Privacy-Preserving Aggregation in Federated Learning: A Survey

0. Abstract

Privacy-Preserving Federated Learning 隐私保护的联邦学习（PPFL）

Privacy-Preserving Aggregation (PPAgg) 隐私保护聚集是PPFL中的关键协议

1. Introduction

Federated Learning: 允许不同的参与者/数据所有者，使用本地数据单独训练ML模型，然后server聚合，构建一个全局的FL模型

Deep leakage from **gradient**，恶意攻击者可以重构用户的数据

隐私保护技术：Homomorphic Encryption HE, Multi-Party Computation (MPC) , Differential Privacy (DP), blockchain, Trusted Execution Environment (TEE)

PPFL需要考虑不同的计算能力和带宽等，也要考虑半诚实的敌手，试着学习诚实参与者的信息而不偏离FL协议

2. Overview and fundamentals of federated learning

2.1 Overview of Federated Learning

Federated learning，不同的participants各自独立用 local data train ML model, then, central server aggregated to construct a FL model.

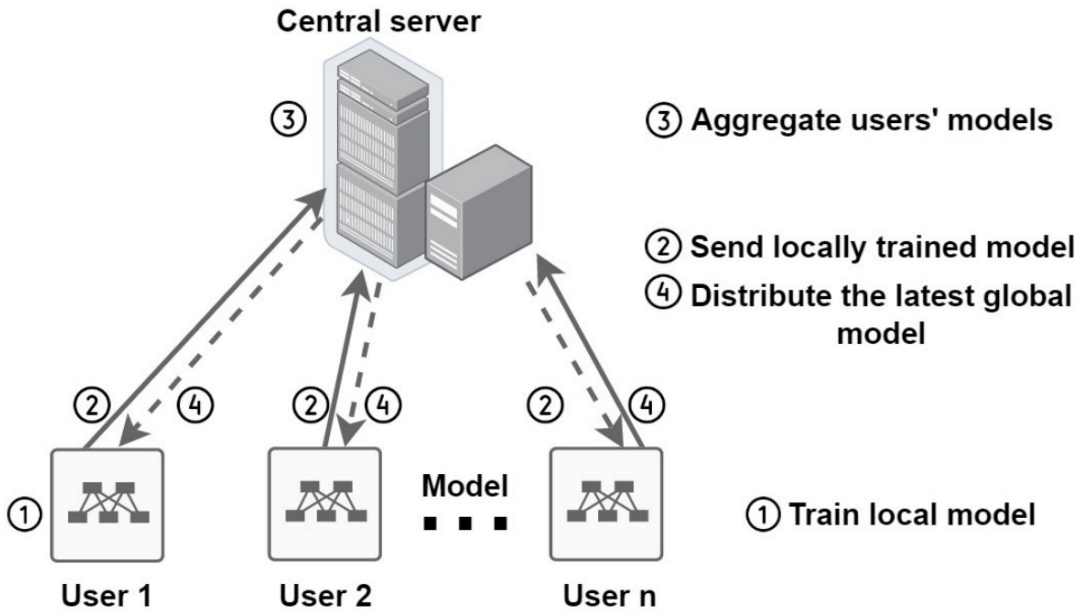
FL scheme works:

(1) Local model training

(2) Model uploading

(3) Model aggregation

(4) Model updating



2.2 Privacy Threats to Federated Learning

information leakage: FL participants, eavesdroppers(窃听者)

Privacy threats in FL （联邦学习中的隐私威胁来源）：

(1) Privacy threats to **user's model** (用户模型的隐私)

(2) Privacy threats to **global model** (全局模型的隐私)

Single adversary or multiple adversary (单个或者多个敌手):
(1) Single adversary: an FL user/ the central server/ the third party

(2) Colluding adversaries: 导致共谋，FL用户与Server和更多的敌手合并导致更大隐私泄露

Capability of adversaries（敌手能力）:

(1) Honest: 诚实制行协议

(2) Passive malious（诚实但好奇/ 半诚实）: 在遵守协议的同时，尝试探索诚实参与者的隐私信息

(3) Active malious: 随时可能违背协议，eg. 操纵身份或发送虚假消息

3. Techniques for privacy-preversing Aggregation

3.1 One-time Pad

One-time Pad (OTP)，发送者**随机生成密钥**并加密，解密时候需要使用与之对应的OTP密钥，OTP的随机性保证了每次加密是unique，并与其他加密没有关系，提供了可证明的无条件安全性。

通常是在有限域F上，使用加法或者乘法的方式完成OTP

OTP提供的是**完美安全性**，但是DP(差分隐私)仍然会泄露数据库的一些统计信息

OTP基于**masking**，可以在Aggregation的时候提供准确的结果，但是DP会不可避免的产生噪声，降低FL性能

注：OTP仅适用于在finite filed上能够保证这种unconditional security.

3.2 Homomorphic Encryption

Homomorphic Encryption (HE)同态加密，可以在加密的数据上进行运算

根据允许的算数运算进行分类：

（1）Partially Homomorphic Encryption (PHE) 半同态

只允许无限次的加法或者乘法这样的一种运算

（2）Somewhat Homomorphic Encryption (SWHE)

允许多种算术运算，但是有限制次数。eg.　一次乘法和无限次加法

（3）Fully Homomorphic Encryption (FHE) 全同态

无限制类型和无限制次数的算术运算

对于加密的global model，FHE将会提供非常好的使用给FL users。没有FHE，则FL users 只能将加密的数据转换为secretly shared format，然后利用MPC协议来训练ML model

在FHE中，基于格的CKKS在PPFL中使用广泛

HE机制也可扩展到threshold或者multi-key version，虽然会降低一些效率，但也进一步提高了安全性

3.3 Secure Multi-Party Computation

Secure Multi-Party Computation (MPC or SMPC)保活一些 homomorphic encryption (HE), Garbled Circuit (GC), Oblivious Transfer (OT), and Secret Sharing Scheme (SSS)

Secret Sharing

（t,n），秘密值s，被分成n份发给n个参与者，只有不少于t个参与者合作才能重构出秘密s，否则，少于t个参与者将得不到关于秘密值s的任何信息。

通常有线性映射加法机制或者Shamir的构造Lagrange多项式

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{t-1} x^{t-1} \bmod p$$

3.4 Differential Privacy

Differential Privacy (DP)

$$\frac{P[\mathcal{M}(\mathcal{X}) \in S]}{P[\mathcal{M}(\mathcal{X}') \in S]} \leq \exp(\epsilon). \qquad (1)$$

隐私预算 ϵ越小，隐私保护水平更高，但是效用也会降低（失去统计性）

一些noise 生成机制：**Gaussian机制**和**Laplace 机制**

根据noise的添加者来分类：

（1）GDP-based

（2）LDP-based

LDP-based PPAgg在FL中是更common和pratical的，但是无法抵抗property inference attacks

3.5 Blockchain

一个block中包含一些transactions以及之前block的hash值，以此来提供链接性和可追踪性

附加到blockchain的每个transaction需要大多数nodes通过共识协议确认

Transparency透明性和Auditability可审性:

每笔交易对所有nodes都是可见的，并可以traced back

3.6 Trusted Execution Environment

Trusted Execution Environment (TEE), processor的一块安全区域，可以用于存储敏感数据或执行代码，在一个被隔离的和受保护的环境中

TEE与REE相对应

需要硬件enclave 的支持： eg. Intel SGX and ARM TrustZone

4. Privacy-Preserving Aggregation Protocols in Federated Learning

4.1 Masking-based Aggregation

Pair-wise masking

SecAGG机制

在server Aggregation 的时候，the masks 自动消失

$$\boldsymbol{y}_i = \boldsymbol{x}_i + \sum_{i<j} \mathrm{PRG}(s_{i,j}) - \sum_{i>j} \mathrm{PRG}(s_{j,i})$$

PRG可以基于seed生成一系列数字，所有的masks会消失

$$\sum_{u_i \in \mathcal{U}} \boldsymbol{y}_i = \sum_{u_i \in \mathcal{U}} \left(\boldsymbol{x}_i + \sum_{i<j} \mathrm{PRG}(s_{i,j}) - \sum_{i>j} \mathrm{PRG}(s_{j,i}) \right) = \sum_{u_i \in \mathcal{U}} \boldsymbol{x}_i$$

Diffie-Hellman(DH)密钥交换协议被用来为每一对(u_i,u_j)协商seed, FL users只需要share seed，可以减小通信开销

优化communication efficiency：通过**quantization techniques**

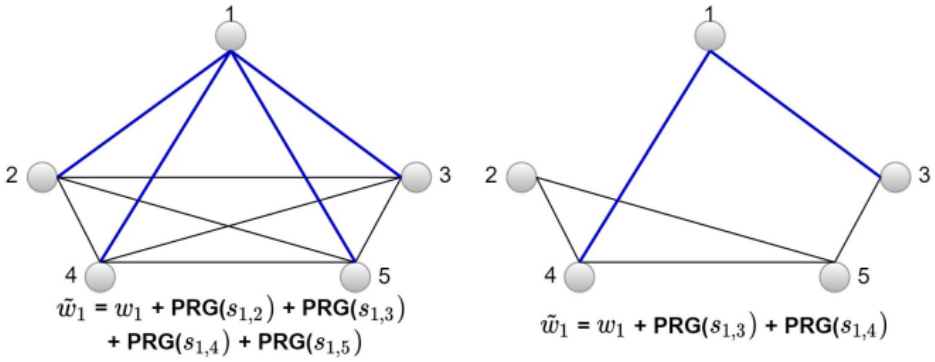
compress user's model : **gradient sparsification**(梯度稀疏化)

减少communication overheads的方法：

只与users的一个subset communication，而不是与所有的users通信

CCESA

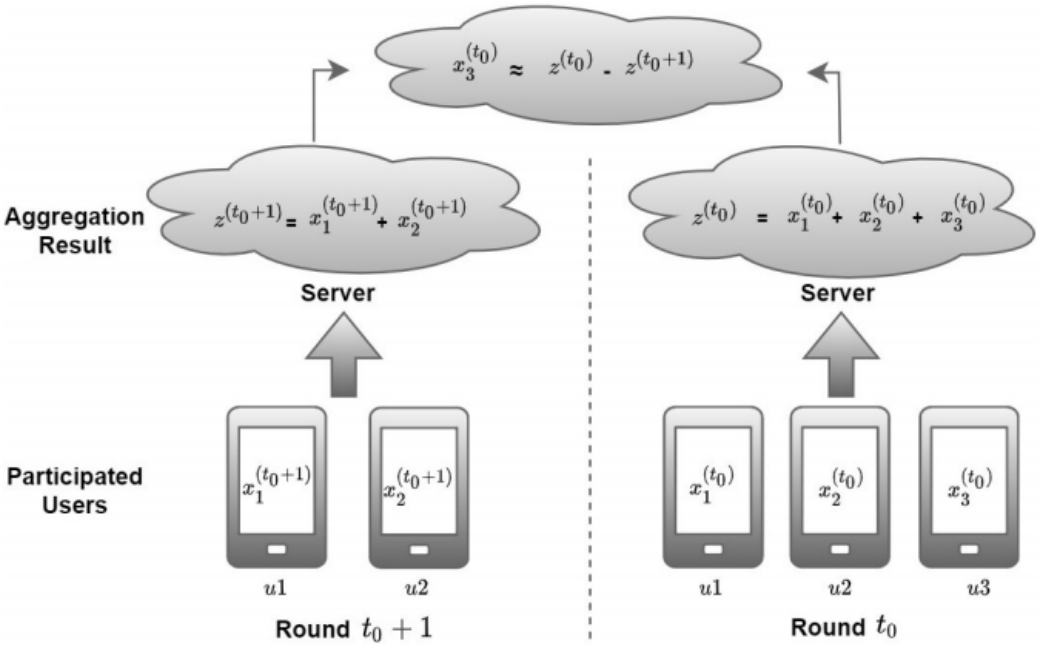
随机稀疏图，每个FL user都是以一定的概率p互相连接，p的选择是安全性水平和协议高效性的trade-off



(a) SecAgg algorithm　　(b) CCESA algorithm

或者替换DH 密钥交换为轻量级的或者非交互的算法

异步的聚合，会导致信息的泄露，由于动态用户的参与。通常是可以使用差分的方法窃取信息，解决方法是从user batch中聚合，而不是一个个 的用户聚合



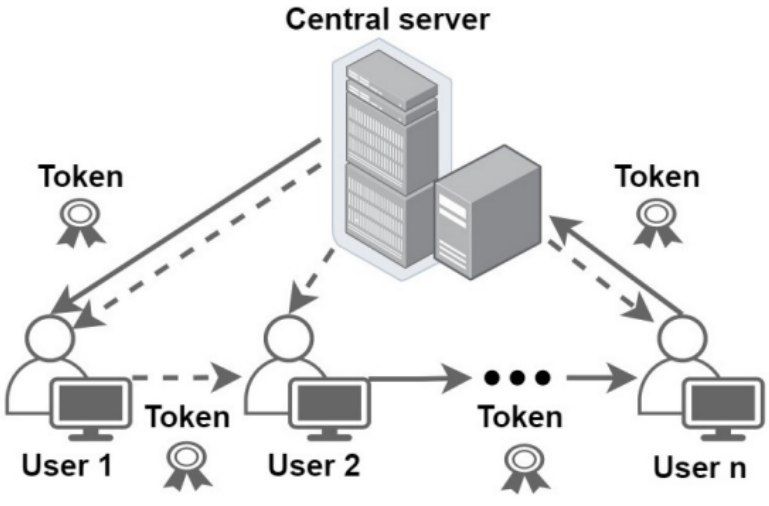
Non-pair-wise masking

pair-wise不适用于large-scale FL system

每个user生成自己的mask，不与其他用户交互，并使用这个mask对自己的vector加密，central server可以在协议中知道所有的mask的和，server就能很轻松的获得所有人的聚合信息

轻量级的mask生成是通过HPRG，同态PRG生成

第一个用户mask自己的model，其他用户链式aggregating自己的model上去，然后最好一个user发送到server，由server进行unmasking



The global model

之前提到的基于mask的方式是为了保护FL users' models或者gradients

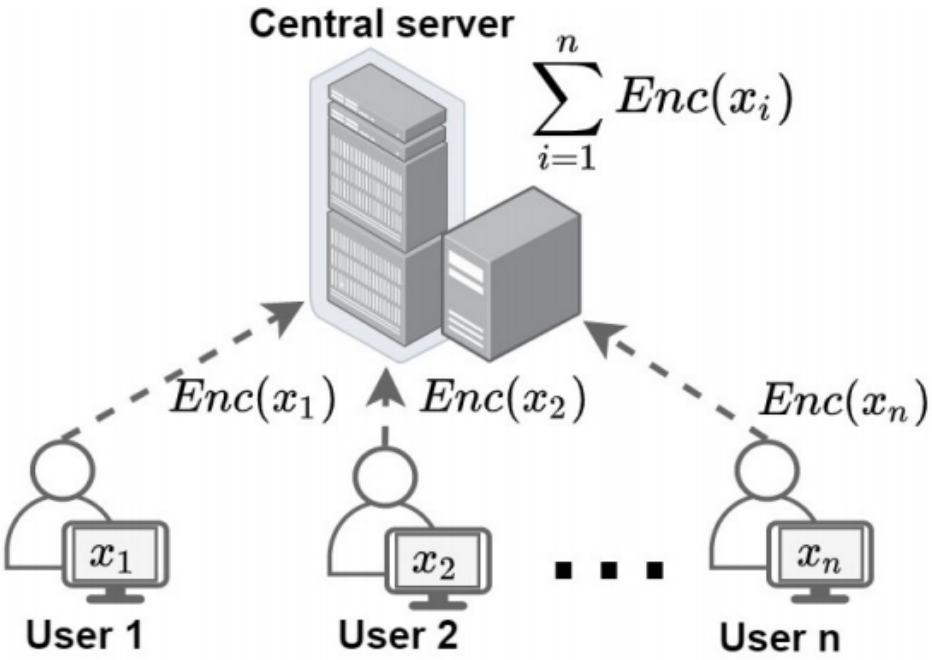
server可以masks自己的global model，然后由所有的FL users在masked model上训练

在模型聚合的时候也可加入noises，保证用户model的隐私

4.2 HE-based Aggregation

FL users在本地encrypt models，发送到central server,在满足加法同态的情况下，server将其累加，然后decrypt就可以得到对应的global model。

此方法中，密钥的管理是最重要的



密钥的管理：

M1 setting

只有FL users知道，central server不知道

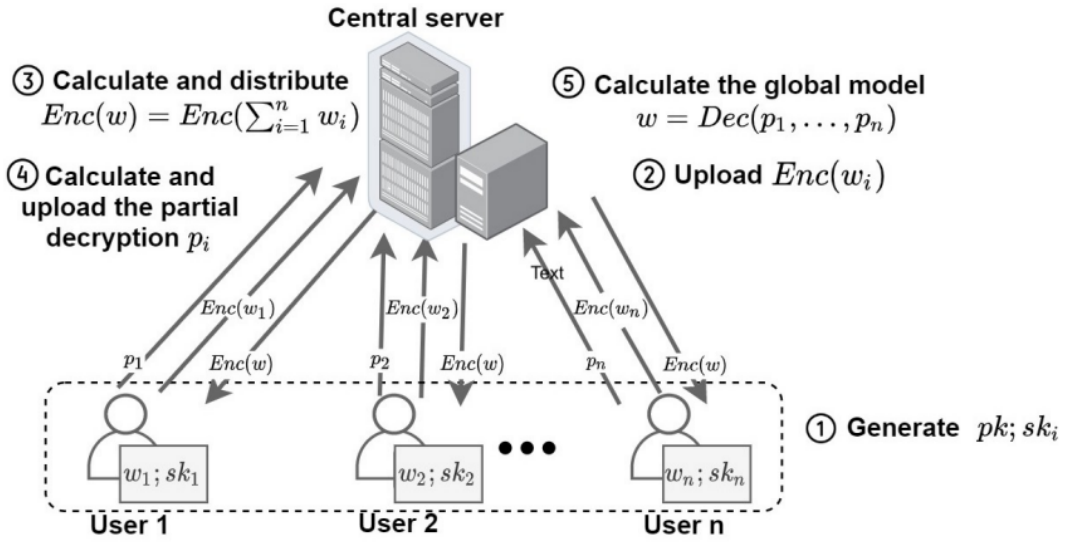
M2 setting

只有central server知道，其他的FL users不知道

M3 setting

在FL 参与者之间划分

半同态PHE



Protecting the global model

FL users train encrypted ciphertext。

需要考虑FHE,基于格的FHE

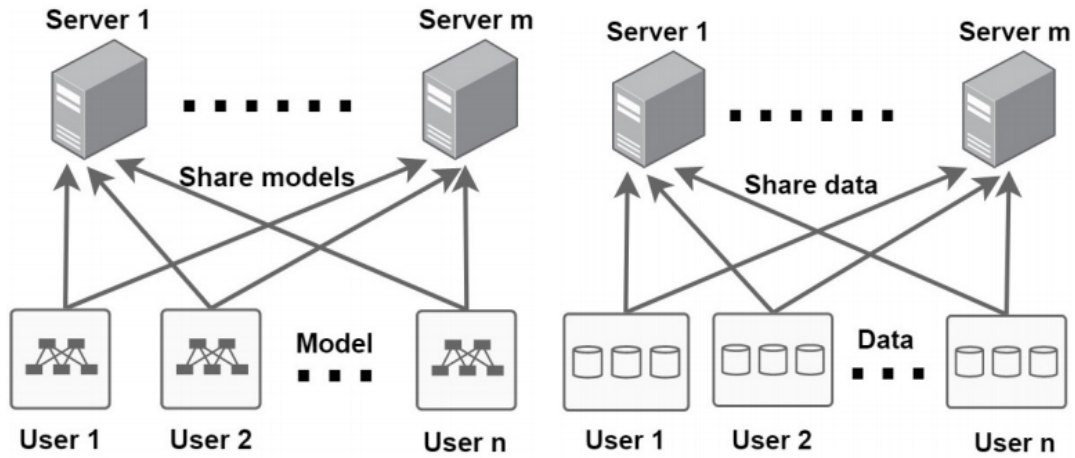
threshold 和multi-key不可避免地会导致更高的开销，在users上去 训练ML models对于深度学习来说是不切实际的

需要权衡privacy guarantee and FL efficiency

4.3 MPC-based Aggregation

Share model

基于MPC的aggregation和train



(a) Aggregation (b) Training

（1）all users 通过MPC支持的协议构建committee

（2）与committee中的其他users分享自己的model，来完成aggregation

Share data

FL用户将本地的数据发送到多个server

MPC的本质限制了它在FL上的应用：

secret sharing会导致较大的通信消耗，在FL中使用MPC，则data或者model必须迁移到很少量的MPC参与者上

而且，MPC需要参与者保持online以及足够的bandwidth，但这是通常mobile或者IoT FL上不具备的

Handle dropped users

secret sharing可以用来处理dropped users以及verify计算结果

只有当一组数量大于阈值的用户验证了聚合结果时，才会确认对聚合结果的正确性的验证。

阈值的设置是关键，需要根据实际情况权衡。阈值大意味着安全性的保障高，但是也导致了协议的可用性会变差。

4.4 DP-based Aggregation

LDP

考虑到没有可信的server，因此使用LDP PPAgg protocols

本地user会先perturb information，然后给server发送扰动值

FedSGD algorithm: 上传扰动的本地gradients，来保护隐私信息

通常就是Gaussian、Laplace 机制添加noise

GDP

LDP的不足：LDP添加了更多的noise，以及可能对model 的可用性造成破坏

Hybrid methods

eg.

- 加法的HE与Gaussian机制的DP ----> HE保护local gradients，DP保护model sharing
- RDP和MPC ---->RDP保护local gradients, MPC保护global models ,前提是数据分布在不同user之间相似且model不是独立同分布的
- MPC和DP 保护client information以及global model
- LDP和function encryption 实现data-level 和 content-level的PP
- Compression和 secure aggregation和DP
- LDP and Shufflfled Model

Trade-off: **information privacy** 和 **model utility**

适当结合PPAgg protocols以及DP protocols可以提高模型的可用性

适当选择privacy budget可以提高DP模型的可用性

4.5 Blockchain-based Aggregation

blockchain提供了**auditability(可审性**）以及**anoonymity**（匿名性）

blockchain-based aggregation protocol:

(1) publisher用一个初始的global model发布FL training task

(2) 每个FL user 获取global model

(3) train local model

(4) 生成并broadcasts(广播)一笔记录自己local model的transaction

(5) 选取consensus(共识)节点，将这些local models聚合，通过consensus protocol得到global model

(6) 新的global model被包含在一个block中，之后被添加到blockchain中，用于下一轮的training

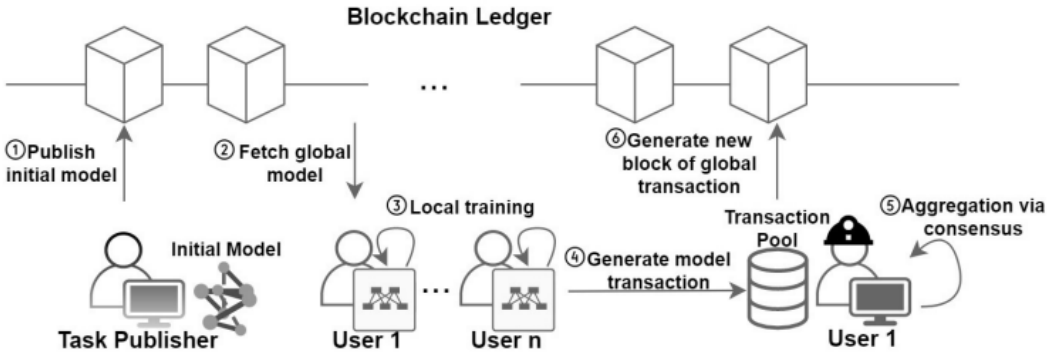


Fig. 9. An illustrative figure of the aggregation in FL based on blockchain.

Integration of PPTs

FL 可能在Aggregation的时候导致信息的泄露，因为用户上传到server的model是明文信息

- CDP
- LDP在user upload之前被扰动
- Paillier crypto systems ，使用threshold

Smart contract

除了PPTs，还可以使用**智能合约**来提高聚合的安全性和效率

4.6 TEE-based Aggregation

TEE-based aggregation protocol：
(1) 所有users加密local trained models，并发送给REE

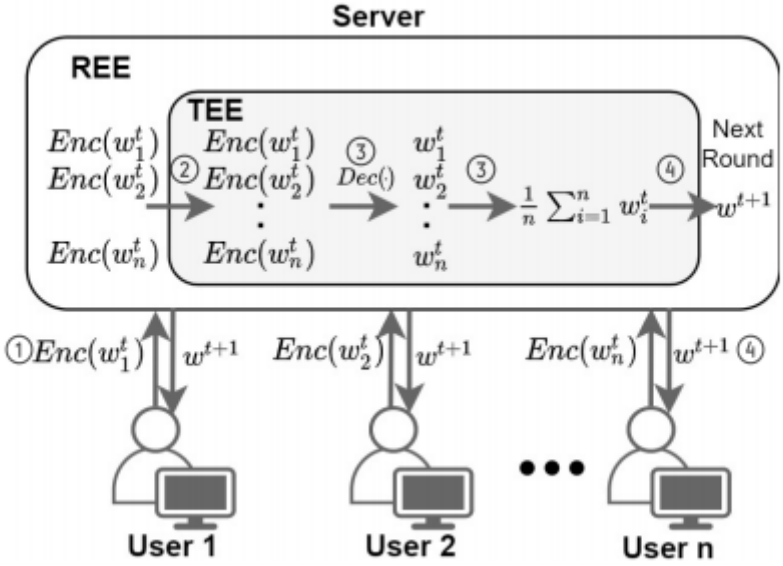
(2) TEE 从REE loads收到的加密后的models

(3) 解密并聚合

(4) TEE输出结果，并发送到REE，分配给所有用户

此外，可以在user 上传到TEE之前使用DP扰动

也可在TEE上传用户model到central server之前，shuffle一次



在TEE中部署ML算法是比较困难的，因为内存受限

此外，**signature**和**verification**机制也是可以与TEE共同使用的

基于TEE的Aggregation是需要格外的硬件成本的

4.7 Discussion

DP-based PPAgg是最轻量级的，但是会时去性能

HE和MPC可以构建更加精确解，但是代价也昂贵

加法HE是比较合适的，直接就能支持Aggregation

在整个FL workflow上使用PP是需要FHE的支持，但是这个是在大型ML上不可行的

MPC比HE更高效，但是需要大量的communition的开销

masking-based PPAgg是较有前景的，可用于cross-device FL

5. Federated Learning Frameworks for privacy-preserving aggregation

6. Challenges and future directions

6.1 Throughput Improvement

对于大规模的网络结构，现有的PPAgg在构建block的时候利用了昂贵的加密技术，需要大量的computation以及communication.

提高吞吐量的方法：

(1) batch encryption

(2) SIMD in HE

(3) parallel hardware: FPGA/ GPU

结合Compression techniques:
Top-k稀疏化，但是需要顺序信息来重构原始的vector

但是这样的按顺序的vector可能会泄露隐私信息，然而如果按顺序的加密信息则会更加困难

6.2 Hybrid Schemes for Stronger Security

poisoning attacks中毒攻击

目的是降低准确度

random attacks 或者诱导FL model输出目标label，通过操纵本地数据来完成有针对性的或者后门攻击

PPAgg协议无法抵抗数据中毒攻击

server可以检测异常，但是这种机制需要访问user的data以及model，这与PPFL是违背的

Byzantine-resilient aggregation算法是符合的，不会访问user的data以及model

inference attacks推理攻击

探测目标的ML模型，获取membership\attribute\data

PPAgg协议通过保护users' models来缓解这个问题

但是LDP-based Aggregation就不能抵抗这种攻击，GDP-based Aggregation只有在牺牲很大的实用性的基础上可以抵抗。

融合其他的一些PPTs，与PPAgg协议共同加强安全性

7 Conclusion

1、overview of federated learning

2、introduced the basic knowledge of supporting tools for constructing PPAgg protocols

3、 reviews and analyses of different constructions of PPAgg protocols in detail to deal with a variety of privacy issues in FL systems

4、outlined existing challenges as well as several directions for future research.