Project Proposal

Maurice Chiu and Pascal Jutras-Dubé

March 28 2023

Motivation

The Mobile Edge Computing and the Internet Of Things have accelerated the production rate of data. In parallel, machine learning models have become widely used in applications running on the Network Edge. Applications usually rely on a centralized data center where the aggregated data can be processed to train these data-hungry models. Unfortunately, these data may be used to identify individuals with very high accuracy, which raises privacy concerns.

Distributed computing can be helpful in mitigating these privacy risks; Federated Learning (FL) is a machine learning approach for which training is distributed across multiple decentralized edge nodes. Each node's raw data is stored locally and only the learned weights need to be transferred to a data center hence protecting the sensitive information of the user.

However, FL alone is not enough and many privacy concerns remain. It can be possible to attack directly the aggregated model's weights. Differential Privacy (DP) addresses the paradox of learning about a population without learning anything about an isolated individual present in the data. That is, DP guarantees by adding randomization into the learning mechanism that the same conclusions can be drawn whether or not a node is in the training dataset.

Goal

As a bonus project, we would like to reproduce the SIGCOMM'21 paper FedRAN: Federated Mobile Edge Computing with Differential Privacy [3]. The authors proposed FedRAN, a mobile edge, a federated learning system that incorporates differential privacy to improve the privacy integrity of sensitive edge information.

Since there was no source code provided, the majority of efforts will be devoted to the implementation of FedRAN. Then, we want to reproduce the experiments and propose further evaluations.

Plan

Part A: Implement FedRAN

The most critical parts of this project consists of the implementation of a distributed edge computing algorithm that operates over a LTE environment.

Controlled Radio-Frequency Environment

We plan to deploy our implemented version of FedRAN in a real controlled radio-frequency LTE environment [2]. However, should we encounter difficulties, we might use a simulator. In this part of the project, we will get the LTE environment working so that multiple simulated mobile devices can exchange simple messages with the FL central server.

Federated Learning

We plan to use IBM's FL library to handle the distributed computing algorithm and the interactions between the aggregator (central server) and the edge nodes (the clients).

FedRAN iteratively trains a global model over multiple connection rounds. In each connection round:

1. The server transfers the current state to the selected clients
2. Each client then makes local computations and inject noise with respect to DP
3. The clients transmit the result back to the central server

The process is repeated until it converges.

Differentially Private Machine Learning

The challenge of this part of the implementation is to make the edge nodes as indistinguishable as possible with differential privacy when running the distributed computing algorithm.

1. Task & Model:
   We will implement a convolutional neural network (CNN) for classifying digits images in Python.

2. Data:
   We will use MNIST, a large database of handwritten digits [4]. We will follow the same data processing procedures as in the FedRAN paper to reshape all the digit images and to partition the data into training and testing sets.

3. Training with Differential Privacy:
   We will train the model with a differentially private version of the stochastic gradient descent algorithm [1, 5] with TensorFlow Privacy or Pytorch Opacus librairies. As in FedRAN, we will optimize a Categorical Cross Entropy objective and add Gaussian noise to the gradients.

Part B: Evaluation and Analysis

1. **Distributed Learning**
   To evaluate whether the federated learning model performs well, we will analyze the relationship between the number of communication rounds and the accuracy of clients' classification.

2. **Preserving User Privacy**
   To evaluate the model's ability to preserve client information, we will run the DP FL training procedures and the non-DP FL training procedures and compare the accuracy of the resulting models.

Deliverable

1. Our written source code (Link to GitHub Repository)

2. Readme file explaining how to run our code

3. Final report describing our progress, deviations from the original plan, the difficulties encountered, list the evaluation and propose future research directions.

4. Video (15 min)

5. 5-min in-class presentation

References

[1] M. Abadi, A. Chu, I. Goodfellow, B. McMahan, I. Mironov, K. Talwar, and L. Zhang. Deep learning with differential privacy. In *23rd ACM Conference on Computer and Communications Security (ACM CCS)*, pages 308–318, 2016.

[2] J. Breen, A. Buffmire, J. Duerig, K. Dutt, E. Eide, M. Hibler, D. Johnson, S. K. Kasera, E. Lewis, D. Maas, A. Orange, N. Patwari, D. Reading, R. Ricci, D. Schurig, L. B. Stoller, J. Van der Merwe, K. Webb, and G. Wong. POWDER: Platform for Open Wireless Data-driven Experimental Research. In *Proceedings of the 14th International Workshop on Wireless Network Testbeds, Experimental evaluation & Characterization*, WiNTECH'20, pages 17–24, New York, NY, USA, Sept. 2020. Association for Computing Machinery. ISBN 978-1-4503-8082-9. doi: 10.1145/3411276.3412204. URL `https://dl.acm.org/doi/10.1145/3411276.3412204`.

[3] A. Gottipati, A. Stewart, J. Song, and Q. Chen. Fedran: Federated mobile edge computing with differential privacy. In *Proceedings of the 4th FlexNets Workshop on Flexible Networks Artificial Intelligence Supported Network Flexibility and Agility*, FlexNets '21, pages 14–19, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450386340.

[4] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, Nov. 1998. ISSN 1558-2256. doi: 10.1109/5.726791. Conference Name: Proceedings of the IEEE.

[5] I. Mironov. Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 263–275, 2017.