Large Scale Machine Learning and Applications Progress Report

Li Zhirong

July 3, 2017

1 Introduction

Due to the demand from industry or research teams for big data analytics, distributed systems for large-scale data processing become very popular in recent years. MapReduce and Hadoop are among the first such systems. However, these two systems are not designed for processing iterative workloads and therefore they are inefficient for jobs such as machine learning and graph analytics. So my project focuses on the research of Distributed or Multi-core Implementation of Variance Reduction based Methods, which overcomes the weakness of some distributed systems such as MapReduce and Hadoop.

2 Completed Research

Since submitting my proposal, I have spent most of my research time obtaining and reading sources. My principal source, Hogwild: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent[Feng Niu, 2011], is a paper that has required much time to read through.

The paper shows that Hogwild! allows processors access to shared memory with the possibility of over-writing each others work. When the associated optimization problem is sparse, meaning most gradient updates only modify small parts of the decision variable, then Hogwild! achieves a nearly optimal rate of convergence.

The second paper I have read is Accelerating Stochastic Gradient Descent using Predictive Variance Reduction. The paper introduces an explicit variance reduction method for stochastic gradient descent which is called stochastic variance reduced gradient(SVRG). For smooth and strongly convex functions, they prove that this method enjoys the same fast convergence rate as those of SDCA and SAG. Moreover, unlike SDCA or SAG, this method does not require the storage of gradients and thus is more easily applicable to complex problems such as structured prediction or neural network learning.

3 Remaining Research

In the next few days, I will talk with TAs about the project and do more research on Distributed or Multicore Implementation of Variance Reduction based Methods. In addition, I will contribute to the programming of the experiment of the research which TAs have done.

4 Conclusions

This progress report shows the status of my research on Distributed or Multicore Implementation of Variance Reduction based Methods. From two papers I have read, I learned some parallelly techniques on Implementation of Variance Reduction. On next stage, I will learn more about how to implement it using programming.

References

[1] F. Niu, B. Recht, C. Re and S. J. Wright, Hogwild!: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent, Computer Sciences Department, University of Wisconsin-Madison, June 2011.

[2] R. Johnson, T. Zhang, Accelerating Stochastic Gradient Descent using Predictive Variance Reduction.