Paper Review: Communication Efficient Distributed Machine Learning with the Parameter Server

Many modern applications of optimization involve very large-scale data, so large that they can no longer be processed on a single workstation running single-threaded codes. A major challenge of distributed optimization algorithms is the inter-machine data communication. In this paper, the authors propose a new algorithm taking advantage of the parameter server and proximal gradient methods to solve a non convex and non smooth optimization problem efficiently. In order to reduce the effect of the asynchronous communication, they introduce two relaxations into the proposed framework. Firstly, they use the dependency graph with bounded delay to represent the task dependency and to relax consistency requirements between tasks. Furthermore, they apply user-defined filters to the parameter server framework, allowing for more fine-grained control of consistency. These filters can transform and selectively synchronize the data communicated in a task, resulting in a better data compression. Moreover, for reducing the sensitivity to data inconsistency, they modify the proximal gradient methods with a block scheme, where only one block of parameters is updated per iteration. This delayed block proximal gradient method makes the communication more efficient and is able to handle high-dimensional sparse data. Assuming that the change in the gradients of loss function is bounded and it is block-Lipschitz continuous, the algorithm is able to converge to a stationary point in expectation.To demonstrate its efficiency, they did experiments for several challenging tasks on the real datasets up to 0.6PB size with hundred billions samples and features. However, this framework can not avoid the use of master nodes either, which is referred to as servers recording and gathering all the information from worker nodes. Since it is easier to be understood and applied, it was the de facto standard for a long time in distributed systems, though it is not as perfect as we expect.