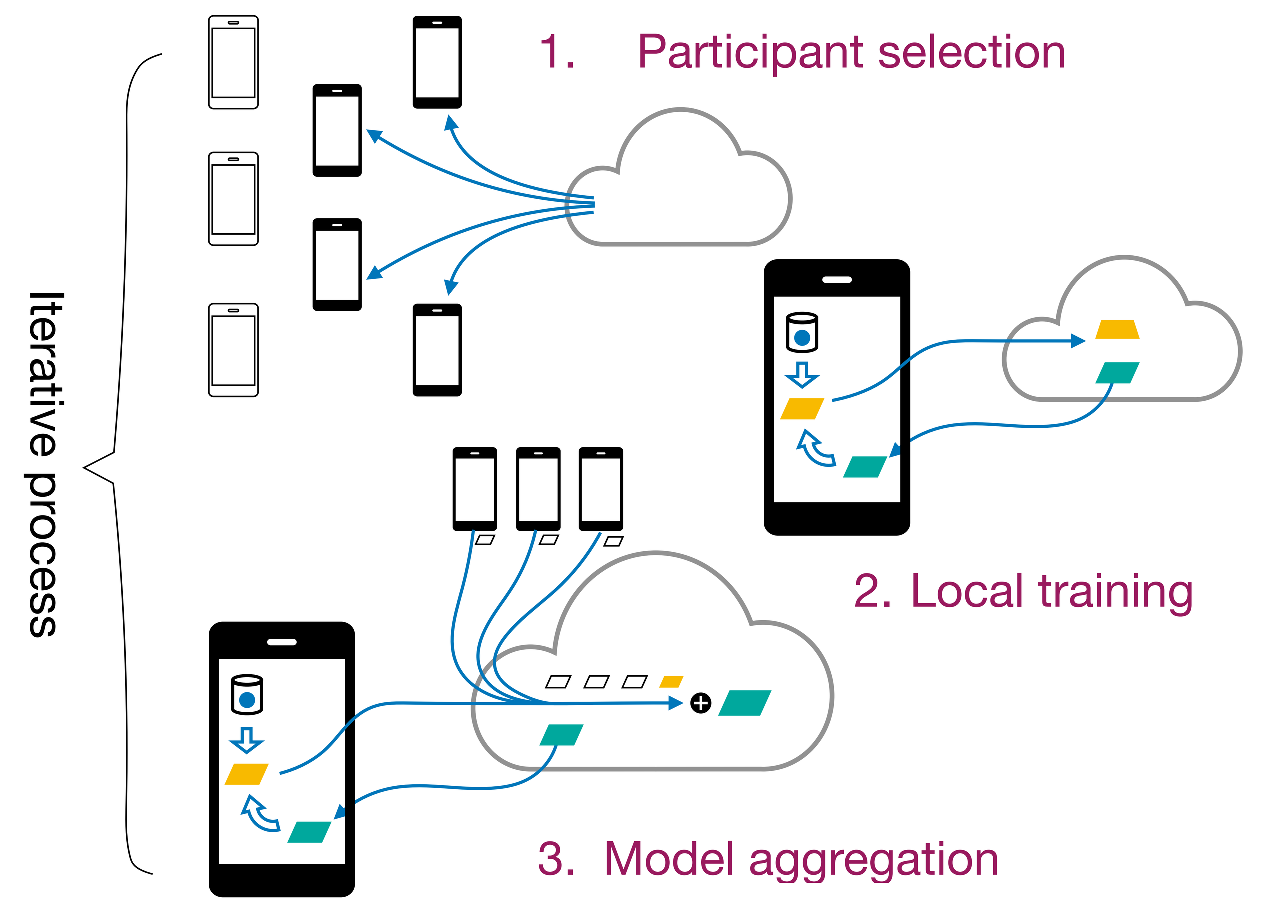
Dordis: Efficient Federated Learning with Dropout-Resilient Differential Privacy

Zhifeng Jiang[1], Wei Wang[1], Ruichuan Chen[2]
[1]HKUST [2]Nokia Bell Labs

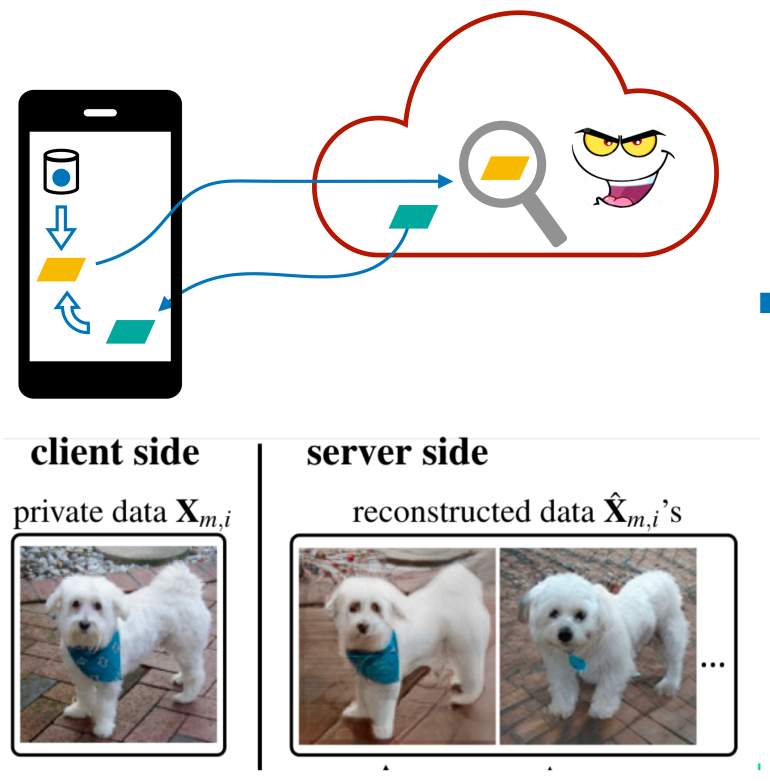
Federated Learning (FL)

FL allows distributed clients to train a shared model without exchanging their private data.



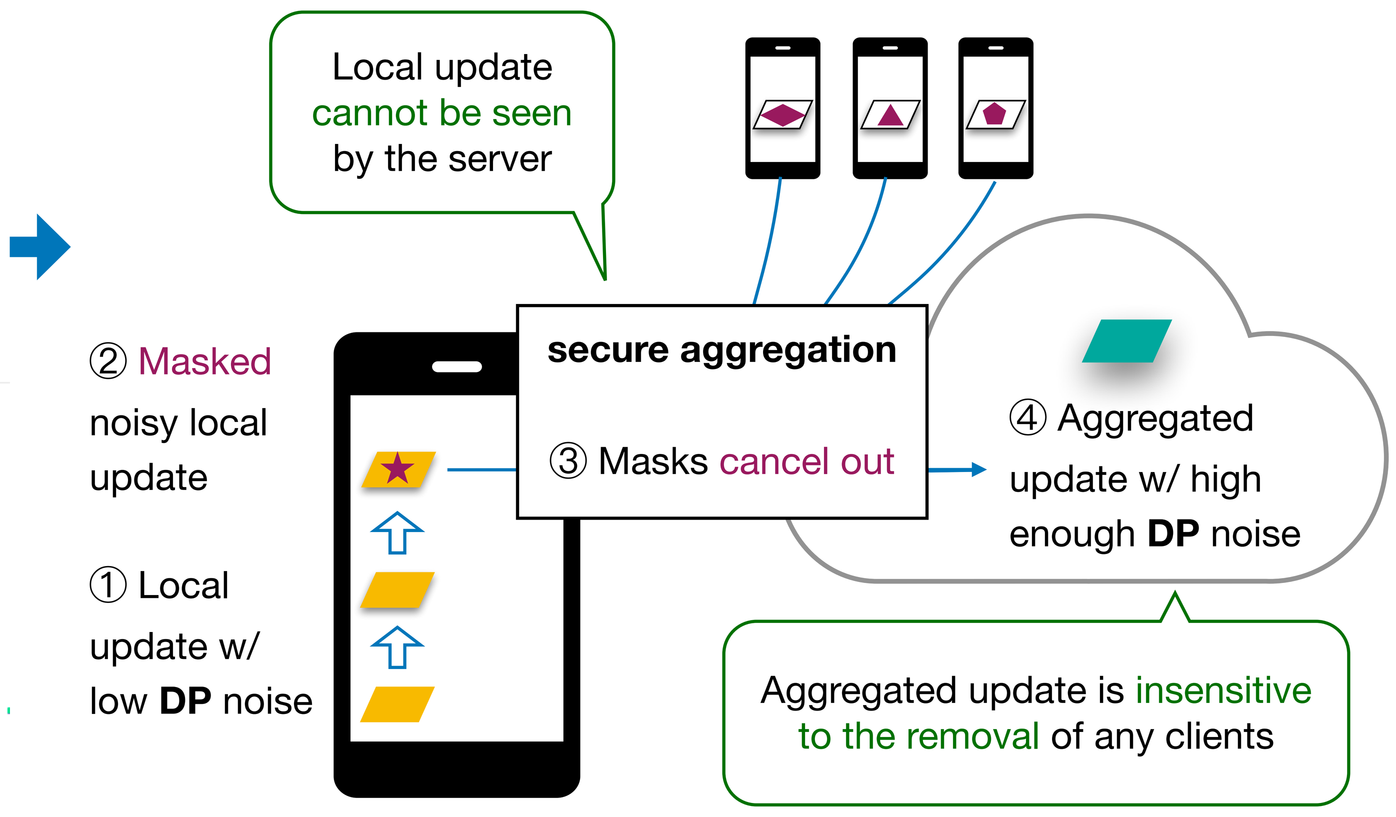
Distributed Differential Privacy (DP) for FL

Local updates can still leak private information.



e.g., data reconstruction attacks[1] are shown to be feasible in FL practice

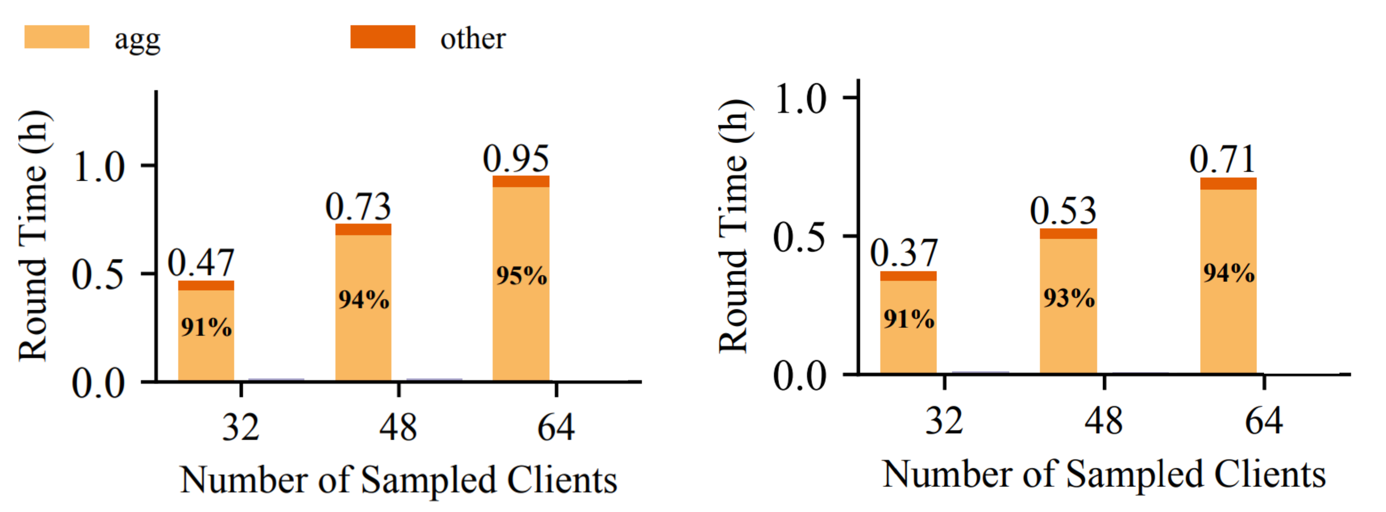
Higher privacy can be achieved by Distributed DP[2,3] = secure aggregation + DP



Pipeline-Parallel Acceleration

Inefficiency of Operations in Secure Aggregation

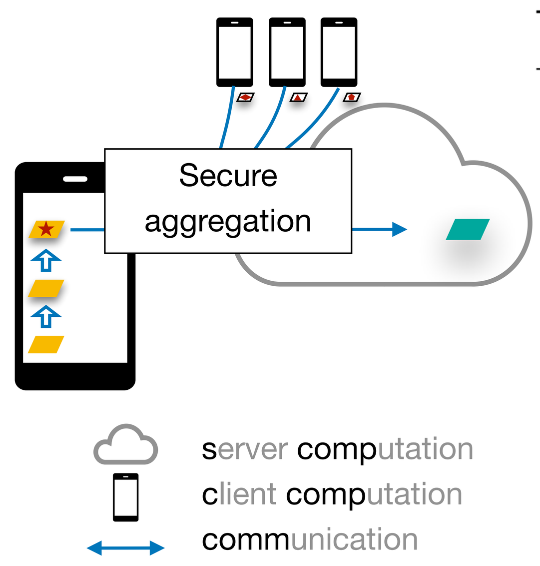
Expensive cryptographic primitives in **SecAgg**[4] dominates run time



SecAgg+[5] improves little in the common small-scale FL practice

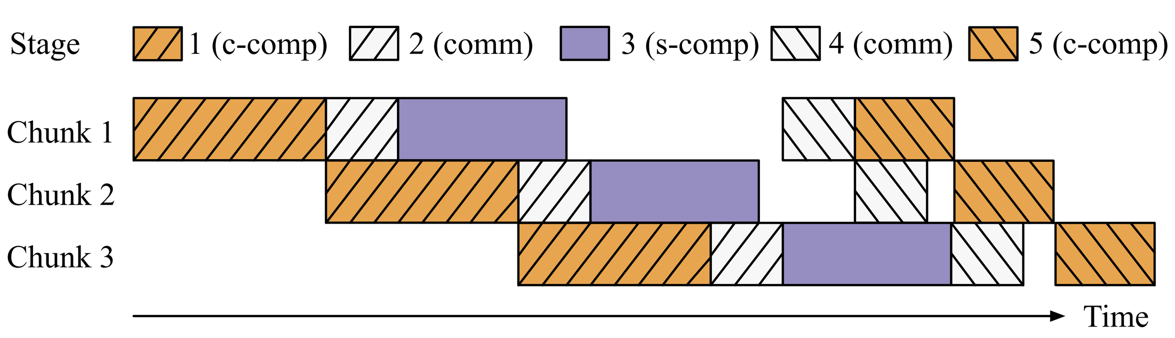
Pipelining Chunk-Aggregation

① Identify dominant resources at each step & group consecutive steps that use the same resource



Step	Operation	Stage (Resource)
1	Clients encode updates.	1 (c-comp)
2	Clients generate security keys.	
3	Clients establish shared secrets.	
4	Clients mask encoded updates.	
5	Clients upload masked updates.	2 (comm)
6	Server deals with dropout.	3 (s-comp)
7	Server computes aggregate update.	
8	Server updates the global model.	
9	Server dispatches the aggregate.	4 (comm)
10	Clients decode the aggregate.	5 (c-comp)
11	Clients use the aggregate.	

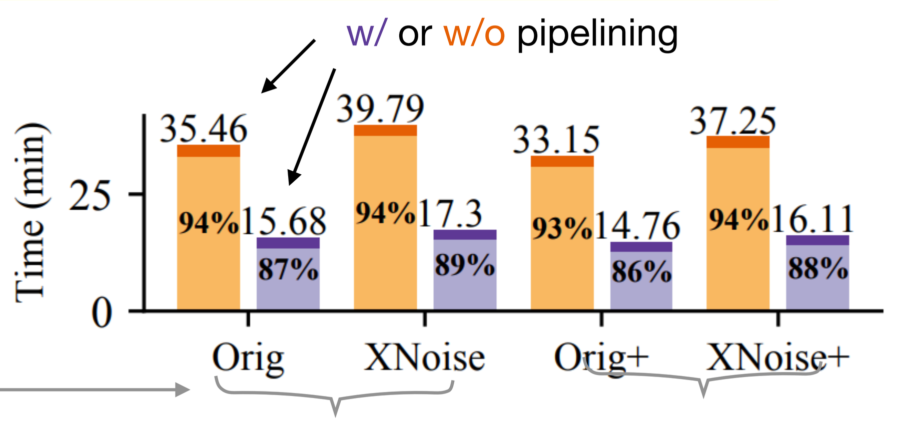
② Pipelining chunk-aggregation with optimal # chunks



Empirically Achieving Up to 2.4× Speedup

Case study: CIFAR10 @ VGG19, dropout rate = 30%

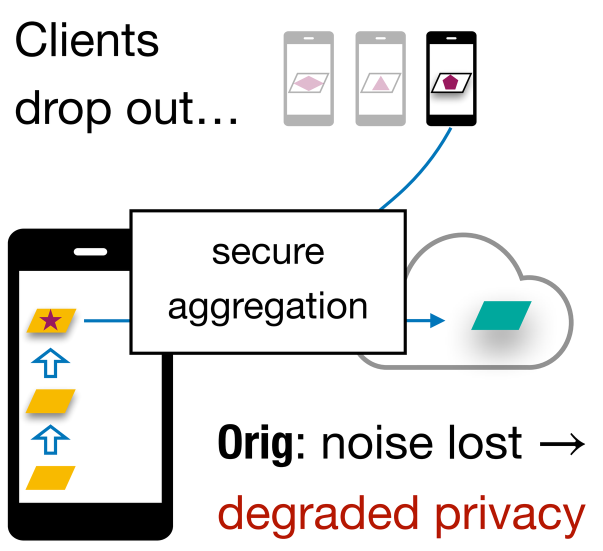
Implemented using **SecAgg** or **SecAgg+**

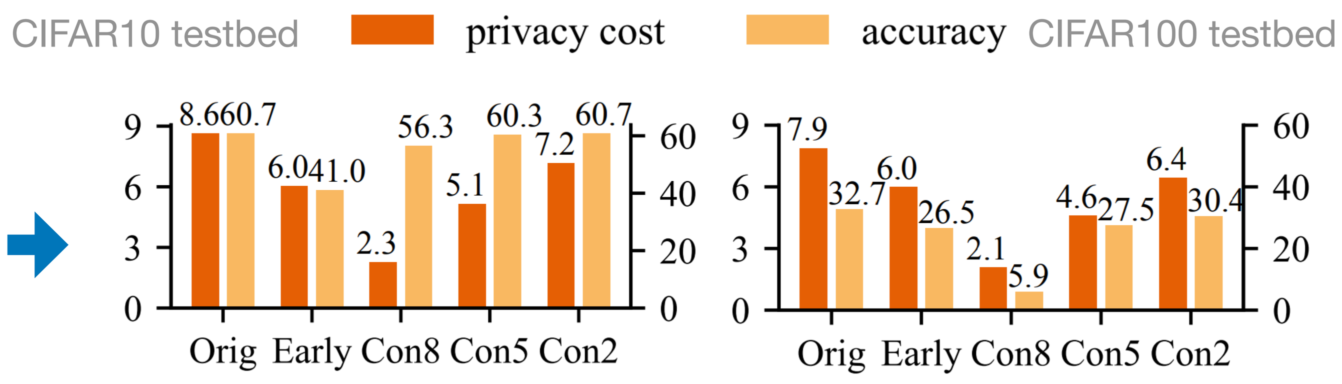


Performance gain in general:
① Grows with model size
② Scales with # participants
③ Insensitive to client dropout

Dropout-Resilient Noise Enforcement

Privacy Issue Caused by Client Dropout



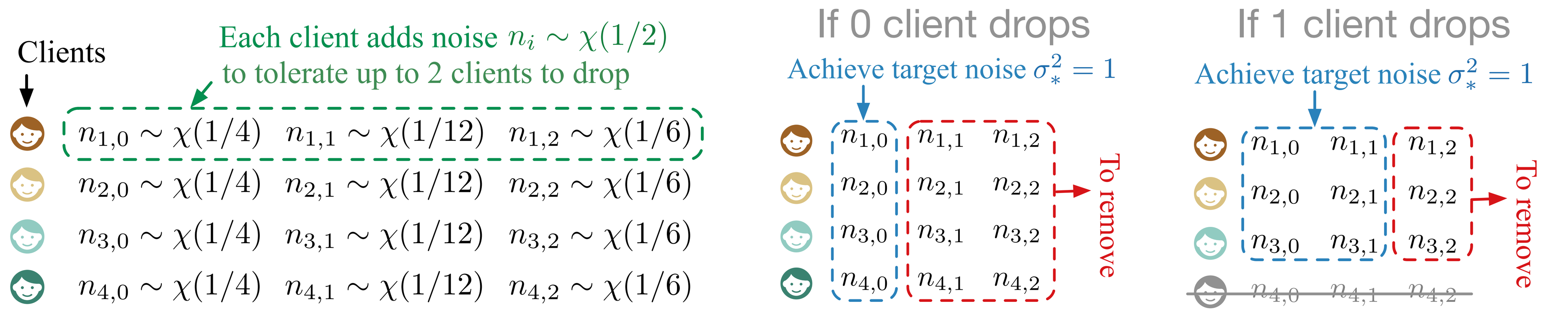


Orig: noise lost → degraded privacy

Simple fixes:
- **Early**: stop when budgets runs out—hurts utility
- **ConX**: add more noise for X0% cases—inflexible

XNoise: Add-Then-Remove Scheme to Enforce Minimum Necessary Noise

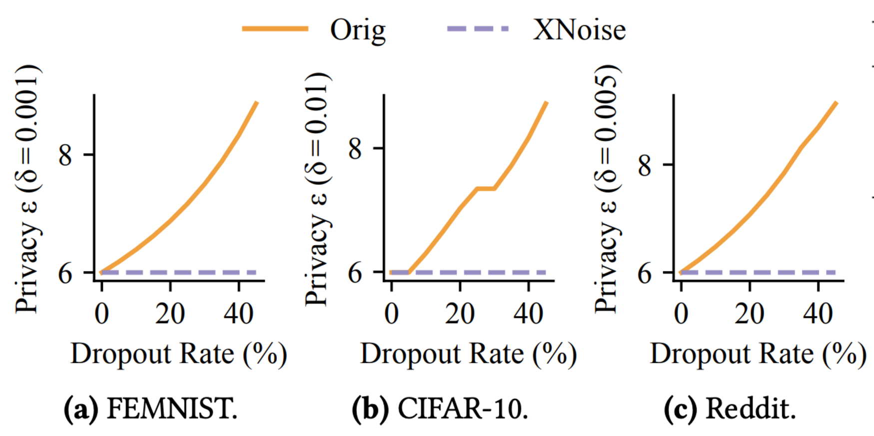
Toy example: # sampled clients = 4, dropout tolerance = 2, necessary noise level = 1



① Add excessive additive noise

② Then remove unnecessary noise precisely

Empirical Effectiveness in the Presence of Simulated Client Dropout



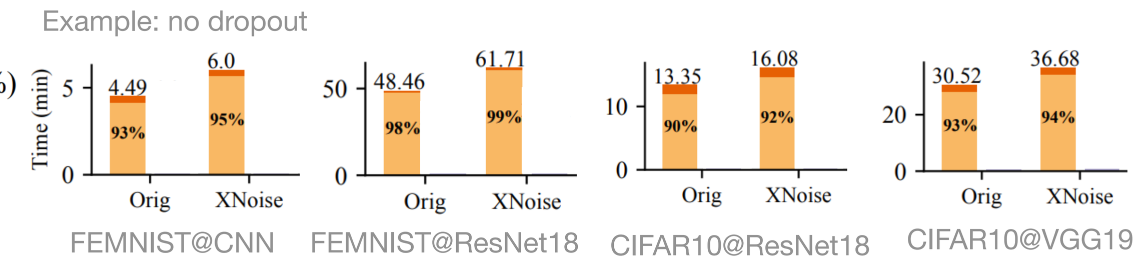
Enforce target privacy ($\epsilon = 6$)

(Also **theoretically provable**)

d	0		10%		20%		30%		40%	
	Ori	XNo	Ori	XNo	Ori	XNo	Ori	XNo	Ori	XNo
F	61.3	61.4	61.4	61.4	61.2	61.4	61.2	61.2	61.4	61.5
C	66.5	66.3	66.7	66.9	66.6	65.7	64.3	65.7	63.8	64.2
R	2169	2142	2158	2179	2286	2285	2294	2317	2299	2329

Preserve model utility

Example: no dropout



Induce acceptable runtime overhead ($\leq 34\,\%$)

Reference

[1] Gradient Obfuscation Gives a False Sense of Security in Federated Learning. *Security, 2023*
[2] The Skellam Mechanism for Differentially Private Federated Learning. *NeurIPS, 2021*
[3] The Distributed Discrete Gaussian Mechanism for Federated Learning with Secure Aggregation. *ICML 2021*
[4] Practical Secure Aggregation for Privacy-Preserving Machine Learning, CCS 2017
[5] Secure Single-Server Aggregation with (Poly) Logarithmic Overhead, CCS 2020

Preprint available at: **Code available at:**

Accepted to appear in the proceedings

Contact: Zhifeng Jiang (zjiangaj@cse.ust.hk)