



HAL
open science

Communication-efficient Federated Learning through Clustering optimization

Hugo Miralles, Tamara Tasic, Michel Riveill

► **To cite this version:**

Hugo Miralles, Tamara Tasic, Michel Riveill. Communication-efficient Federated Learning through Clustering optimization. SophI.A. Summit, Nov 2021, Biot, France. hal-03479640

HAL Id: hal-03479640

<https://hal.inria.fr/hal-03479640>

Submitted on 17 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Communication-efficient Federated Learning through Clustering optimization

Hugo Miralles^{1,2}, Tamara Tomic², Michel Riveill¹

¹ Université Côte d'Azur, Inria, CNRS, I3S, France

² Orange Labs, France

We study the problem of model personalization in Federated Learning (FL) with non-IID (Independent and Identically Distributed) data collected at nodes in a network, under the network communication cost constraints. Classical FL collaboratively trains a unique global model. If data is statistically heterogenic (non-IID), personalized models for groups of nodes with similar statistics have been shown to provide better performances compared to FL [1].

We propose a Clustered Federated Learning approach that provides a trade-off between identifying models that are more adapted to nodes locally, under communication cost constraints. Our method identifies clusters of nodes with similar data statistics, which improves the local model accuracy. In particular, it aims at finding the cluster structure, cluster heads and a set of model weights (one per cluster) that minimize an objective function composed of two terms: a classical multi-task optimization term and a communication cost regularization. Local model updates represent proxy values of the local data distributions (statistically similar train sets have similar updates) where similar updates are aggregated together [2,3,4].

Our algorithm has two phases: initialization and cluster optimization. During the initialization, nodes collaboratively train a global initial model. The cluster head nodes are identified and nodes are clustered based only on the communication cost minimization [5]. The cluster optimization phase starts by applying the Hierarchical Agglomerative Clustering on a distance metric composed of two terms: the cosine dissimilarity between the locally computed model updates of two nodes, and the communication cost of grouping two nodes in the same cluster. In parallel, respective cluster heads are also optimized. The clusters are organized in a tree hierarchy. At each round, the cluster heads verify if a new cluster optimization is needed based on the model update values. If required, the same method is applied to further create sub-clusters.

We evaluate our method on several non-IID settings generated from MNIST dataset, while simulating the communication cost at each round. We show that our algorithm improves the quantity of nodes reaching 99% of accuracy (from 48% to 72%) and can reduce the overall communication cost by 35%. Finally, it is able to adapt the cluster structure in case of new conditions (new network nodes or time-evolution of local data distribution) by a tree structure search.

- [1] Y. Mansour, M. Mohri, M., J. Ro and A. T. Suresh, Three approaches for personalization with applications to federated learning. *ArXiv*, 2020.
- [2] F. Sattler, K. R. Möller and W. Samek, Clustered federated learning: Model-Agnostic distributed multi-Task optimization under privacy constraints. *ArXiv*, pp. 1–13, <https://doi.org/10.1109/tnnls.2020.3015958>, 2019.
- [3] C. Briggs, Z. Fan and P. Andras, Federated learning with hierarchical clustering of local updates to improve training on non-IID data. *ArXiv*, 2020.
- [4] X. Tang, S. Guo and J Guo, Personalized Federated Learning with Clustered Generalization. In *Proceedings of ACM Conference (Pre-print)* (Vol. 1, Issue 1). Association for Computing Machinery. <http://arxiv.org/abs/2106.13044>, 2021.
- [5] M. S. H. Abad, E. Ozfatura, D. Gündüz and O. Ercetin, Hierarchical federated learning across heterogeneous cellular networks. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2020-May(MI)*, pp. 8866–8870, <https://doi.org/10.1109/ICASSP40776.2020.9054634>, 2020.

