Experiment plan

Thomas Werthenbach

April 2023

1. Scalability

Question: in which section should these results be located?

1. Number of participants
   - We perform simulations to demonstrate FoolsGold's inability to scale to millions of devices.
   - No actual federated learning simulation required, we can execute the aggregation function on dummy data, the required time does not change.
2. Communication bottleneck at server
   - Show accuracy against time in training for both our improved algorithm and FoolsGold.

   Alternative:
   - Average time per round against number of participants.

2. Show that FoolsGold performs badly in decentralized learning compared to our algorithm

- Show that for certain attack topologies, FoolsGold performs well, and for some performs badly.

3. Different network topologies

1. Spread attacks. We vary the density of Sybil attack edges (which are as spread out in the network as possible). Density of 1 means that every honest node in the network is connected to an attack edge. Density of 0.5 means that half of the honest nodes are connected to an attack edge (where the distance between the attack edges is maximized).
2. Local attacks. Given an honest node as a target, we can vary the amount of attack edges around this node (vary the variance in normal distribution to map the amount of attack edges per node based on the amount of hops separated from the target).

Note: mathematical definitions are coming soon. Note: For both these attacks, we can use different network topologies for the honest node distribution (not yet sure how to approach this):

- Random geometric graph
- Fully connected graph
- Single line graph

4. Comparison with different techniques

- FedAvg
- Krum
- Multi-Krum
- Median
- FoolsGold (SGD/AVG)
- Our algorithm (SGD/AVG)
- Some more?

5. Effect on different datasets

Show the results of multiple algorithms on different datasets:

- CIFAR-10: Classification of coloured real-world objects/animals (10 output classes). LeNet.
- EMNIST: Classifcation of handwritten uppercase/lowercase letters, and numbers ($26 + 26 + 10 = 62$ output classes). Single log softmax layer.
- SVHN: Classification of coloured real-world house numbers (10 output classes). LeNet.
- FashionMNIST: Classification of different typesof clothing (10 output classes). Single log softmax layer.

6. Effect of Non-IID-ness

We can vary the degree to which data is distributed Non-IID using the alpha parameter on the Dirichlet distribution.