ECE209AS - Winter 2022
UCLA

“Hariram, Veeramani”
Review for Paper Titled:

“SecureML- A System for Scalable Privacy Preserving Machine Learning”

1- Summary

This work addresses the privacy problem where data owners operating in a two-server model distribute their data among two non-colluding servers and train various models on the joint data using Two-Party Computation(2PC). It proposes two new techniques to support secure Arithmetic operations on shared decimal numbers and develops MPC-friendly alternatives to expensive, non-linear activation functions such as Sigmoid and Softmax which are said to be orders of magnitude faster than the state-of-the-art Privacy preserving learning implementations and scale very well to millions of data samples with thousands of features. This work presents Efficient protocols for Privacy preserving Machine Learning techniques utilizing Linear Regression, Logistic Regression and neural network training using the Stochastic Gradient Descent Method.

2- Novelty/Contribution

This work proposes novel contributions to Arithmetic on Shared numbers, MPC-friendly activation functions and Vectorizing protocols recognizing that the main source of inefficiency in previous secure 2PC systems is the bulk of computations happening inside Boolean circuits (mainly Yao’s garbled circuit). It proposes an effective strategy for multiplication wherein shared decimal numbers between two parties are represented as integers, then perform multiplication on these integers using offline-generated multiplication triplets and then have each part truncate its share of the product so that a fixed number of bits represent the fractional part without affecting the accuracy of the model. It proposes new MPC activation function, which can be seen as the sum of two ReLu Functions and replace the softmax function by combination of ReLu, addition and a single division. Furthermore, it offers a customized solution between Arithmetic sharing and Yao Sharing that significantly reduces computation cost by minimizing rounds of interactions and required number of Oblivious Transfers (OT). Application of Vectorization techniques in a shared setting is shown to have improved the performance of the OT generated multiplication triplets by a factor of 4x and the Linearly Homomorphic Encryption (LHE) based generation improved by a factor.

3- Evaluation

The system implemented in C++(100x faster compared to Arithmetic’s in GMP or NTL library) , Eigen Library is used to handle matrix operations, OT’s and garbled circuits are implemented using the EMP toolkit and Cryptosystem of DGK is used for LHE. Amazon EC2 with 60GB RAM is chosen as platform with a LAN network with 0.17ms delay and 1GB/s bandwidth metric and with a WAN network of 72ms latency and 9 MB/s bandwidth. Thus, these experiments, capture the scenarios where two servers have a high bandwidth/low-latency network connection but otherwise are not controlled/administered by the same party. The evaluation is carried on using MNIST dataset and it reports consistent performance improvements both in accuracy and speed as mentioned above.

4- Unique Strength/Weakness

This is one of the works, describing very elaborately the treatment that they provided to arrive at the results. I like the extreme finer details presented along with thorough information about the assumptions and background.