Accepted Manuscript

Title: HotML: A DSM-based Machine Learning System for Social Networks

Author: Yangyang Zhang Jianxin Li Chenggen Sun Md Zakirul Alam Bhuiyan Weiren Yu Richong Zhang

PII: S1877-7503(17)30345-9
DOI: http://dx.doi.org/doi:10.1016/j.jocs.2017.09.006
Reference: JOCS 759

To appear in:

Received date: 1-4-2017
Revised date: 26-8-2017
Accepted date: 12-9-2017

Please cite this article as: Yangyang Zhang, Jianxin Li, Chenggen Sun, Md Zakirul Alam Bhuiyan, Weiren Yu, Richong Zhang, HotML: A DSM-based Machine Learning System for Social Networks, *<![CDATA[Journal of Computational Science]]>* (2017), http://dx.doi.org/10.1016/j.jocs.2017.09.006

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

HotML: A DSM-based Machine Learning System for Social Networks

Yangyang Zhang[a], Jianxin Li[a], Chenggen Sun[a], Md Zakirul Alam Bhuiyan[b], Weiren Yu[a], Richong Zhang[a]

[a]School of Computer Science and Engineering, Beihang University, Beijing, 100191, China

[b]Department of Computer and Information Sciences, Fordham University, New York, 10458, USA

Highlights:

1. We proposed HotML, a novel distributed machine learning system for social networks based on a recently proposed high-performance DSM system, Grappa. HotML contains many important components that covers the whole pipeline of machine learning for social networks.
2. The design of parameter server component is greatly improved by decoupling the PS servers and workers physically. The dedicated parameter server is introduced to maximize the computing resources, and improve server throughput as well as the overall performance of HotML.
3. High-level data abstraction and user interfaces are designed. Data is abstracted as *GlobalTable* along with high-level data flow operations like *map* / *reduce* to improve the ease of use. *GlobalTable* could be used to load and do feature engineering on training data easily.
4. Lightweight task scheduling is introduced to overlap computation with network communication and better utilize computing resource and network resource. In HotML, training process of ML algorithms for social networks will be splitted into many small tasks to reduce network waiting time.
5. Flexible consistency models are designed to boost the convergence of machine learning algorithms for social networks. SSP is implemented to relax the consistency to speed up convergence of ML algorithms for machine learning. In addition, SSPPush is introduced in servers to leverage idle network bandwidth to push global parameters to workers in advanced to reduce SSP waiting time. SSPDrop is designed in worker to drop trivial parameter updates to reduce network communication.
6. HotML provdies two checkpoint methods designed for fault tolerance. One is server-side distributed checkpoint method that is transparent to the workers. The other one is worker-side checkpoint method that is more flexible and useful for user-implemented checkpoint policies, fine-tuning.
7. The load balancer is introduced to balance workloads among workers to deal with straggler problem especially in a cluster of machines with different computing resources.
8. Experiments are conducted systematically to demonstrate the performance of the proposed system. Experimental results show that HotML can reduce networking time by about 50%, and achieve up to 1.9x performance compared to the popular ML system, Petuum.

Accepted Manuscript

Download English Version:

https://daneshyari.com/en/article/6874373

Download Persian Version:

https://daneshyari.com/article/6874373

Daneshyari.com