Trust and scalability through randomized communication graphs

Jan Ramon

Abstract: While machine learning can potentially extract highly valuable models from personal data, there is a growing awareness of the privacy risks. Therefore, there is an increasing interest in privacy-preserving decentralized learning, a form of federated or distributed machine learning where sensitive data stays on its owner's device and only encrypted messages are exchanged between the data owners with the goal of collaboratively learning a model without revealing any sensitive information.

One can't always assume all participants in such effort are honest. In the presence of malicious parties, many existing cryptographic approaches need every data owner to communicate with every other data owner, causing a poor scaling.

We are interested in random communication graphs. Drawing a random graph connecting participants is an efficient way to generate a graph with a high probability of containing a spanning tree of the subgraph induced by the honest participants. In the PhD work of Cesar Sabater[1], we showed that using a random communication graph we can in time logarithmic in the number of participants perform secure privacy-preserving aggregation in the presence of malicious participants. This presentation will briefly review this idea and then further explore the properties of random communication nnetworks.

[1] Cesar Sabater, Aurelien Bellet, Jan Ramon. An Accurate, Scalable and Verifiable Protocol for Federated Differentially Private Averaging, Machine Learning, 2022.