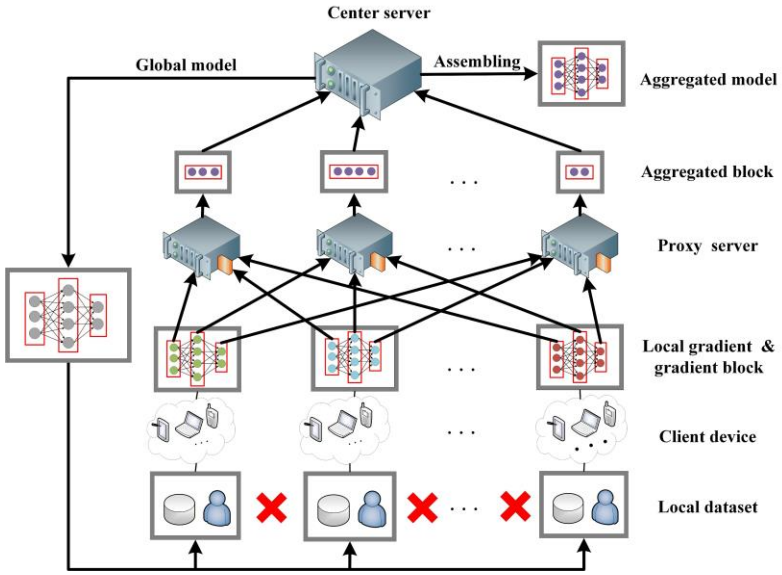
FedDAA: A robust federated learning framework to protect privacy and defend against adversarial attack

Shiwei LU, Ruihu LI, Wenbin LIU

Frontiers of Computer Science, DOI: 10.1007/s11704-023-2283-x

Problems & Ideas

- Problems of the security of federated learning:
  - Defense against adversarial attack and privacy protection are two different secure defense methods.
  - Existing defense against adversarial attack and privacy protection methods cannot be used in the common federated learning framework at the same time .
- Ideas: A new federated learning framework is designed to protect user privacy without encryption or noise, and defend against adversarial attack.



A federated learning framework with model decomposition, aggregation and assembling (FedDAA)

Main Contributions

- Contributions:
  - We propose a novel framework for federated learning with model decomposition, aggregation and assembling (FedDAA) and design a training algorithm to complete federated training in FedDAA.
  - We give a method to quantify the privacy leakage in DLG attack
  - We explore how to apply Byzantine-robust aggregated rules and model verification in FedDAA. In addition, a defense scheme with similarity measurement is proposed in FedDAA to detect and remove backdoor.

Method	Parameter setting	Model accuracy	SSIM	Reconstructed image	Original image
DP with Gaussian noise[24]	N ~(0,0.01)	0.934	0.176		
	N ~(0,0.05)	0.902	0.037		
DP with Laplace noise[25]	L ~(0,0.01)	0.927	0.138		
	L ~(0,0.05)	0.883	0.018		
FedDAA(ours)	$\lambda = 0.8$	0.952	0.924		
	$\lambda = 0.7$	0.954	0.857		
	$\lambda = 0.6$	**0.952**	**0.014**		

The performance comparison of differential privacy and FedDAA on privacy protection