Project Proposal - AFCFL: Adaptive Fairness Compensation-based Federated Learning

Asif Ahmed Utsa
Mobaswirul Islam
0423052090@grad.cse.buet.ac.bd
0424052053@grad.cse.buet.ac.bd
Bangladesh University of Engineering & Technology
Dhaka, Bangladesh

1 PROBLEM STATEMENT

Federated Learning (FL) is a way for many devices or organizations to train one shared global model without sharing raw data with a central server. Although the data remain private, the final model may become *unfair* or the whole training process can be *biased*. The reasons are:

(1) **Heterogeneous Environment**
Slow or busy devices are picked less often for training rounds, so their data have less influence on the model.

(2) **Non Identical Dataset**
Difference between the sizes of clients' dataset also have the overall model. Because the clients with larger dataset plays more impact on the overall training.

(3) **Accuracy gap**
Clients with similar data get good local accuracy, but clients with rare or very different data get poor accuracy.

FCFL is a new method that reduces *accuracy gap* by keeping a fairness queue for every client. However, FCFL still uses fixed settings chosen by hand and does not directly address the "chance to join" issues.

Project Goal: We will use FCFL as a starting point, but we aim to build a small extension that:

- adjust FCFL's two key settings (α for queue growth, and r for random selection) automatically during training, and
- add another rule so every client gets roughly the same chance to take part in training.

We want to show that we can cut the accuracy gap *and* make participations more equal, while keeping the average accuracy high.

2 MOTIVATIONS

Fairness in FL is wider than just closing the accuracy gap. A practical system should work well for every client, give each device a fair chance to take part and avoid pushing all the workload onto one group of machines. Our project focuses on these reasons for the extra steps we propose:

(1) **Adaptive settings**
Training dynamics change over time because the data seen by the model evolve and the pool of active devices is never the same from one round to the next. Letting the algorithm tune α (queue growth) and r (random share) on its own removes manual grid search, shortens deployment time, and makes the method self-correcting if conditions drift after launch. Also, manually searching the grid to find values of α and r is time-consuming.

(2) **Selection fairness**
Equal accuracy means little if some clients rarely join. A second queue that rewards long-waiting clients balances participation, spreads communication cost, and keeps the training data stream rich in diversity, which can also improve overall accuracy in the long run.

(3) **Room for other goals**
The queue idea is flexible. Once it works perfectly, new fairness terms—such as **demographic**, **energy**, or **privacy fairness**—can be plugged easily.

3 EXPERIMENT PLAN

We break the project into the following steps:

(1) **Rebuild FCFL**
Re-implement the original accuracy queue with fixed α and r, and confirm the reproduced results on a quick non-IID split so that later changes can be compared fairly.

(2) **Add adaptive rules**
We will then try to implement define the values of α and r adaptively during training time.

(3) **Insert selection queue.**
We will try to keep track of the clients in which round who are getting selected or not. Based on that, we will try to give prioritize he less selected ones.

(4) **Run experiments.**
We will then run experiments on our initially written FCFL and Adaptive FCFL to compare whether the new one can actually compete with the previous one or not.

AML Project, October 2024, CSE, BUET