A Serverless Approach to Federated Learning Infrastructure Oriented for IoT/Edge Data Sources (Student Abstract)

Anshul Ahuja, Geetesh Gupta, Suman Kundu
Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur, Rajasthan - 342037, India
{ahuja.2, gupta.15, suman}@iitj.ac.in

Abstract

The paper proposes a Serverless and Mobile relay based architecture for a highly scalable Federated Learning system for low power IoT and Edge Devices. The aim is an easily deployable infrastructure on a public cloud platform by the end user and democratize the use of federated learning.

Introduction

Federated learning (FL) is a decentralized machine learning (ML) technique developed for mobile devices [1] and privacy protection. In FL system, ML model is trained at several geographically separated devices. It removes the necessity of data transfer to central cloud servers, in-turn reduces the bandwidth requirement and privacy breach. Usually, a central server is kept to orchestrate the learning process. It accumulates & distribute the learned model among different nodes. In (Bonawitz et al. 2019), a scalable FL production system has been developed for mobile devices. Recently, FL based personalized intelligent Internet of Things (IoT) system design is studied by (Wu, He, and Chen 2020). However, these works mostly considered for mobile. A scalable solution for low power IoT devices is equally important as several low power IoT device, generating huge amount of data, is being used for various purposes. These devices lack resources to run ML algorithms and bandwidth to connect with cloud. To solve the latter, mobile relaying is used by (Manzoor et al. 2018). Can we judiciously integrate the relay techniques with FL to have a scalable FL for IoT?

In the present work, we developed a Serverless Orchestration Architecture for low energy IoT devices, where a connected mobile device is used as Federation node and Serverless act as a central orchestration facilitator. The architecture will preserve the privacy concerns as the data is trained on personal mobile devices. The learned model is transferred to Serverless Orchestrator in order to aggregate and distribute it among others.

[1] https://ai.googleblog.com/2017/04/federated-learning-collaborative.html

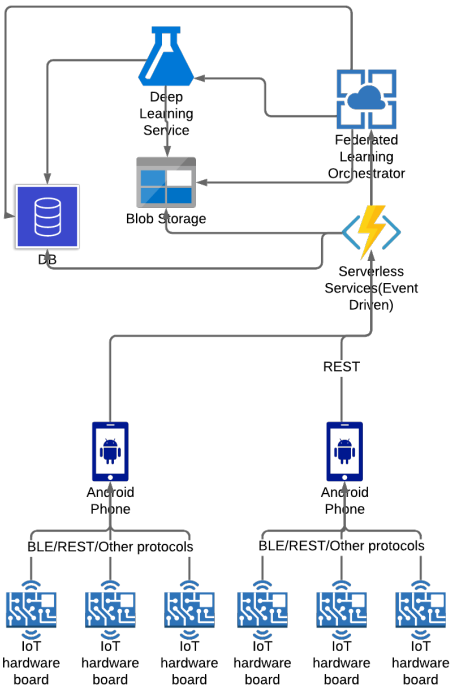

Figure 1: Architecture Diagram. Arrow heads represents the data movement from one component to another.

The Architecture

The proposed architecture is shown in Figure 1. Each component is explained in this section.

Machine Learning Pipeline

IoT devices need pervasive and secure connections for transferring aggregated data to a remote central cloud server where the collective learning takes place. A monolithic TensorFlow Federated [2] service would serve as the main ML model to train on the federated data obtained from various devices. This service in the infrastructure would be deployed on an independent IaaS Virtual Machine hosted either on-premise or on a public cloud platform.

[2] https://www.tensorflow.org/federated

Highly Available Geo-replicating Database

A highly available database is required to maintain state of connected devices. This database should support horizontal scaling with Geo-replication. A few choices are Azure Cosmos DB, AWS DynamoDB etc. It may help to keep all the services easily deployable on public clouds.

Blob Storage (Hot)

A blob storage solution, such as Azure Blob storage and AWS S2 buckets, operating in a Hot tier will act as an intermediary for storing the data received from the various devices. It ensures fast data access for the ML Pipeline.

Serverless Federated Orchestration Infrastructure

Serverless functions based on Web Sockets would perform the work of an orchestrator for facilitating various steps of the FL protocol. It also perform the various phases of FL. Serverless is an optimum choice here as it provides unlimited scaling and is easy to deploy on a public cloud platform. Each aspect of the orchestration would be divided into separate functions which would be separated logically from each other. These tasks are:

Device Monitoring It will monitor the devices and fetch data whenever the edge devices meet certain predefined conditions (e.g., While a mobile device is charging, manually or externally triggered for commercial enterprise applications). It may also be a time based periodic trigger. Once the devices are connected to the Serverless infrastructure, the function would make an entry with all the information of the connection in the highly available Database corresponding to the connected device.

Device Selection With the data collected in the high availability database, a randomization protocol will select devices from the currently connected devices. The selection of devices will be based on various parameters to optimize the training outcomes like: higher priorities for the devices those had not participated in last many training cycles. Devices are preferred to be from different Geo-locations to enhance the quality and diversity.

Data Aggregation Using the federated averaging approach of federated learning, this function would collect the predicted data-set from the various devices selected in the device selection. The data would be streamed directly to the Blob for further processing by the Machine Learning pipeline.

Trigger Pipeline The process can be triggered in two different ways. First, using a time based trigger for disconnecting the pipeline after an amount of time. The other trigger can be a threshold based on the number of new blobs (of data) received.

Model Update After the model has been retrained using the federated data on the ML pipeline, the updated model is pushed back to the devices by this function.

Edge Devices

These devices have the ML model for performing predictions on the device. These are mainly mobile phones which depends on IoT/Local devices stated next for the data. Any other capable edge device can be used.

IoT/Local devices and Real world Applications

These are the end data capturing devices. We listed some of the example applications where we think this infrastructure can prove useful:

Wearables/Smart Watches: The concern for privacy of medical (e.g., heart rate, ECG) and location data collected through these devices has been growing nowadays. The federated infrastructure can be extended here to perform predictions from data on a mobile whilst updating the model periodically via the federated infrastructure ensuring there is no privacy breach of sensitive data.

Home Assistant Devices: Similarly, the concern of the 24×7 audio data collected and being streamed to the servers is a huge privacy concern. Since these devices are extremely low on computing, an edge relayed model can help ensure that the data is not continuously being streamed and saved by the service provider.

CCTV & Cameras: Various applications for CCTV cameras are very popular now. It contains features like NEST cameras with face detection and other intelligent functions inbuilt. But many times these predictions are made at the cloud which leads to possible breach of security. Our FL approach helps in ensuring the user data is safe whilst providing improved services.

Conclusion

We explored the ideas and concepts of Federated Learning principles and their applications for IoT/Edge devices. As a solution to the existing problems of connectivity and limited hardware capacity of IoT devices, we proposed an architecture with Mobile Relaying data from IoT devices to Serverless infrastructure hosted on public cloud. This enhances the ML aspects and preserve the privacy of data within the edge layer.

References

Bonawitz, K. A.; Eichner, H.; Grieskamp, W.; Huba, D.; Ingerman, A.; Ivanov, V.; Kiddon, C. M.; Konecny, J.; Mazzocchi, S.; McMahan, B.; Overveldt, T. V.; Petrou, D.; Ramage, D.; and Roselander, J. 2019. Towards Federated Learning at Scale: System Design. In *SysML 2019*.

Manzoor, A.; Porambage, P.; Liyanage, M.; Ylianttila, M.; and Gurtov, A. 2018. DEMO: Mobile Relay Architecture for Low-Power IoT Devices. In *2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, 14–16.

Wu, Q.; He, K.; and Chen, X. 2020. Personalized Federated Learning for Intelligent IoT Applications: A Cloud-Edge Based Framework. *IEEE Open Journal of the Computer Society* 1: 35–44.