About securing federated learning

Phillip Holzmann Dorian Reineccius
Michael Nüsken
b-it, University of Bonn
Friedrich-Hierzebruch-Allee 6,
D-53115 Bonn
nuesken@bit.uni-bonn.de
https://crypto.bit.uni-bonn.de/~nuesken/

33th Crypto Day, 17 September 2021

Federated learning gets more popular eg. for image classification, as it promises better performance, offline application and privacy. This technique allows participants to train their own models locally and one central entity combines all participants' achievements into one common model. The participants synchronize their models with the central one regularly by uploading their gradients and getting back the updated weights. This way all participants benefit from the training in the entire pool. However, Hitaj, Ateniese & Pérez-Cruz (2017) show that a malicious player can attack federated learning by acting as a usual participant. By manipulating the training she can use a Generative Adversarial Network (GAN) to reveal training images used by other participants. Our incentive is to prevent this attack by using homomorphic encryption.

Fully homomorphic encryption (FHE) allows computations on plaintexts by operating with ciphertexts only. To counteract the attack we need to hide the model data from the participants while still allowing them to obtain the classification answers. We need to keep in mind that the participants still have to perform computations involving the sensitive local training data. In contrast to the plain text scenario, additional communication between server and client is needed, in order for the client to decrypt his results.

Though we get strong limitations on the performance caused by the homomorphic operations this shows that secure federated learning is indeed possible. For more details, consider Holzmann & Reineccius (2021).

References

Briland Hitaj, Giuseppe Ateniese & Fernando Pérez-Cruz (2017). Deep Models Under the GAN: Information Leakage from Collaborative Deep Learning. *CoRR* **abs/1702.07464**. URL http://arxiv.org/abs/1702.07464.

Phillip Holzmann & Dorian Reineccius (2021). About federated learning. Lab on Cryptography 2021. URL https://crypto.bit.uni-bonn.de/teaching/21ss/lab/.

crypto day matters 33 (2021)
Beiträge vom 33. Kryptotag, 17.09.2021, online.
doi:10.18420/cdm-2021-33-33