Towards Efficient and Practical Federated Learning
Ahmed M. Abdelmoniem
Queen Mary University of London, UK

Abstract

Federated learning (FL) is increasingly becoming the norm for training models over distributed and private datasets. In this paper, we describe our efforts which are motivated by the urgent need for exploring the promising prospects and imminent challenges towards the facilitation of the adoption of federated learning for service providers. The prospects are motivated by the growing interest and momentum towards the adoption of privacy preservation and 5G/6G technologies. Whereas, the challenges that hinder wide FL adoption are mainly the data, resource and user heterogeneity as well as communication overhead. Resource heterogeneity is challenging in particular because the training resources and conditions are not under the control of the central server. Specifically, FL-trained models in practice require a significant amount of time (days or even weeks) because FL tasks execute in highly heterogeneous environments where devices only have widespread yet limited computing capabilities and network connectivity conditions.

Introduction

We present our initial efforts focused on studying training FL models in heterogeneous environments. We note that heterogeneity can have a detrimental impact on both the model quality and fairness. Then, we present our efforts towards designing efficient yet practical mitigation methods to limit the impact of system heterogeneity in FL settings.

Federated Learning in A Nutshell

Federated Learning (FL) is a recent paradigm that has sparked great research interest to enable learning collaboratively over distributed datasets [1, 2, 3]. In FL, with the help of a central aggregation server, the end devices train a global model on their private data without transferring the private data over the network. Even though, the paradigm was initially proposed by the industry (e.g., [4, 3]) to solve a practical problem, it has seen tremendous interest from academia (e.g., [5, 6]).

Empirical Study of Heterogeneity [7]

![Fig. 1: Model Quality with System Heterogeneity](image)

AQFL: Adaptive Model Quantization [8]

AQFL is an adaptive model quantization technique that varies the quantization level in proportion to the resources available at clients’ end-devices. This approach reduces both the computational and communication overhead of the training process, which helps with increasing the percentage of devices submitting their model updates within the deadline.

![Fig. 3: Performance with vs without AQFL](image)

RELAY: Resource Efficient Federated Learning [9]

RELAY maximizes existing FL systems’ statistical, system, and resource efficiency with little modification to the existing FL solutions. It decouples the collection of participant updates from aggregation into an updated model. It also intelligently selects among available participants that are least likely to be available in the future.

![Fig. 4: RELAY vs State-of-the-Art Systems](image)

Summary

Our main findings towards studying and mitigating the effects of heterogeneity are:
1. Our empirical study shows that system heterogeneity can result in degradation of 5X on average for model quality [7].
2. AQFL is an adaptive model quantization method which can homogenize the device training time and maintain the quality and fairness [8].
3. RELAY is a system enabling status-aware aggregation and intelligent participant selection to improve resource efficiency and model quality [9].

References