CS352 Cryptography and Applications Lab Project Proposal

Differentially Private Federated Learning with Homomorphic Encryption mechanism

Team: (Names in ascending order of roll numbers)
Mayur Bhat (181CO132), Sukruth N Bhat (181CO154), Videh Raj Nema (181CO158)

Introduction

Federated Learning (FL) enables organizations to collaboratively train a Machine Learning (ML) model by aggregating local gradient updates from each client without sharing privacy-sensitive data. In other words, a trusted curator aggregates parameters optimized in a decentralized manner by multiple clients. The resulting model is then distributed back to all the clients. However, this protocol is sensitive to differential attacks as the distributed model can be analyzed and a client's contribution during training, and information about their data can be revealed. As a solution to this problem, we implement an algorithm for client-sided differential privacy-preserving federated optimization [1, 2]. This incorporates both data-level as well as client-level differential privacy (dp) for the system.

However, during the aggregation process in federated learning, information about the client can be revealed. To overcome this, we use Homomorphic Encryption (HE) mechanisms [3, 4, 5]. HE allows the clients to mask the local gradient updates by enabling computations to be performed on encrypted data. **HE ensures that performing operations on encrypted data and decrypting the results is equivalent to performing analogous operations without any encryption**. We intend to show applications in machine learning problems [1, 2] and depending on the time left also intend to apply it in Machine Learning as a Service (MLaaS) [6].

Objectives

- Learn about algorithms used for homomorphic encryption. Implement it efficiently.
- Learn about differential privacy concepts and use them to make the data client-anonymous.
- Implement federated learning techniques using the above two mechanisms to ensure privacy.
- Implement the above approach for machine learning problems in a federated setting and demonstrate (given the time constraint) it on MLaaS platforms.

Project Deliverables

Demonstrate a secure usage of a client's data in a federated setting (datasets specified in [1, 2] and MLaaS platforms) by utilizing principles from differential privacy and homomorphic encryption.

References

1. Robin C. Geyer, Tassilo Klein, Moin Nabi, "Differentially Private Federated Learning: A Client Level Perspective", 2017 (link).
2. H. Brendan McMahan, Daniel Ramage, Kunal Talwar, Li Zhang, "Learning Differentially Private Recurrent Language Models", 2018 (link).
3. Chengliang Zhang, Suyi Li, Junzhe Xia, Wei Wang, Feng Yan, Yang Liu, "BatchCrypt: Efficient Homomorphic Encryption for Cross-Silo Federated Learning", 2020 (link).
4. Abbas Acar, Hidayet Aksu, A. Selcuk Uluagac, Mauro Conti, "A Survey on Homomorphic Encryption Schemes: Theory and Implementation", 2017 (link).
5. Payal V. Parmar, Shraddha B. Padhar, Shafika N. Patel, Niyatee I. Bhatt, Rutvij H. Jhaveri, "Survey of Various Homomorphic Encryption algorithms and Schemes", 2014 (link).
6. Mauro Ribeiro, Katarina Grolinger, Miriam A.M. Capretz, "MLaaS: Machine Learning as a Service", 2015 (link).