Federated Machine Learning in Privacy-Sensitive Settings

Anastasia Pustozerova

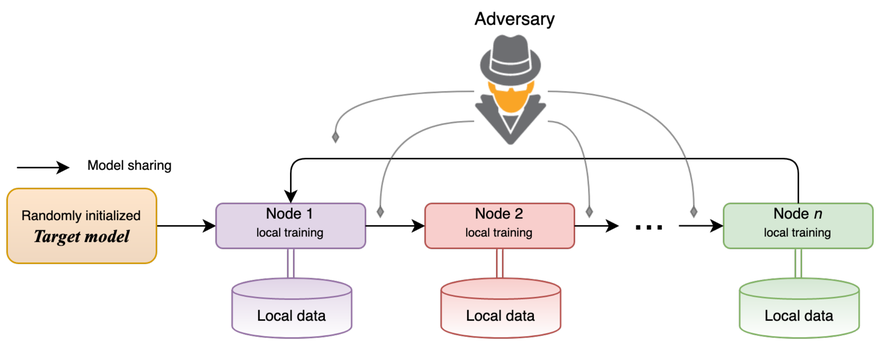
COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Federated learning allows performing machine learning over distributed data while *preserving privacy* of data owners. Each data holder independently and locally trains a machine learning model on her own data and then shares the model with other participants of the federated learning process, so other parties can proceed training on their own data, or aggregate several models to a global one. Federated learning addresses the issue of *data locality and sensitivity* and also enables using *computational power of distributed systems*, closer to the place where the data is originating. However, models, which are exchanged during the federated learning process, can leak information about their training data. In this work, we

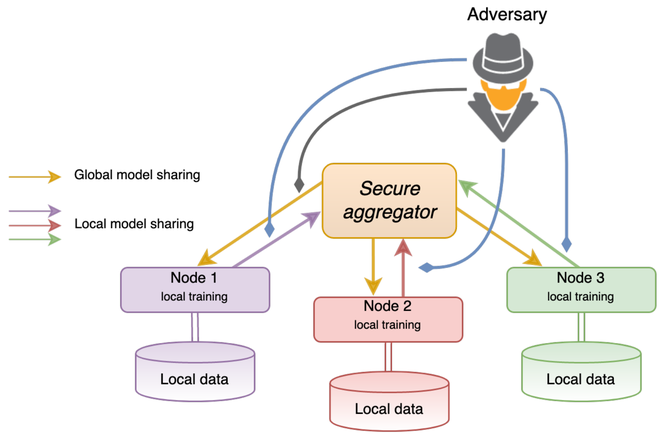
- evaluate privacy risks in federated learning by performing membership inference attack,
- propose mitigation strategies to improve privacy properties of federated learning,
- develop guidelines for federated learning allowing to maintain effectiveness of the models while preserving privacy of the data.

Parallel federated learning vs. Sequential Federated learning



Sequential federated learning

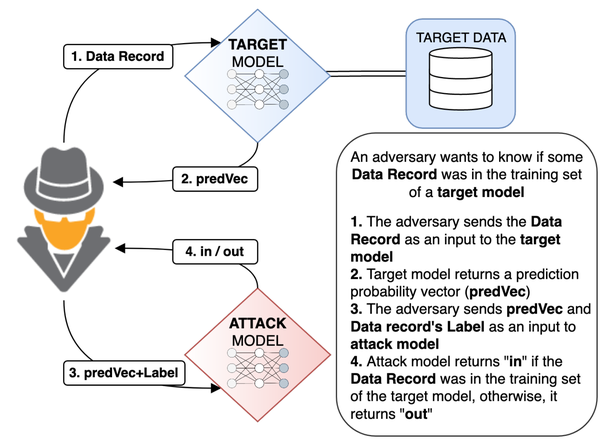
A randomly initialized model is locally trained at the first client and then passed to the next node in the sequence. After completing a full round of n nodes, the model is passed again to the first node for repeating the training process.



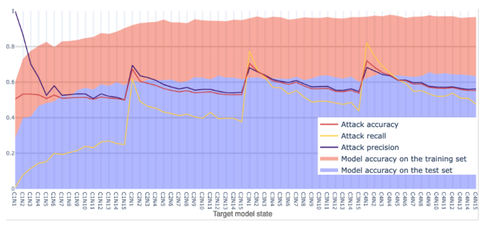
Parallel federated learning

An *aggregator* initializes a global model with random weights and shares it to every node in the setting. Each node trains the model in parallel on its local data and then returns it to the secure aggregator. From the locally trained models, a new global model is aggregated and shared to the clients for the following training cycle.

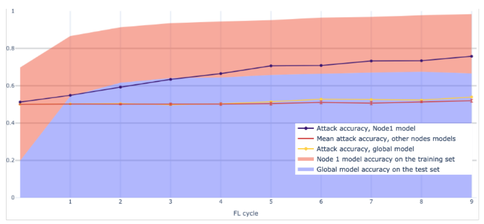
Membership inference attack



Attack evaluation

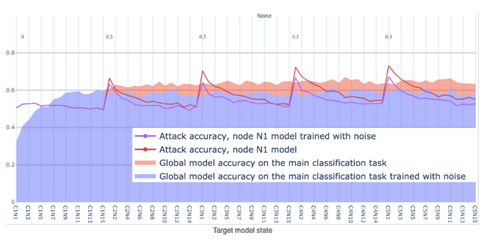


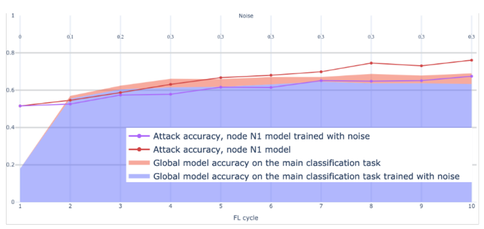
In **sequential federated learning**, the attack on node N1 data has the highest accuracy while performing membership inference on the shared model right after training at node N1.



In **parallel federated learning**, the attack on node N1 data has higher accuracy while performing membership inference on the model trained locally at N1, than attacking global model, or local models from other nodes.

Mitigation evaluation





In both **sequential** and **parallel** federated learning adding, noise to the training data allows to mitigate the risks of membership inference attack. However, the noise should be properly chosen to not cause loss in effectiveness of the global model on the classification task.

Conclusion

- Federated learning allows to avoid data transferring while training machine learning models without loosing in models effectiveness
- The models shared during federated learning process can leak information about their training data, e.g. when attacker performs membership inference attack
- The membership inference attack accuracy can be reduced by adding noise to the training data.

References:

1. A. Pustozerova, and R. Mayer, "Information Leaks in Federated Learning", Workshop on Decentralized IoT Systems and Security (DISS), San Diego, CA, USA, 2020
2. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-efficient learning of deep networks from decentralized data.," 20th International Conference on Artificial Intelligence and Statistics, (AISTATS), Fort Lauderdale, FL, USA, 2017.
3. S. Truex, L. Liu, M. Gursoy, L. Yu, and W. Wei, "Demystifying membership inference attacks in machine learningas a service,"IEEE Transactions on Services Computing, 2019.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien Ein Fonds der Stadt Wien
Europäische Kommission
FWF Der Wissenschaftsfonds.
netidee OPEN INNOVATIONS

Das Kompetenzzentrum SBA Research wird im Rahmen von COMET – Competence Centers for Excellent Technologies durch BMVIT, BMDW, das Land Wien gefördert. Das Programm COMET wird durch die FFG abgewickelt.