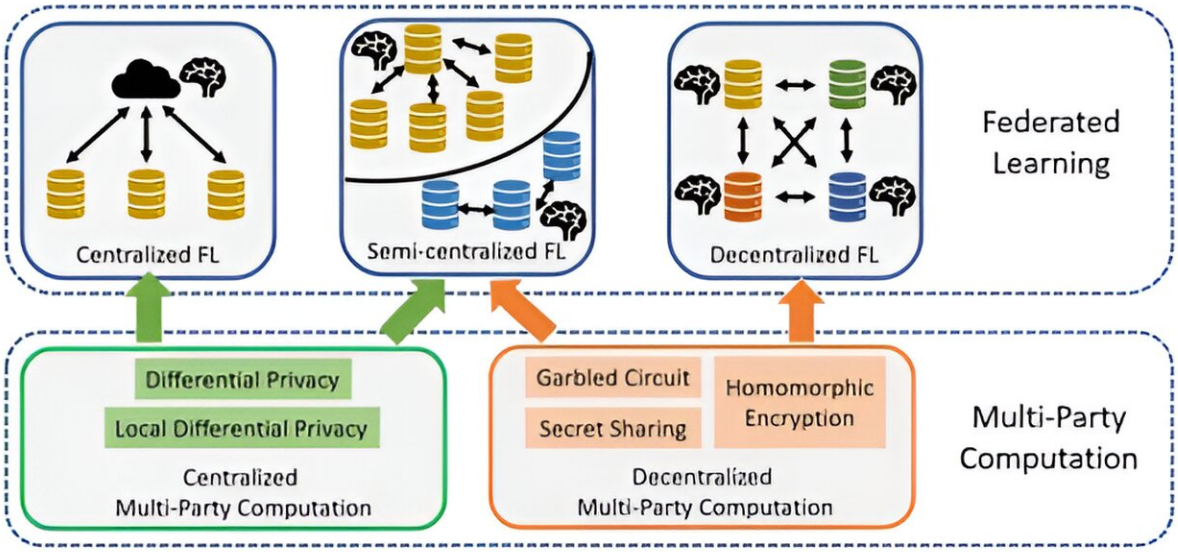
A survey on federated learning: A perspective from multi-party computation

February 28 2024



The taxonomy of federated learning from a perspective of multi-party computation. Credit: *Frontiers of Computer Science* (2023). DOI: 10.1007/s11704-023-3282-7

Federated learning (FL) has emerged as a popular machine learning paradigm which allows multiple data owners to train models collaboratively without sharing their raw datasets. It holds potential for a wide spectrum of analytics applications on sensitive data.

For example, federated learning has been applied to medical big data

analysis such as disease prediction and diagnosis without revealing the patients' private medical information to third-party services. It has also been exploited by banks and insurance companies to train an accurate machine learning model for risk assessment or customer recommendation.

Federated learning enables collaborative model training without sharing raw datasets among data owners by decomposing the training procedure into local training and model aggregation. A paper describing a survey on federated learning has been published in the journal *Frontiers of Computer Science*.

Each data owner performs local training on its own data partition and only communicates intermediate results e.g., gradients for model aggregation at either a centralized server or other data owners. Federated learning with a central server to coordinate the model aggregation is called centralized FL, while model aggregation in a peer-to-peer manner is known as decentralized FL.

Centralized FL imposes high computation workload to the server, whereas decentralized FL involves excessive communication among peers. Consequently, semi-centralized FL has recently been proposed to balance the computation and communication cost by conducting clustered or hierarchical model aggregation.

We focus on federated learning with privacy guarantees. Note that exchanging intermediate results (e.g., gradients) rather than raw datasets may still leak privacy. Accordingly, extra techniques are compulsory for secure communication and computation during federated learning.

Of particular interest is multi-party computation, a generic and fundamental category of techniques that takes multi-party private inputs for aggregated computation without revealing the private data of each

party. Common multi-party computation techniques include garbled circuit, secret sharing, homomorphic encryption, differential privacy, and so on.

Recent years have witnessed a surge to enhance the privacy of federated learning via multiparty computation.

More information: Fengxia Liu et al, A survey on federated learning: a perspective from multi-party computation, *Frontiers of Computer Science* (2023). DOI: 10.1007/s11704-023-3282-7

Provided by Higher Education Press