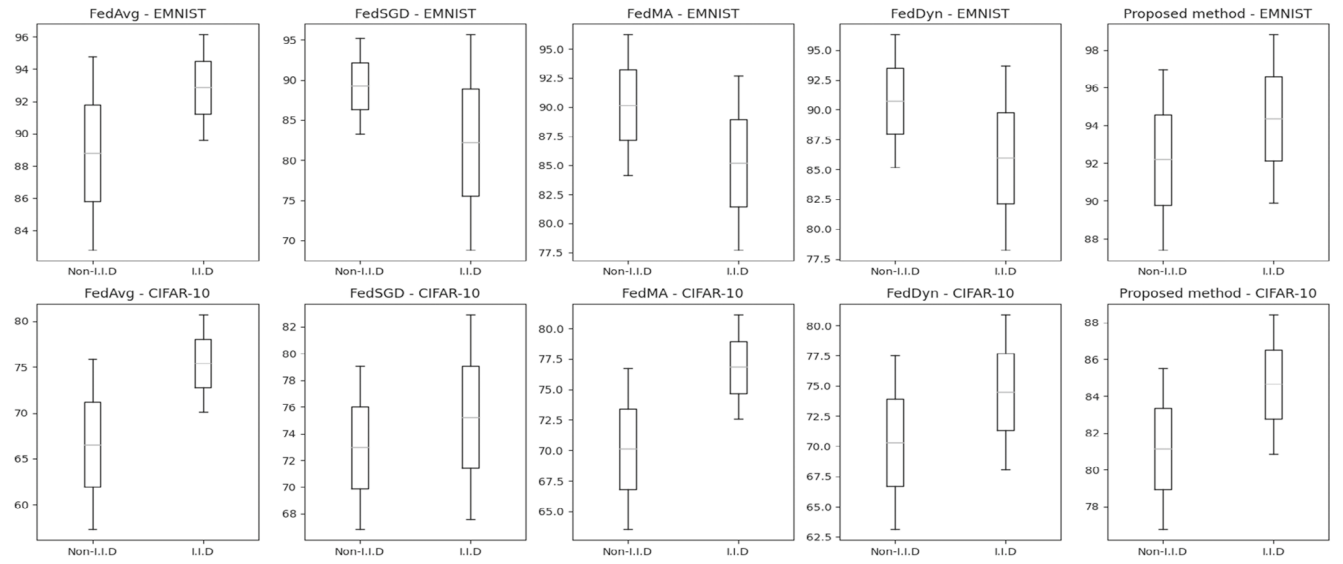

FIGURE 5 Comparison of Algorithms' Performance on EMNIST and CIFAR-10 Data sets: Non-IID versus IID.

TABLE 6 Comparison of SOTA model accuracies: EMNIST and CIFAR-10.

Models	EMNIST accuracy	CIFAR-10 accuracy
FedSGD+deep neural network (DNN)[26]	75%	54%
FL + K-nearest neighbor (KNN)[27]	65%	62%
FedDyn+DP + convolutional neural networks (CNN)[28]	-	82%
FedAvg+recurrent neural network (RNN)[29]	89%	-
FedAvg+CNN[30]	96%	81%
Proposed model	96.29%	82.88%

TABLE 7 Accuracy comparison between different scenarios and privacy budgets.

	Privacy budgets	$\varepsilon = 2$		$\varepsilon = 4$		$\varepsilon = 6$		$\varepsilon = 8$	
Distribution	Algorithm	r-1	r-2	r-1	r-2	r-1	r-2	r-1	r-2
Non-IID	FedAvg	0.528	0.59	0.652	0.726	0.67	0.748	0.719	0.76
	HDP-FL	0.561	0.634	0.684	0.784	0.693	0.802	0.74	0.821
IID	FedAvg	0.56	0.608	0.689	0.715	0.711	0.738	0.735	0.79
	HDP-FL	0.609	0.636	0.699	0.747	0.735	0.79	0.751	0.835

EMNIST and a competitive range of 82–88% on CIFAR-10. These findings underscore the proposed model's potential, particularly its robust performance on CIFAR-10, warranting further in-depth comparisons with existing FL models to comprehensively assess its efficacy.

The scenario-based illustrative evaluation investigates HDP-FL and FedAvg algorithms in the context of cross-IoT device FL. The clients are limited to one or two training rounds, and each client exclusively handles two distributions of data from the MNIST data set is presented in Table 7. HDP-FL consistently outperforms FedAvg in both Non-IID and IID distributions across different privacy levels (ε =2, 4, 6, 8). The rationale behind the higher accuracy enhancement observed in scenarios where clients engage in two training rounds (r-1,2) are attributed to the personalized data transformation incorporated in HDP-FL. This approach better adapts to and models the local training data, consequently mitigating the heterogeneity induced by DP in such settings.

The algorithm demonstrates its efficacy through detailed parameter tuning for differential privacy, illustrating the impact of noise on privacy loss over training rounds. Analyses of noise effects on accuracy and privacy, comparative assessments against existing techniques, performance benchmarking, and scenario-based evaluations consistently highlight the algorithm's superiority. The HDP-FL approach effectively manages the trade-off between privacy and utility, making it a promising solution for secure knowledge sharing in IoT settings. The extensive experimental evaluation and referenced findings collectively establish the applicability and effectiveness of HDP-FL in preserving privacy while ensuring high accuracy in FL.

Expanding our methodology to larger data sets and diverse applications poses several challenges. First, scalability becomes a concern as handling extensive data may strain computational resources, potentially causing delays in real-time updates. Balancing privacy and utility at a larger scale becomes more complex, affecting model accuracy due to stronger privacy measures. Adapting the model effectively to different data distributions across diverse contexts is also challenging. Communication among numerous clients or managing larger data volumes risks latency and resource-heavy processes. To ensure scalability without compromising accuracy, simplifying the algorithm's complexity is crucial. Guaranteeing the model's adaptability across varied data sets and contexts while addressing infrastructure readiness and security concerns is essential. Overcoming these challenges requires innovative algorithmic approaches, and a careful balance between privacy and utility.

6 | CONCLUSION

In conclusion, the findings showcased the method for safe and trustworthy information exchange among IoT devices while safeguarding data privacy, promoting effective learning methods across decentralized networks. The proposed

HDP-FL showed great potential in managing the balance between privacy and usefulness, hinting at its broader application in addressing diverse challenges in FL. Experiments performed on EMNIST and CIFAR-10 data sets showed the excellent performance of this methodology concerning the accuracy and resilience of the model. We propose that our hybrid DP and FL approach can be extended to address additional challenges in FL, such as data heterogeneity, effective communication, and model customization. To improve security in big data handling through FL, the study suggests standardizing security measures, fine-tuning privacy settings, and finding ways to streamline communication for better data protection. Future work could focus on developing optimized privacy-preserving mechanisms tailored for specific industry applications, such as implementing the hybrid approach in healthcare for secure collaborative diagnostics across hospitals while ensuring patient data confidentiality.

AUTHOR CONTRIBUTIONS

Conceptualization: O.I.K. and A.S.R. *Data curation*. S.P. and D.S. *Formal analysis*: S.P. *Funding acquisition*: M.S.S., W.E. *Investigation*, A.S.R. *Project administration*: O.I.K., M.S.S., and W.E.; *Resources*: S.P. and D.S. *Software*: A.S.R. and D.S. *Validation*: O.I.K. and X.X. *Visualization*: S.P. and O.I.K *Writing—original draft*: S.P., A.S.R., M.S.S, and W.E. *Writing—review and editing*: O.I.K.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Oshamah Ibrahim Khalaf https://orcid.org/0000-0002-4750-8384
Mhd Saeed Sharif https://orcid.org/0000-0002-4008-8049

REFERENCES

1. Li T, Sahu AK, Talwalkar A, Smith V. Federated learning: challenges, methods, and future directions. *IEEE Signal Process Mag*. 2020;37(3):50-60. doi:10.1109/MSP.2020.2975749
2. Jahromi AN, Karimipour H, Dehghantanha A. An ensemble deep federated learning cyber-threat hunting model for industrial internet of things. *Comput Commun*. 2023;198:108-116. doi:10.1016/j.comcom.2022.11.009
3. Ponomareva N, Hazimeh H, Kurakin A, et al. How to dp-fy ml: a practical guide to machine learning with differential privacy. *J Artif Intell Res*. 2023;77:1113-1201. doi:10.1613/jair.1.14649
4. Guendouzi BS, Ouchani S, Assaad HE, Zaher ME. A systematic review of federated learning: challenges, aggregation methods, and development tools. *J Netw Comput Appl*. 2023;220:103714. doi:10.4236/jis.2023.142008
5. Zhang L, Shen L, Ding L, Tao D, Duan LY. Fine-tuning global model via data-free knowledge distillation for non-iid federated learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE; 2022:10174-10183.
6. Liu L, Truex S, Liu L, Chow KH, Gursoy ME, Wei W. LDP-Fed: federated learning with local differential privacy. *Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking*. Association for Computing Machinery; 2020:61-66. doi:10.1145/3378679.3394533
7. Kairouz P, McMahan HB, Avent B, et al. Advances and open problems in federated learning. Foundations and trends®. *Mach Learn*. 2021;14(1–2):1-210. doi:10.1561/2200000083
8. Bonawitz K, Kairouz P, Mcmahan B, Ramage D. Federated learning and privacy. *Commun ACM*. 2022;65(4):90-97.
9. Xue X, Palanisamy SK, Manikandan A, et al. A Novel partial sequence technique based Chaotic biogeography optimization for PAPR reduction in eneralized frequency division multiplexing waveform. *Heliyon*. 2023;9:e19451. doi:10.1016/j.heliyon.2023.e19451
10. AbdulRahman S, Tout H, Ould-Slimane H, Mourad A, Talhi C, Guizani M. A survey on federated learning: the journey from centralized to distributed on-site learning and beyond. *IEEE Internet Things J*. 2020;8(7):5476-5497. doi:10.1109/JIOT.2020.3030072
11. Liu X, Huang H, Xiao F, Ma Z. A blockchain-based trust management with conditional privacy-preserving announcement scheme for VANETs. *IEEE Internet Things J*. 2019;7(5):4101-4112. doi:10.1109/JIOT.2019.2957421
12. Palanisamy SK, Abdulsahib GM, Khalaf OI, Ajitha SS, Wong W-K, Pan S-H. Design of Artificial Magnetic Conductor based stepped V-shaped printed multiband antenna for wireless applications. *Int J Adv Soft Comput Appl*. 2023;15(3):100-116. doi:10.15849/IJASCA.231130.07
13. Bhowmick A, Duchi J, Freudiger J, Kapoor G, Rogers R. Protection against reconstruction and its applications in private federated learning. *arXiv preprint arXiv:1812.00984*. 2018. doi:10.48550/arXiv.1812.00984

14. Singh S, Shukla KK. Privacy-preserving Machine Learning for Medical Image Classification. *arXiv preprint arXiv:2108.12816*. 2021. doi:10.48550/arXiv.2108.12816
15. Wang YZ, Luan TH, Zhang N, Li F, Chen T, Cao H. Secure and efficient federated learning for smart grid with edge-cloud collaboration. *IEEE Trans Industr Inform*. 2021;18(2):1333-1344. doi:10.1109/TII.2021.3095506
16. Xue X, Abdulsahib GM, Khalaf OI, et al. Soft computing approach on estimating the lateral confinement coefficient of CFRP veiled circular columns. *Alex Eng J*. 2023;81(15):599-619. doi:10.1016/j.aej.2023.09.053
17. Abdulsahib GM, Hassan HJ, Khalaf OI. A modified bandwidth prediction algorithm for wireless sensor networks. *J Inf Sci Eng*. 2024;40(1):177-188.
18. Dash B, Sharma P, Ali A. Federated learning for privacy-preserving: a review of PII data analysis in Fintech. *Int J Softw Eng Appl*. 2022;13(4):1-13.
19. Premkumar M, Ashokkumar SR, Mohanbabu G, Jeevanantham V. Jayakumar S security behavior analysis in web of things smart environments using deep belief networks. *Int J Intell Netw*. 2022;3:181-187. doi:10.1016/j.ijin.2022.10.003
20. Anand D, Khalaf OI, Abdulsahib GM, Chandra GR. Identification of meningioma tumor using recurrent neural networks. *J Autonom Intell*. 2024;7(2):653. doi:10.32629/jai.v7i2.653
21. Nair AK, Sahoo J, Raj ED. Privacy preserving federated learning framework for IoMT based big data analysis using edge computing. *Comput Stand Interfaces*. 2023;86:103720. doi:10.1016/j.csi.2023.103720
22. Li T, Zaheer M, Liu KZ, Reddi SJ, McMahan HB, Smith V. Differentially Private Adaptive Optimization with Delayed Preconditioners. *arXiv preprint arXiv:2212.00309*. 2022. doi:10.48550/arXiv.2212.00309
23. Hu E, Tang Y, Kyrillidis A, Jermaine C. Federated learning over images: vertical decompositions and pre-trained backbones are difficult to beat. *Proceedings of the IEEE/CVF International Conference on Computer Vision 2023*. IEEE; 2023:19385-19396. doi:10.48550/arXiv.2309.05505
24. Wang H, Yurochkin M, Sun Y, Papailiopoulos D, Khazaeni Y. Federated learning with matched averaging. *arXiv preprint arXiv:2002.06440*. 2020. doi:10.48550/arXiv.2002.06440
25. Acar DAE, Zhao Y, Navarro RM, Mattina M, Whatmough PN, Saligrama V. Federated learning based on dynamic regularization. *arXiv preprint arXiv:2111.04263*. 2021. doi:10.48550/arXiv.2111.04263
26. Chinnasamy P, Wong WK, Raja AA, Khalaf OI, Kiran A, Babu JC. Health recommendation system using deep learning-based collaborative filtering. *Heliyon*. 2023;9(12):e22844. doi:10.1016/j.heliyon.2023.e22844
27. Kang J, Xiong Z, Niyato D, Zou Y, Zhang Y, Guizani M. Reliable federated learning for mobile networks. *IEEE Wirel Commun*. 2020;27(2):72-80. doi:10.1109/MWC.001.1900119
28. Chen D, Xie LJ, Kim B, et al. Federated learning based mobile edge computing for augmented reality applications. *In 2020 International Conference on Computing, Networking and Communications (ICNC)*. IEEE; 2020:767-773. doi:10.1109/ICNC47757.2020.9049708
29. Zhao Z, Wang J, Hong W, Quek TQ, Ding Z, Peng M. Ensemble federated learning with non-IID data in wireless networks. *IEEE Trans Wirel Commun*. 2023;99:1. doi:10.1109/TWC.2023.3309376
30. Abdulsahib GM, Selvaraj DS, Manikandan A, et al. Reverse polarity optical orthogonal frequency division multiplexing for high-speed visible light communications system. *Egypt Inform J*. 2023;24(4):100407. doi:10.1016/j.eij.2023.100407

How to cite this article: Ibrahim Khalaf O, S.R A, Algburi S, et al. Federated learning with hybrid differential privacy for secure and reliable cross-IoT platform knowledge sharing. *Security and Privacy*. 2024;e374. doi: 10.1002/spy2.374