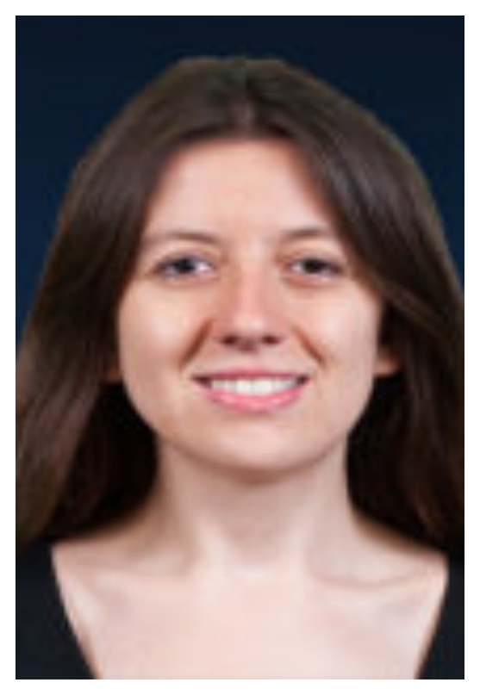
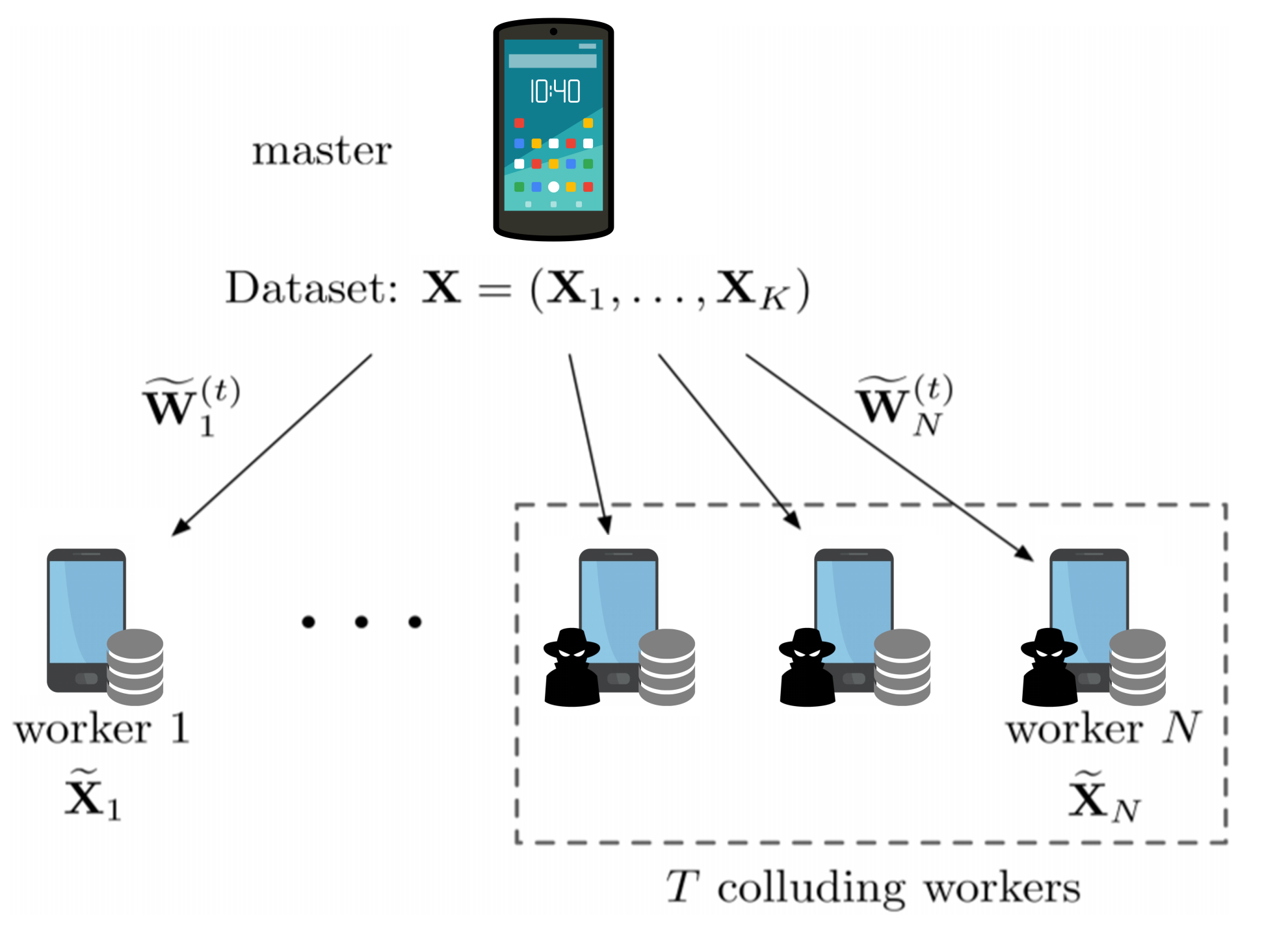
Secure Distributed Coded Computations for IoT : An Information Theoretic and Network Approach

PIs: Yingying Chen[1], Salim El Rouayheb[2], Hulya Seferoglu[3]
[1,2]Rutgers University, [3]University of Illinois at Chicago
[1]http://www.winlab.rutgers.edu/~yychen/, [2]http://eceweb1.rutgers.edu/~csi/, [3]http://nrl.ece.uic.edu/



Background

- The number of mobile IoT devices is increasing and estimated to reach billions in the next few years.
- Data collected by IoT devices will grow at exponential rates.
- By 2022 about 75% of all data will need analysis and action computed by heterogenous networks with varying latency, data volume, bandwidth, cost, data sovereignty and compliance.
- Distributed Computing: Tasks in an IoT device could be offloaded to other connected devices including sensors, mobile devices, and/or servers in close proximity



Challenges

- Heterogeneity, resource time variance, and mobility: addressed in C3P
- Security: addressed in SC^3
- Privacy: addressed in PRAC

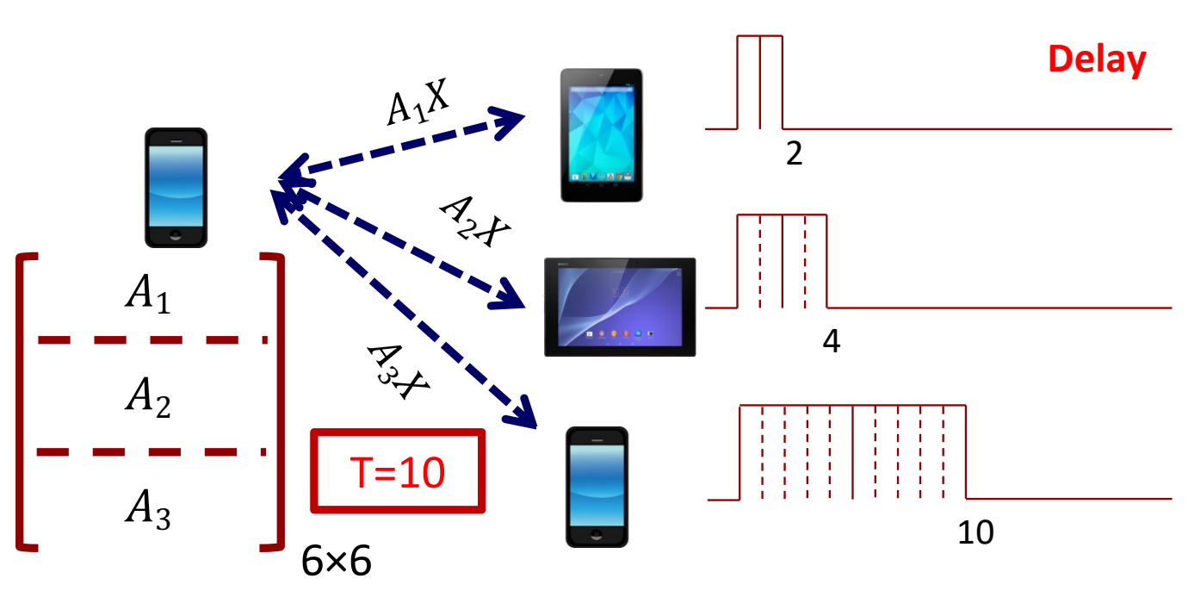
Scientific Impact

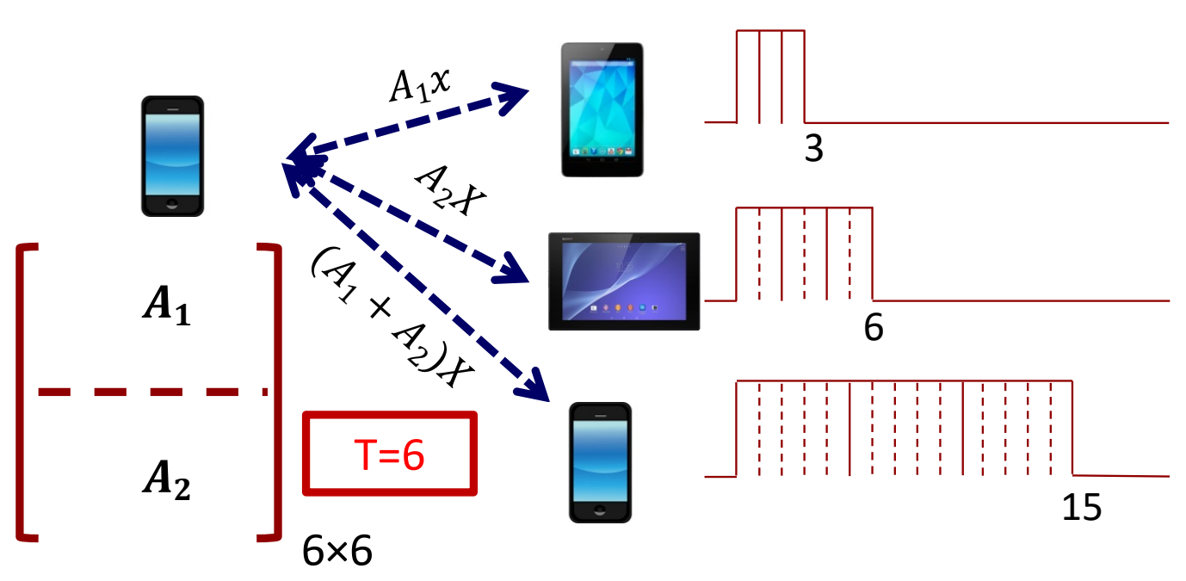
- Secure distributed coded computations are capable of providing information theoretic security across a scalable and heterogenous IoT network

Dynamic Heterogeneity-Aware Coded Cooperative Computation (C3P) at the Edge

Goal: Calculation of matrix multiplication $y = Ax$ using 3 workers

- Trivial Approach:
  - A is divided into 3 submatrices with equal size.
- **Coded Computation:**
  - A is divided into 2 submatrices with equal size.
  - 3 coded tasks are generated from the 2 submatrices
- Advantage of coded computation:
  - Higher reliability
  - Smaller delay
  - Lower communication cost





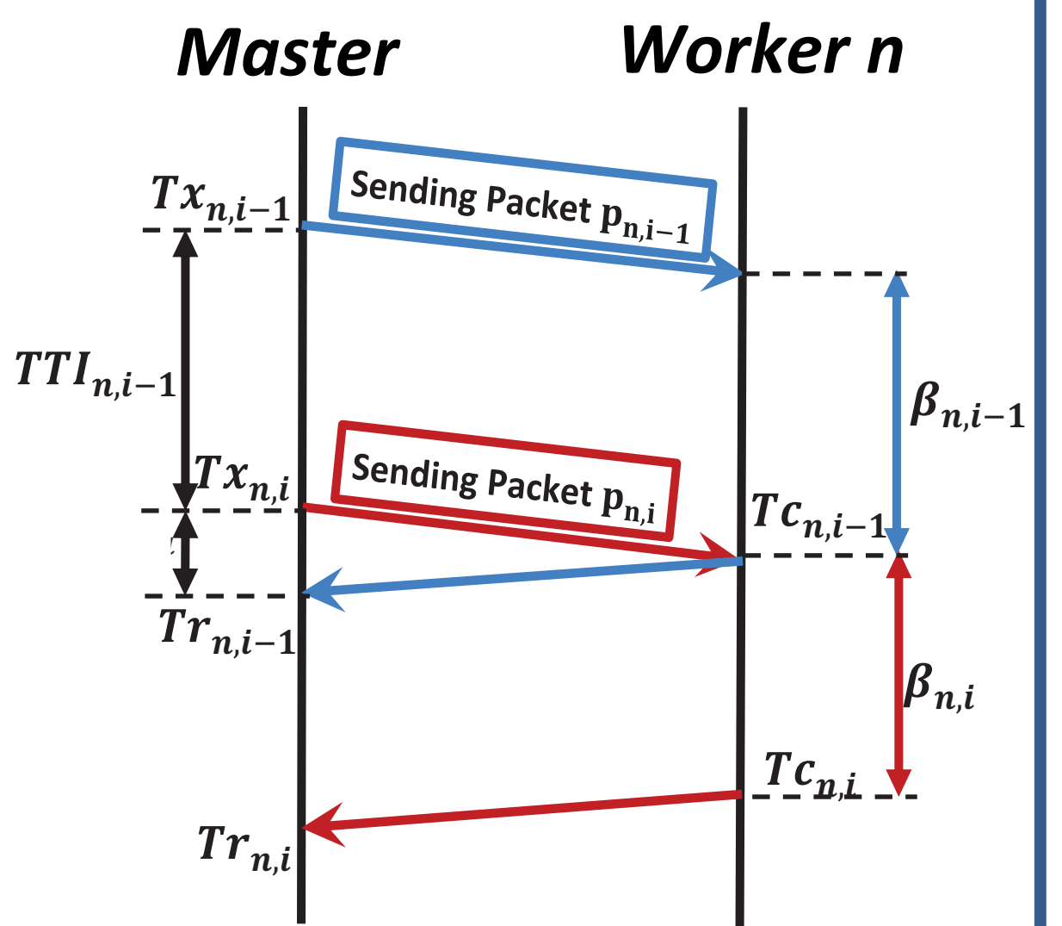
C3P Approach:

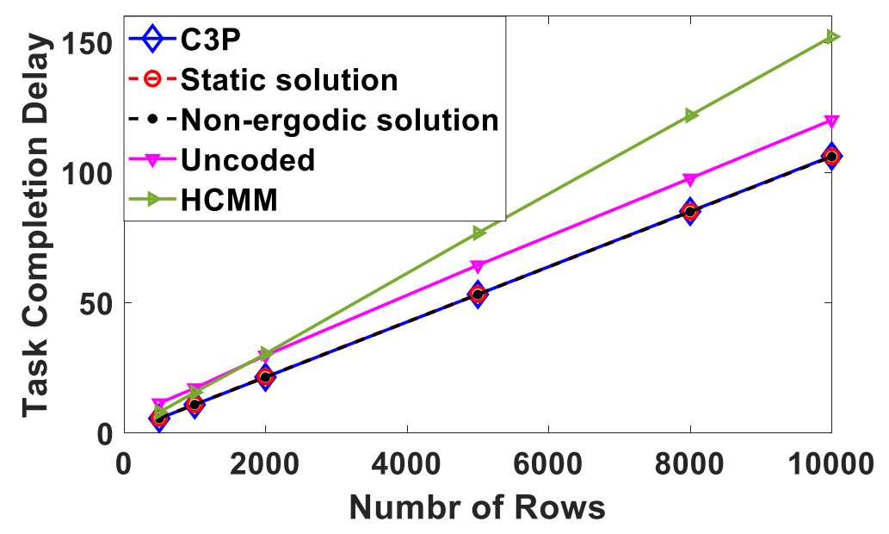
- Inspired by ARQ mechanism, master transmits packets to workers dynamically
- Fountain codes is used due to their rateless property, low encoding and decoding complexity, and low overhead

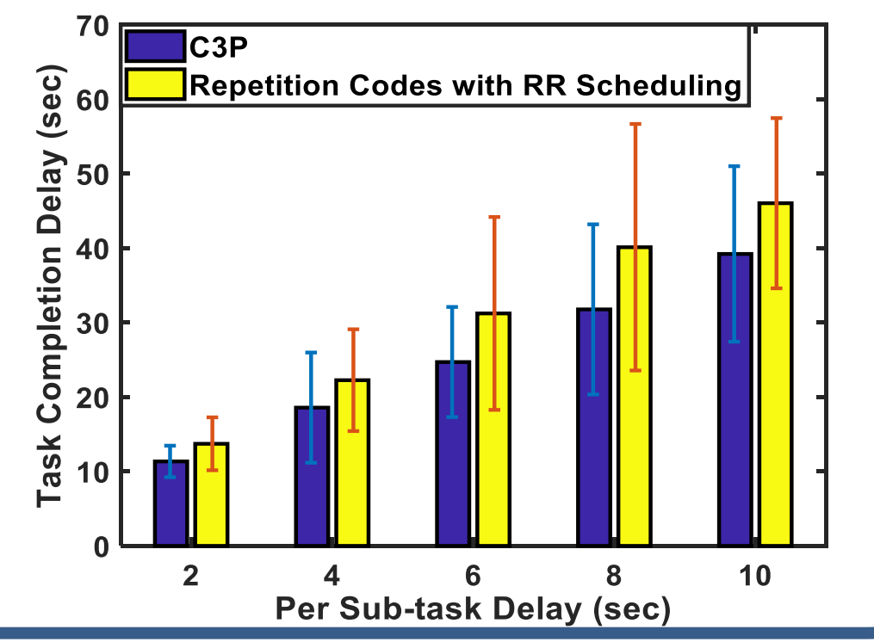
C3P Algorithm:

- Divide matrix A into R rows
- Apply Fountain codes on rows to create packets
- Send packets iteratively with $TTI_{n,i} = \min(\mathbb{E}[\beta_{n,i}], Tr_{n,i} - Tx_{n,i})$ until $R + \epsilon$ computed packets are received at the master collectively from all workers, where ϵ is the coding overhead
- In C3P, $\mathbb{E}[\beta_{n,i}]$ is estimated using runtimes of previously received packets: $\mathbb{E}[\beta_{n,i}] \approx \frac{\sum_{n=1}^{m_n} \beta_{n,i}}{m_n}$

- C3P improves task completion via both simulations and in a testbed consisting of real Android-based smartphones
- The efficiency of C3P in terms of resource utilization is higher than 99% in practice



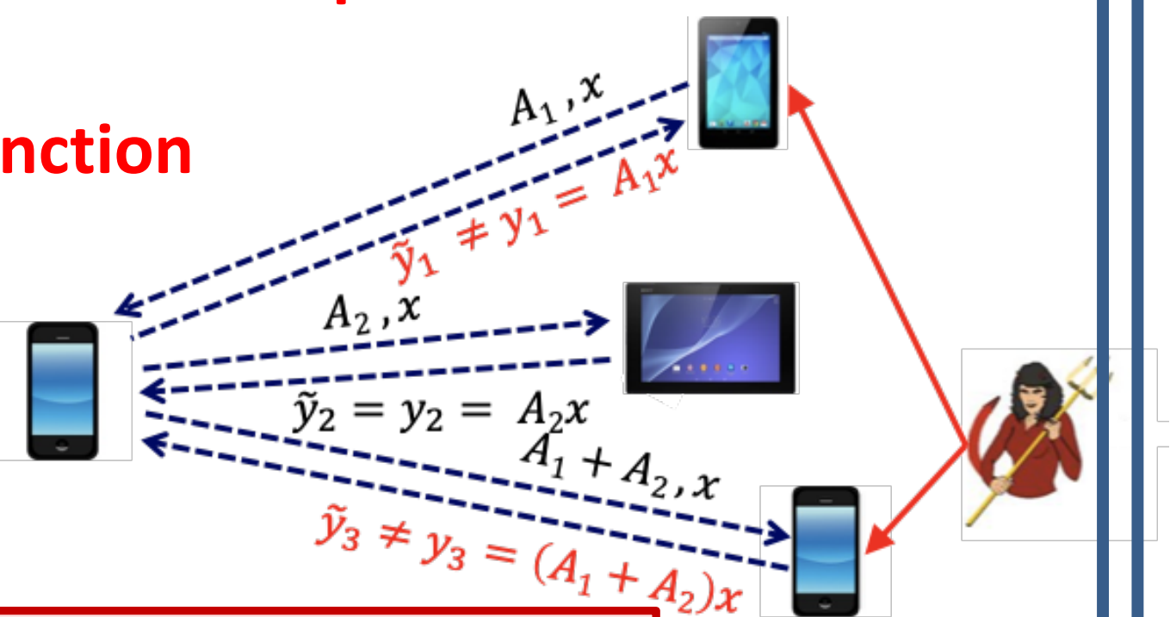




Secure Coded Cooperative Computation (SC^3) at the Heterogeneous Edge against Byzantine Attacks

Problem: Byzantine attacks; workers can corrupt their offloaded tasks

Solution: Use homomorphic hash function to check the integrity of the received results from each worker in a computationally efficient way:



Homomorphic Hash Function: $h(a) \triangleq mod(g^{mod(a,q)}, r)$

g is a number in $\mathbb{F}_r$ s.t. $g = b^{(r-1)/q}$ for a random selection of $b \in \mathbb{F}_r, b \neq 1$

Prime number randomly selected from $\mathbb{F}_\phi$

Prime number s.t. $q|(r-1)$

Theorem: If worker w_n is not malicious, *i.e.*, $\tilde{y}_{n,i} = y_{n,i}, \forall i$, then $\alpha_n = \beta_n$ for a nonzero integer c_i at the master

$$\underbrace{\alpha_n = h\left(\sum_{i=1}^{Z_n} c_i\, \tilde{y}_{n,i}\right)}_{\text{Calculated using the received result from } w_n} \qquad \underbrace{\beta_n = mod\left(\prod_{j=1}^{C} h(x_j)^{mod\left(\sum_{i=1}^{Z_n} c_i p_{n,i,j}, q\right)}, r\right)}_{\text{Calculated using the local info}}$$

Z_n: Number of packets sent to w_n

C: Number of columns of matrix A

$p_{n,i,j}$: jth element of the ith packet sent to w_n

	Light-Weight Integrity Check (LW)	Heavy-Weight Integrity Check (HW)
Property	$c_i \in U\{1, -1\}$	$c_i \in \mathbb{F}_q$
Probability of attack detection	$1 - \left(\frac{\tilde{Z}_n!}{2^{\tilde{Z}_n}\left((\tilde{Z}_n/2)!\right)^2}\right)$	$1 - \frac{1}{q}$
Computation Complexity	$O(CM(r)\log_2 q)$	$O(CZ_nM(\phi))$

Theorem: The attack detection probability of $\log_2 q$-round LW is equal to the attack detection probability of HW. However, the computational complexity of $\log_2 q$ -round LW is lower than HW if the following condition is satisfied:

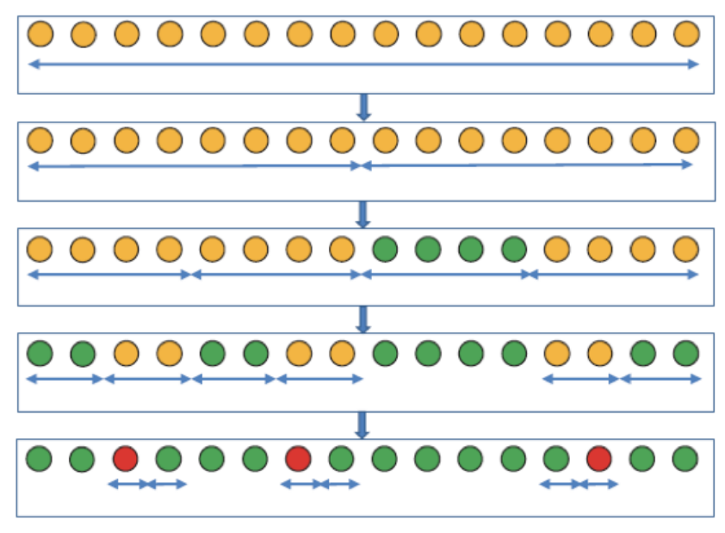
$$Z_n \geq \frac{M(r)}{M(\psi)}(\log_2 q)^2$$

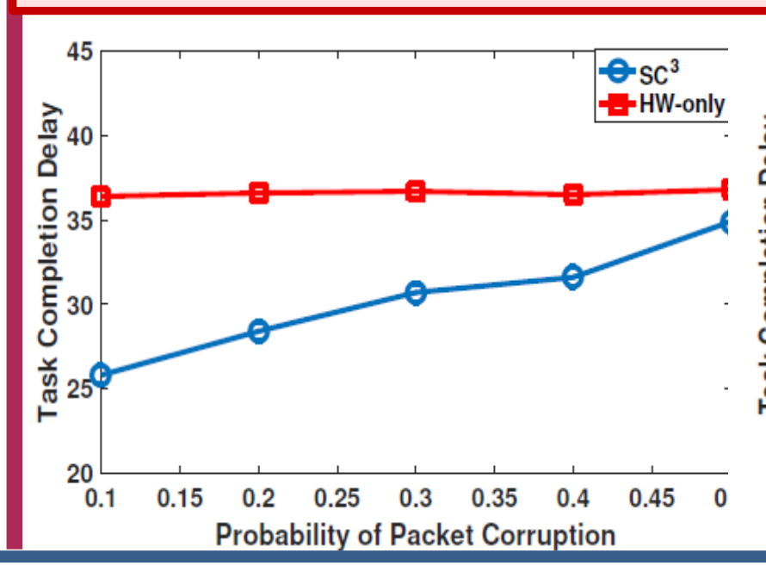
SC^3 Algorithm

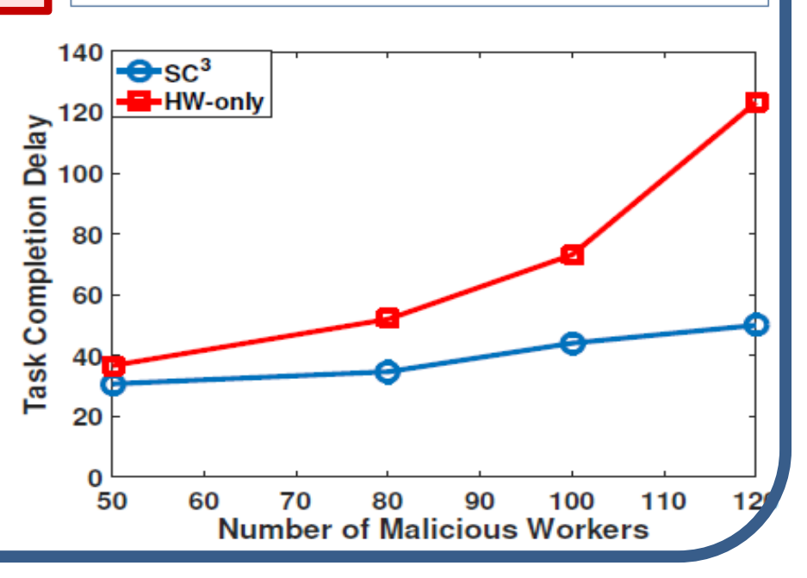
```
1: $V = 0$
2: while $V < R + \epsilon$
3: Determine the time period $T$ as the time interval during which $R + \epsilon - V$ computed packets are received from all workers collectively.
4: for $n = 1: N$ do
5: Create the set $\mathcal{Z}_n$ consisting of packets received from worker $w_n$ during the time period $T$
6: $V_{add} = 0$
7: Apply the attack detection module on $\mathcal{Z}_n$ and set $V_{add}$ as the number of packets labeled as verified.
8: Update $V$ as $V + V_{add}$
9: if $V \geq R + \epsilon$ then
10: Stop the process and use $R + \epsilon$ packets labeled as verified for Fountain decoding.
```

Attack detection module:

- Phase 1: Apply LW and discard all packets of $\mathcal{Z}_n$ if $\alpha_n \neq \beta_n$ and go to phase 2 if $\alpha_n = \beta_n$
- Phase 2 (Efficient attack recovery): Use a binary search algorithm and apply $\log_2 q$-round LW or HW (based on Theorem 2) to detect corrupted packets, where the number of corrupted packets is small

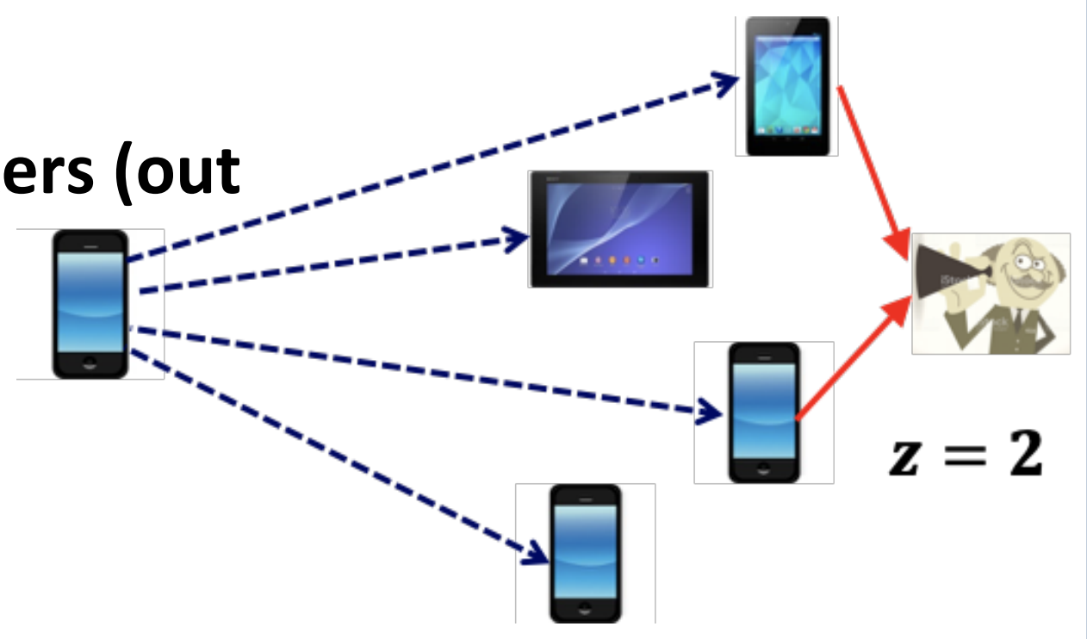






PRAC: Private and Rateless Adaptive Coded Cooperative Computation

Problem: Eavesdropping attacks; z workers (out of n) collude to spy on data A



Solution: Mask data with coded keys and send the masked data along with keys dynamically to the workers:

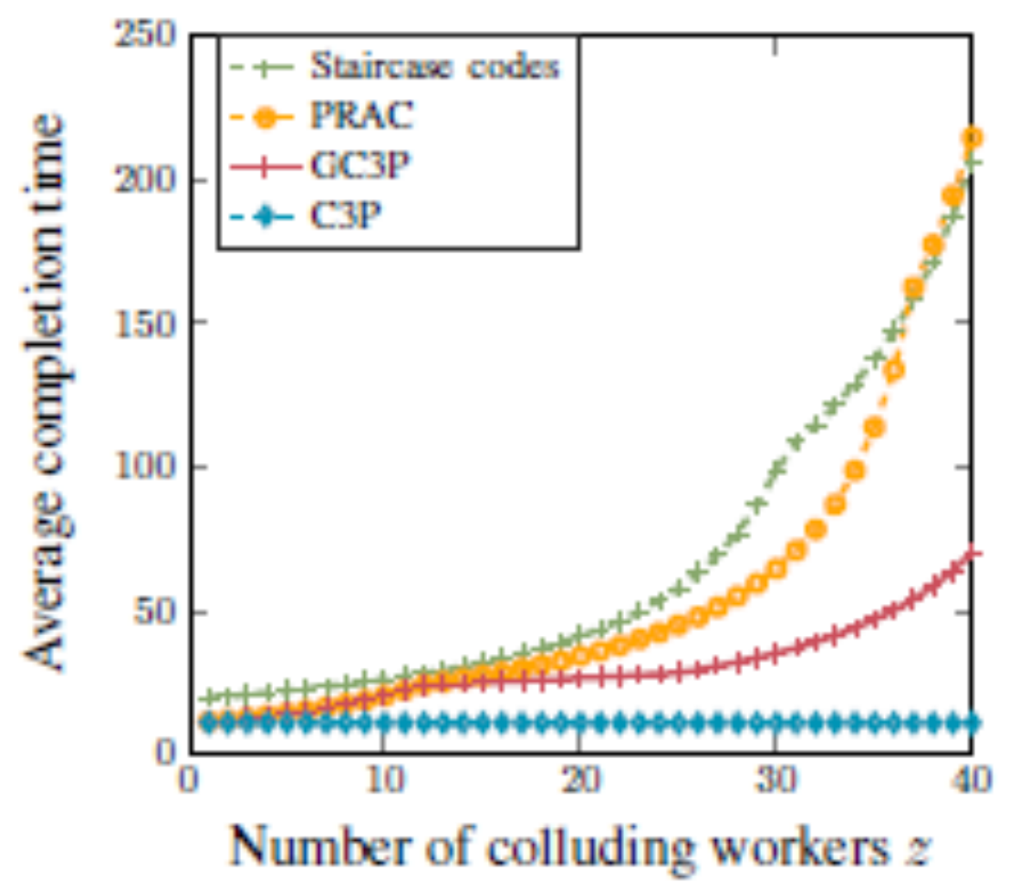
Time	Worker 1	Worker 2	Worker 3	Worker 4
1	$R_{1,2}$	$R_{1,1}$	$A_1 + A_2 + R_{1,1} + R_{1,2}$	$A_3 + A_4 + R_{1,1} + 2R_{1,2}$
2				$R_{2,1}$
3	$R_{2,2}$			
4		$A_1 + A_3 + R_{2,1} + R_{2,2}$	$A_1 + A_5 + R_{2,1} + 2R_{2,2}$	
5	$R_{3,1}$	$R_{3,2}$		$A_3 + A_5 + R_{3,1} + R_{3,,2}$

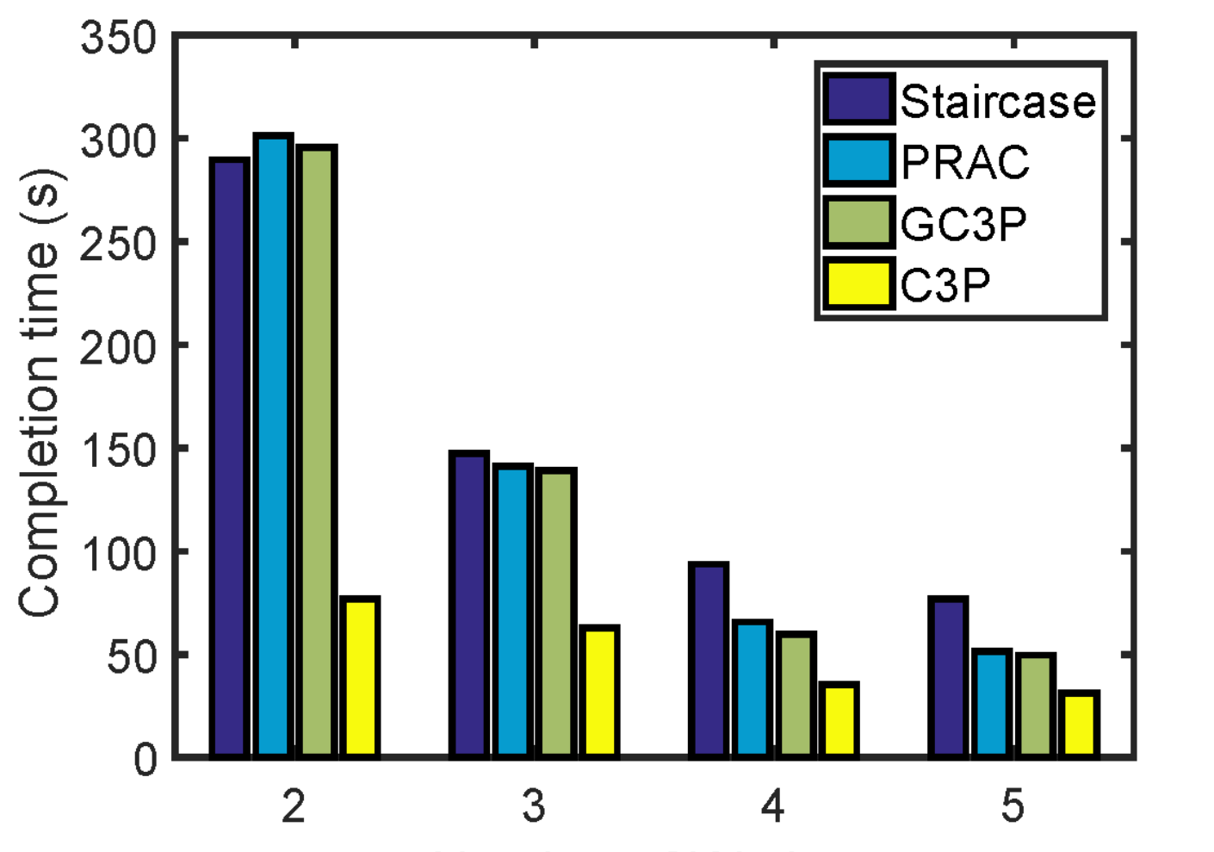
PRAC Outline:

- PRAC uses C3P for offloading packets to workers
- For each round t, the master generates z random keys $R = [R_{t,1}, R_{t,2}, \ldots, R_{t,z}]$
- For each worker n at round t, $R_{t,j}$ is transmitted if $j \leq z$ and a fountain coded packet masked with a coded key (e.g. $A_1 + A_5 + R_{2,1} + 2R_{2,2}$) is transmitted if $j > z$, where j is the number of workers with the current round of t
- For the master to decode one coded computed packet, it should receive the result of computation for the z keys (used for masking

Theorem: PRAC is a rateless real-time adaptive coded computing scheme that satisfies the information theoretic privacy for a given $z < n$ using the minimum required number of keys.

Theorem: PRAC achieves the optimum completion time required for any private distributed linear computation (i.e. speed of the $(z + 1)^{st}$ fastest worker.





Broader Impact – Societal Wellbeing

The obtained results on secure coded computations have the potential to transform the design of next-generation IoT networks where security must be a prime requirement rather than an afterthought

Broader Impact – Education and Outreach

The validation component and the collaboration with AT&T Labs will enable fast technology transfer. The educational plan will include: (i) integration of proposed research in course development; (ii) guided tours for undergraduate students to Winlab at Rutgers with possible summer internships; and (iii) workshop organization

Broader Impact – Quantifiable Impact

The proposed research could facilitate new designs of next-generation IoT networks and applications, ranging from health monitoring, to smart cities and driverless cars, to name a few

The 4th NSF Secure and Trustworthy Cyberspace Principal Investigator Meeting
October 28-29, 2019 I Alexandria, Virginia