EAI Endorsed Transactions
on Energy Web Research Article **EAI.EU**

Provision for Energy: A Resource Allocation Problem in Federated Learning for Edge Systems

Mingyue Liu[1,*], Leelavathi Rajamanickam[2] and Rajamohan Parthasarathy[3]

[1,2]SEGi University, Centre for Software Engineering, Faculty of Engineering, Built Environment Information Technology
[3]SEGi University, Centre for Software Engineering, Centre for Network Security and IoT, Faculty of Engineering, Built Environment Information Technology

Abstract

The article explores an energy-efficient method for allocating transmission and computation resources for federated learning (FL) on wireless communication networks. The model being considered involves each user training a local FL model using their limited local computing resources and the data they have collected. These local models are then transmitted to a base station, where they are aggregated and broadcast back to all users. The level of accuracy in learning, as well as computation and communication latency, are determined by the exchange of models between users and the base station. Throughout the FL process, energy consumption for both local computation and transmission must be taken into account. Given the limited energy resources of wireless users, the communication problem is formulated as an optimization problem with the goal of minimizing overall system energy consumption while meeting a latency requirement. To address this problem, we propose an iterative algorithm that takes into account factors such as bandwidth, power, and computational resources. Results from numerical simulations demonstrate that the proposed algorithm can reduce energy consumption compared to traditional FL methods up to 51% reduction.

Keywords: Energy Efficiency, Resource Allocation, Federated Learning

Received on 28 December 2023, accepted on 25 June 2024, published on 03 July 2024

doi: 10.4108/ew.6503

1. Introduction

There has been significant growth in mobile data in recent years, much of it generated in real-time and distributed to edge devices such as smartphones and sensors [1] [2]. Artificial intelligence (AI) technology is widely used to process this mobile data and support various services, such as computer vision and the internet of vehicles [3]. A common practice is to train AI models using elastic cloud computing, which allows operators to achieve optimal performance by accessing large-scale datasets. However, this process poses challenges due to privacy concerns [4], network congestion [5], and service latency [6]. Federated learning (FL) in the edge framework offers a solution to these issues. FL implements distributed machine learning at the network edge, where clients, or edge devices, train local models with their private data and only share parameters like model weights [7] [8] [9]. An FL server is used to aggregate these models into a global model and broadcast updates to each edge device. After several iterations, accuracy is achieved, and the training process is completed. FL avoids the need for data uploads and enables rapid access to real-time data, thus reducing pressure on communication resources and lowering service latency. It is a promising distributed learning algorithm that is likely to be applied in future internet of things systems [10, 11, 12, 13, 14, 15].

Wireless devices, such as those that communicate through cellular networks, benefit greatly from the use of distributed learning frameworks. These frameworks allow for the training of locally collected data using a shared learning model [16, 17, 18]. However, edge devices such

*Corresponding author. Email: mingyue2022@126.com