CERIAS 2020 Annual Security Symposium ★ September 29 & 30 Purdue University, West Lafayette, Indiana

PURDUE UNIVERSITY Discovery Park

Byzantine-Resilient Distributed Optimization

Kananart Kuwarananchareon, Lei Xin, Shreyas Sundaram
School of Electrical and Computer Engineering, Purdue University

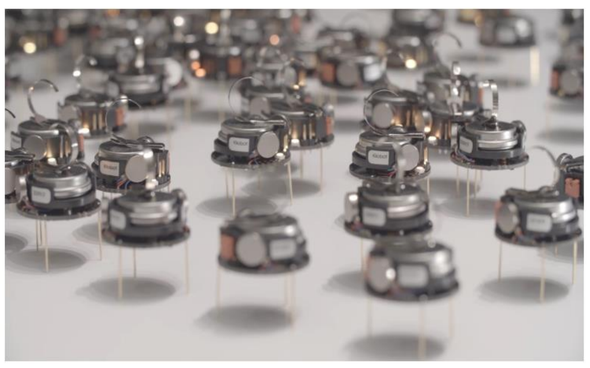
1. The Standard Distributed Optimization Problem

- **Problem:** A network of n agents collaboratively wish to find the minimizer of the sum of their cost functions using an iterative algorithm:

$$\min_{x\in\mathbb{R}^d} \frac{1}{n}\sum_{v_i\in V} f_i(x)$$

where $v_i \in \mathcal{V}$: agent i, $f_i(x)$: function of agent i.

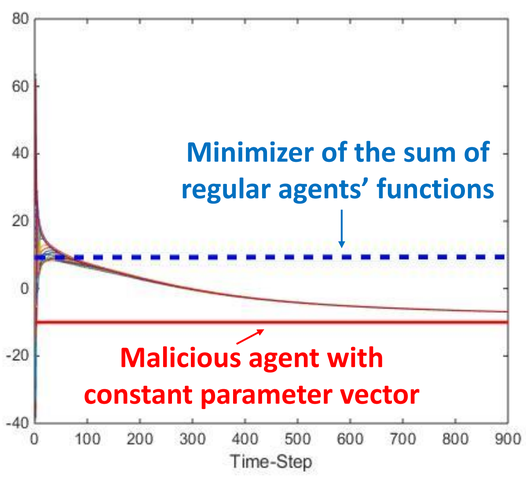
- **Assumptions:**
  - Each function $f_i(x)$ is convex, with non-empty and bounded set of minimizers
  - Each function has bounded subgradients
  - Each agent can only exchange information with its neighbors in the network
- **Applications:** Machine Learning, Power Systems, and Robotic Networks.



Cassidy, Josh, and Johanna Varner. "Can A Thousand Tiny Swarming Robots Outsmart Nature?" *KQED*, 21 July 2015, www.kqed.org/science/131005/can-a-thousand-tiny-swarming-robots-outsmart-nature.

2. The Resilient Distributed Optimization Problem

- A single malicious (or "Byzantine") agent can arbitrarily affect the computed value.



HAYARDENY, E. (2020, March 13). The Hidden Dangers of Malicious Bots [Digital image]. Retrieved September 11, 2020, from https://www.cpomagazine.com/cyber-security/the-hidden-dangers-of-malicious-bots/

- **Additional Assumption:** There are at most F Byzantine agents in the neighborhood of any regular agent ("F -local model").
- **Resilient distributed optimization:** Find a decision parameter that approximately minimizes the sum of functions of regular agents:

$$\min_{x\in\mathbb{R}^d} \frac{1}{|R|}\sum_{v_i\in R} f_i(x)$$

where R is the set of regular agents.

3. Our Contribution

- ❑ Proposed a **scalable** Byzantine-resilient distributed optimization algorithm for multi-dimensional convex functions.
- ❑ Proved convergence to a bounded region that contains the true minimizer irrespective of any Byzantine agents' behaviors, without any statistical assumptions on the functions, when the set of Byzantine agents and network satisfy certain properties.

Email: {kkuwaran, lxin, sundara2}@purdue.edu

4. Our Strategy

Iterative algorithm to compute an approximate minimizer of sum of regular agents' functions:

a) **Reference Point Calculation:** All regular agents run a resilient consensus algorithm with their local minimizers to compute a reference point $\hat{x}$

b) **Initialization:** Each regular agent $v_i \in R$ initializes its parameter vector $x_i[0]$ to be the minimizer of its local function $f_i(x)$

At each iteration k, each regular agent $v_i \in R$:

1. **Information Exchange:** Broadcasts its own parameter vector $x_i[k]$ to, and receives the parameter vectors $x_j[k]$ from, its neighbors $v_j \in \mathcal{N}_i$ in the network
2. **Distance Filter:** Removes F parameter vectors that are farthest from the reference point $\hat{x}$
3. **Component-Wise Filter:** Further removes parameter vectors that are extreme in any coordinate (i.e., contain a value in some coordinate that is in the highest F or lowest F values in that coordinate over all parameter vectors remained from the previous step)
4. **Consensus + Gradient Steps:** Takes the average of the remaining parameter vectors, and updates the result by moving along the direction of negative subgradient of the function evaluated at the average:

$$x_i[k+1] = \underbrace{\sum_{v_j\in S_i[k]} \frac{1}{|S_i[k]|} x_j[k]}_{\text{Consensus Step: average of the remaining parameter vectors}} - \underbrace{\eta[k] g_i[k]}_{\text{Gradient Step}}$$

$S_i[k]$: Set of agents whose parameter vectors are not discarded
$\eta[k]$: Pre-determined step-size
$g_i[k]$: A subgradient of f_i evaluated at the average

***Theorem*:**

Suppose the step-size used in gradient descent is diminishing but sums to infinity, and the network has enough redundancy*.

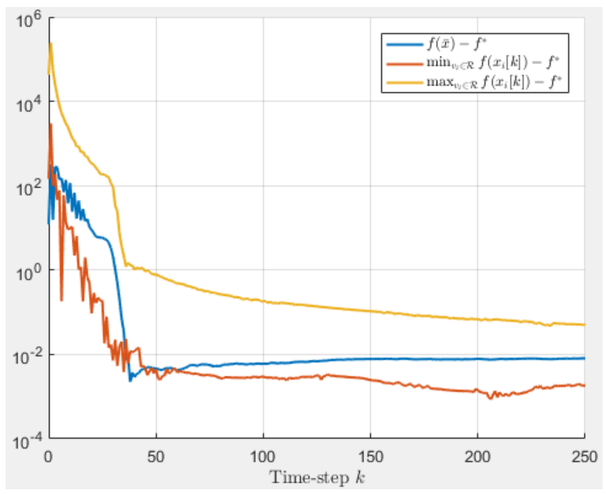
Then, regardless of the actions of any F-local set of Byzantine agents, all regular parameter vectors will asymptotically reach consensus and converge to a ball centered at $\hat{x}$ with a bounded radius*.

*Explicit characterization of amount/type of redundancy and the distance-to-optimality are provided in our paper.

Reference:

K. Kuwarananchareon, L. Xin, and S. Sundaram, "Byzantine-Resilient Distributed Optimization of Multi-Dimensional Functions." Proceedings of the American Control Conference, Denver CO, 2020.

5. Simulation Results



Blue line: Gap between the cost evaluated at the average of regular parameter vectors and the optimal cost
Orange line: Gap between the cost evaluated at the parameter vector that gives lowest value and the optimal cost
Yellow line: Gap between the cost evaluated at the regular parameter vector that gives highest value and the optimal cost