FedCav: Contribution-aware Model Aggregation on Distributed Heterogeneous Data in Federated Learning

Hui Zeng¹, Tongqing Zhou¹, Yeting Guo¹, Zhiping Cai¹, Fang Liu²

¹ College of Computer, National University of Defense Technology
² School of Design, Hunan University
Outline

• Background

• Observation and Problem Statement

• Our Solution

• Experiments

• Conclusion
Background

Two main problems in AI

data privacy
data island
Federated Learning

- Federated Learning (FL) allows multiple distributed devices to cooperatively train models in parallel while preserving data privacy.
Data Heterogeneity Problem in FedAvg

- Data Heterogeneity
  - Global non-IID
  - Local class imbalanced

- Problem
  - Slow convergence
  - Low training performance
Observations of FL with heterogeneous data

Setup -- Observation experiment

Dataset: MNIST

Training Model: LeNet-5

Table 1. Three different types of data distribution

<table>
<thead>
<tr>
<th>Notations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IID &amp; balanced</td>
<td>Global IID and Local class balanced</td>
</tr>
<tr>
<td>non-IID &amp; balanced</td>
<td>Global non-IID and Local class balanced</td>
</tr>
<tr>
<td>non-IID &amp; imbalanced</td>
<td>Global non-IID and Local class imbalanced, the variance of each class is $\sigma$</td>
</tr>
</tbody>
</table>
Observation Results

- Results
  - **Slow** convergence
    - balanced: 5~10
    - imbalanced: 20~35
  - **Low** training performance
    - balanced: ~95%
    - imbalanced: 80% ~ 93%
Problem Statement

• Observation analysis
  Heterogeneous Data
  FedAvg
  Slow convergence
  Low training performance

Data size ≠ Contribution

• Reality

Data size ≠ Contribution

Which is more valuable?

G-board

Animal pictures
Our Solution

FedCav workflow

Detection mechanism

Contribution-aware model aggregation
Contribution-aware model aggregation

Measure the potential contribution

Algorithm 1: FedCav model aggregation algorithm

Input: Number of clients $n$, sample ratio $q$
Output: The expected optimization result of global model $w_{opt}$
1. initialize global model $w_0$;
2. for each communication round $t = 1, 2, \ldots$ do
   3. for each client set $P_t$ in parallel do
      4. for each client $i \in P_t$ do
         5. $w_{t+1}, f_i(w_t) \leftarrow \text{LocalUpdate}(w_t)$;
      6. Clip the $f_i(w_t)$:
         7. $f_i(w_{t+1}) = \min(\max(f_i(w_{t+1}), \text{mean}(f_i(w_t)))), 1)$;
      8. $w_{t+1} = \sum_{i=1}^{n} \text{softmax}(f_i(w_{t+1})) w_{t+1}$;
   9. return $w_{opt}$

Algorithm 2: Function LocalUpdate

Input: Global model parameter $w_t$, local epoch $t$, local mini-batch size $B$.
Output: Local updated model $w_{t+1}$, inference loss $f_i(w_{t+1})$.
1. Compute the loss based on local data $d_i$;
2. $f_i(w, d_i) \leftarrow f_i(w, d_i)$;
3. Initialize local model $w_t \leftarrow w_t$;
4. $b \leftarrow \text{split } d_t \text{ into batches of size } B$;
5. for each local epoch $e = 1, 2, \ldots$ do
   6. for each batch $b \in B$ do
      7. $w_{t+1} \leftarrow w_{t+1} - \eta f_i(w_{t+1}, b) ;$
   8. return $w_{t+1}, f_i(w_{t+1})$.
Model replacement attack

(Simple & Low cost)

Cloud
Healthy global model

Download
Convergent

Attacker Malicious model

Cloud

\[ \text{Update} \approx 0 \]

Upload

\[ \text{Update} = \]

Clients

\[ \ldots \]

Attacker Malicious model

\[ + \text{Updates} \]
Detection mechanism

Comparing the inference loss with historical value.

The final decision is consistent with the majority.

Reversing the current model to cached one.
### Setup -- Dataset & Models

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Samples</th>
<th>Training Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNIST</td>
<td><img src="image1" alt="Samples" /></td>
<td>LeNet-5[1]</td>
</tr>
<tr>
<td>FMNIST (Fashion-MNIST)</td>
<td><img src="image2" alt="Samples" /></td>
<td>9-layers CNN</td>
</tr>
<tr>
<td>CIFAR-10</td>
<td><img src="image3" alt="Samples" /></td>
<td>Resnet18[2]</td>
</tr>
</tbody>
</table>


Results -- Classification performance

Table 4: Average classification accuracy under different levels of data heterogeneity (varying $\sigma$) on three datasets. Here we list the accuracy performance of different methods after the learning process gets convergence.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>$\sigma = 300$</th>
<th>$\sigma = 600$</th>
<th>$\sigma = 900$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNIST</td>
<td>0.9333</td>
<td>0.9391</td>
<td>0.9365</td>
</tr>
<tr>
<td>FMNIST</td>
<td>0.8447</td>
<td>0.8459</td>
<td>0.8621</td>
</tr>
<tr>
<td>CIFAR-10</td>
<td>0.4612</td>
<td>0.4644</td>
<td>0.4686</td>
</tr>
</tbody>
</table>

Results -- Classification accuracy with dynamic environment

Figure 4: Classification accuracy with dynamic data distribution adjustment controlled by factor $\alpha$ on three dataset. Results on different datasets are painted, wherein FedCav shows a generally stable and superior performance.

Fewer rounds to achieve a certain accuracy
Results -- Impact of Clip strategy

Figure 5: Training process with four different algorithms on three datasets. Comparing the difference of whether it is necessary to use the Clip strategy. The curve shows that FedCav without Clip occurs great up-and-down oscillation.
Results -- Impact of model replacement attack

Figure 6: Part of training process of FedCav without detection and FedAvg after the model replacement attack on three datasets.
Results -- Reverse to cached one

- Three different strengths of attacks
- Reverse the current model to cached one
Conclusion

• **Problem**: Address the data heterogeneous problem in FL

• **Observation**: Heterogeneous data cause the slow convergent and low training accuracy.

• **Key idea**: Contribution-aware model aggregation. Aggregate with the contribution not the data size. Also consider malicious attack.

• **Evalutation**: Higher classification accuracy (~2.4%) than baselines on average, fewer communication rounds (~34%) to achieve convergence.
Thank you!

Hui Zeng\textsuperscript{1}, Tongqing Zhou\textsuperscript{1}, Yeting Guo\textsuperscript{1}, Zhiping Cai\textsuperscript{1}, Fang Liu\textsuperscript{2}

\textsuperscript{1} College of Computer, National University of Defense Technology

\textsuperscript{2} School of Design, Hunan University

Email: zenghui116@nudt.edu.cn