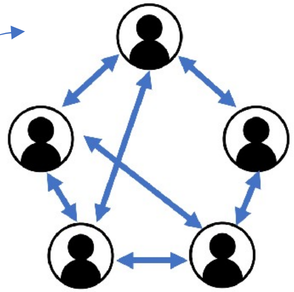
Enhancing Federated Learning with Robust Aggregation and Sequential Orchestration

Ayoub Saidane

UC Berkeley, SkyLab

Research Internship - April 2024 to August 2024

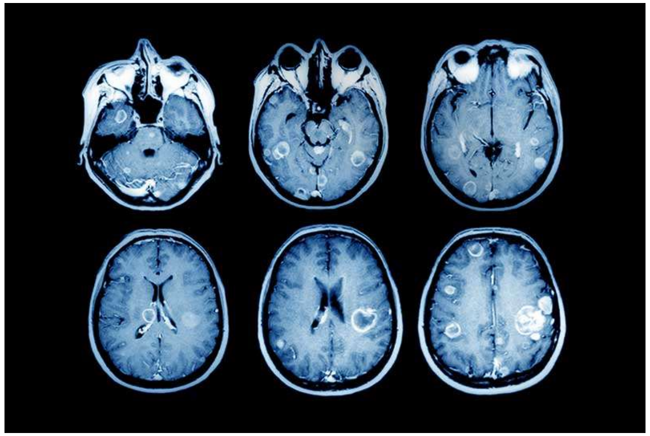
Introduction



Federated Learning of Gboard Language Models with Differential Privacy

Zheng Xu* Yanxiang Zhang* Galen Andrew Christopher A. Choquette-Choo
Peter Kairouz H. Brendan McMahan Jesse Rosenstock Yuanbo Zhang

Google



Introduction

Difficulties

- Clients can be malicious
- Data distribution is heterogeneous

Byzantine Machine Learning: A Primer

RACHID GUERRAOUI, NIRUPAM GUPTA, AND RAFAEL PINOT,
École polytechnique fédérale de Lausanne, Switzerland

Accelerating Federated Learning via Sequential Training of Grouped Heterogeneous Clients

ANDREA SILVI, ANDREA RIZZARDI, DEBORA CALDAROLA, BARBARA CAPUTO, AND MARCO CICCONE
Dipartimento di Automatica e Informatica (DAUIN), Politecnico di Torino, 10129 Turin, Italy

Convergence Analysis of Sequential Federated Learning on Heterogeneous Data

Yipeng Li and Xinchen Lyu *
National Engineering Research Center for Mobile Network Technologies
Beijing University of Posts and Telecommunications
Beijing, 100876, China
{liyipeng, lvxinchen}@bupt.edu.cn

Contributions

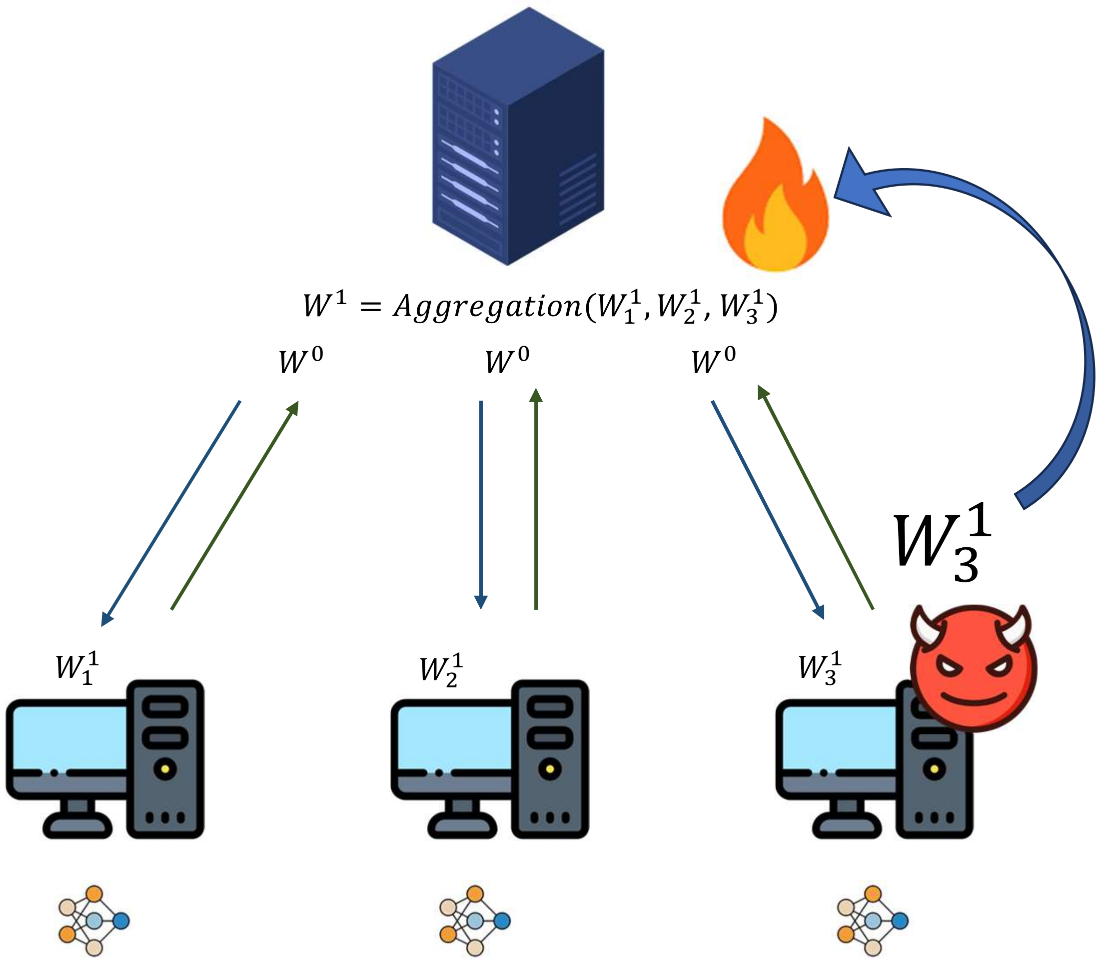
Robust Aggregation

- Description of aggregation rules: CC, CGE, MeaMed, CwTM, CwMed.
- Experiments and performance comparision

Sequential Training

- Implementation of fully sequential and hybrid sequential-parallel algorithms.
- Experiments in homogeneous/heterogeneous settings.

Parallel training



Byzantine Attacks Tested

- **Sign Flipping Attack**
  - Malicious clients invert the signs of their model updates.
  - Goal: Disrupt model convergence by misleading the aggregation process.
- **Label Flipping Attack**
  - Clients corrupt their training data by flipping labels (e.g., class 0 to class 1).
  - Impact: Causes the global model to learn incorrect associations, degrading accuracy.
- **Gaussian Attack**
  - Clients add Gaussian noise to their model updates.
  - Effect: Introduces randomness, slowing convergence or leading to suboptimal solutions.

Aggregation Rules Summary

n: number of clients
f: number of malicious clients

- **Centered Clipping (CC)**:
  - Clips model weights to a threshold c to limit the influence of extreme values.
  - Formula:

$$v_m \leftarrow v_{m-1} + \frac{1}{n} \sum_{i \in [n]} (w_i - v_{m-1}) \min\left\{1, \frac{c}{\|w_i - v_{m-1}\|}\right\}$$

- **Comparative Gradient Elimination (CGE)**:
  - Sorts model weights by their norm, and outputs the average the top $n - f$ weights, filtering out potential outliers.
  - Formula:

$$\mathrm{CGE}(w_1, \ldots, w_n) = \frac{1}{n-f} \sum_{i=1}^{n-f} w_{p(i)}$$

- **Mean around Median (MeaMed)**:
  - Computes the average of the $n - f$ closest elements to the median coordinate-wise, reducing the influence of outliers.
  - Formula:

$$[\mathrm{MeaMed}(w_1, \ldots, w_n)]_k = \frac{1}{n-f} \sum_{i \in C_k} [w_i]_k$$

- **Coordinate-wise Trimmed Mean (CwTM)**:
  - Sorts each coordinate of the model weights, then averages after trimming the top f values.
  - Formula:

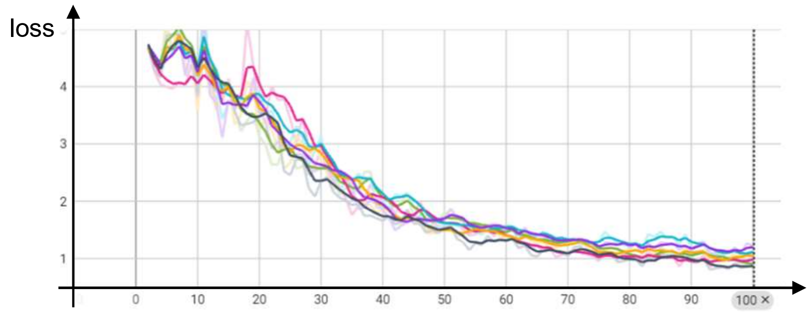
$$[\mathrm{CwTM}(w_1, \ldots, w_n)]_k := \frac{1}{n-2f} \sum_{j=f+1}^{n-f} [w_{p_k(j)}]_k.$$

- **Coordinate-wise Trimmed Median (CwMed)**:
  - Computes the median for each coordinate of the model weights, offering robustness to outliers.
  - Formula:

$$[\mathrm{CwMed}(w_1, \ldots, w_n)]_k := \mathrm{Median}([w_1]_k, \ldots, [w_n]_k).$$

Experimental Setup

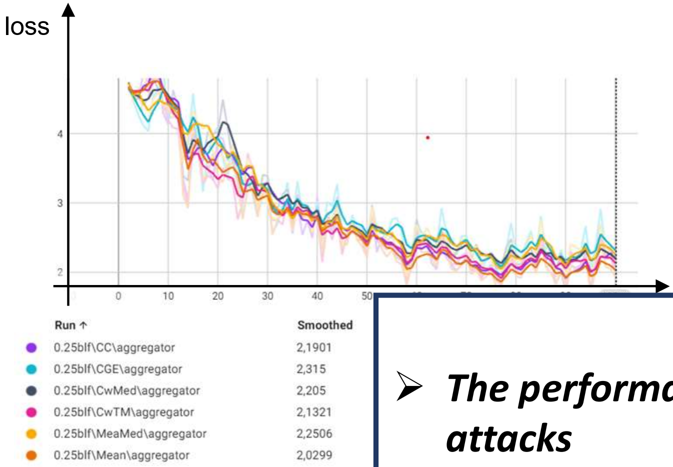
- **Framework:** Implemented an extension of FedScale, an open-source FL framework.
- **Dataset: FEMNIST**
  - Federated version of MNIST with 62 classes (digits, uppercase, and lowercase letters).
  - Simulates a realistic FL scenario with non-IID data distribution across clients.
- **Model Architecture: ResNet-18**
  - Chosen for its balance between depth and computational efficiency.
- **Training Configuration:**
  - 100 rounds of training with evaluation every 5 rounds.
  - Each round involves 50 participants with local training using a batch size of 20 over 5 local steps.
  - Learning rate set to 0.05.

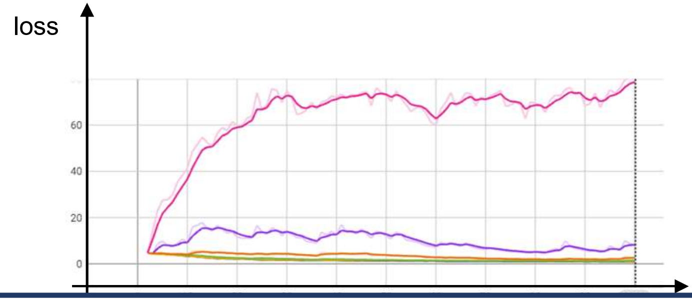


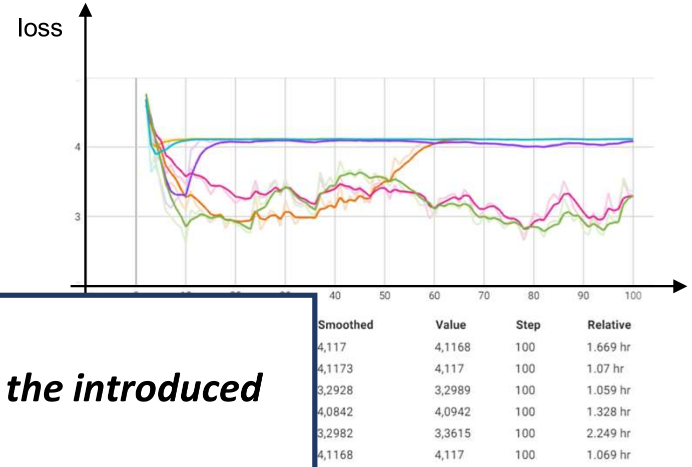
Run ↑	Smoothed	Value	Step	Relative
0b\CC\aggregator	0,9048	0,8806	100	1.698 hr
0b\CGE\aggregator	0,9826	1,0164	100	1.079 hr
0b\CwMed\aggregator	1,1137	1,1524	100	1.083 hr
0b\CwTM\aggregator	1,0463	1,0334	100	10.64 hr
0b\MeaMed\aggregator	1,1974	1,2162	100	2.194 hr
0b\Mean\aggregator	0,8655	0,8713	100	1.02 hr

Label flipping | Gaussian | Sign flipping

25% byzantine

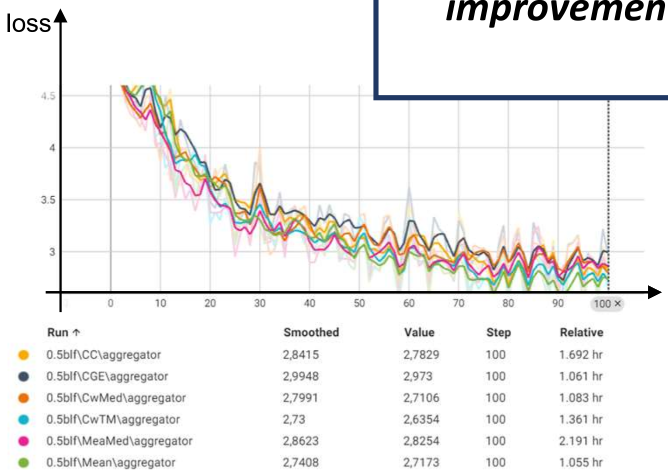


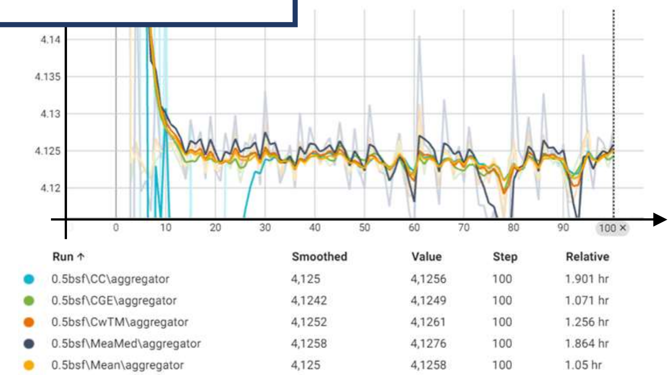




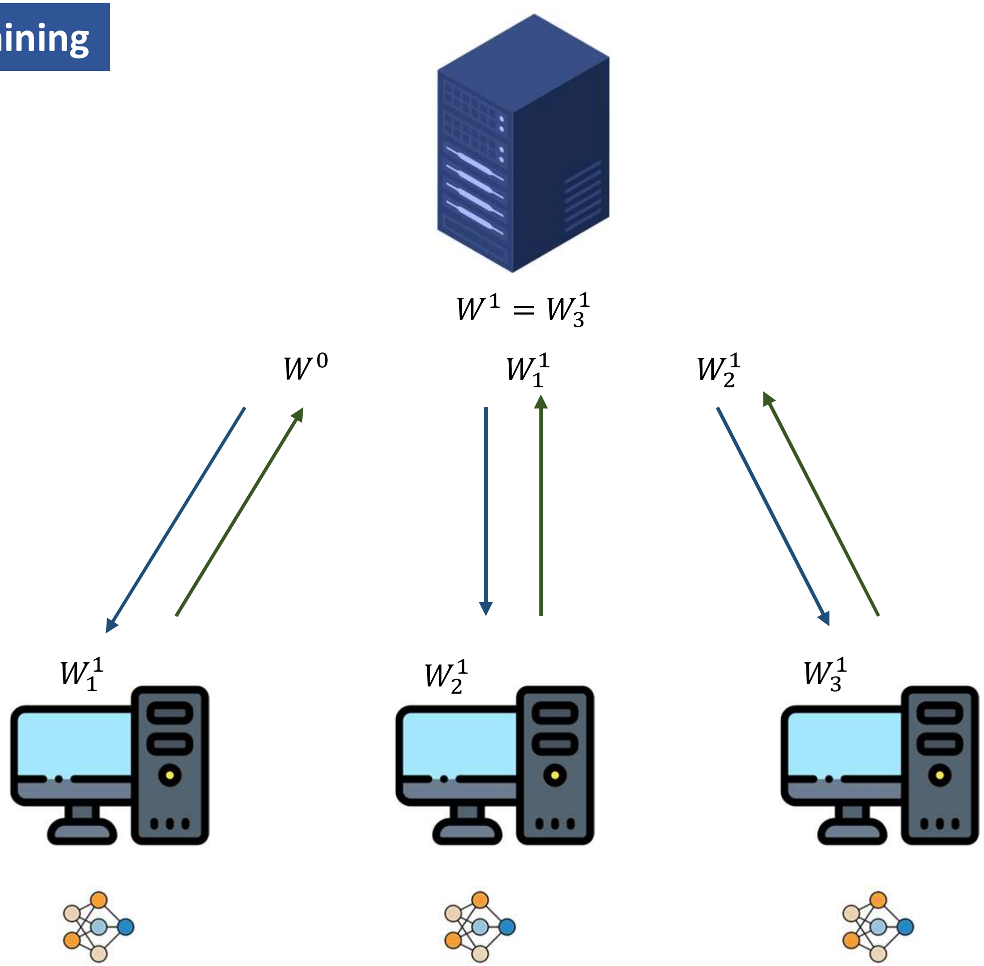
- ➢ ***The performance of the aggregation rules depends on the introduced attacks***
- ➢ ***Centered Clipping seems to be a promising technique for further improvement***

50% byzantine

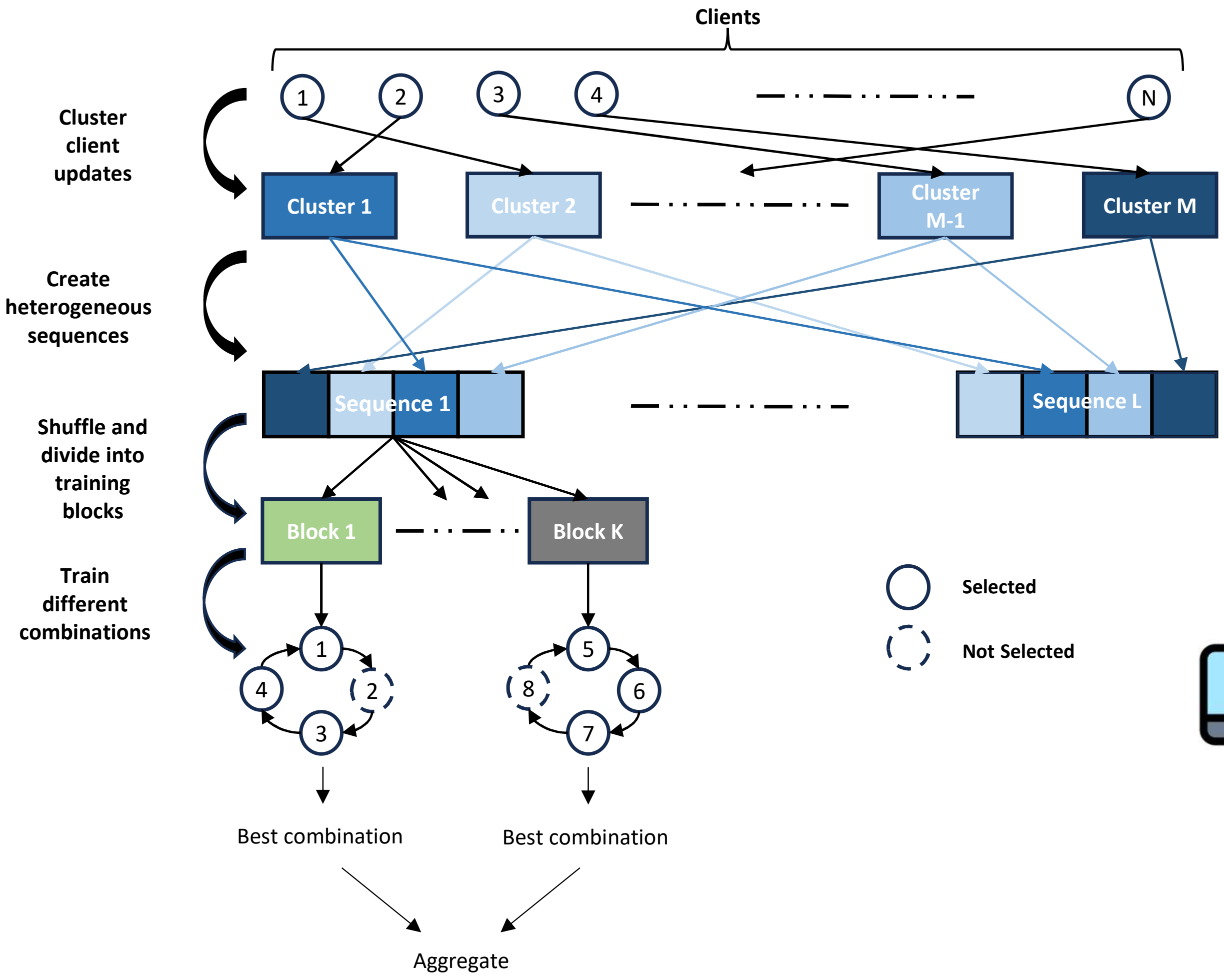


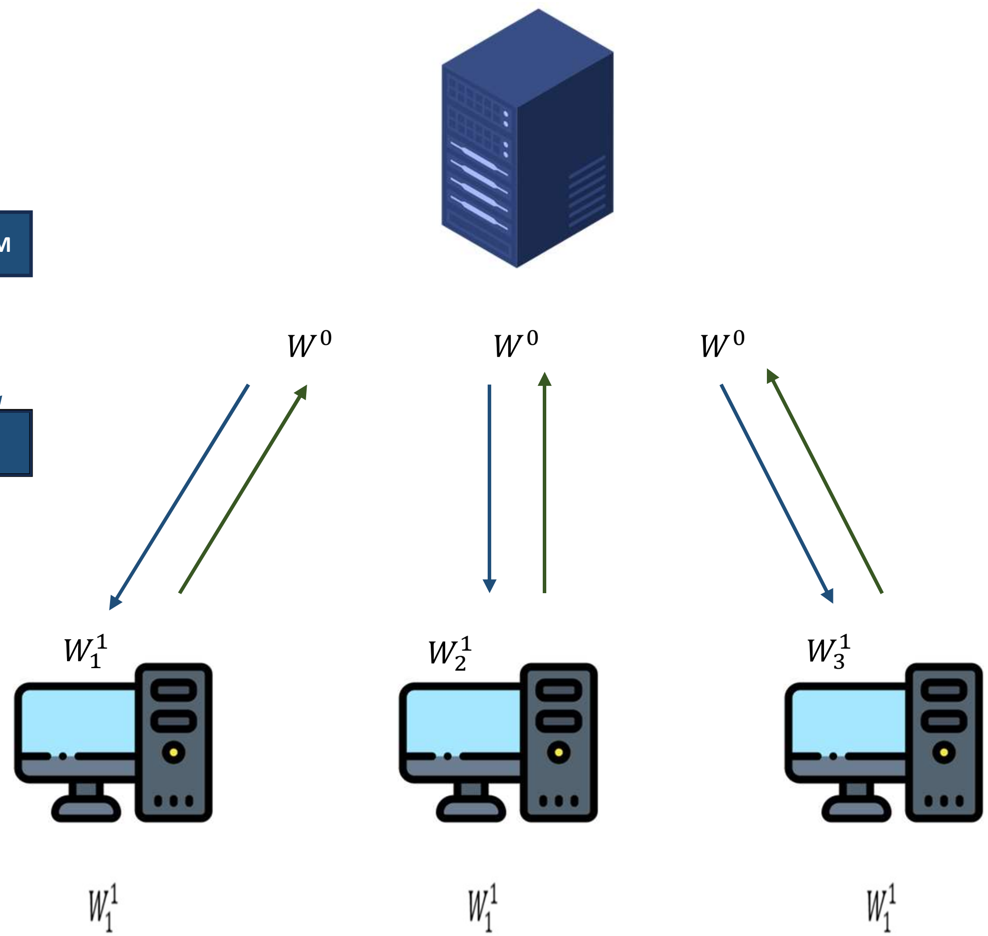


Fully sequential training



Sequential training



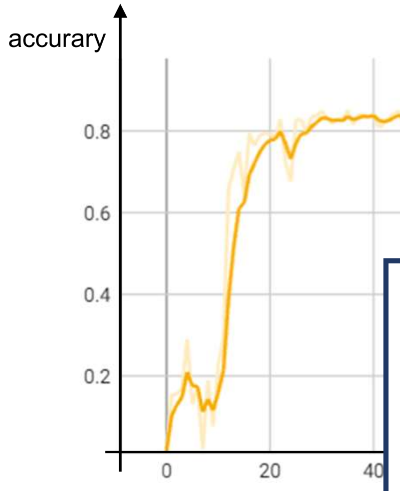


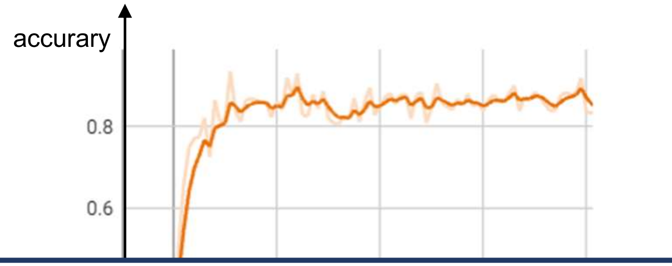
Experimental Setup

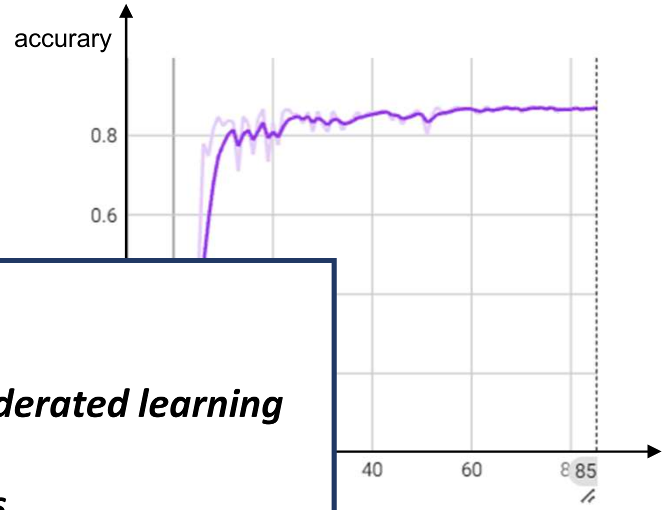
- **Dataset: FEMNIST**
  - Trained ResNet-18 on FEMNIST with homogenous and heterogeneous label distributions across clients.
  - Labels allocated based on Gaussian distribution with mean μ_n and constant standard deviation σ.
  - Three different σ values tested to compare training performance.
- **Training Settings**
  - 12 participating clients.
  - Local training for 1 epoch with a learning rate of 0.01 over 100 rounds.
- **Performance Metrics**
  - Global Accuracy
  - Individual Accuracy

	Parallel	Fully sequential	Sequential

Homogeneous

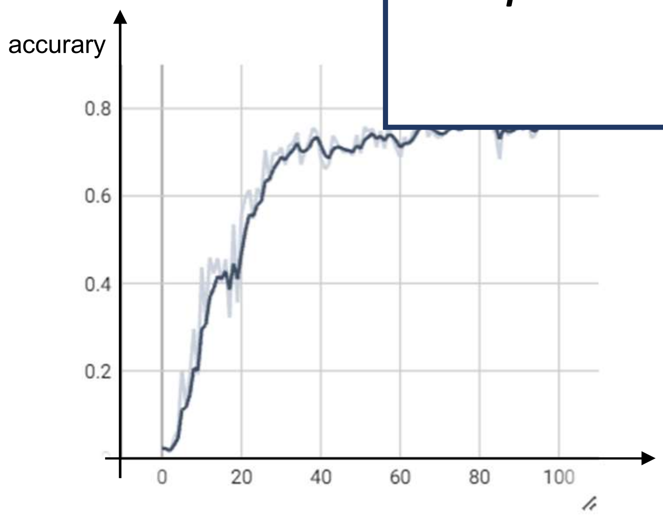


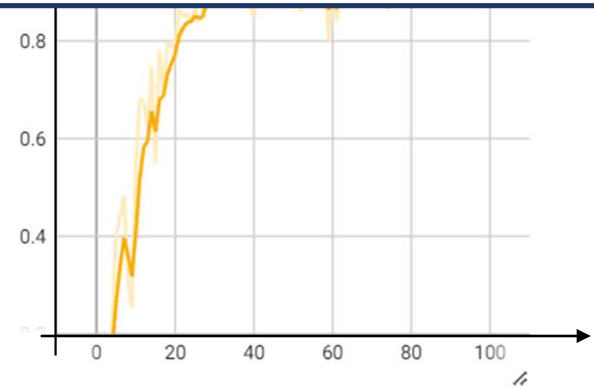


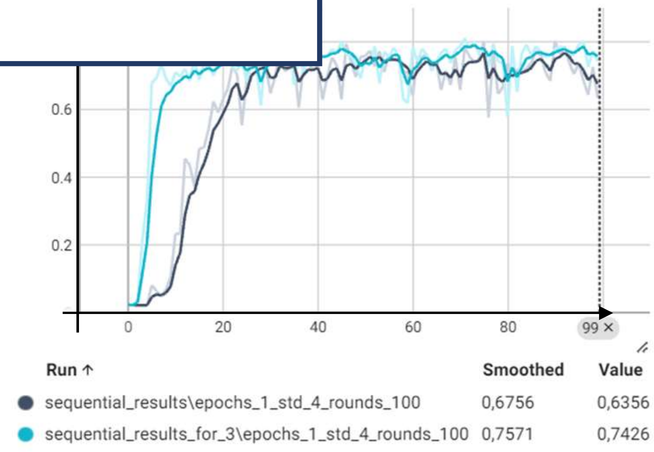


- ➢ ***Heterogeneity is a challenging problem to solve in federated learning***
- ➢ ***Sequential approches seem to give promissing results***

Heterogeneous







Conclusion

Summary of Findings

- Robustness of CC
- Effectiveness of sequential training in heterogeneous data environments.

Recommendations

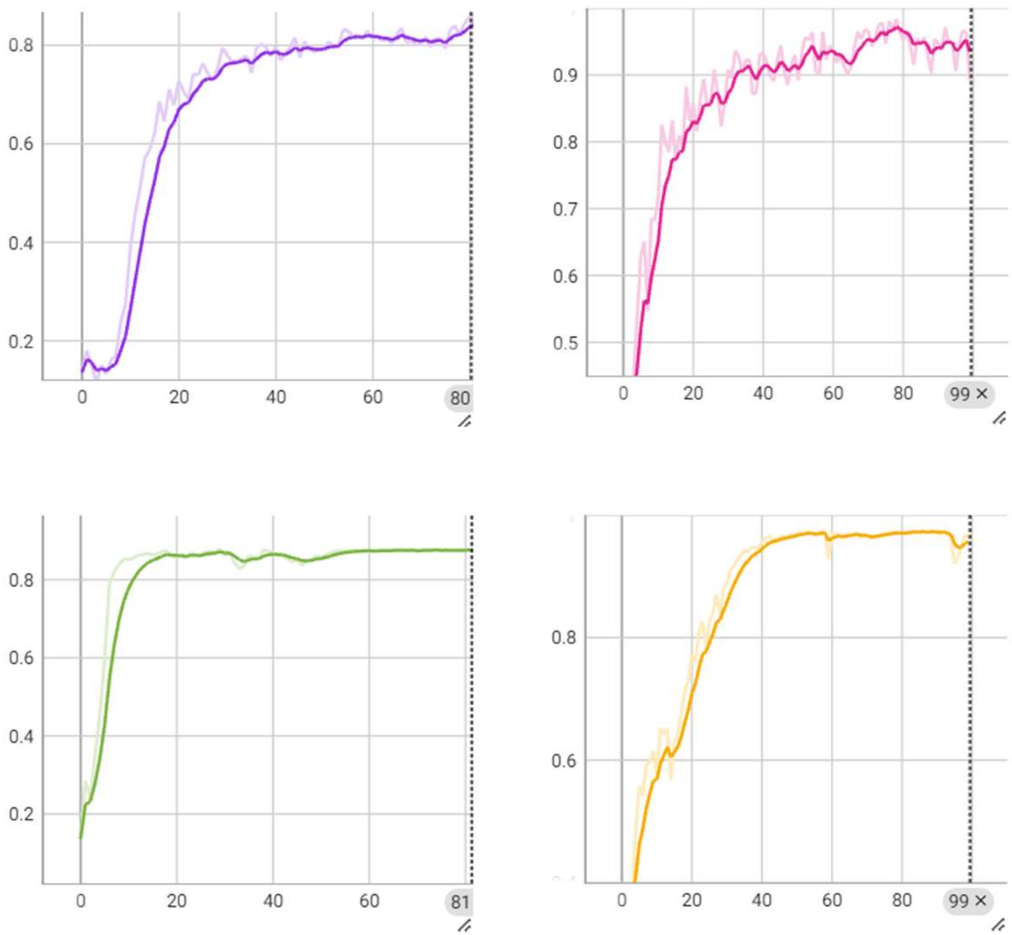
- Adoption of CC in Byzantine-prone environments.
- Further research on sequential training optimization.

individual

parallel | Sequential

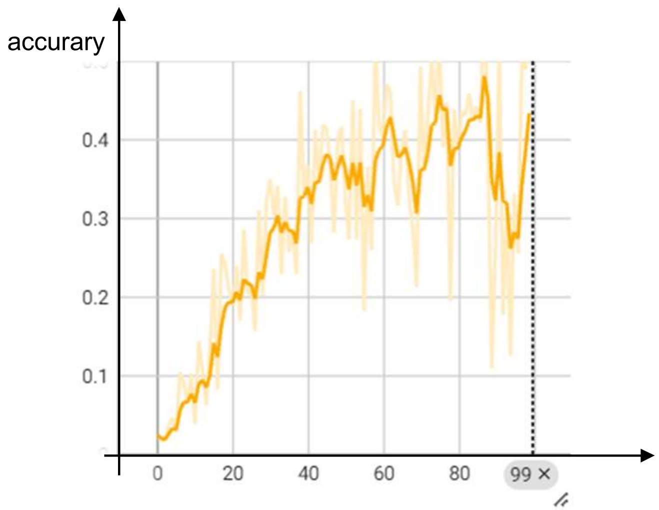
homogeneous

Heterogeneous

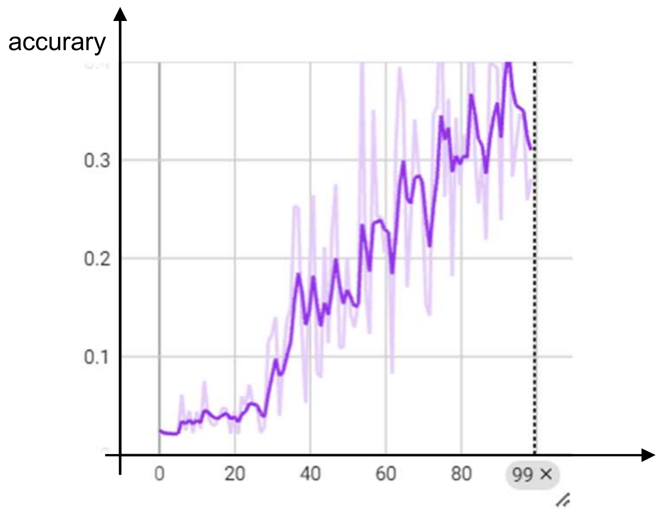


Label Flipping

Parallel



Fully sequential



Sequential