Exploring the Impact of Attacks on Ring AllReduce

Jiayu Wang, Peng Liu, Zehua Guo∗
Beijing Institute of Technology
China

Sen Liu
Peng Cheng Lab, Fudan University
China

Chao Yao
Shaanxi Normal University
China

ABSTRACT

Distributed Machine Learning (DML) is widely used to accelerate the training of the deep learning model. In DML, Parameter-Server (PS) and Ring AllReduce are two typical architectures. Recently, observing that many works address the security problem in PS, whose performance can be greatly degraded by malicious participation during the training process. However, the robustness of Ring AllReduce, which can solve the communication bandwidth problem in PS, to the malicious participant is still unknown. In this paper, we design a series of experiments to explore the security problem in Ring AllReduce, and reveal it can also suffer from the malicious participant.

ACM Reference Format:

1 INTRODUCTION

Deep learning has been successfully applied to many applications in recent years. However, the high cost in computation and memory of deep learning still hinder its implements. To solve this problem, Distributed Machine Learning (DML) has been proposed by using many machines simultaneously. Based on different topological architectures, Parameter Server (PS) and Ring AllReduce are two typical DML methods. In PS, a central server is used to coordinate the training of different participates (usually named workers in DML). The server first sends a model to each worker, then each worker computes the gradient/parameter based on its local data, and sends the gradient/parameter back to the server. The server finally updates the model by the collected gradients/parameters from the workers. The process will be repeated many times, so the communication bandwidth is a heavy burden for the central server. Ring AllReduce [3], which deploys the worker in a ring topology, is proposed to solve the problem. In Ring AllReduce, each worker only needs to receive parameters from one worker and send parameters to another worker. Consequently, the communication bandwidth of Ring AllReduce is greatly saved.

Training a distributed machine learning model is not safe due to cyberattacks. The malicious workers, which are controlled by the attacker, can send the wrong gradients or parameters to the server. An example is presented in Figure 1. In the example, the gradients estimated by the correct workers have a similar direction, the one from the malicious worker can be different from the correct ones. In this way, a wrong estimation of the global gradients is produced, leading to degraded performance or divergence of the model.

Based on the scheme in Figure 1, different attacks are proposed recently [2]. To prevent PS from the attack, many works are also proposed. [1] kicks out the malicious workers to prevent PS from attack. [4] aggregates the gradients of each worker, rather than briefly calculate the average of them, to obtain robust global gradient.

However, existing works mainly focus on the security problem in PS. To the best of our knowledge, the impact of malicious workers on the performance of Ring AllReduce is still unknown yet. Ring AllReduce does not have a central server, and gradients are processed as they arrive at workers instead of being aggregated centrally at the server. Thus, none of the workers can work from a global perspective, and existing defending methods for PS cannot be directly used. In this paper, we address the security problem in Ring AllReduce. We design several experiments to test the influence of malicious workers when training the DML models in Ring AllReduce. The experimental results show that the malicious workers can reduce the accuracy of the models, and even can lead to the models disconverged in the worst case.

(a) The calculation of the gradients with malicious workers.
(b) The calculation of the gradients with malicious workers.

Figure 1: The calculation of the gradients with malicious workers can be highly different with no malicious workers. It can hurt training process with malicious workers.

2 EXPERIMENT DESIGN

2.1 Overview of Ring AllReduce

Training the Ring AllReduce architecture contains many iterations. In each iteration, each worker firstly calculates the local gradients based on the local data. Then, the model is updated by two processes, Scatter-Reduce and Allgather.

In Scatter-Reduce, each worker sends the gradients to the fixed adjacent worker and receives the gradients from another fixed adjacent worker. After \( N - 1 \) steps (where \( N \) is the number of total workers), each worker has parts of the global updated model. Then Allgather is adopted to update the global model. After \( N - 1 \) steps, each worker gets the updated global model. Figure 2 shows the scheme of one iteration in Ring AllReduce.
because that, the proportion of attackers in participants is reduced in this situation. In fact, if there are more attackers, the decrease of the training accuracy and the divergence of the training model will be more obvious.

4 CONCLUSION AND FUTURE WORK

In this paper, we explore the security problem in Ring AllReduce architecture, and find that the malicious workers indeed affect the model’s performance. In future, we will evaluate the impact of existing attack methods on complex training models, datasets, and other processes in Ring AllReduce.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grants 6202019 and 6202066, Beijing Institute of Technology Research Fund Program for Young Scholars, Fundamental Research Funds for the Central Universities under Grant 1301032207, and the project "PCL Future Greater-Bay Area Network Facilities for Large-scale Experiments and Applications (LZC0019)."

REFERENCES