Anonymous and efficient authentication scheme for privacy-preserving distributed learning

Mini Project

Subhajit Ghosh (2102041)

Indian Institute of Information Technology Guwahati

Objective

- To provide an overview of the anonymous and efficient authentication scheme for privacy-preserving distributed learning.
- This proposed work has helped to alleviate some of the drawbacks of privacy-preserving in distributed learning or has innovated new technologies for various applications.
- I attempted to identify the paper’s problems while also comprehending the methodology (that is, the working model and algorithms).

Introduction

- A tremendous amount of data has been generated and transmitted to cloud servers due to the integration of machine learning (ML) and cloud computing, which has been developed rapidly in the past decade.

  To resolve this issue, the concept of distributed learning (DL) is used.

- In this, the participating devices first train the ML model locally with the raw data, and once computed, it sends the model parameters to the cloud. The cloud then aggregates into a global ML model.

- To tackle this challenge of privacy violations , deploying **encryption techniques** or **differential privacy (DP)** are two primary approaches.
- Encryption techniques, such as homomorphic encryption and secure multiparty computation, are employed to design secure data exchange and aggregation protocols.

  Encryption techniques is not generalised for all ML algorithms in DL.

- Differential privacy (DP) is deployed to obfuscate information by introducing random noise to the raw data or model parameters. DP-based schemes are generalised for all ML algorithms in DL.

  The randomness reduces data utility, bringing undesired accuracy degradation to ML models.

  Model convergence with longer latency.

The proposed scheme is based on the fact that if the participants in the DL system are anonymous, then although the raw data is recovered, the adversary cannot link the data to the corresponding participant's real identity, thereby achieving privacy preservation.

- In this scheme, it deploys anonymous authentication.
- The new generalised certificateless signature scheme PCLS, which is built on pairing-based cryptography, The designed signature scheme has an anonymous and efficient authentication protocol that supports batch verification (AAPBV).

Problem statement

- The major problems addressed in this paper are:

  Adversary is able to derive the raw data (user identity) by analysing the obtained machine learning models.

  Significantly reduces the time consumption in batch verification, achieving high computational efficiency.

Proposed Methodology

System Overview

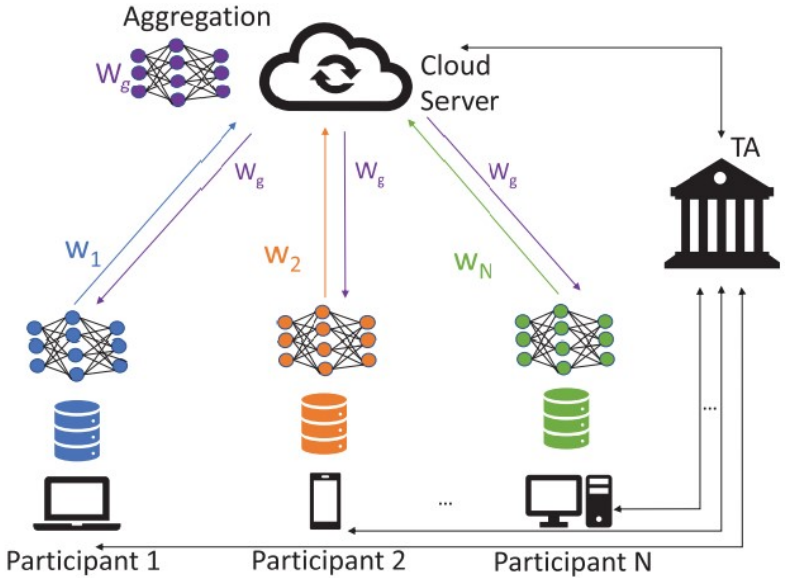
- Participants: It refer to individuals that are equipped with intelligent devices, such as laptops , mobiles.
- Cloud Server: It receives the model parameters from participants and aggregates into a global ML model.
- Trusted Authority (TA): TA is a powerful system manager that is responsible for system initialization, participant registrations, key generation, identity management, and so on.

Proposed Methodology

Model Flow

- Step 1: The cloud server initializes a global model and distributes the model to all participants.
- Step 2: After receiving the global model, each participant continues to train the model on its local dataset and obtains the updated model parameters Wn. Wn is then uploaded to the cloud server.
- Step 3: The cloud server aggregates the model parameters. The global ML model is corrected with Wg and sent to the participants for next-round training.
- Step 4: Repeat Steps 2, 3 until the global model converges.

Proposed Methodology



Proposed Methodology

Pairing based certificateless signature scheme (PCLS)
It lays down a design basis for the proposed anonymous authentication protocol.

- Instead of setting the signer's real identity as the partial public key, the hash value of the signer's real identity is the partial public key, protecting the real identities from disclosure.
- Modified the signature generation and message verification to support batch verification, such that the authentication is more efficient.

Proposed Methodology

Pairing based certificateless signature scheme (PCLS)

It involves following steps:

- Setup
- Key Extract
- Sign
- Verify

Proposed Methodology

Pairing based certificateless signature scheme (PCLS)

Setup

- TA publishes (I, p, P, e, G1, G2, H1, H2, Y_{TA}) as the system parameters.
- Private key X_{TA} is kept by TA and remains as a secret.

Key-Extract

- The signer randomly selects X' as its partial private key and calculates Y' = X'P as its partial public key.
- The signer computes Y" = H1(I D) as the other partial public key and sends Y" to TA, where ID is its true identity.
- Given Y", TA then sets X" = X_{TA} Y" and sends to the signer through a secure channel. The signer uses (X', X") as its full private key and (Y, Y") as its full public key.

Proposed Methodology

Pairing based certificateless signature scheme (PCLS)

Sign

k = H2(m $|| Y' || Y'' || Y_{TA}, rP)$

U = X' Y_{TA} -X"k

The signature is = (k, U) on message m.

Verify

After receiving the message m with signature , the verifier checks

e(Y' , Y_{TA}) = e(U, P)e(Y", $Y_{TA})^k$

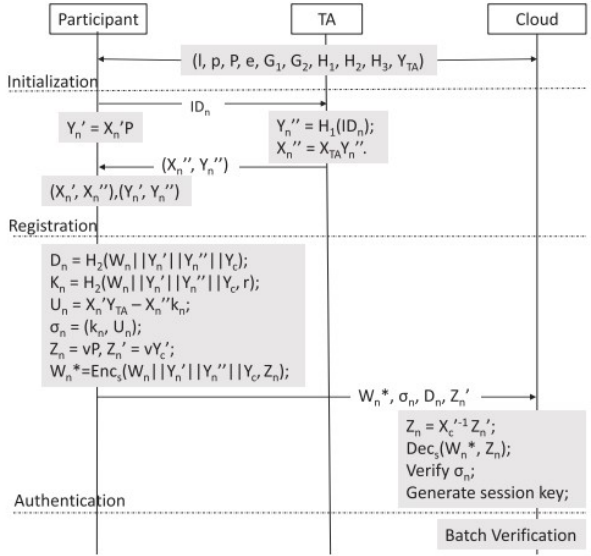
Proposed Methodology

Anonymous authentication protocol AAPBV

The proposed AAPBV involves four stages

- Initialization stage
- Registration stage
- Authentication stage
- Batch verification stage

Proposed Methodology



Result

Protocols	Sign	Verify
Liu et al. [11]	$5T_m$	$T_e + 4T_m + T_p$
Zhang at al.[12]	$9T_m$	$6T_m$
Yang et al. [14]	T_e+T_m	$T_e + 5T_p$
Shen et al. [15]	$T_e + T_m + T_p$	$T_e + T_m + 2T_p$
Proposed	$4T_m$	$T_e + 2m + 3p$

COMPARISON OF THE COMPUTATIONAL TIME WITHOUT BATCH VERIFICATION

Result

Protocols	Sign	Verify
Liu et al. [11]	75.875	80.26
Zhang at al. [12]	136.575	91.05
Yang et al. [14]	16.675	91.8
Shen et al. [15]	34.735	52.795
Proposed	60.7	86.03

COMPARISON OF THE COMPUTATIONAL TIME WITHOUT BATCH VERIFICATION

Result

Protocols	Time
Yang et al. [14]	$5(n+1)T_m+5T_p$
Shen et al. [15]	$nT_e + nT_m + 2nT_p$
Proposed	$nT_m + 3T_p$

COMPUTATIONAL TIME OF BATCH VERIFICATION

Conclusions

- The proposed protocol achieves efficient batch verification with a slight sacrifice of computational efficiency in single verification. In addition, benefiting from the designed DBV algorithm, the global model converges quickly under the proposed AAPBV protocol, further improving the computational efficiency in DL.
- The proposed protocol guarantees various security properties while supporting batch verification. The protocol provides high efficiency and enhanced security with a small sacrifice in computational costs in single authentication, making it more practical for large-scale DL systems.

Thank you!