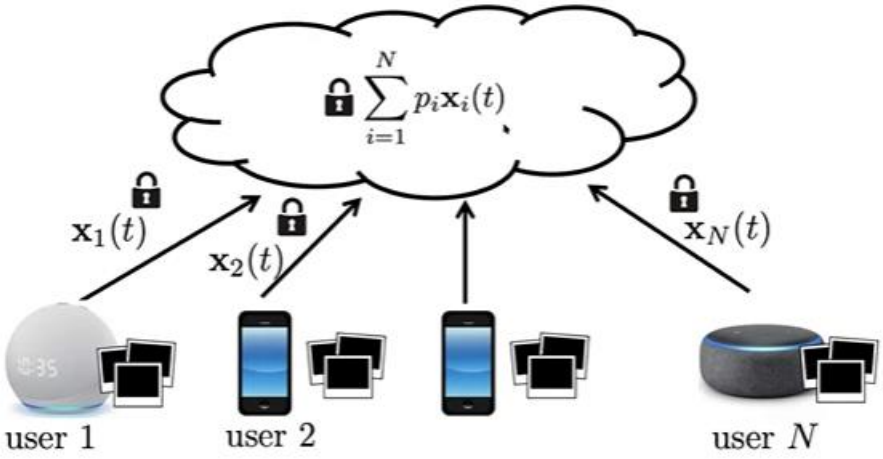
Federated Learning

Main Principle: Train Locally – Average Globally



Foundations

- Model Aggragation
- Data Heterogenity
- No/Weak Labels – Unsupervised FL

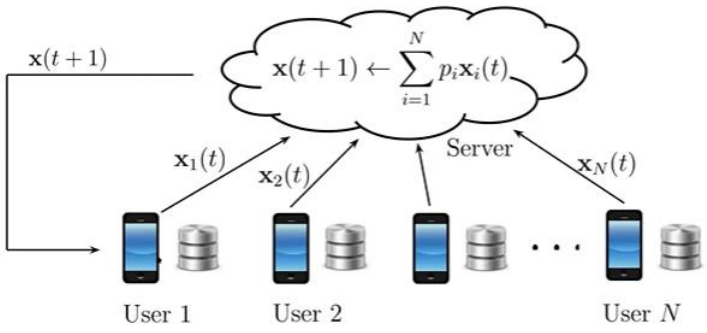
Scalability

- Resource Constrained FL (Small edge models, large server model)
- Convergence: 1K users to 1M Users
- Federated neural architecture search.

Trustworthiness

- Secure and resilient model aggregation
- Adversarial users (data/model poisoning)
- Leveraging trusted computing environments

Ensuring Privacy by avoiding Data Movement from the users



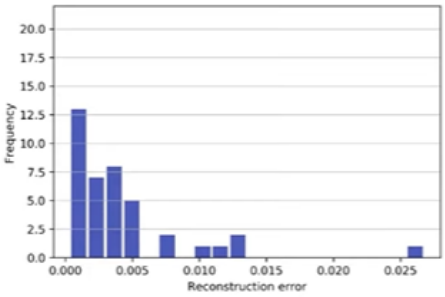
Secure aggregation is a **multi-round secure MPC problem with user dropouts.**

Partial user participation leads to **privacy leakage.**

- Random selection may reveal all individual models.
- Exp
  - N=40 users
  - MNIST dataset with non iid distribution
  - K=8 users are selected at random at each round
  - The server estimates the individual gradients through least-squares.

Reconstruction error of user i at time t

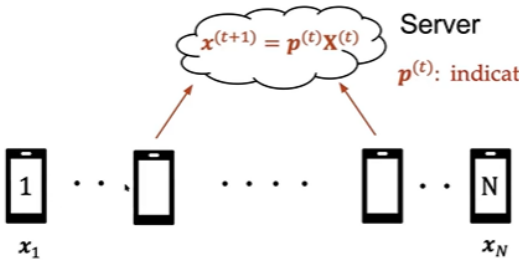
$$\frac{\left\|\boldsymbol{x}_i^{(t)}-\widehat{\boldsymbol{x}}_i^{(t)}\right\|^2}{\left\|\boldsymbol{x}_i^{(t)}\right\|^2}$$



reconstruction error after 600 rounds

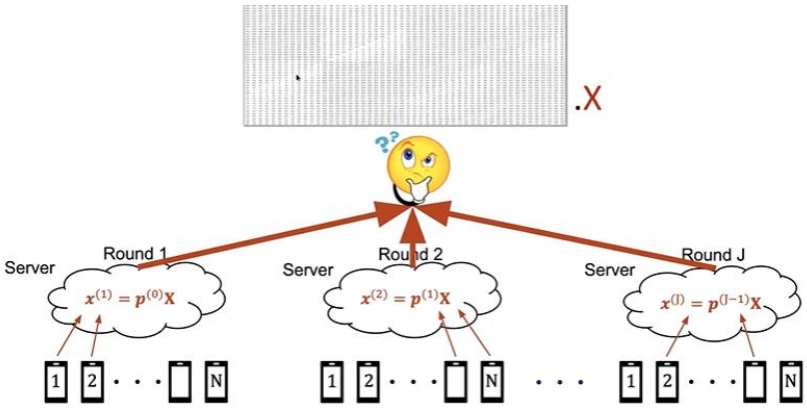
reconstruction error < 0.005 (for many users)

Federated averaging with partial user participation



- Participation matrix $\boldsymbol{P}^{(J)} = \begin{pmatrix} \boldsymbol{p}^{(0)} \\ \vdots \\ \boldsymbol{p}^{(J-1)} \end{pmatrix} \in \{0,1\}^{J \times N}$, J: number of rounds

Method 1:Multi-round Privacy (T)



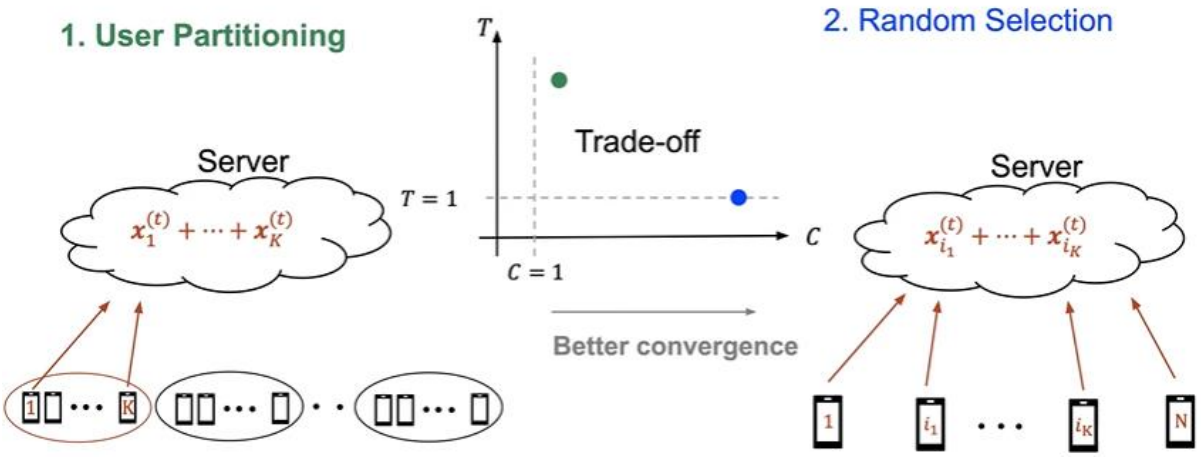
$a_1 \sum_{j\in\mathcal{S}_1} \boldsymbol{x}_j + a_2 \sum_{j\in\mathcal{S}_2} \boldsymbol{x}_j + \cdots + a_n \sum_{j\in\mathcal{S}_n} \boldsymbol{x}_j$, where $|\mathcal{S}_i| \geq T$.

Baseline

1. User partitioning
   - Large multi-round privacy ***T=group size***
   - In many rounds, however, no groups are available.
2. Random Selection
   - Small multi-round privacy ***T=1***
   - Any subset of available users can be selected in any round.

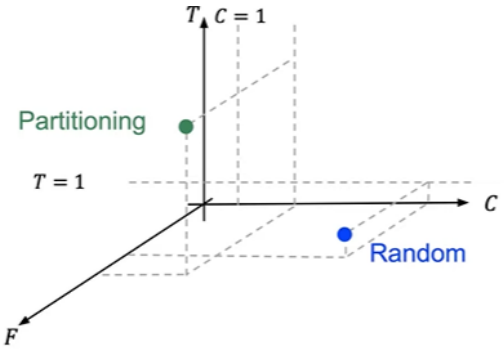
Method 2: Average Aggregation Cardinality (C)

C = Average number of participating users over all rounds



Method 3: Aggregation Fairness Gap (F)

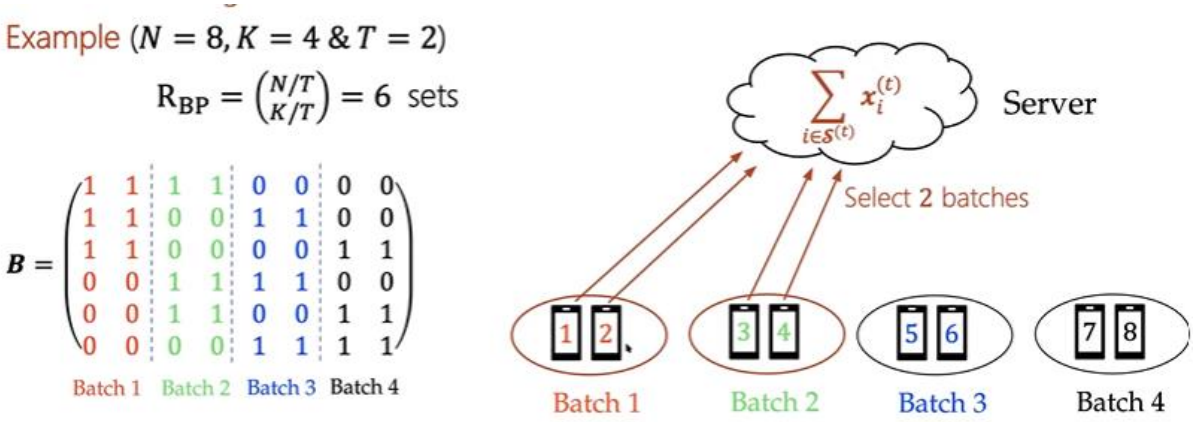
- Aggregation Fairness Gap F
- F = max. average participation frequency – min. average participation frequency



Proposed Approach: Multi Round Sec Agg

1. **Batch Partitioning**
   - **Idea:** Partition users into T-user batches; allow selection of any K/T available batches

- **Input**: N, K<= N, 1<= T <= K
- **Output**: A family of K User sets satisfying the multi-round privacy T
  - This Family is represented by a matrix B.



2. **Available batch selection to guarantee fairness.**
   - **Idea:** Select based in the minimum frequency of participation
   - **Input:** Set of available users at round t and B
   - **Output:** Set of users that will participate at round t.