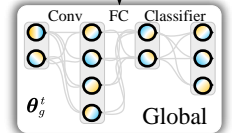


$$\theta_g^{t+1} = \frac{1}{|C^t|} \sum_{i \in C^t} \frac{m_i^t}{m^t} \theta_i^t$$

5) Global Aggregation

1) Global Model Distribution



2) Model decoupling

1. Computing OBP Score

$$I_o(\theta_i^{t-1,k}; \mathcal{D}) = |(\theta_i^{t-1,k} - \theta_g^{t,k}) \cdot \theta_i^{t-1,k}|$$

2. Thresholding

3. Global Model Bipartition

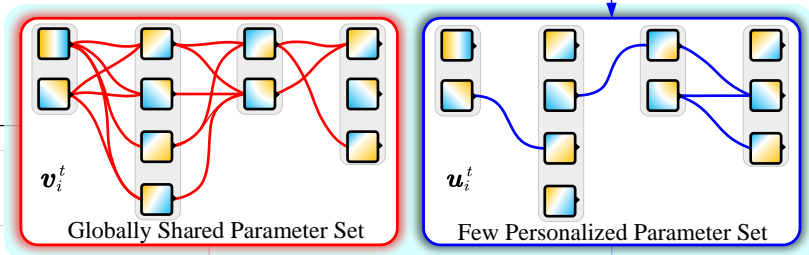
$$\mathcal{K}(\mathbf{u}_i^t) = \{k \mid I_o(\theta_i^{t-1,k}; \mathcal{D}) > \tau\}$$

$$\mathcal{K}(\mathbf{v}_i^t) = \{k \mid I_o(\theta_i^{t-1,k}; \mathcal{D}) \leq \tau\}$$

Server

Clients

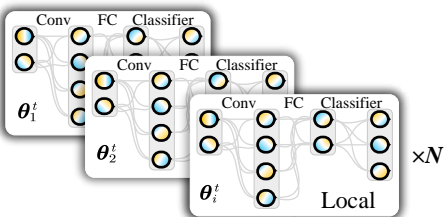
2) Model decoupling



3) Local Merging

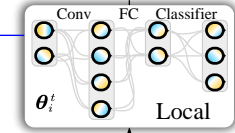
$$\tilde{\theta}_i^t = \mathbf{u}_i^t \cup \mathbf{v}_i^t$$

$\{\theta_i^t\}_{i \in C^t}$



Upload to Server

backup



4) Local Training

$$\theta_i^t = \tilde{\theta}_i^t - \eta \nabla_{\theta} \mathcal{L}(\tilde{\theta}_i^t; \mathcal{D}_i)$$

