ZIP-DL: Low-Cost Privacy-Preserving Decentralized Learning using correlated noises

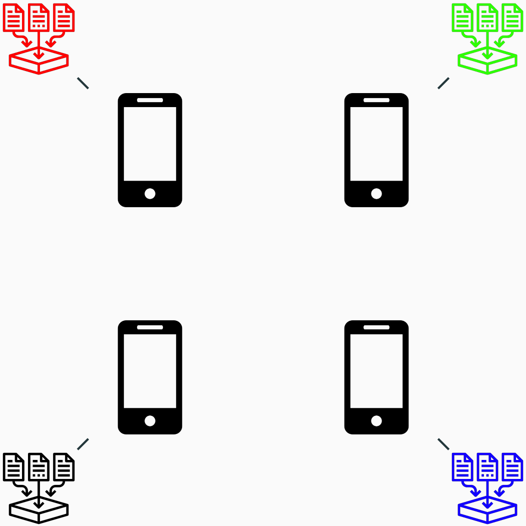
July 15, 2025

Dimitri Lerévérend — Davide Frey — Romaric Gaudel — François Taïani
Sayan Biswas — Anne-Marie Kermarrec — Rafael Pires — Rishi Sharma

PETS 2025 — [Artifact available]

Context: Decentralized learning & Privacy

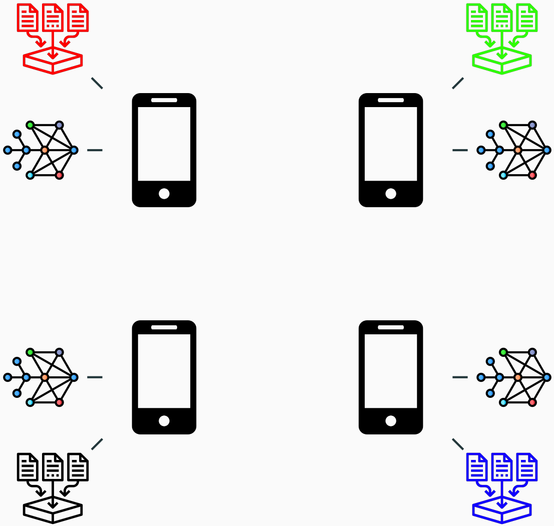
Decentralized Learning (DL)



DL main caracteristics

- Heterogeneous data
- Model exchanges
- Communication graph W
- No central server

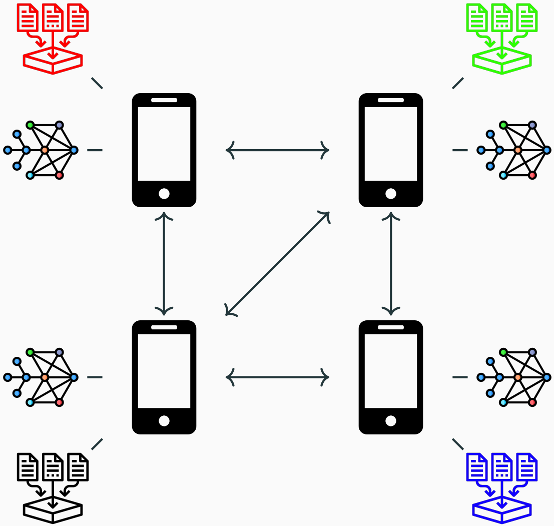
Decentralized Learning (DL)



DL main caracteristics

- Heterogeneous data
- Model exchanges
- Communication graph W
- No central server

Decentralized Learning (DL)

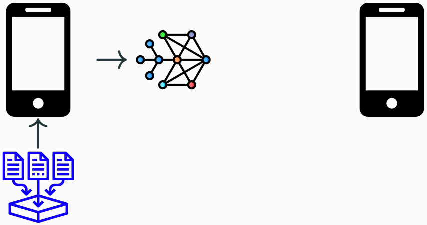


DL main caracteristics

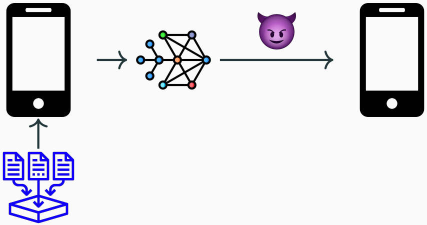
- Heterogeneous data
- Model exchanges
- Communication graph W
- No central server

Privacy attacks — Shared models leak private information

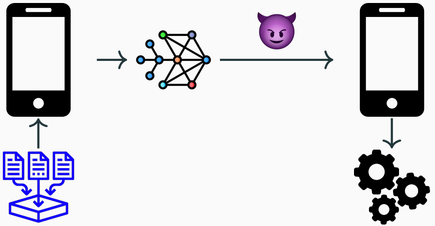
Privacy attacks — Shared models leak private information



Privacy attacks — Shared models leak private information



Privacy attacks — Shared models leak private information

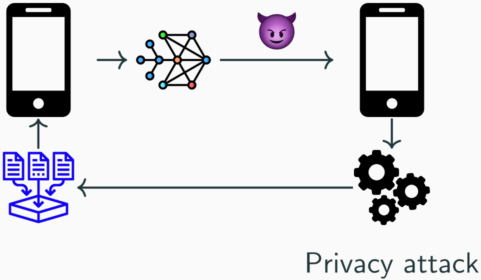


Privacy attack

Privacy attacks

- Membership Inference Attacks (MIA)
- Gradient inversion attacks
- Attribute inference attacks

Privacy attacks — Shared models leak private information

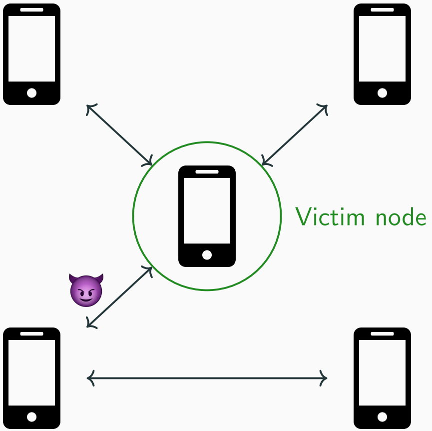


Privacy attacks

- Membership Inference Attacks (MIA)
- Gradient inversion attacks
- Attribute inference attacks

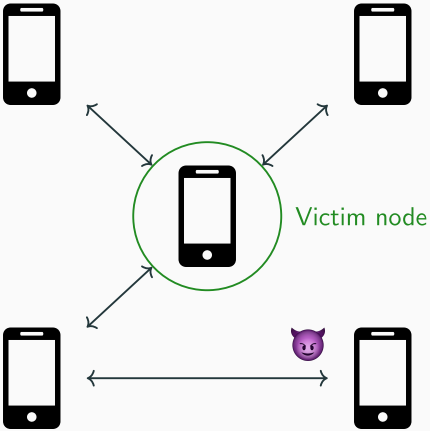
$\Longrightarrow$ Attackers can infer information about the private local data.

Attacker model — Pairwise Network Differential Privacy



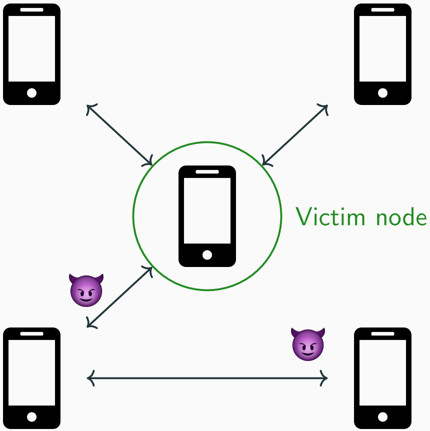
[2] E. Cyffers, M. Even, A. Bellet, and L. Massoulié, Muffliato: Peer-to-Peer Privacy Amplification for Decentralized Optimization and Averaging, NeurIPS, Dec. 2022.

Attacker model — Pairwise Network Differential Privacy



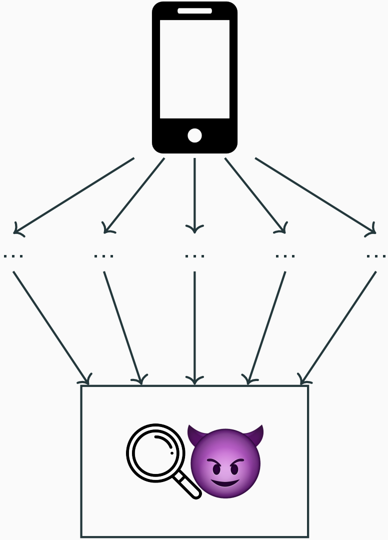
[2] E. Cyffers, M. Even, A. Bellet, and L. Massoulié, Muffliato: Peer-to-Peer Privacy Amplification for Decentralized Optimization and Averaging, NeurIPS, Dec. 2022.

Attacker model — Pairwise Network Differential Privacy



[2] E. Cyffers, M. Even, A. Bellet, and L. Massoulié, Muffliato: Peer-to-Peer Privacy Amplification for Decentralized Optimization and Averaging, NeurIPS, Dec. 2022.

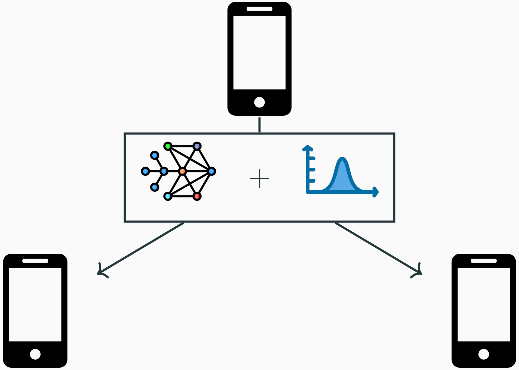
PNDP: attacker model



Attacker capabilities

- Honest-but-curious
- A (set of) node(s) in the process

Muffliato [2]



[2] E. Cyffers, M. Even, A. Bellet, and L. Massoulié, Muffliato: Peer-to-Peer Privacy Amplification for Decentralized Optimization and Averaging, NeurIPS, Dec. 2022.

Muffliato — Limitations

Muffliato — Convergence rate

$$\mathcal{O}\left(\frac{\sigma^2}{nT} + \frac{A}{T^2} + r_0 e^{-T}\right)$$

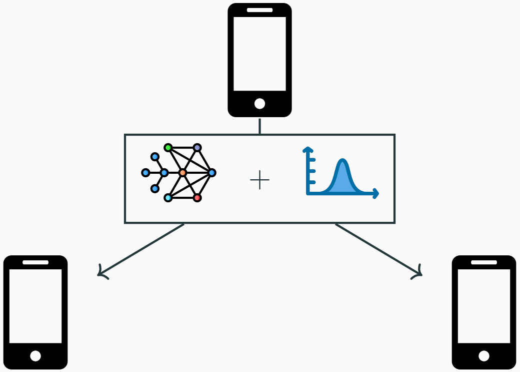
with σ^2 the variance of the LDP noise.

[2] E. Cyffers, M. Even, A. Bellet, and L. Massoulié, Muffliato: Peer-to-Peer Privacy Amplification for Decentralized Optimization and Averaging, NeurIPS, Dec. 2022.

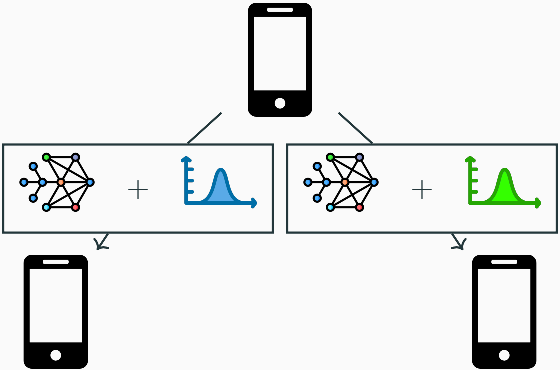
Can we leverage the properties of PNDP to bridge this utility-privacy gap?

Our approach: Zip-DL

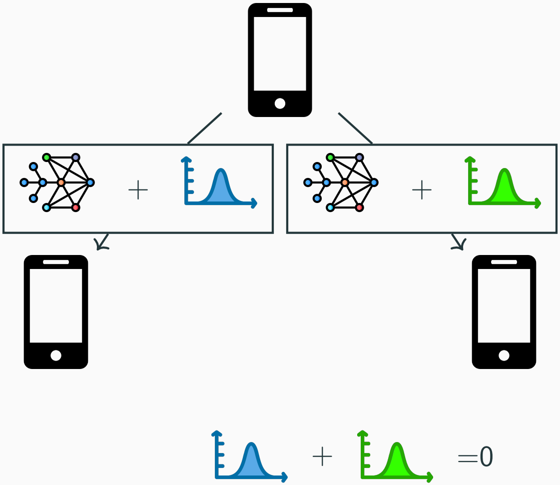
Zip-DL: Adding correlated noise



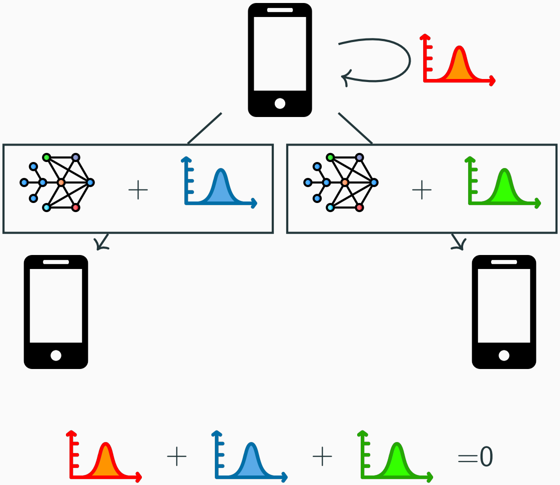
Zip-DL: Adding correlated noise



Zip-DL: Adding correlated noise



Zip-DL: Adding correlated noise



Results

Results

Theoretical results

Core result — Convergence

Muffliato (Reminder)

$$\mathcal{O}\left(\frac{\sigma^2}{nT} + \frac{A}{T^2} + r_0 e^{-T}\right)$$

with σ^2 the variance of the LDP noise.

Our solution: Zip-DL

For any number of iterations T, $\frac{1}{2W_T}\sum_{t=0}^{T} w_t(\mathbb{E}\left[f(\bar{x}^{(t)})\right] - f^*) + \frac{\mu}{2} r_{T+1}$ is bounded:

$$\mathcal{O}\left(\frac{1}{nT} + \frac{\alpha\left(A + A'\sigma^2\right)}{T^2} + r_0 e^{-T}\right)$$

where $f^* = f(x^*)$, $r_t = \mathbb{E}\left[\left\|\bar{x}^{(t)} - x^*\right\|^2\right]$, $A' = \frac{1}{n}\sum_{a,v=1}^{n} d_a \frac{(d_v-1)^2}{d_v}$, $\alpha = \frac{16-4p}{2(16-7p)}$

Core result — PNDP

Zip-DL Privacy guarantees:

- T iterations of $\textsc{Zip-DL}$ satisfies $\left(\alpha, \epsilon_{a \to v}^{(T)}\right)$-PNDP
- More details in the paper

Results

Experimental evaluation

Evaluation setup

Metrics

- Utility: Test loss/accuracy
- MIA: Loss attack/classifier attack
- Communication overhead

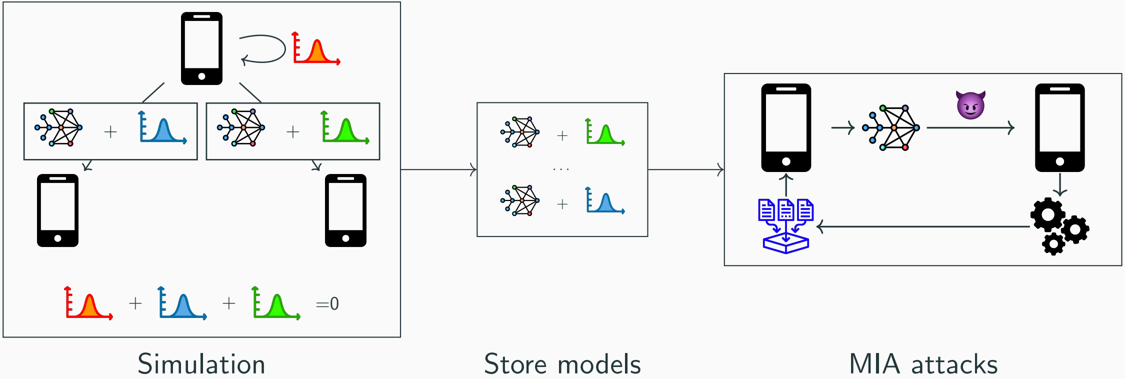
Dataset & models

- MovieLens — Matrix factorization
- Cifar — Resnet-18 $\rightarrow$ details in paper.

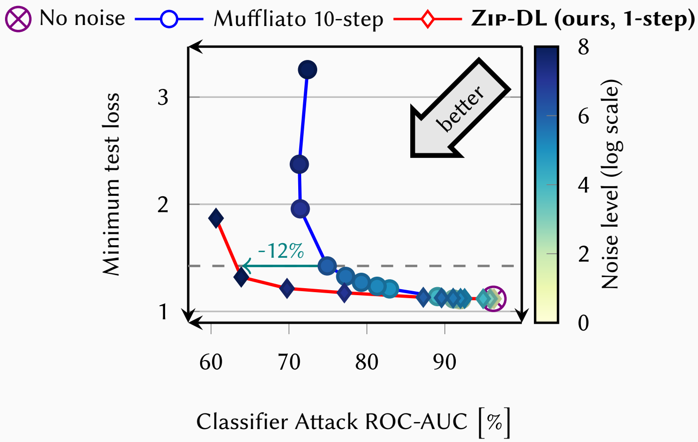
Decentralized simulation settings

- 100 nodes
- 5-regular network
- NIID data

Evaluation pipeline

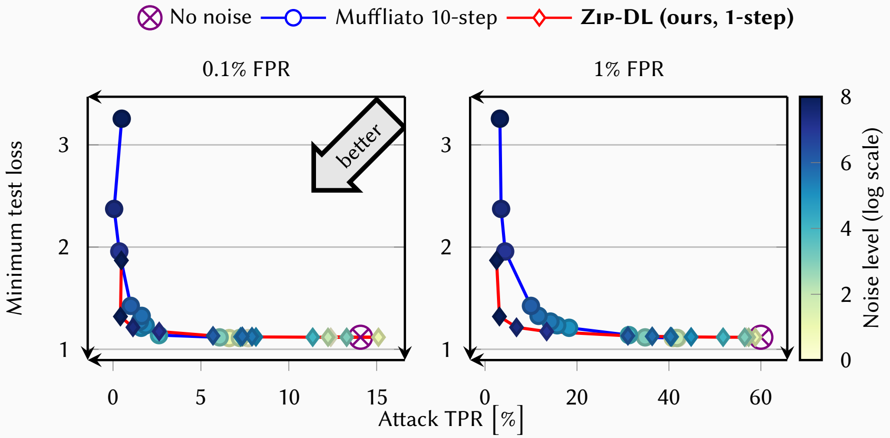


Experimental results — Privacy-utility tradeoff



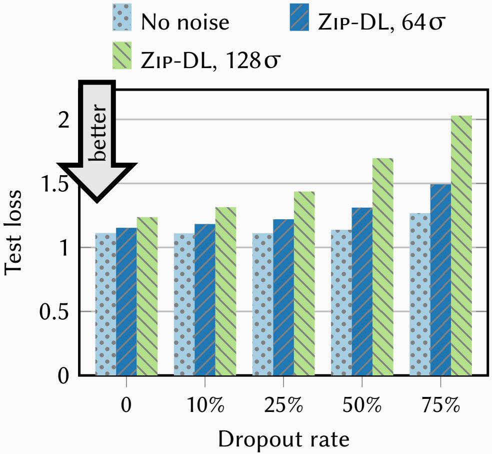
$\Longrightarrow$ Zip-DL has a better privacy-utility tradeoff.

Experimental results — TPR at low FPR



$\Longrightarrow$ ZIP-DL has better privacy-utility tradeoff at 1% FPR, comparable at 0.1% FPR.

Experimental results — Dropout



$\Longrightarrow$ Zip-DL is resilient to moderate dropout, even at high noise level σ.

Conclusion

ZIP-DL

- Privacy-preserving using correlated noise addition
- Improves convergence by a factor $\frac{1}{T}$ compared to LDP
- Efficient practical utility-privacy tradeoff with PNDP guarantees

Future works

- Extend to more colluding attackers
- Exploit time correlation to reduce the attack surface.

Paper:

Code:

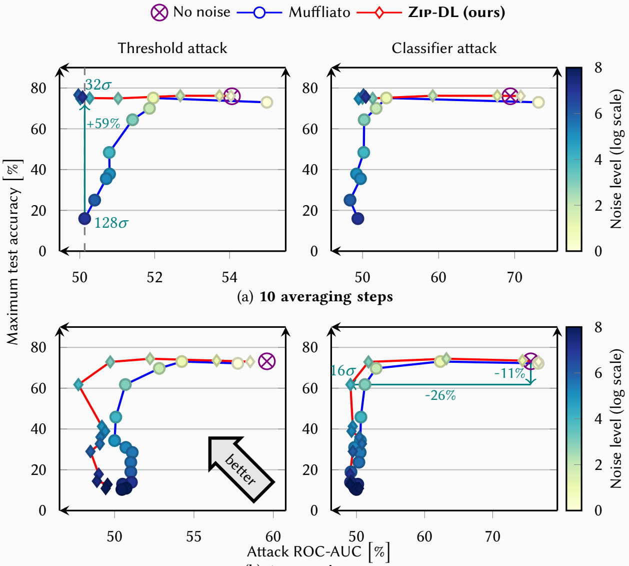
My website:

Appendix

Appendix

Additional Experimental results

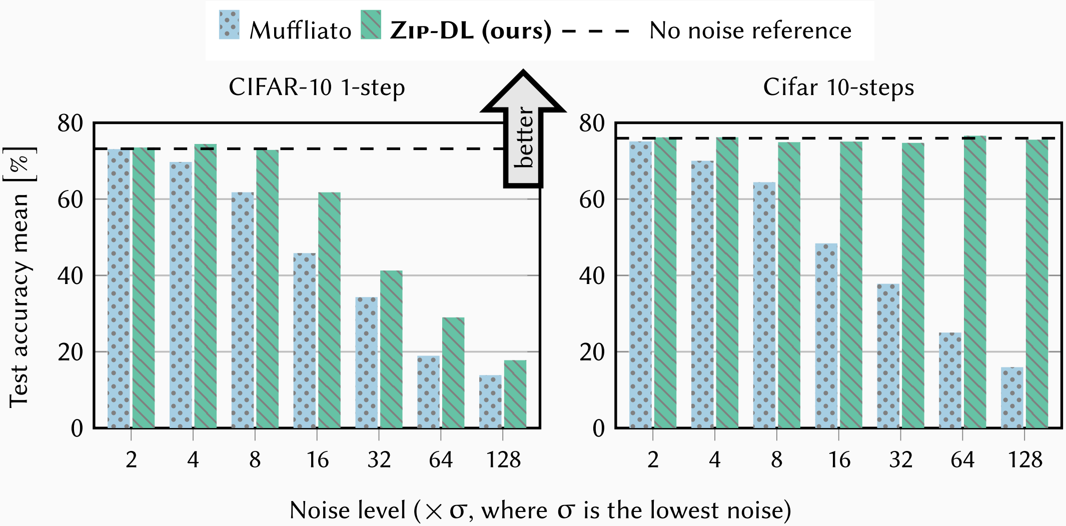
Experimental results — CIFAR dataset



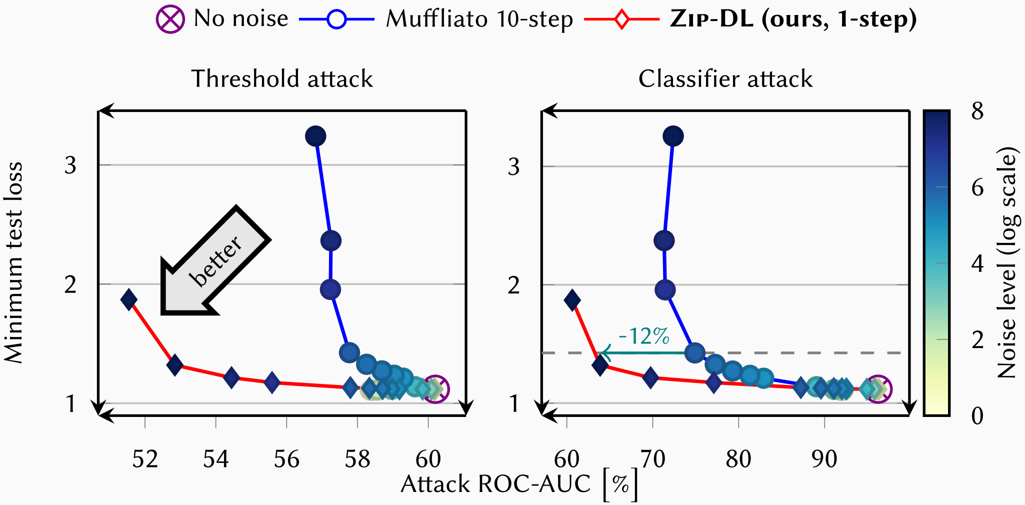
(a) **10 averaging steps**

(b) **1 averaging step**

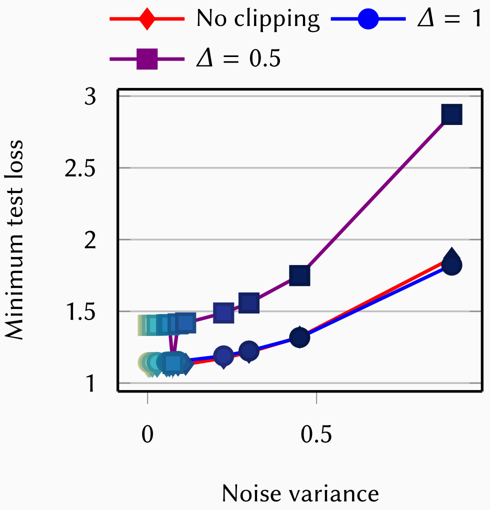
Experimental results — CIFAR dataset accuracy



Experimental results — MovieLens attacker tradeoff



Gradient clipping



Appendix

Evaluation

Experimental setup

Simulation

- Datasets: CIFAR-10, MovieLens
- 128 nodes.
- Resnet, Recommandation matrix.

Evaluation

- Multiple noise levels σ.
- Baselines: **Muffliato**, no noise.
- Two attacks: **threshold attack** and **classifier attack**.

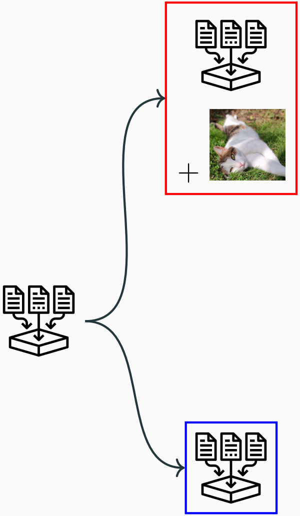
Formalizing privacy loss

Formalizing privacy loss

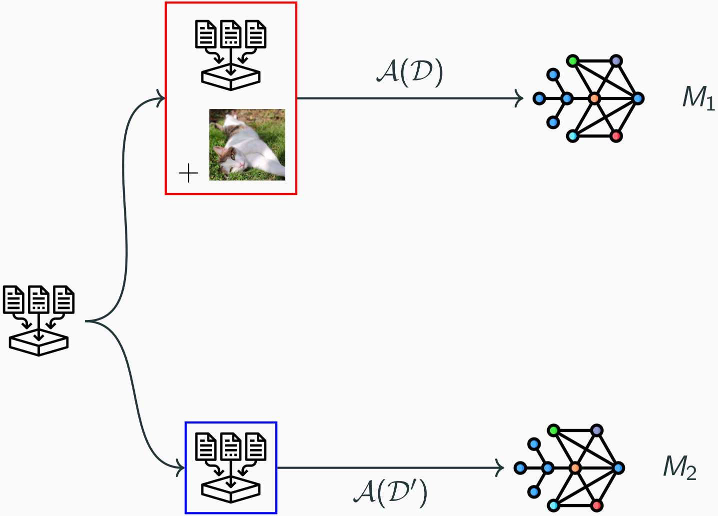
Differential privacy and variants

Existing solution: Differential Privacy (DP)

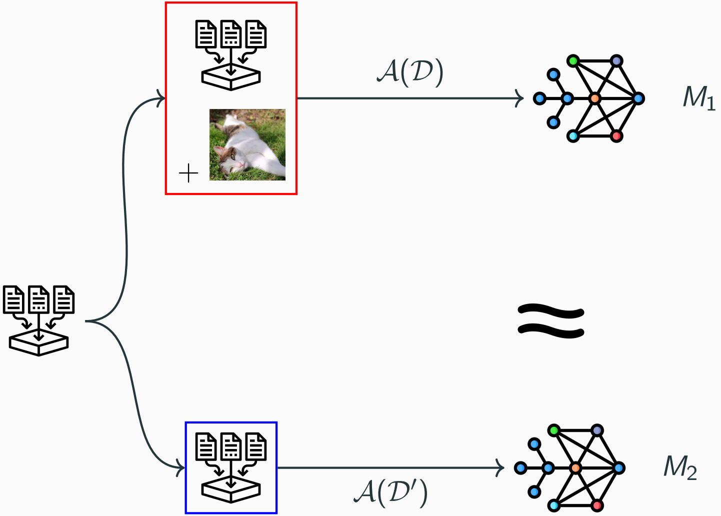
Existing solution: Differential Privacy (DP)



Existing solution: Differential Privacy (DP)



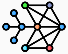
Existing solution: Differential Privacy (DP)



DP-SGD: small illustration

SGD:

Model

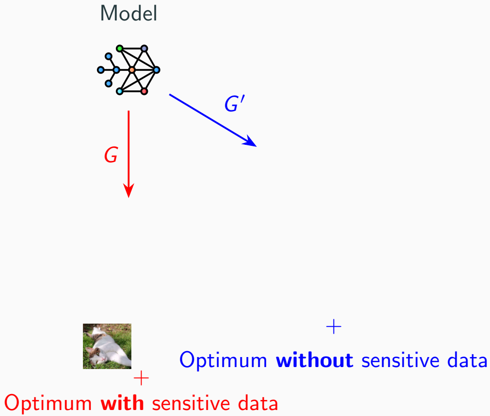


Optimum **without** sensitive data

Optimum **with** sensitive data

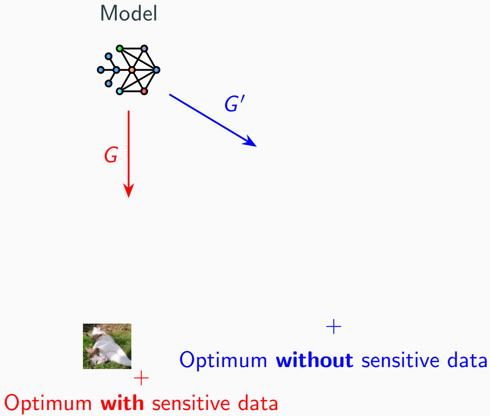
DP-SGD: small illustration

SGD:

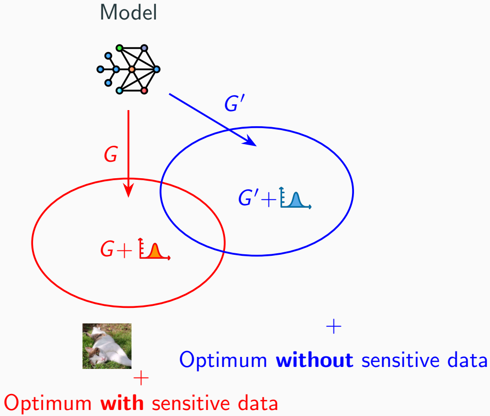


DP-SGD: small illustration

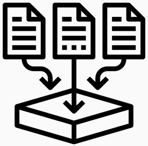
SGD:



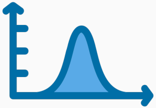
DP-SGD:



Noise addition & Local considerations

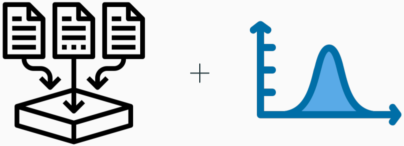


+

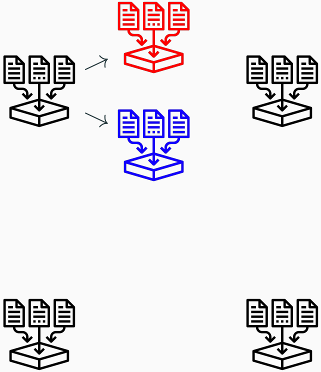


DP

Noise addition & Local considerations

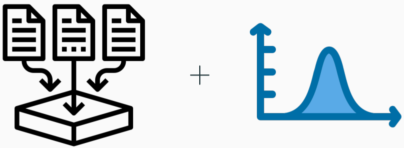


DP

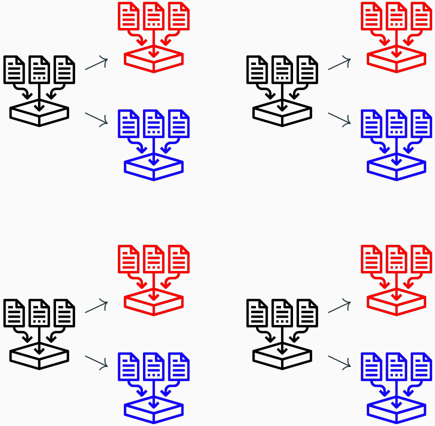


Local DP

Noise addition & Local considerations

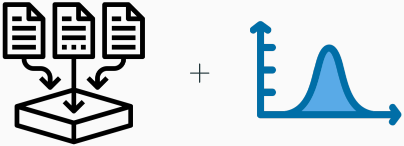


DP

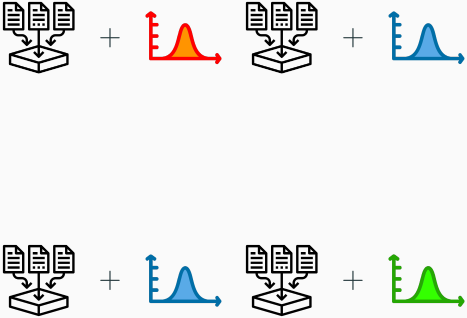


Local DP

Noise addition & Local considerations



DP



Local DP

Privacy & local considerations

(α, ε)-Rényi-DP

- $D_\alpha\left(\mathcal{A}(\mathcal{D}) \| \mathcal{A}(\mathcal{D}')\right) \leq \varepsilon$
- $\mathcal{A}(\mathcal{D})$: Distribution of result of algorithm $\mathcal{A}$ on dataset $\mathcal{D}$

Pairwise Network DP [3]

- $D_\alpha\left(\mathcal{O}_v(\mathcal{A}(\mathcal{D})) \| \mathcal{O}_v(\mathcal{A}(\mathcal{D}'))\right) \leq g(a, v)$.
- $\mathcal{O}_v(\mathcal{A}(\mathcal{D}))$: Distribution of the information received by v.
- Consider spatial information.

[3] Cyffers et al, Muffliato, NeurIPS 2022.

Core result — PNDP

Zip-DL Privacy guarantees:

T iterations of $\textsc{Zip-DL}$ satisfies $\left(\alpha, \epsilon_{a\to v}^{(T)}\right)$-PNDP with:

$$\epsilon_{a\to v}^{(T)} \leq \frac{2\alpha\gamma^2\Delta^2}{L+4\gamma^2L^2}\sum_{t=0}^{T-1}\sum_{\substack{\hat{v}\in\hat{V}\\ \hat{w}\in\hat{\Gamma}_{\hat{v}}^{(t)}}}\frac{(2+4\gamma^2L)^t-1}{\left(\left(\tilde{W}\tilde{C}\right)^{(t)}\tilde{\Sigma}_{\tilde{Y}^{(t)}}{}^{T}\left(\tilde{W}\tilde{C}\right)^{(t)}\right)_{\tilde{w},\tilde{w}}},$$

Formalizing privacy loss

Privacy issues in decentralized learning

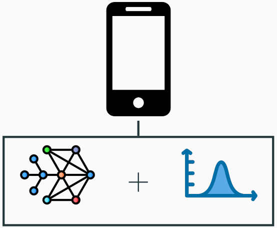
Baseline: Muffliato[1]

Muffliato

- Noisy local model.

[1] Cyffers et al, Muffliato, NeurIPS 2022.

Baseline: Muffliato[1]



Muffliato

- Noisy local model.

Limitations

- Performances degradation.

[1] Cyffers et al, Muffliato, NeurIPS 2022.

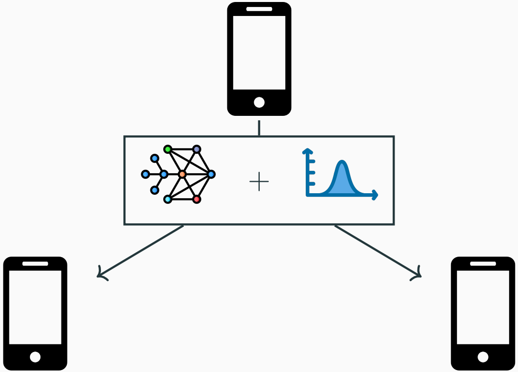
Baseline: Muffliato[1]

Muffliato

- Noisy local model.

Limitations

- Performances degradation.



[1] Cyffers et al, Muffliato, NeurIPS 2022.

Baseline: Muffliato[1]

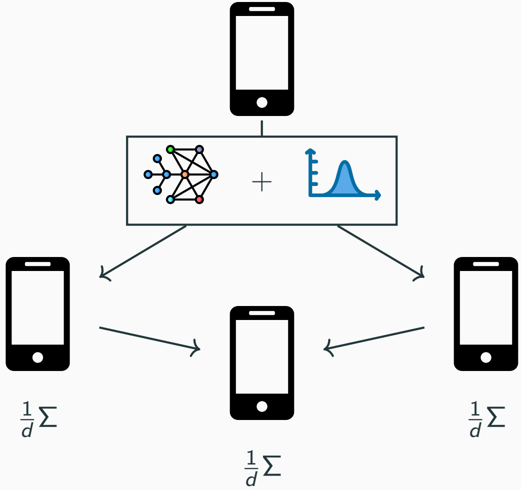
Muffliato

- Noisy local model.
- Multiple unnoised averaging rounds.

Limitations

- Performances degradation.
- Multiple averaging rounds.

[1] Cyffers et al, Muffliato, NeurIPS 2022.



Additional details

Additional details

Assumptions

Assumptions

Expected consensus rate

$$\mathbb{E}_{W^{(t)}}\left[\left\|W^{(t)}X - \bar{X}\right\|_F^2\right] \leq (1-p)\left\|X - \bar{X}\right\|_F^2.$$

L-smoothness

$$\left\|\nabla F_i(x', \xi) - \nabla F_i(x, \xi)\right\| \leq L\left\|x - x'\right\|.$$

Noise structure

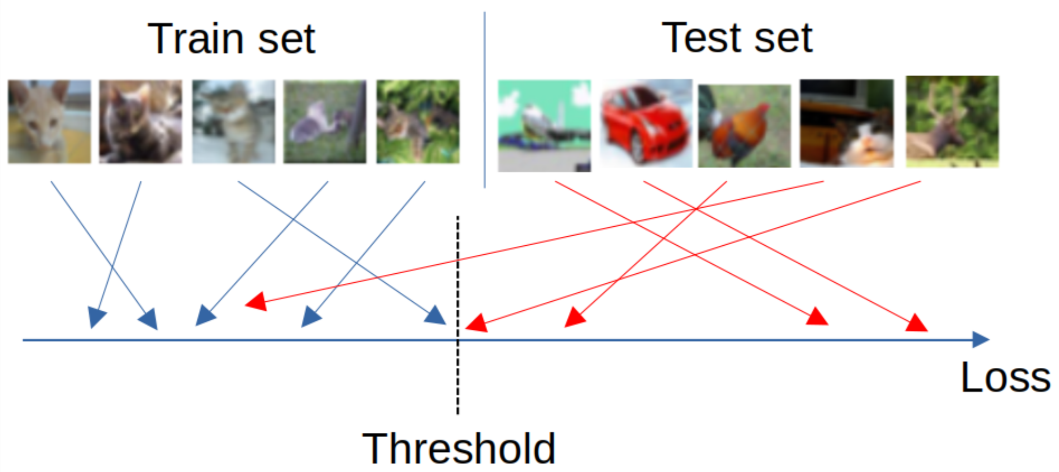
For all $i \in \mathcal{V}$, for all data sample ξ_i and model x, if we consider a noise $Z \sim \mathcal{N}(0, \Sigma)$, then we have:

$$\nabla F_i(x + Z, \xi_i) \sim \mathcal{N}(\nabla F_i(x, \xi_i), L\Sigma)$$

Additional details

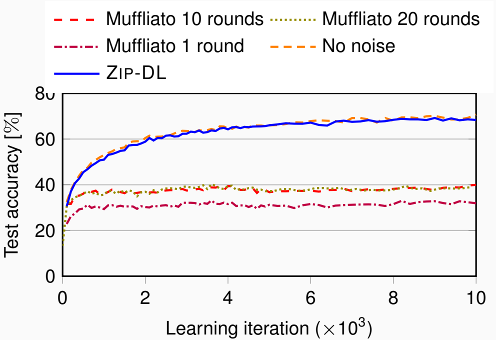
Privacy attacks

Threshold attack



Multiple rounds of communication

8σ



32σ