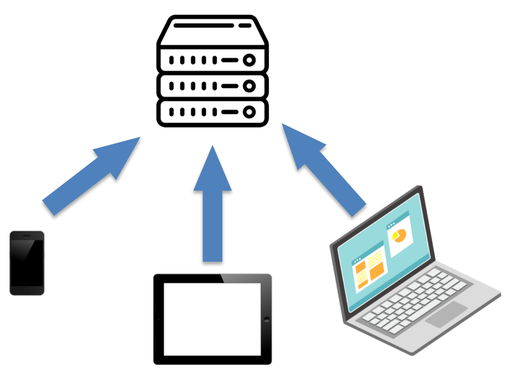
Privacy Preserving Wireless Federated Learning

PI: Ravi Tandon, University of Arizona

NSF Project ID: 1715947
https://www.nsf.gov/awardsearch/showAward?AWD_ID=1715947

Motivation

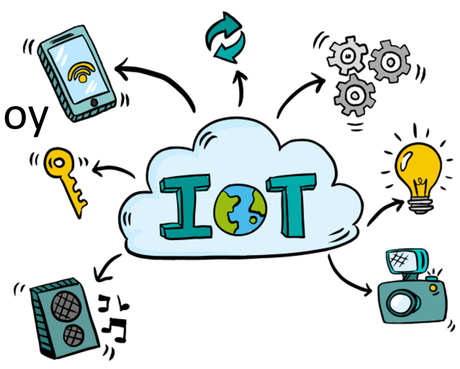
Large-scale Personalized Data
(Data Privacy is a key concern)



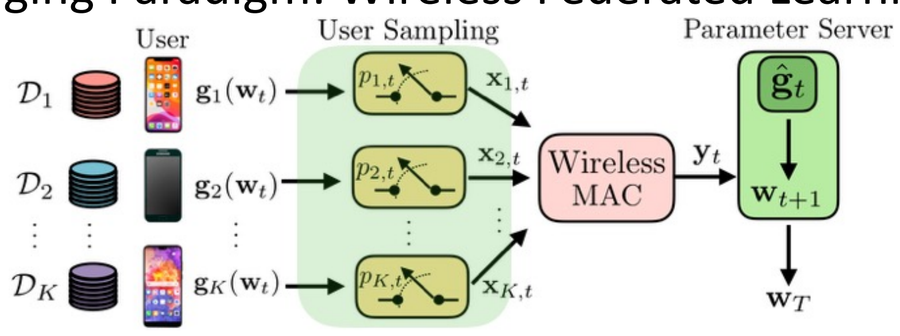
Distributed Systems
(Require Communication Efficiency)

Scientific Impact

- Distributed systems which deploy ML with privacy constraints:
  - IoT Systems, Edge Networks
  - Healthcare Networks
  - Autonomous Vehicle Systems



Emerging Paradigm: Wireless Federated Learning



- FedSGD Optimization:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} F(\mathbf{w}) \triangleq \frac{1}{|\mathcal{D}_{\text{total}}|} \sum_{k=1}^{K} \sum_{i=1}^{|\mathcal{D}_k|} f_k((\mathbf{u}_i^{(k)}, v_i^{(k)}); \mathbf{w})$$

(ith feature, ith label)

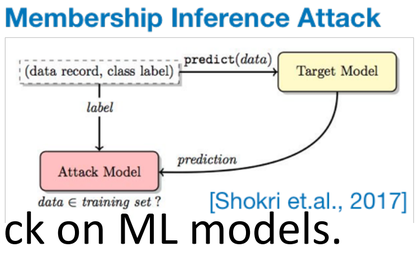
$$\mathbf{w}_{t+1} = \mathbf{w}_t - \eta_t \left[\frac{1}{|\mathcal{K}_t|} \sum_{k \in \mathcal{K}_t} \mathbf{g}_k(\mathbf{w}_t) + \mathbf{z}_t \right], \quad \mathbf{z}_t \sim \mathcal{N}(0, \sigma_z^2 \mathbf{I}_d)$$

(kth Loss function)

Step size; Estimate of full gradient; Gaussian noise

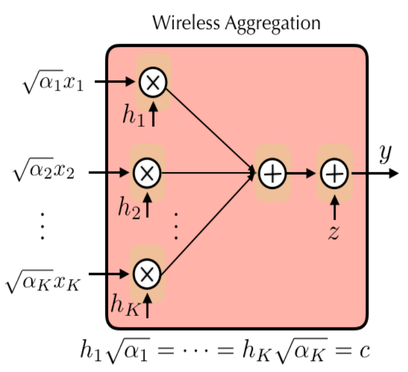
Challenges

- Data Privacy:
  - Known attack on ML models.

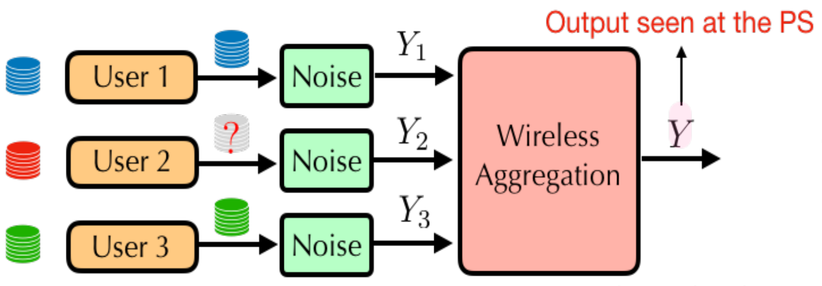


[Shokri et.al., 2017]

- Communication Efficiency:
  - Digital vs Analog Schemes
  - Power alignment for Analog: requires coordination between users.
  - Also, requires perfect channel knowledge at transmitters.



Contributions



Noise due to local perturbation and channel noise

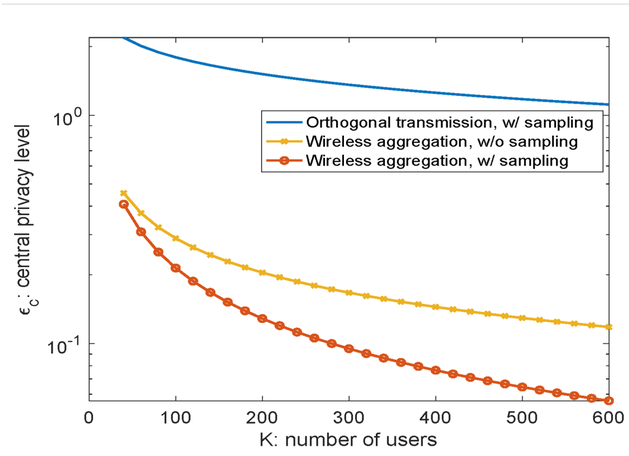
- Received signal at PS:

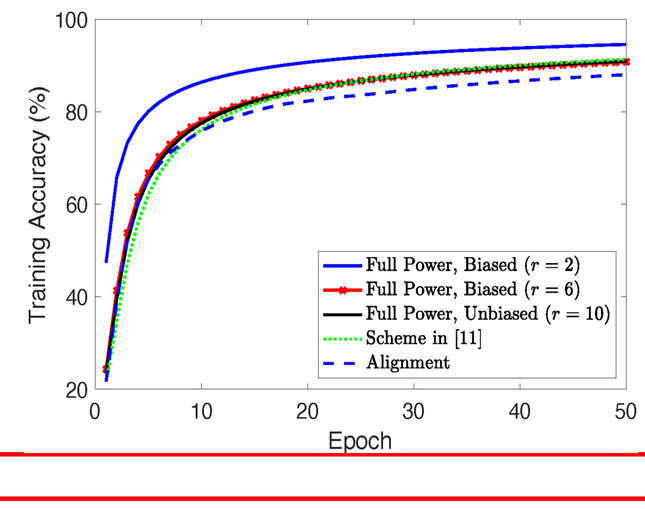
$$\mathbf{y}_t = \sum_{k \in \mathcal{K}_t} h_{k,t} \sqrt{\alpha_{k,t}} \mathbf{g}_k(\mathbf{w}_t) + \sum_{k \in \mathcal{K}_t} h_{k,t} \sqrt{\beta_{k,t}} \mathbf{n}_{k,t} + \mathbf{m}_t$$

Estimate of full gradient; Set of sampled users at iteration t

- Power allocation:

- Studied the impact of privacy, Communication, & statistical benefits of Analog Wireless Federated Learning
- **Coordination Benefits:** Alignment of signals from users is not necessary for FedSGD to converge.
- Carefully chosen bias speeds up the convergence when using full power.
- Coordination not always needed.
- **Privacy Benefits:** Wireless aggregation sums up individual noise from users to provide strong differential privacy leakage scaled as $O(1/\sqrt{K})$.
- The combination of user sampling and wireless aggregation provides improved privacy leakage scaled as $O(1/K^{3/4})$.





Broader Impact on Society

- The proposed research advances our fundamental understanding of modern privacy-preserving machine learning systems and emerging paradigms such as over –the-air ML.
- Many applications in industry including IoT, location-based services, mobile health, etc.

Broader Impact on Education and Outreach

- New graduate-level course(s) developed on the topic of distributed and private machine learning.
- Several Ph.D. and M.S. students trained on the topic.
- Delivered Seminars at UA REU summer programs.

Broader Impact and Broader Participation

- Fundamental understanding of privacy-preserving machine learning.
- Exploiting new benefits from wireless channels which can have impact in 6G and beyond.
- Engagement with NSF IUCRC (BWAC)