Abstract

We study the limits of communication efficiency for function computation in collocated networks within the framework of multi-terminal source coding theory. With the goal of computing a desired function of sources at a sink node, nodes interact with each other through a sequence of error-free, network-wide broadcasts of finite-rate messages. For any function of independent sources, we derive a computable characterization of the set of all feasible message coding rates - the rate region - in terms of single-letter information measures. We show that when computing symmetric functions of binary sources, the sink node will inevitably learn certain additional information which is not demanded in computing the function. This conceptual understanding leads to new improved bounds for the minimum sum-rate. The new bounds are shown to be order-wise better than those based on cut-sets as the network scales. The scaling law of the minimum sum-rate is explored for different classes of symmetric functions and source parameters.

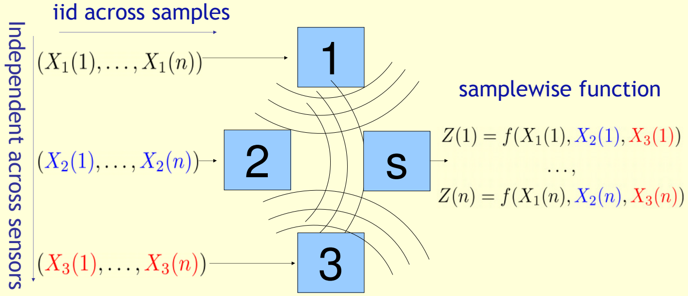
Bounds for Interactive Computation in Collocated Networks

Nan Ma[1]

Joint work with Prakash Ishwar[1] and Piyush Gupta[2]

[1] ECE Dept., Boston University [2] Bell Labs, Alcatel-Lucent

1. General Interactive Computation Problem



- m sensors. n iid source samples per sensor, indep. across sensors
- Samplewise function f computed at a sink node
- A sequence of noiseless broadcast for r rounds. Everyone "hears" everyone else. Transmission schedule: (1, 2, 3, 1, 2, 3, …, r rounds)
- Minimum sum-rate $R_{sum,r}$ = min (total # bits per sample) s.t. Pr(block computation error) → 0 as n → infinity

Objectives: To understand the

- Interplay between multi-round interaction and function structure
- Scaling behavior of the min sum-rate wrt #sensors m

2. Complete Solution to the General Problem

- Single-letter characterization of the entire rate-region
- Single-letter characterization of the minimum sum-rate:

$$R_{sum,r} = \min_{U^t} I(X^m; U^t)$$

mutual information *"source" rvs.* *"message" aux. rvs.*

s.t. U^t satisfy the following Markov chain and conditional entropy constraints:

$$\forall j \in [1,t], k = (j \bmod m),$$
$$U_j - (U^{j-1}, X_k) - (X^{k-1}, X^m_{k+1}),$$
$$H(f(X^m)|U^t) = 0$$

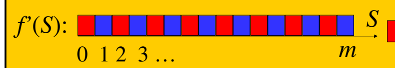
3. Computing Symmetric Functions of Binary Sources

- $X_1, X_2, X_3, \ldots$ independent Bernoulli sources
- $f(X^m) = f'(S)$ where $S = \sum_{i=1}^m X_i$
- Maximal f'-monochromatic intervals: $\{ [a, b] \}$

Color = value of function f'

4. Inevitability of Information Leakage!

Example: PARITY function



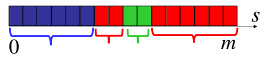
For any zero-error code, for each sample, can prove:

Surprise 1: The messages will inevitably reveal the exact value of S to the sink.

Surprise 2: The messages will inevitably reveal which sensors observe 1 and which observe 0.

- The sink WILL learn all the source samples, if it can compute their PARITY!

Key Lemma:

For zero-error block codes (Pr (block error) = 0)	For single-letter characteriz. (Pr (blk. error) → 0)
Given the messages, for *each sample i*: (i) Sink can locate $S(i)$ to within a *single* maximal monochrom. interval $[a_i, b_i]$,  (ii) Sink can *identify* a_i sensors observing 1 and $(m\text{-}b_i)$ sensors observing 0.	Given U^t satisfying $\forall j \in [1,t], k = (j \bmod m)$, $U_j - (U^{j-1}, X_k) - (X^{k-1}, X^m_{k+1})$, $H(f(X^m)|U^t) = 0$ with probability one: (i) S is in a *single* maximal monochromatic interval $[a, b]$ (ii) Can *identify* a X's which are 1 and $(m\text{-}b)$ X's which are 0.

"Collocated network + function structure" reveals more information than intended

5. Some Fascinating Implications

- New lower and upper bounds for min sum-rate $R_{sum,r}$
- For type-threshold fns. (e.g., MIN, MAX) of Ber(p) sources $R_{sum,r}(m) = \Theta(1)$ (compared with zero error: $\Theta(\log(m))$)
- For all symmetric functions of iid Ber(½) sources:
  $$\tfrac{1}{2} R_{sum,1} \leqslant R_{sum,r} \leqslant R_{sum,1}$$
- Scaling laws of $R_{sum,r}(m)$ in various other cases; in star nets, multiple rounds changes scaling order

New lwr. bnds orderwise better than cut-set bnds

Example: To compute MIN for iid Ber(½) sources, as m grows

- Cut-set bound: →0 (useless!)
- With new lwr bnd: $R_{sum,r}(m) = \Theta(1)$ (tight scaling!)

[1] N. Ma, P. Ishwar, P. Gupta, "Information-theoretic bounds for multiround function computation in collocated networks," International Symposium on Information Theory, Seoul, Korea, 2009.