The Influence of Placement on Transmission in Distributed Computing of Boolean Functions

IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Lucca, Italy, September 10-13, 2024.

Ahmad Tanha, Derya Malak   Communication Systems Department, EURECOM, France.   {tanha, malak}@eurecom.fr

Motivation

- Offloading computations over servers [1-6]
- Nonlinear coding of distributed sources [8]
- Efficient usage of scarce resources [1-6]
- Evaluating complex functions [1-6]
- Reducing the communication cost [2-6]

Related Works

- Distributed source coding [7-9]:
  Compression for function computation
- Coded computing [2-6]:
  Adding redundancy to tolerate stragglers, Storage-communication-computation tradeoffs

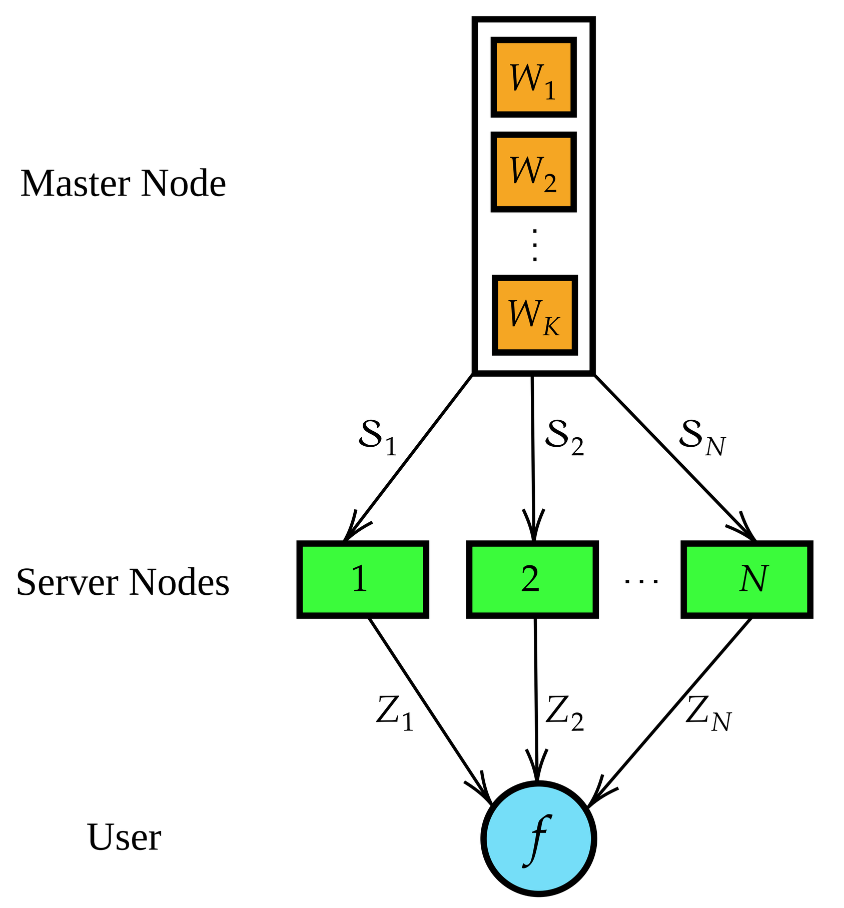
What is missing?

The **placement-transmission tradeoff**.

Our novel approach:

- captures any Boolean function
- is sensitivity-based
- adapts any placement (uncoded, acyclic)

Sensitivity-based computation



Exploiting the polynomial representation:

$$f(\mathbf{W}) = \bigoplus_{\mathcal{P}\subseteq[K]} c_{\mathcal{P}} \prod_{k\in\mathcal{P}} W_k$$

for some subsets $\mathcal{P}$ of K datasets and coefficients $c_{\mathcal{P}} \in \mathbb{F}_2$.

The function f :

- **Nonlinear** (computationally complex)
- Any (low/high) **degree**
- **Placement-dependent** complexity

$\Pi = \arg\min |Z| = \arg\min \sum_{n=1}^{N} |Z_n|$?

Placement with the minimum number of transmissions?

Phases of Distributed Computing

Dataset placement:

$$\rho_n : \mathbb{F}_2^K \longrightarrow \mathbb{F}_2^M \ , \ \forall n \in [N] \ ,$$
$$S_n = \rho_n(\mathbf{W}) \subseteq \{\mathbf{W}\} \ , \ |S_n| = M \ , \ \forall n \in [N] \ .$$

Encoding and Transmissions:

$$E_n^f : \mathbb{F}_2^M \longrightarrow \mathbb{F}_2^{|Z_n|} \ ,$$
$$Z_n = E_n^f(S_n) = \{z_{ni} \,|\, i \in [|Z_n|]\} \ , \ \forall n \in [N] \ .$$

Decoding:

$$D : \mathbb{F}_2^{|Z|} \longrightarrow \mathbb{F}_2 \ , \ Z = \{Z_n \,|\, n \in [N]\} \ .$$

System model and key assumptions

$(K, N, M, \mathcal{S}, f)$ distributed computing scheme:

$$f(\mathbf{W}) = \bigoplus_{n=1}^{N} f_n \ , \quad f_n = \prod_{k\in\mathcal{K}_{\mathcal{P}_n,d}} W_k \ ,$$

where $\mathcal{K}_{\mathcal{P},d}$ is a subset $\mathcal{P}$ with cardinality d of K datasets:

$$\mathcal{K}_{\mathcal{P},d} \triangleq \{\mathcal{P} \subseteq [K] \mid |\mathcal{P}| = d\} \ ,$$

where we assume:

- $d = M$ (the degree of each subfunction)
- $K = NM$ (the number of datasets)

Key definitions

Definition 1. *A $(K, N, M, \mathcal{S}, f)$ distributed computing scheme is called achievable if the function f can be recovered in an error-free manner by the user, i.e., $D(Z) = f$, where the set of transmissions*

$$Z = \{Z_n \,|\, n \in [N]\}$$

is possibly a nonlinear combination of the encoded data $Z_n = E_n^f(S_n)$, $n \in [N]$, which is placement-dependent and function-aware.

The total number of transmissions by all servers:

$$T^{(\mathcal{S})} \triangleq |Z| = \sum_{n=1}^{N} |Z_n| \ .$$

Definition 2. *The joint sensitivity of $f(\mathbf{W})$ to the set $\mathcal{S}$ of subsets of datasets is defined as*

$$\mathrm{Sen}_{\mathcal{S}}(f, \mathbf{W}) = \sum_{n=1}^{|\mathcal{S}|} \mathrm{Sen}_{\mathcal{S}_n}(f, \mathbf{W}) \ ,$$

where

- $\mathrm{Sen}_{\mathcal{S}_n}(f, \mathbf{W}) = 1_{f(\mathbf{W}\oplus e_{\mathcal{S}_n}) \neq f(\mathbf{W})}$.
- $e_{\mathcal{S}_n} \triangleq \bigoplus_{\{k | w_k \in \mathcal{S}_n\}} e_k$.

Definition 3. *The joint influence of the datasets of $\mathcal{S}_n$ on the function f is defined as*

$$\mathrm{Inf}_{\mathcal{S}_n}(f) = \mathbb{P}[f(\mathbf{W} \oplus e_{\mathcal{S}_n}) \neq f(\mathbf{W})]$$
$$= \mathbb{E}_{\mathbf{W}}[\mathrm{Sen}_{\mathcal{S}_n}(f, \mathbf{W})] \ .$$

Main results

Definition 4. *The average joint sensitivity of $f(\mathbf{W})$ to the set*

$$\mathcal{S} = \{S_n | n \in [N]\}$$

of all possible datasets specified by ρ_n given a cache size constraint M is given as follows:

$$\mathrm{as}_{\mathcal{S}}(f) = \mathbb{E}_{\mathbf{W}}[\mathrm{Sen}_{\mathcal{S}}(f, \mathbf{W})] = \sum_{n=1}^{|\mathcal{S}|} \mathrm{Inf}_{\mathcal{S}_n}(f) \ .$$

Lemma 1. *The joint influence of multiple datasets in a subset $\mathcal{S}_n$ with an arbitrary size from a product subfunction $f_n(\mathbf{W}) = \prod_{k\in\mathcal{K}_{\mathcal{P},d}} W_k$ of degree d equals the influence of each dataset on f_n, i.e.,*

$$\mathrm{Inf}_{\mathcal{S}_n}(f) = \mathrm{Inf}_k(f) = \frac{1}{2^{d-1}} \ .$$

Multiplication reduces the joint influence.

Lemma 2. *Let $f = f_n \oplus f_{n'}$, where $n \neq n'$. Consider two different subsets of datasets denoted by $\mathcal{S}_1 = \{W_k | W_k \in \mathcal{I}_{f_n}\}$ and $\mathcal{S}_1' = \{W_k | W_k \in \mathcal{I}_{f_n} \cup \mathcal{I}_{f_{n'}}\}$. We then have:*

$$\mathrm{Inf}_{\mathcal{S}_1'}(f) \geq \mathrm{Inf}_{\mathcal{S}_1}(f) \ ,$$

where the set of variables included in the subfunction f_n is

$$\mathcal{I}_{f_n} \triangleq \{W_k \,|\, f_n = \prod_{k\in\mathcal{K}_{\mathcal{P}_n,d}} W_k\}, \quad \forall n \in [N] \ .$$

Addition increases the joint influence.

Theorem 1. (A communication-optimal placement configuration.) *Given a $(K, N, M, \mathcal{S}, f)$ distributed computing scheme, the average joint sensitivity and the total number of transmissions are lower bounded by*

$$\mathrm{as}_{\mathcal{S}}(f) \geq \mathrm{as}_{\mathcal{S}^*}(f) = \frac{N}{2^{M-1}}, \ T^{(\mathcal{S})} \geq T^{(\mathcal{S}^*)} = N,$$

respectively, corresponding to $\mathcal{S}^* = \{\mathcal{S}_n^* \,|\, n \in [N]\}$, where $\mathcal{S}_n^* = \{W_k \,|\, W_k \in \mathcal{I}_{f_n}\}$.

Conclusions and future works

- The impact of placement on transmissions
- Communication-optimal placement
- Nonlinear Boolean functions
- Multi-user nonlinearly-separable case

References

[1] [Dean-Ghemawat, 08] [2] [Li et al., 18]
[3] [Yan-Yang-Wigger, 22]
[4] [Wan et al., 22] [5,6] [Khalesi-Elia, 23, 24]
[7,8] [Malak, 23, 24] [9] [Feizi-Médard, 14]

EURECOM Sophia Antipolis   erc   FRANCE 2030