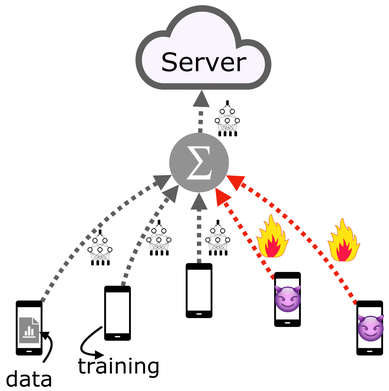
Robust Aggregation for Federated Learning

Krishna Pillutla, Sham Kakade, Zaid Harchaoui

University of Washington · Google Research · Harvard

Robust federated learning

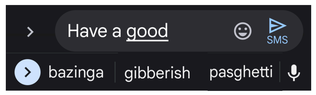


Our Setting

Some clients are adversarial:

- Data poisoning
- Model poisoning

$\Longrightarrow$ Poor predictions for **all** clients



Objective $\min_{w} \sum_{i \in \mathcal{D}} F_i(w)$

$\mathcal{D}$: Set of "inlier" clients

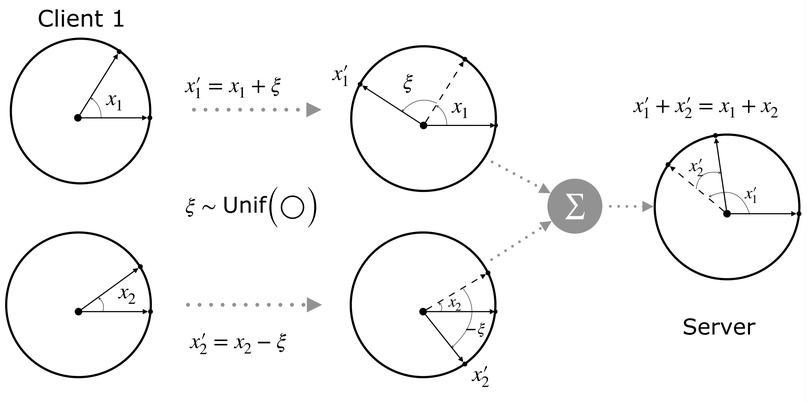
F_i: Loss on client i

Goals of the work

A robust aggregation approach that is

1. Communication-efficient
2. Implementable via **secure summation**

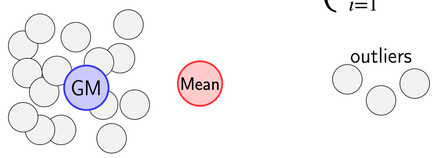
Server only sees $x_1', x_2' \sim \text{Unif}(\bigcirc)$ but calculates the correct sum $x_1 + x_2 = x_1' + x_2'$



[Bonawitz et al. CCS (2017), Bell et al. CCS (2020)]

RFA: Geometric Median Aggregation

$$\text{GM}(w_1, \cdots, w_m) = \arg\min_z \left\{ \sum_{i=1}^{m} \|z - w_i\|_2 \right\}$$

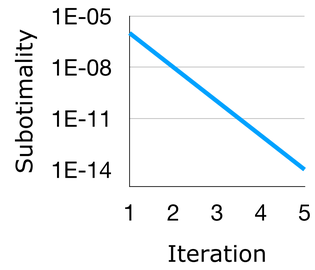


Robustness: Breakdown point= 1/2

Weiszfeld's Algorithm:

$$\beta_{i,t} = 1/\max\{\|z_t - w_i\|_2, \nu\}$$

$$z_{t+1} = \frac{\sum_i \beta_{i,t} w_i}{\sum_i \beta_{i,t}}$$



Theory (Least Squares)

Suppose $Y_i = X_i^\top w_i^\star + \xi_i$ where $\xi_i \sim \mathcal{N}(0, \sigma^2)$

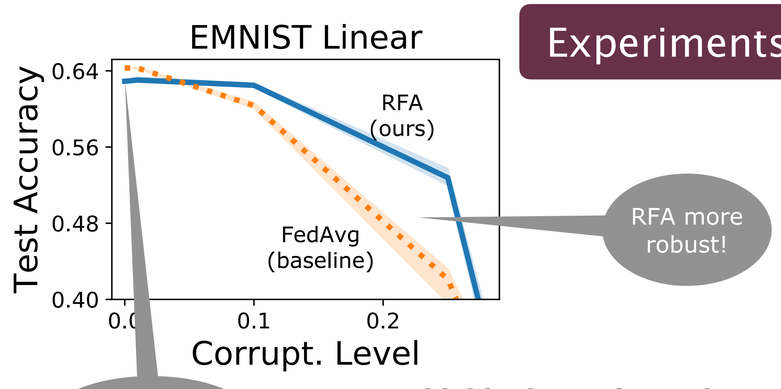
Theorem: Assume that $F(w)$ is strongly convex, $\|X_i\| \le 1$ and # local steps $\propto 2^t$. Let $\mathscr{E}$ denote the event that $\ge 1/2 + c/2$ non-corrupted devices are chosen (out of the $1/2 + c$ available) in each round.

Then, RFA with ε-approximate GM satisfies

$$\mathbb{E}\left[\|w_t - w^\star\|^2 \middle| \mathscr{E}\right] \lesssim \frac{\|w_0 - w^\star\|^2}{2^t} + \frac{1}{c^2}\left(d\sigma^2 \frac{t}{2^t} + \frac{\varepsilon^2}{m^2} + \Omega_X^2 \Omega_Y^2\right)$$

Optimization error · Statistical error · GM Approx. Error · **Heterogeneity Error**

Experiments



Unavoidable due to lower bounds of robust mean estimation

[Chen, Gao, Ren. Annals of Stat. (2018)]

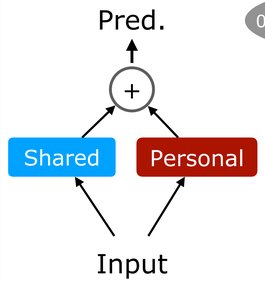
Handling Heterogeneity

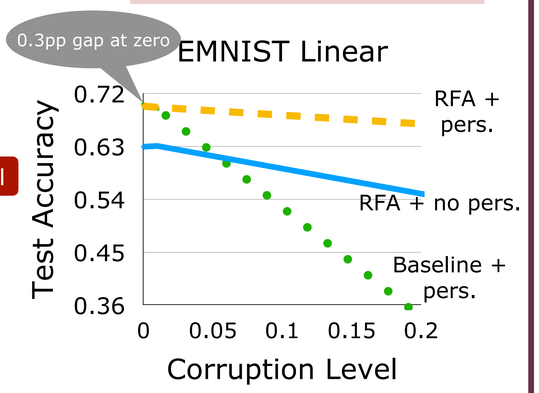
Personalize parts of the model

$$\min_{u, v_1, \cdots, v_n} \sum_{i \in \mathcal{D}}^{n} F_i(u, v_i)$$

u: shared parameters

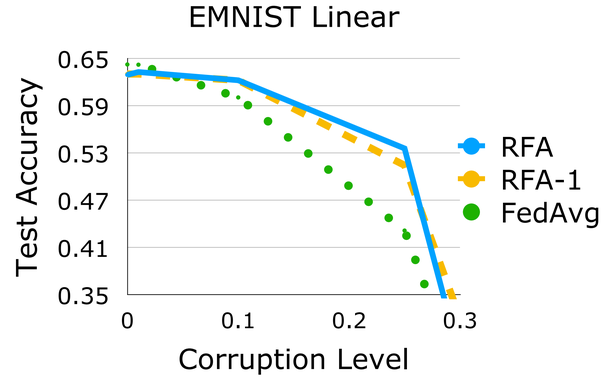
v_i: personal parameters





Improving the Communication Cost

Single Weiszfeld iteration is also robust!



Software for the geometric median

Install: **pip install geom-median**

Documentation: github.com/krishnap25/geom_median

krishnap25

KrishnaPillutla

krishnap25.github.io

Code

SCAN ME