github.com/PermiJW/signSGD-with-Majority-Vote

signSGD with Majority Vote is Communication Efficient and Byzantine Fault Tolerant

Jeremy Bernstein Jiawei Zhao Kamyar Azizzadenesheli Anima Anandkumar

1. Motivation

Distributed deep learning can be limited by the **time cost of communicating gradients** and the chance of **network faults** for large machine counts. We studied a simple algorithm to solve these problems.

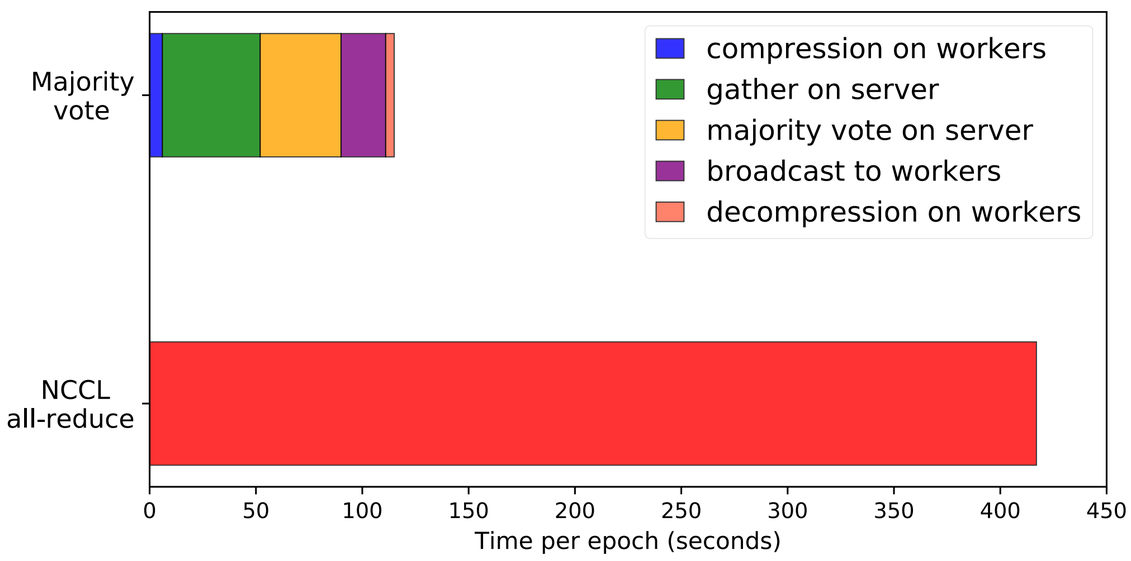
2.Algortithm

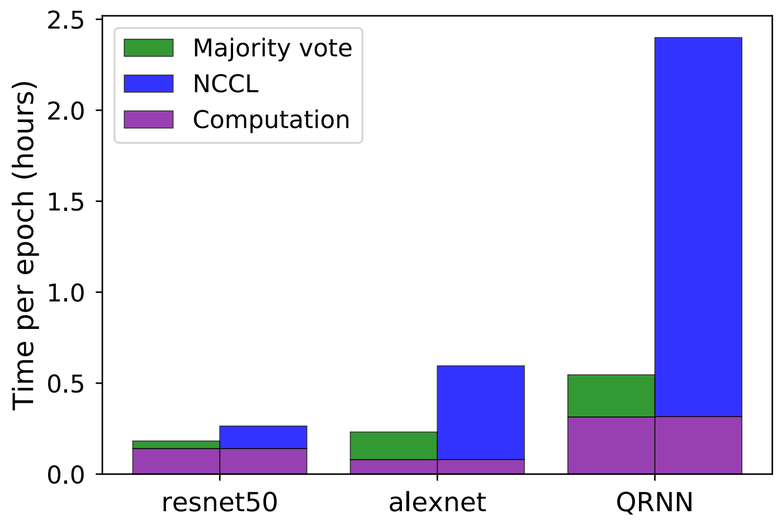
Worker → Server:

sign(g)

Server → Workers:

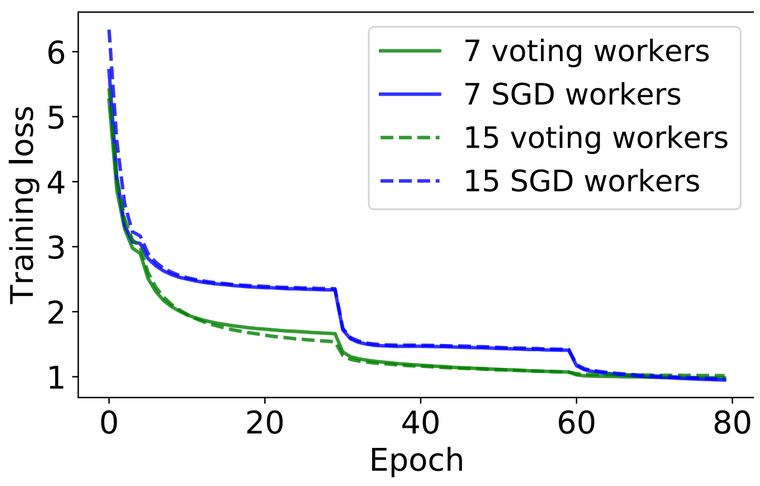
sign[∑ sign(g)]

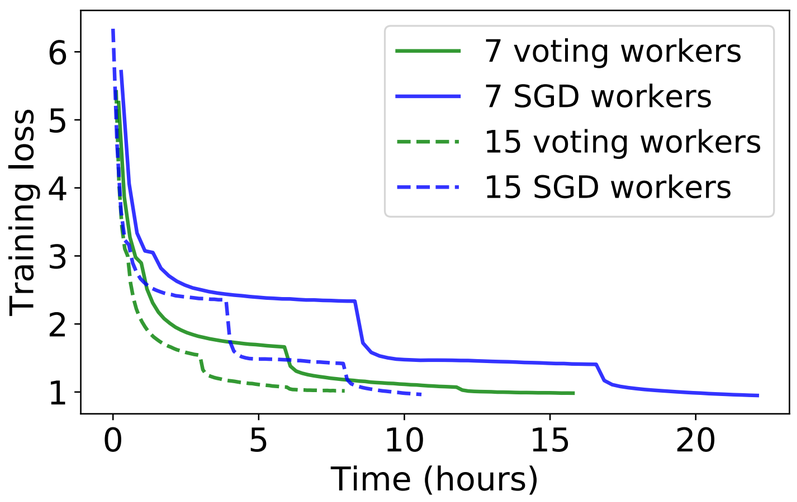


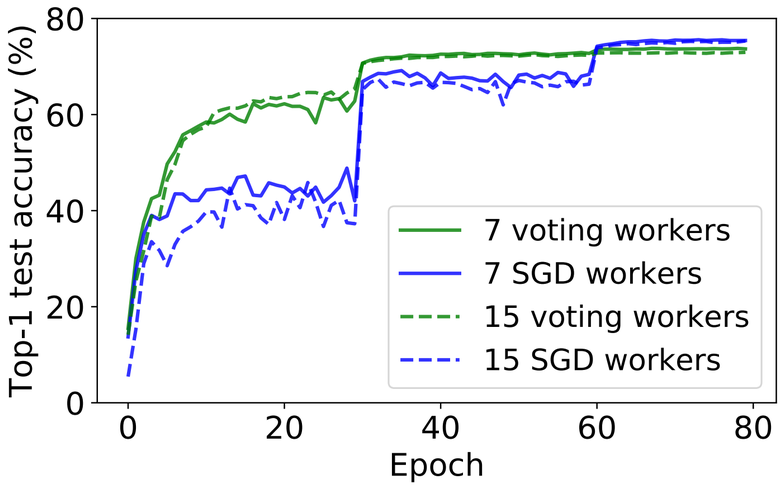


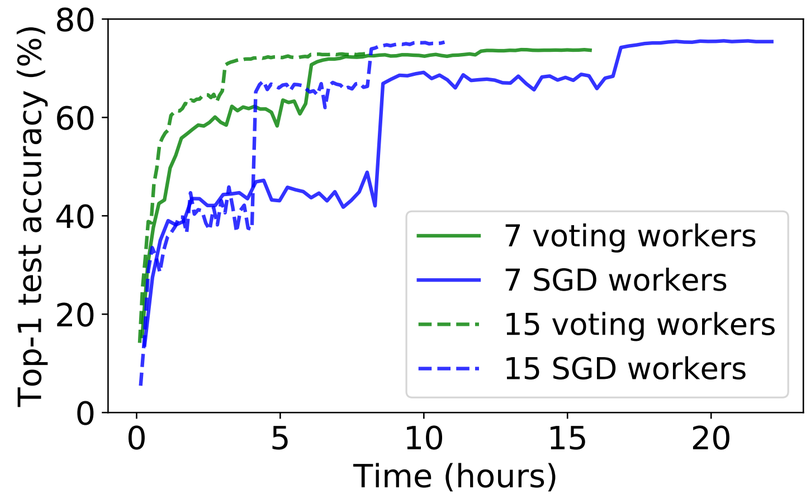
3. ImageNet communication experiments

We built our distributed training system in **PyTorch**. We **benchmarked against NCCL** (the state of the art collective communications library).









4. ImageNet robustness experiments

Adversarial workers invert their sign gradient estimate. We are **robust to < 50% adversaries.**