

Seminar

Byzantine-resilient distributed training

The training for many machine learning algorithms is based on iterative gradient descent. In the case of large data sets, the training is carried out distributedly on computation clusters. However, computation errors and transmission errors as well as malicious nodes have a negative impact on the training process, and may lead to an erroneous model. Especially so-called Byzantine failures pose a challenge for distributed training. Therefore, an important research objective are techniques for robustness against Byzantine failures. The goal of this seminar work is to understand and analyze a coding-based approach to Byzantine-resilient distributed training.

Main paper:

- Chen, Lingjiao, Hongyi Wang, Zachary Charles & Dimitris Papailiopoulos (2018). DRACO: Byzantine-Resilient Distributed Training via Redundant Gradients. ArXiv:1803.09877 [Cs, Math, Stat].

Related papers:

- Tandon, Rashish, Qi Lei, Alexandros G. Dimakis & Nikos Karampatziakis (2017). Gradient Coding. ArXiv:1612.03301 [Cs, Math, Stat].
- Chen, Y., Su, L. & Xu, J. (2017). Distributed statistical machine learning in adversarial settings: Byzantine gradient descent. Proceedings of the ACM on Measurement and Analysis of Computing Systems, 1(2), 1-25.
- Blanchard, P., El Mhamdi, E. M., Guerraoui, R., & Stainer, J. (2017). Machine learning with adversaries: Byzantine tolerant gradient descent. Advances in Neural Information Processing Systems, 30.

Prerequisites

- Knowledge in coding theory
- Proficiency in linear algebra
- Basic knowledge in machine learning and gradient descent algorithms

Advisors

Luis Maßny