Dynamic Federated Learning for Heterogeneous Learning Environments

Eric Samikwa [1] Torsten Braun [1]

[1]Communication and Distributed Systems, Institute of Computer Science, University of Bern

Motivation

- Massive influx of data generated by edge devices and the Internet of Things (IoT).
- Machine learning models trained on this data can provide valuable insights and predictions, leading to better decision-making and intelligent applications.
- Federated Learning (FL) enables remote devices to collaboratively train models without sharing sensitive data, preserving user privacy and reducing overhead.
- Heterogeneous learning environments significantly limit its performance and hinder its real-world applications.

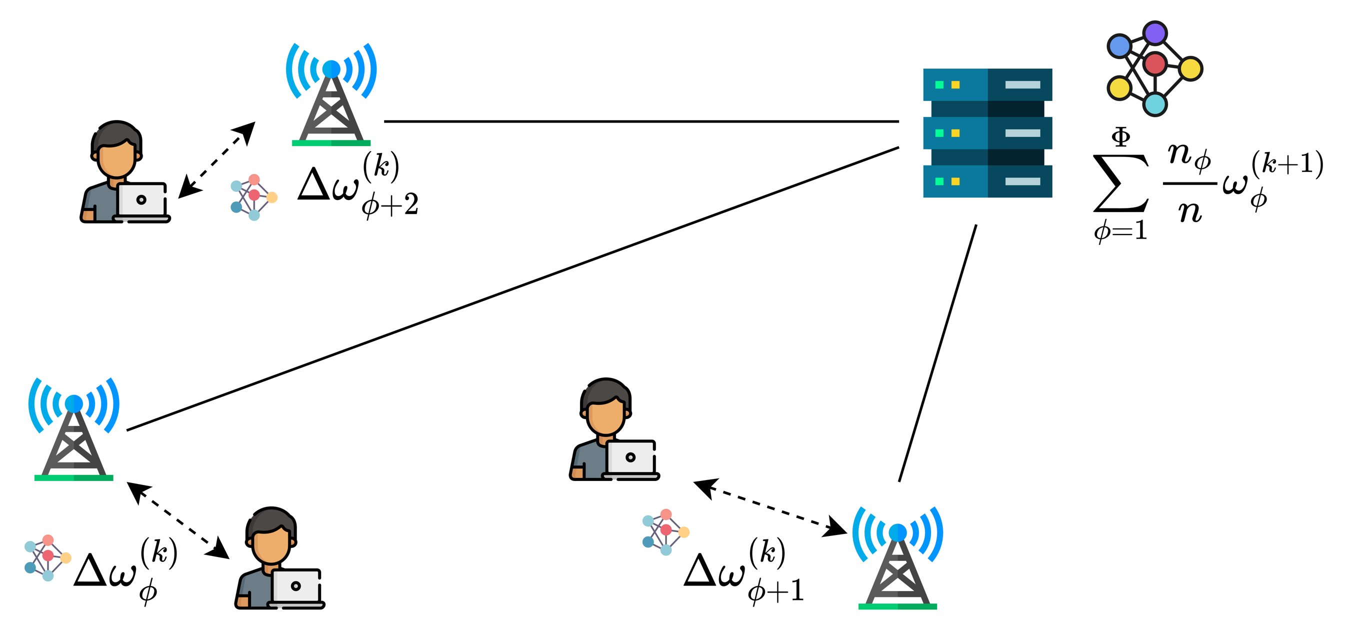

Figure 1. Federated Learning with Φ users contributing to a global model.

Heterogeneous Learning Environments

- The *statistically heterogeneous* (usually non-independent identically distributed) training data in geographically distributed clients can deteriorate the FL training performance, such as the accuracy and learning stability.
- The *heterogeneous computing and communication resources* in IoT devices often result in unstable training processes that slow down the training of a global model and affect energy consumption.
- Most existing approaches address only the unilateral side of the heterogeneity issue, either the statistical or the resource heterogeneity.
- The resource heterogeneity among various devices does not necessarily correlate with the distribution of their training data.

Contributions

- Dynamic Federated Learning (DFL) addresses the joint problem of data and resource heterogeneity in FL under heterogeneous learning environments.
- DFL combines resource-aware split computing of deep neural networks and dynamic clustering of training participants based on the similarity of their sub-model layers.

Resource-Aware Federated Learning

- The allocation of the FL training tasks on resource-constrained participants is adjusted to match their heterogeneous capabilities in terms of computing and communication resources.
- Resource-capable participants carry out the classic FL training only transmitting trained weights.

$$\Delta_s^{(k)}(\phi) = \frac{W_\phi}{\xi}\left(\delta_s^{(k)}(\phi) + \delta_s^{(k)}(C) + \frac{V_s}{B_\phi^{(k)}} + \frac{V'_s}{B'^{(k)}_\phi} + \Theta^{(k)}\right) + \lambda \cdot \sum_{L=1}^{s} \frac{|\omega_L|}{B_\phi^{(k)}} \quad (1)$$

- Sensitivity parameters β_i consider the tradeoffs in minimizing the training, time, energy, and loss.

$$U_k(s_k, D) = \beta_1 \cdot \Delta_k(s_k) + \beta_2 \cdot E_k(s_k) + \beta_3 \cdot \mathcal{L}(D, \omega) \quad (2)$$

Similality-Based Layerwise Aggregation

- Utilizing centered kernel alignment (CKA) for determining the similarity of neural network layers to address the data heterogeneity without requiring direct access to data.
- Layerwise sub-model aggregation is carried out within clusters of similar training participants.

$$\mathrm{CKA}(XX^T, YY^T) = \frac{\|Y^T\|X\|_F^2}{\|XX^T\|_F\|YY^T\|_F} \quad (3)$$

Applications

- FL deployments in geographically distributed clients (edge computing devices) that possess non-iid data.
- Smart healthcare: DFL can potentially be used in healthcare applications where data privacy is crucial.
- Resource-constrained heterogeneous IoT: Industrial IoT, environmental monitoring, multimedia IoT, etc.

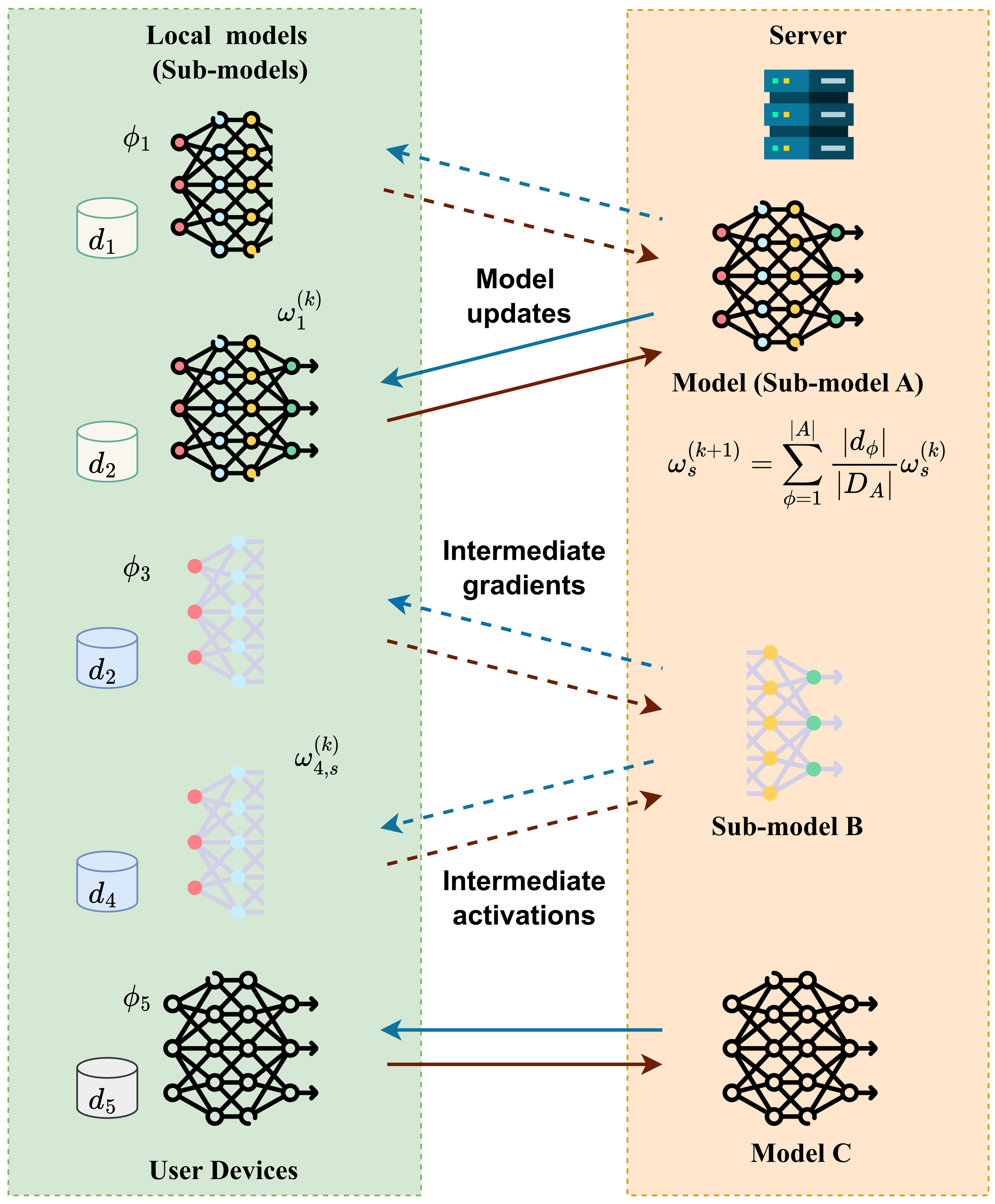

Figure 2. DFL via resource-aware split computing and similarity-based layerwise model aggregation.

Experimental Results

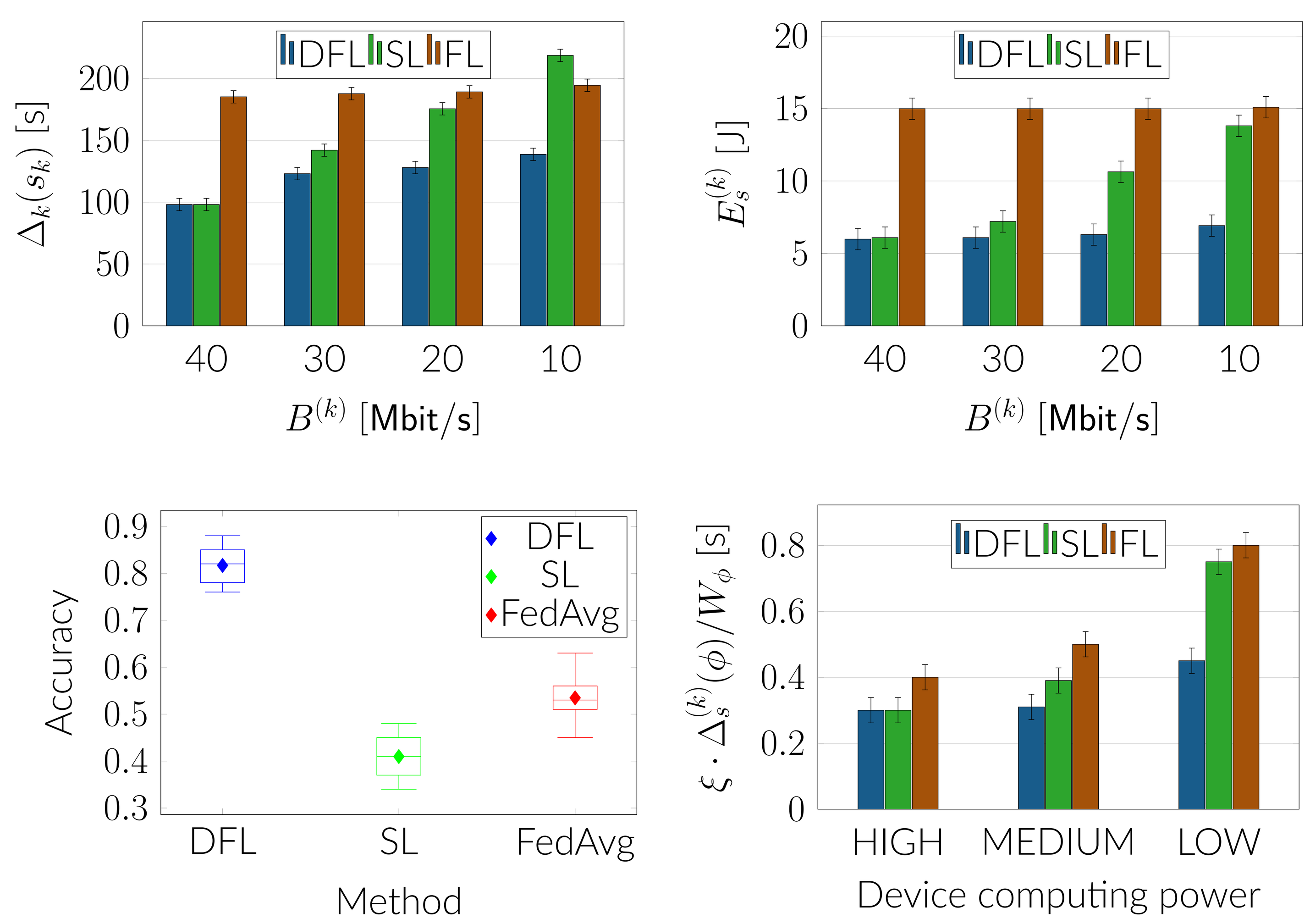

Figure 3. Average training time $\Delta_k(s_k)$, energy consumption $E_s^{(k)}$ per round, and model accuracy for non-iid data over variable computing and network resources.

- DFL improves training performance (i.e., training time up to 48%, accuracy up to 30%, and energy consumption up to 61.4%) in heterogeneous learning environments with both data and resource heterogeneity.

Conclusions and Future Work

- FL enables scalability, user privacy, low response time, less cloud dependency, etc.
- Heterogeneous learning environments limit its performance and its real-world applications.
- DFL improves training performance (i.e., training time, accuracy, and energy consumption) in heterogeneous learning environments.

References

[1] Simon Kornblith, Mohammad Norouzi, Honglak Lee, and Geoffrey Hinton. Similarity of neural network representations revisited. In *International Conference on Machine Learning*, pages 3519–3529. PMLR, 2019.

[2] Eric Samikwa, Antonio Di Maio, and Torsten Braun. Ares: Adaptive resource-aware split learning for internet of things. *Computer Networks*, 218:109380, 2022.

[3] Xiaofei Wang, Yiwen Han, Victor CM Leung, Dusit Niyato, Xueqiang Yan, and Xu Chen. Convergence of edge computing and deep learning: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 22(2):869–904, 2020.

{eric.samikwa, torsten.braun}@unibe.ch