

DRPS : Efficient Disk-Resident Parameter Servers for Distributed Machine Learning

Zhen SONG, Yu GU, Zhigang WANG, Ge YU

Frontiers of Computer Science, DOI: [10.1007/s11704-021-0445-2](https://doi.org/10.1007/s11704-021-0445-2)

Problems & Ideas

- Problems of Existing PS-based Distributed Systems
 - Large-scale Models: can reach even one trillion in industry;
 - Memory Constraints: renting large-scale cloud servers is economically infeasible for research teams and small companies;
 - High I/O Costs: expanding directly to disk induces enormous I/O;
 - Unified Parallel Control: cannot have the “predictability”
- Ideas: Efficient Disk-Resident Parameter Servers (DRPS)
 - Efficient Index Structure and Multi-objective Partition for Models
 - Worker-selection Parallel Model of Computation (WSP)

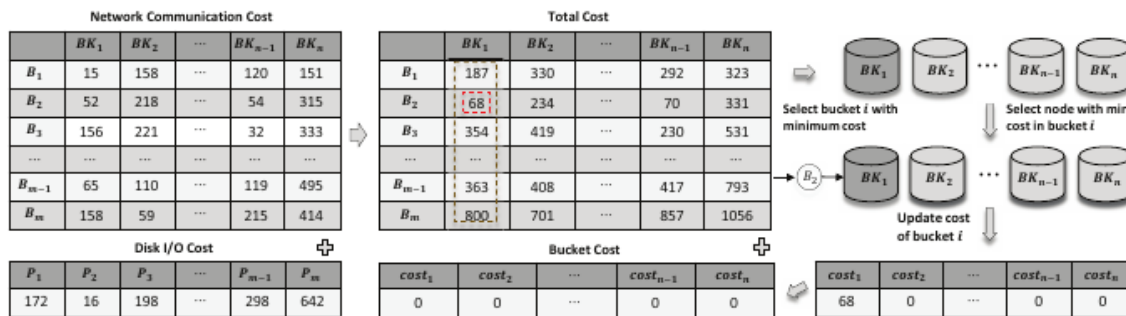


Fig.1 Multi-objective Partition

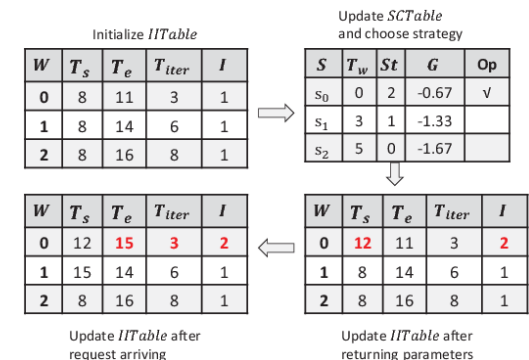
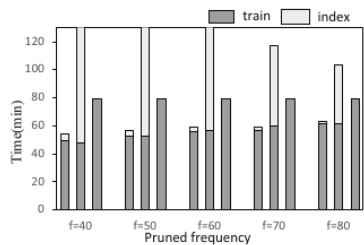


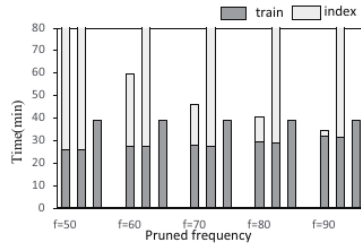
Fig.2 Worker-selection Parallel Model

Main Contributions

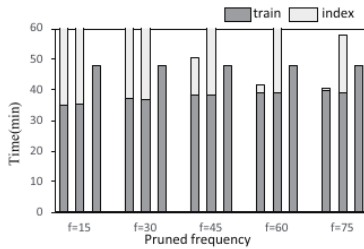
- Compared with Qin et.al's index
- Compared with Traditional Parallel Control



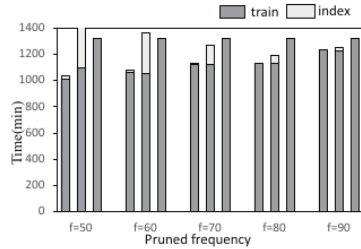
(a) LoR on Avazu-CTR



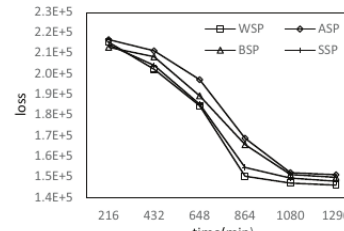
(c) LoR on Classification



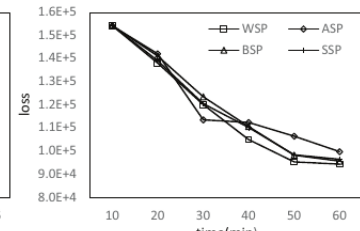
(b) LMF on MovieLens



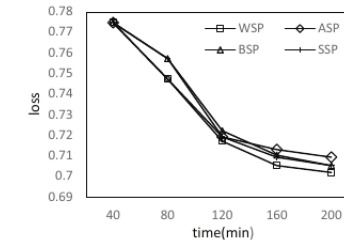
(d) LMF on Matrix



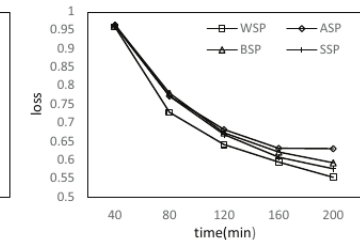
(a) LMF on Matrix



(b) LMF on MovieLens



(c) LoR on Classification



(d) LoR on Avazu-CTR

Ablation Experiments

DRPS	Indexing	Partitioning	BSP	ASP	SSP	WSP	$T_{Mat}(min)$	$T_{Mov}(min)$	$T_{Clas}(min)$	$T_{Aza}(min)$
✓			✓				1468	93	260	273
✓	✓		✓				1268	82	232	238
✓	✓	✓	✓				1125	75	214	227
✓	✓	✓		✓			1135	78	232	236
✓	✓	✓			✓		979	62	212	219
✓	✓	✓				✓	864	50	182	212