Federated Learning with Differential Privacy and Secure Multiparty Computation - Checkpoint I

1 Introduction

The task was to implement Federated Learning (FL) for the CIFAR-10 dataset, enhancing it with privacy-preserving techniques such as Differential Privacy (DP) and Secure Multiparty Computation (SMC). The goal was to study how these techniques affect the overall performance of the model in terms of accuracy and privacy. FL allows multiple clients to collaborate in training a shared model without centralizing their data, while DP ensures that individual data cannot be inferred from model updates, and SMC ensures secure aggregation of models from different clients.

This report details the steps taken to develop the system, challenges faced, and the outcomes of integrating DP and SMC in FL.

2 Task Overview

The task can be broken down into three primary stages:

1. **Basic Federated Learning Implementation:** The initial step was to implement a simple FL system without any privacy enhancements.

2. **Incorporating Differential Privacy:** The next step involved adding Gaussian noise to the data used in training to achieve differential privacy.

3. **Adding Secure Multiparty Computation:** Lastly, SMC was introduced to secure the aggregation process and further strengthen the privacy guarantee.

The overall objective was to compare the performance of these methods and assess the trade-offs between privacy and accuracy.

3 Federated Learning: Initial Implementation

3.1 Dataset and Model

The CIFAR-10 dataset was chosen for the task. It consists of 60,000 32x32 color images in 10 different classes, with 50,000 training images and 10,000 test images. Basic transformations such as normalization and conversion to tensor format were applied.

A simple Convolutional Neural Network (CNN) was used, consisting of two convolutional layers followed by max pooling, and two fully connected layers for classification. This CNN is simple enough to train on small amounts of data locally at each client, which fits the federated learning setup.

3.2 Federated Learning Procedure

In the basic FL setup, the CIFAR-10 dataset was split into multiple subsets, each representing data on a different client. Each client trained its local model independently for a few epochs and sent the model updates to a central server. The server performed federated averaging (FedAvg), which averages the parameters (weights) of the models from all clients to update the global model. This process was repeated over several rounds.

3.3 Results and Challenges

In the initial FL implementation, the test accuracy of the global model after several rounds of training across five clients reached an average of around 16%. This relatively low accuracy can be attributed to the following challenges:

- **Data Distribution:** The data was split among clients, and each client had access to a limited portion of the dataset, affecting the model's ability to generalize.

- **Simple Model Architecture:** The CNN used was basic and may not have been sufficient to capture complex patterns in the CIFAR-10 dataset.

4 Differential Privacy Implementation

4.1 Differential Privacy Concept

Differential Privacy (DP) aims to provide a strong privacy guarantee by adding noise to the model updates or data, ensuring that the influence of any single data point on the model's output is minimized. In this task, Gaussian noise was added to the data during training at each client to obscure the individual contributions of client data points.

4.2 DP in Federated Learning

To apply DP, a function was created to add Gaussian noise to the data, controlled by parameters ϵ () and δ (). These parameters govern the amount of noise added, with smaller values of ϵ providing stronger privacy guarantees at the expense of model accuracy.

The local training at each client was performed with noisy data, and then the model updates were sent to the central server for aggregation using the same FedAvg process.

4.3 Results and Trade-offs

After integrating DP, the test accuracy dropped to around 10%. This decline in performance was expected due to the noise added to the data. The key challenges here were:

- **Noise-Accuracy Trade-off:** While the noise ensures that individual data points are protected, it also degrades the quality of the model's predictions. Finding the right balance between privacy (ϵ) and accuracy was difficult.

- **Epsilon Tuning:** Fine-tuning ϵ is crucial. Stronger privacy (lower ϵ) adds more noise, further reducing accuracy.

5 Secure Multiparty Computation

5.1 Secure Multiparty Computation Overview

Secure Multiparty Computation (SMC) is a cryptographic technique used to compute a function across multiple parties while keeping each party's input private. In this task, SMC was applied during the aggregation phase, ensuring that the central server could aggregate client models without directly accessing their updates.

5.2 SMC in Federated Learning

In the basic FedAvg algorithm, the central server directly averages the model parameters from each client. However, to make this process more secure, SMC was implemented by securely aggregating the parameters using mathematical transformations. This allows model updates to be aggregated without revealing individual client contributions.

5.3 Results

SMC did not significantly impact the model's accuracy, but it strengthened the privacy guarantees of the system. After applying SMC with DP, the accuracy remained around 10%, consistent with the results from DP alone.

The key benefit of SMC is enhanced security during aggregation, but it did not introduce further degradation in accuracy compared to DP alone.

6 Comparison of Results

Model Type	**Test Accuracy (%)**
Basic Federated Learning	16
Federated Learning with DP and SMC	10

The results clearly show that adding privacy-preserving techniques like DP reduces the model's accuracy. The drop in accuracy from 16% (basic FL) to 10% (FL with DP and SMC) highlights the challenge of maintaining performance while ensuring privacy. SMC, while beneficial for secure aggregation, did not further affect accuracy but provided an additional layer of privacy during the aggregation phase.

7 Challenges and Observations

7.1 Accuracy Degradation

The most significant challenge was the drop in accuracy after introducing DP. The noise added to protect client data inevitably hindered the model's ability to learn effectively. Achieving the right balance between privacy and accuracy remains a difficult challenge.

7.2 Privacy-Accuracy Trade-off

The balance between privacy (ϵ) and model performance is a core issue in FL with DP. High privacy guarantees require more noise, which directly affects the model's predictions. Adjusting these parameters requires careful consideration of the privacy requirements of the system and the acceptable accuracy levels.

7.3 Computational Overhead

Although SMC provides stronger privacy guarantees, it introduces additional computational complexity during aggregation. However, this was not the primary bottleneck in the task, as the computational overhead was manageable with the scale of the CIFAR-10 dataset and the small number of clients.

8 Conclusion and Future Work

In this project, a Federated Learning system with Differential Privacy and Secure Multiparty Computation was implemented to evaluate the trade-offs between privacy and model performance. The results showed that while DP and SMC provide enhanced privacy guarantees, they come at the cost of accuracy, with the test accuracy dropping from 16% to 10%.

In conclusion, while privacy-preserving techniques are essential in modern FL systems, achieving an optimal balance between privacy and performance remains an ongoing challenge.

Note: we didn't implement functional encryption because of the huge computation time. Moving forward, we will implement a database system with a hashing algorithm to save computation time for checkpoint II.