Department of Electrical and Computer Engineering

Title: ***«Distributed neural network training via independent subnets»***

Anastasios Kyrillidis
Noah Harding Assistant Professor
Computer Science Dept., Rice University

Wednesday, 2 March 2022, 12:00 – 13:30
Auditorium, Department of Electrical and Computer Engineering, Karamanli & Lygeris Str.

Abstract:

Distributed machine learning (ML) can bring more computational resources to bear than single-machine learning, thus enabling reductions in training time. Distributed learning partitions models and data over many machines, allowing model and dataset sizes beyond the available compute power and memory of a single machine. In practice though, distributed ML is challenging when distribution is mandatory, rather than chosen by the practitioner. In such scenarios, data could unavoidably be separated among workers due to limited memory capacity per worker or even because of data privacy issues. There, existing distributed methods will utterly fail due to dominant transfer costs across workers, or do not even apply.

We propose a new approach to distributed fully connected neural network learning, called independent subnet training (IST), to handle these cases. In IST, the original network is decomposed into a set of narrow subnetworks with the same depth. These subnetworks are then trained locally before parameters are exchanged to produce new subnets and the training cycle repeats. Such a naturally ``model parallel'' approach limits memory usage by storing only a portion of network parameters on each device. Additionally, no requirements exist for sharing data between workers (i.e., subnet training is local and independent) and communication volume and frequency are reduced by decomposing the original network into independent subnets. These properties of IST can cope with issues due to distributed data, slow interconnects, or limited device memory, making IST a suitable approach for cases of mandatory distribution. This talk will provide results on MLPs, ResNets, CNNs, efficient pretraining tasks, GCNs as well as some theoretical guarantees.

Biography:

My name is Anastasios Kyrillidis, Noah Harding Assistant Professor at the Computer Science department at Rice University. Prior to that, I was a Goldstine PostDoctoral Fellow at IBM T. J. Watson Research Center (NY), and a Simons Foundation PostDoc member at the University of Texas at Austin. I finished my PhD at the CS Department of EPFL (Switzerland). I got my M.Sc. and Diploma from Electronic and Computer Engineering Dept. at Technical University of Crete (Chania). My research interests include (but not limited to):
- Optimization for machine learning
- Convex and non-convex algorithms and analysis
- Large-scale optimization
- Any problem that includes a math-driven criterion and requires an efficient method for its solution.