James Salamy, Ayush Sharma, Manya Ghobadi and Muriel Médard



FlexEnt: Entropy Coding to Curb Stragglers in Large-Scale Distributed Machine Learning



Large scale machine learning workload [ImageNet]

Distributed training using ringallreduce aggregation [Horovod] Cloud datacenter with fat tree topology [SIGCOMM'08]



Why wait for Stragglers?

Key Idea Choose Highest Entropy Images



System Proposal

- Primary Workers (Pi): Train on all data
- Shadow Workers (Si): Train on high entropy data

Entropy Graph Colouring

1) Dataset



2) Detect eye features



Selection of Entropy-Rich Data

- Combine multiple feature colours together

Joint Entropy Ranking

- Exploit concavity of joint entropy





Future Directions

- How to use the results from CNN training filters and other internal results or labelling to select features?

- What is the best ratio of stragglers for efficient protection?
- Can we use shadows to entirely replace primaries even when they haven't straggled?