Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

Overview of Federal Learning and Privacy Protection

Ziyue Tian

Academe of Electrical and Information Engineering, Northeast Petroleum University, Daqing, China

210601140511@stu.nepu.edu.cn

Abstract. In recent years, federated learning technology has developed rapidly and been widely used in the field of data processing. This paper makes a comprehensive discussion on the privacy protection methods in federated learning,and makes a detailed analysis of three basic methods:data encryption,data perturbation and trusted hardware-based,introduces the principle of each method,and objectively analyzes their performance in practical applications. This paper gives a visual example to compare the principle of secure multi-party computing, which is easy for readers to understand.Aiming at homorphic, this paper first introduces the background of this method and explains the algorith with block diagram,and then compares and analyzes the advantages and disadvantages of somewhat homorphic encryption and fully homorphic encryption. The specific denoising mechanisms are divided into three types:Laplace mechanism,Gauss mechanism and exponential mechanism.This paper summarizes these three mechanisms according to the types of data that need to be denoised.Finally,the paper systematically expounds the privacy protection methods based on trusted hardware ,lists two typical schemes,TrustZone and SGX,and analyzes how they work. In this paper,the hot spots and development prospects are prospected.

Keywords: federated learning; privacy protection; secure multi-party computing; homorphic encryption; differential privacy.

1. Introduction

Federated learning grew out of the main idea presented by McMahan in a 2016 paper in Google Seattle:bringing model training to edge data.That is,the original data of each participant is saved locally,not exchanged or transmitted,and only upload the locally updated models[1].Its primary purpose is to protect privacy,and it is a type of machine learning that supports large-scale participants,with multiple parties collaborating to solve model updates under the coordination of a federated server or service provider.It is mainly in the use of reverse derivative optimization machine learning caculations.The method uploads local gradient update values in each round of training and then from the center[2].The server collects and updates the maintenance center model together with weighted summation.Federated learning has obvious distributed characteristics,and each user node has the right to update the global model.Federated learning has already been applied to word prediction for Android cell phones,and the model is uploaded to the server when the phone is charged.

This article mainly introduces the common privacy protection methods in federated learning and evaluates their performance.From the perspective of protection means,there are three main forms of privacy protection.The first is based on data encryption cryptographic protection method to protect the data,this form is widely used.The methods used are homomorphic encryption,secure multi-party computing and so on. The second type is based on data disturbance,the specific methods is differential privacy， which blurres the original data by adding noise.And the third type is a privacy information protection technology based on trusted hardware.The common technology is the trusted execution environment,which protects the data security by creating a secure and isolated execution environment (Figure 1).

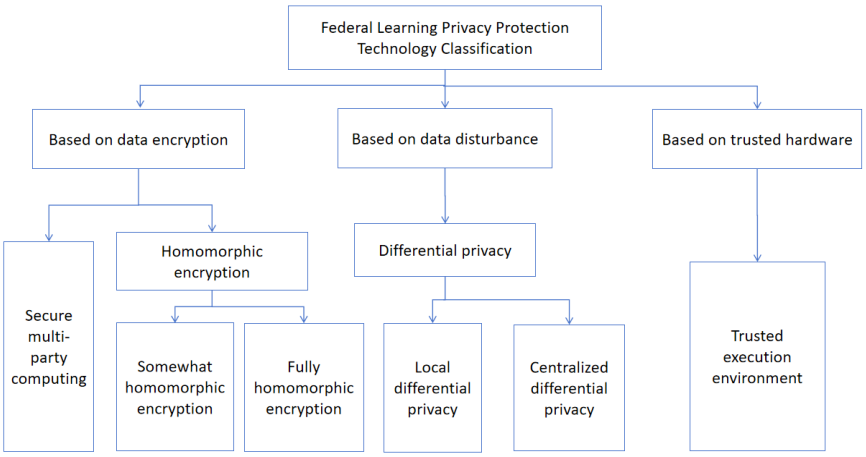

Figure 1. Classification of privacy protection technologies (Photo/Picture credit :Original)

2. Privacy protection method based on data encryption

2.1. Secure multi-party computing

Secure multi-party computing is a cryptographic technique that allows multiple parties holding private data to compute a specific result together [3].This specific result is obtained by combining the data of each participant,and during the entire process,the participant is not informed of any data other than himself,and the nature,type or other privacy related information of the data entered by the participant is not disclosed to the public [2]. Only when the entire caculation process is completed can participants receive the output they need.This result may be a judgment based on the above private data,or an approved transaction.In short,MPC technology can collect distributed confidential information,analyze it,and then output the results. In this way,no single participant knows the underlying truth,thus greatly reducing the possibility of a single participant reconstructing or leaking confidential information.

In order to understand how the privacy protection method of secure multi-party computation works,let's take a simple example to simulate the computation process.Suppose there are two employees A and B in a company,and they want to know whether their salary is the same as each other's while chatting,but they do not want to tell each other their actual salary.There is a way around this problem without the involvement of third parties.A can prepare a number of different mailboxes,and then paste on each mailbox all the possible salary levels of the company's employees,and take the key to the mailbox corresponding to his salary.Then B prepares some pieces of paper and puts the piece of paper with "Yes" on it into the mailbox corresponding to his salary,and puts the other pieces of paper with "no" on it.When A opens the mailbox,if the note written "No",it means that B's salary is different from A.In this way,they successfully exchange information whlie protecting their privacy:neither of them knows who is earning more or less.Their secrets are safe.Of course,the veracity of the output depends on honesty of the participants[4].Therefore,an effective MPC must meet two requirement (Figure 2).The first is that no one can infer the confidential information of each participant from the execution of the agreement,that is,the results of the caculation will not give the holder any indication of the private information held by the participant.The second is that the information uploaded by each participant must be true and reliable.

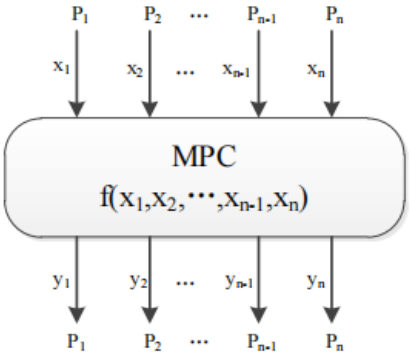

Figure 2. Secure multiparty computing (Photo/Picture credit :Original)

2.2. Homomorphic encryption

Homomorphic encryption is another cryptographic privacy protection method based on data encryption technology. The concept of homomorphic encryption was first proposed by Ron Rivest et al. In 1978 to solve the problem of processing data without touching the data[2].The application of homomorphic encryption technology can directly perform multiple multiplication or addition operations on the encrypted data without decrypting the encrypted data,and then decrypt the final result after the operation is completed.The result is the same as the result of the operation on unencrypted data [5].

Typically,homomorphic encryption uses the public key for data encryption,and only the owner of the data has the private key to decrypt the data [1]. During this time,third-party providers can perform certain types of operations on the data without viewing it.Therefore,it is considered as a special encryption mechanism that can protect security and privacy issues without affecting operational efficiency.Take cloud computing as an example.In homomorphic encryption technology,if a user wants to query some information on a cloud server,the data will undergo the following processing process.First,the data owner encrypts the data and stores it in a cloud server.When users wants to query relevant information,they need to send the information to the cloud server.The cloud server uses a homomorphic encryption algorithm to predict the encrypted data and help users find the information to be queried without knowing the specific content.The cloud server then returns the encrypted prediction to the user. Finally,the user uses the private key to decrypt the predicted results,and can view the research content,while protecting their privacy (Figure 3).

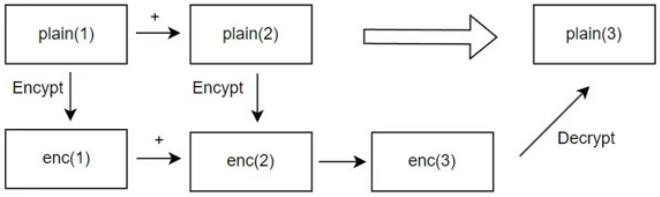

Figure 3. Homomorphic encryption (Photo/Picture credit :Original)

In summary, homomorphic encryption has these advantages. Using homomorphic encryption technology allows enterprises to securely utilize cloud computing and storage services,which ensures the security of available data.Businesses do not have to rely on cloud services to protect their private data while retaining their ability to perform computations.At the same time, homomorphic encryption technology allows third parties to participate in the business of processing sensitive data without worrying about the risk of sensitive information leakage.

According to different types and times of operations, homomorphic encryption can be divided into Somewhat homomorphic encryption and full homomorphic encryption.

2.2.1. Somewhat homomorphic encryption

Somewhat homomorphic encryption supports a finite number of operations or a single class of operations.For example, Paillier supports addition between ciphertexts but not multiplication between

ciphertexts. BGN supports unlimited addition and multiplication between ciphertexts.This means that the depth of their encryption is limited.Therefore,somewhat homomorphic encryption is generally not used as an independent privacy computing scheme,but more as a security enhancement measure for federated learning algorithms.

2.2.2. Fully homomorphic encryption

Fully homomorphic encryption means that the encryption process and subsequent operations are not restricted by any constraints, and the privacy protection reaches an idealized level. However,due to the need to consume a lot of computing time, it can not reach the practical standard at present.Besides,fully homomorphic encryption only supports addition and multiplication operations, and some scenarios that require more complex computational methods,such as the root operation,or some mathematical statistics methods,there is often a greater loss of precision. Table I. Summary of different homomorphic encryption techniques.

Table 1. Summary of different homomorphic encryption techniques

	advantage	**disadvantage**
Somewhat homomorphic encryption	Through homomorphic encryption,users can safely use cloud services to process data,let third parties participate in the business of processing sensitive data,and do not worry about the risk of privacy leakage	Only a limited number of operations or a single operation can be performed on the ciphertext,so the encryption depth is limited
Fully homomorphic encryption		The caculation is slow and inefficient.Currently,only addition and multiplication are support for ciphertext

2.3. The disadvantages of homomorphic encryption algorithm

Compared with the plaintext caculation,the ciphertext caculation is N orders of magnitude slower.Taking fully homomorphic encryption as an example,on a general-purpose chip,it takes 100,000times longer to compute a ciphertext than it does to compute the same plaintext.In addition,since the space required to store a ciphertext is much larger than the plaintext,this further reduces the computational efficiency.So,if you want to use homomorphic encryption technology,you also have to have a large enough memory to store data.In addition,due to the limitations of the technical level,the current encrypted data only supports a limited number of arithmetic operations,namely addition and multiplication,and can not be performed at a higher level.

3. Privacy protection method based on data disturbance

3.1. Differential privacy

Differential privacy is a privacy protection method based on data disturbance,which distorts sensitive data by adding random noise to the original data,so that the outside world can not get private information.Differential privacy introduces a privacy loss or privacy budget parameter known as “epsilon” that controls the amount of noise and randomness added to the raw data [6]. The parameter episilon controls the trade-off between privacy and data accuracy:in general,the higher the episilon,the more accurate the data,but the less privacy[6].The realization of differential privacy in federated learning can be divided into three steps:clipping,aggregation and strengthening. First,the participant computes the model gradient locally,and then sends the gradient to the server.After the server receives the gradient,the gradient clipping is performed first, and then the aggregation is performed [7]. The purpose of clipping is to limit the gradient range and prevent the gradient of some participants from

being too large,thus affecting the aggregation effect.After aggregation,the server adds noise to the gradient,then sending the gradient with noise to all participants. Participants receive the noisy gradient and use it to update their model.

Differential privacy is mainly applied to the scenario where the data of multiple clients is summarized into a third-party calculation (Figure 4). Take a simple example:a local government wants to count the income of local residents in the economic census to calculate per capital income level. At this point,statistical departments can use differential privacy technology to protect the real income of residents.Suppose the income of n residents is S_i .Here we introduce a random noise X with an expected value of 0.When residents declare their income,the random noise X_i is automatically added to the income S_i ,that is,the income data for each resident shown in the statistics department's system is $R_i = S_i + X_i$. In the process of caculating the average income,the noise value of each reported income data is based on the same random distribution and is sampled independently.When the number of inhabitants is largely enough, n is large enough, $\frac{1}{n}\sum_{i=1}^{n} X_i = 0$ .Thus, $\sum_{i=1}^{n} S_i$ can be obtained by directly calculating $\sum_{i=1}^{n} R_i$, and the actual per capita income can be calculated indirectly. Differential privacy can be implemented locally or globally[6][7].

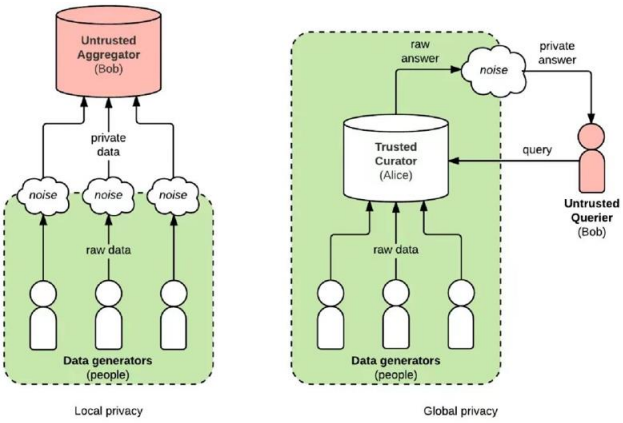

Figure 4. Model of differential privacy [6]

3.2. Classification of differential privacy by application scenario

3.2.1. Centralized Differential Privacy

Centralized Differential Privacy (CDP) is a privacy protection method wherein the original data is centralized to a trusted data center, where a differential privacy algorithm is applied to process the data, and the processed data is then released [7]. This approach ensures data privacy through centralized management, yet it also poses limitations and challenges.

On one hand, CDP relies on trusted third-party data collectors to manage and protect the data. However, finding fully trustworthy third-party data collectors is difficult in real-world scenarios. For instance, in the medical domain, hospitals may need to collaborate with third-party data collectors to share patient data for research purposes, raising concerns among patients about potential misuse or leakage of their sensitive information. On the other hand, CDP is vulnerable to risks of data center single points of failure. If a data center suffers an attack or data breach, it could lead to the potential exposure of privacy information of a large number of users, posing a threat to individual privacy.

3.2.2. Local Differential Privacy

Local Differential Privacy (LDP) is an emerging privacy protection method where the responsibility for data privacy is shifted from the data center to each user[6]. In LDP, users process and protect their personal data themselves instead of centralizing it to a data center for processing. The hallmark of this approach is that users randomize their data before uploading it, ensuring the privacy of individual data.

For example, consider a social media application aiming to collect statistics on user clicks on specific advertisements while preserving user privacy. In LDP, each user randomizes their click behavior, such as by adding random noise, before uploading the processed click count to the server[6]. This allows the server to obtain statistical information on advertisement clicks while safeguarding user privacy.

One of the advantages of LDP is its ability to significantly reduce the risk of data leakage[6]. Since data processing and randomization occur at the user's end, it becomes challenging for attackers to infer specific information about the original data even if they gain access to the processed data. However, compared to CDP, LDP may introduce some level of data distortion due to the introduction of random noise during processing[6].

3.3. Several mechanisms to implement data disturbance

The differential privacy technology mainly includes Laplacian mechanism,Gaussian mechanism and exponential mechanism,among which the first two are mainly suitable for the noise of numerical data,and the last one is suitable for the noise of non-numerical data.

3.3.1. Laplace mechanism

The idea of Laplace's mechanism is to add an independent zero-mean Laplace distribution of noise to a numerical query f,the size of which depends on the privacy budget and the sensitivity of the query function.For a function f(x) that can output a numerical result, F(x) is defined as follows to satisfy differential privacy: $F(x)=f(x)+Lap(s/\mathcal{E})$[6].Where,s is the sensitivity of f,which refers to when the input changes from data set x to adjacent data set x',the output change of f Lap(s) represents the Laplacian distribution sampling with a mean of 0 and a reduction coefficient of s.

3.3.2. Gaussian mechanism

The Gaussian mechanism is also applied to the numerical queery f, its idea is to add an independent zero mean Gaussian noise distribution on the basis of f. The Gaussian mechanism approximates a real-valued function f:D→R with a differential privacy mechanism.The specific Gaussian mechanism adds calibrated Gaussian noise to the sensitivity S_f of the function data set.The sensitivity is defined as the absolute distance $\| f(d)-f(d')\|_2$ between two inputs d,d' that are close to each other.

3.3.3. Exponential mechanism

If needed to add noise to non-numerical discrete results,it should apply an exponential mechanism.The exponential mechanism is defined as :a random algorithm M is provided,the input is a data set D,the output threshold is RANGE,and the output value is an entity object $r\epsilon$RANGE. q(D,r) is the availability function and Δq is the sensitivity of the availability function.If algorithm M selects from RANGE and outputs r with a probability proportional to $\exp(\frac{\mathcal{E}q(D,r)}{2\Delta q})$,then algorithm M provides Ɛ-differential privacy protection.The differential privacy protection mechanism of the index mechanism can be well applied to the voting system that does not need to master the specific votes.While protecting the privacy of voters,it can intuitively understand the preference degree of the public.

For example,a class is running for class leader,and the candidates are A,B,C,D.All the classmates will use a voting software to vote,based on the voting results to determine the class leader.In order to ensure the privacy of user voting,differential privacy of exponential mechanism can be used to meet the requirements.Express mathematically,the voter is data set D,the output field RANGE{A,B,C,D},and the number of votes is taken as the availability function q,then q(D,A)=30 and sensitivity Δq=1,caculate the value of $\exp(\frac{\mathcal{E}q(D,r)}{2\Delta q})$ under the given Ɛ,and finally normalize to find the probability of each candidate being elected.When the differential privacy budget of Ɛ=1 is added,it can be found that the vote of A is returned with A has 92.4% probability of being elected.The smaller the value of Ɛ,the higher the degree of privacy protection and the worse the data availability.When Ɛ=0,user voting is indistinguishable,privacy protection is the highest,and data availability is lost.

4. Privacy protection method based on trusted hardware

The core idea of trusted execution environment (TEE) is isolation, that is, the execution environment and the operation system are seperated and run in parallel [7]. It provides a safe and trusted space for the privacy data information that needs security protection to ensure the safety and integrity of the program and data in the environment.Its security is guaranteed by the relevant hardware mechanisms.Within the specific area provided by the CPU for the TEE,the execution of data and code is completely isolated from the external environment.Applications within the TEE can normally access the full function of the CPU and memory,but are not affected by other programs in the operating system [7].

In application,TEE depends on different implementation schemes provided by specific technology manufacturers.They differ in the mechanisms they use to achieve isolation,as well as in the functionality they support.Currently,the most representative solutions are ARM's TrustZone and Intel's SGX.

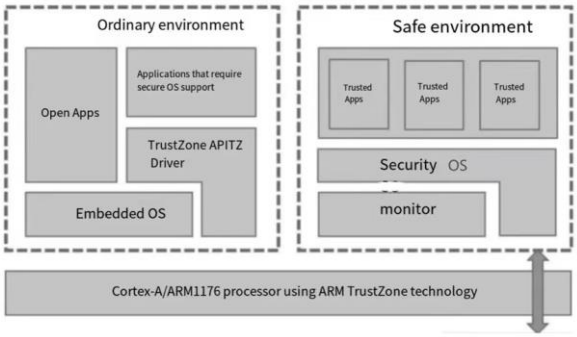

Figure 5. Trusted execution environment [7]

4.1. TrustZone

TrustZone divides the hardware and software resources of a system into Secure World and Normal World.The Secure World has higher execution permissions.All operations that need to be kept secret are executed in the Secure World.The rest of the general operations are performed in the Normal World and do not have access to the Secure World.Developers implement specific security features by developing more trusted applications using API provided by secure operating systems.The execution of trusted applications needs to establish a chain of trust through verification.When a program wants to enter the security environment ,the verification operating system needs to check its security,and only the programs that pass the inspection can enter the security environmnt.The safe environment is converted to the normal environment by a mode called Monitor Mode [8].

4.2. Software Guard Extension

Software Guard Extension is an instruction set extension introduced by Intel in 2013.SGX allows an application to create an enclave by seperating out a region in its address space,and only the code inside the enclave can access the memory region where the enclave resides[9].Even the operating system can not affect the code and data inside the enclave.The security boundary of an enclave contains only the CPU and itself.A CPU can run multiple enclaves independently of each other,preventing the security of the entire system if one enclave is damaged.The entire trust mechanism of SGX is focused on the CPU, so the deployment of SGX is relatively simple,as long as the code definition in the enclave can be deployed.

5. Conclusion

This article summarizes several privacy protection methods commonly used in federated learning,and specifically introduces the methods to implement them,as well as the advantages and disadvantages of each method.The existing mainstream privacy protection methods are applicable to different

scenarios.This article compares some different protection methods,so that readers can clearly understand the scope of application of the above methods and make reasonable choises in specific applications.

Although federated learning offers many ways to protect users' privacy,there are still many problems that need to be improved.For example,the running speed of some methods is slow and inefficient,resulting in the failure to promote the use of a large scale;Some methods take up a lot of storage space during implementation;There are also ways to solve the problem of high communication overhead.

Federated learning,as a new technology from its emergence to development,is gradually improving and maturing,and has broad application prospects.Future research on federated learning may be more inclined to improve the efficiency on the basis of ensuring the accuracy of the model,extend it to larger scale environments,and further improve the technology.

Reference

[1] H. Wang, Y. Liang, L. Li, R. Li. Survey on privacy-Preserving Mechanism in Federated learning. Modern Computer, vol. 28(14), pp. 1-12, 2022.

[2] S. Xiong, D. He, Z. Wang, R. Du. A Review of Federated Learning and its Security and Privacy Protection. Computer Engineering, pp. 1-17, 2024.

[3] G. Wu. Secure multi-party computing privacy protection and data security in communication networks. Modern Transmission, vol. 2023(04), pp. 51-54.

[4] X. Xiao, Z. Tang, B. Xiao, K.-L. Li. A Survey on Privacy and Security Issues in Federated Learning. vol. 46(5), pp. 1019-1044, 2023.

[5] X. Liu, F. Xu, Z. Ma, M. Yuan, H. Qian. Research on Privacy Protection Technology in Federated Learning. Journal of Information Security Research, vol. 10(03), pp. 194-201, 2024.

[6] Y. Xie. Federal Learning Privacy Protection Method based on Local Differential Privacy. Information Technology and Informatization, vol. 08, pp. 160-163, 2023.

[7] R. Xu, L. Dai, D. Xia, et al. Research on Centralized Differential Privacy Algorithm for Federated Learning. Netinfo Security, vol. 24(1), pp. 69-79, 2024.

[8] G. Fan, P. Dong. Research on Trusted Execution Environment Building Technology Based on TrustZone. Netinfo Security, vol. 2016 (3), pp. 21-27.

[9] J. Cui, Z. Cai, K. Liu. A survey on SGX isolation technology. J. Huazhong Univ. of Sci. & Tech. (Natural Science Edition), vol. 52(2), pp. 1-15, 2024.