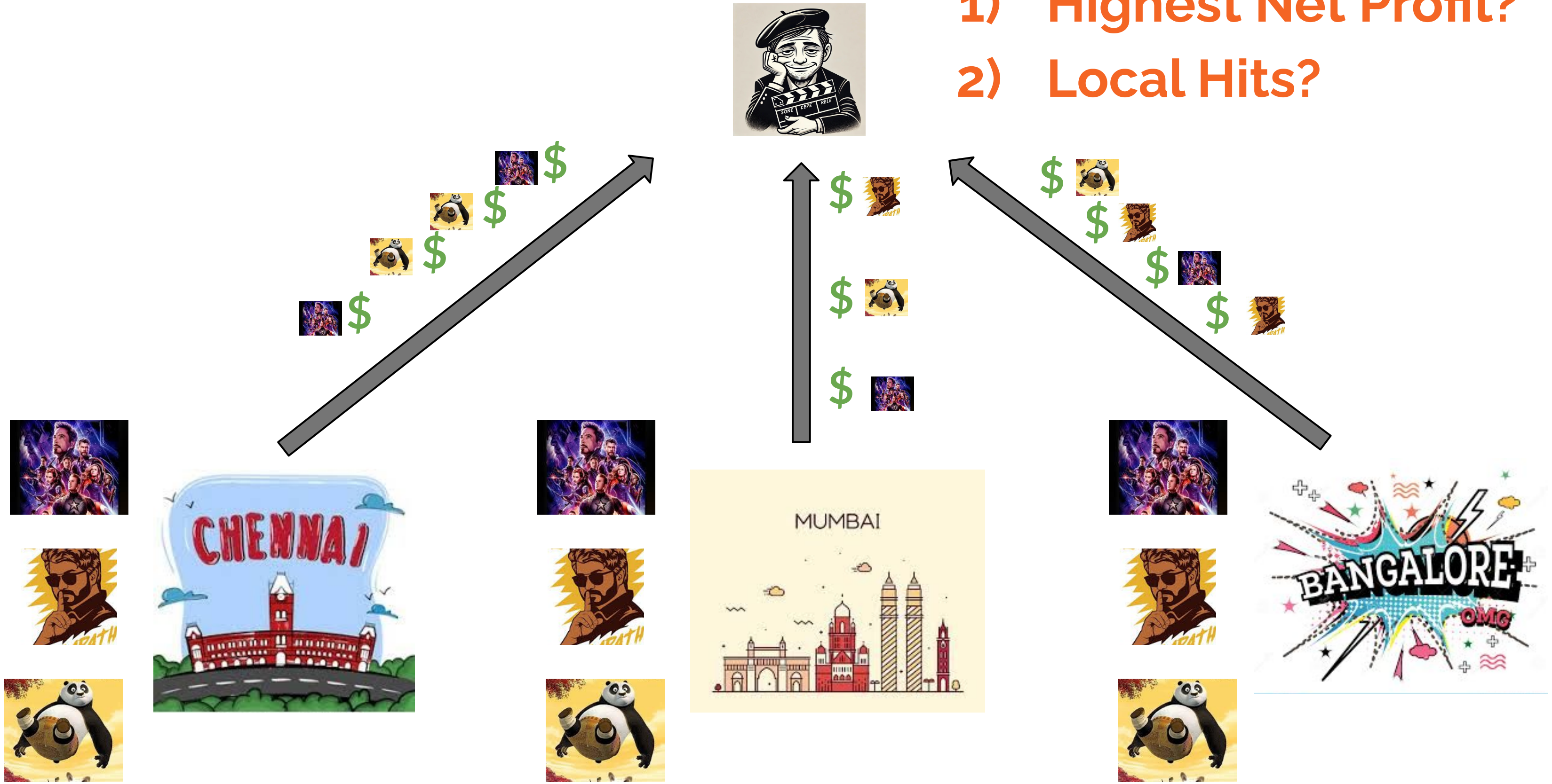
Almost Cost-Free Communication in Federated Best Arm Identification

Kota Srinivas Reddy, P. N. Karthik, and Vincent Y. F. Tan

Presented By:
Riya Mahesh (EE21B112)
Vamsi Krishna Chilakamarri (EE21B153)

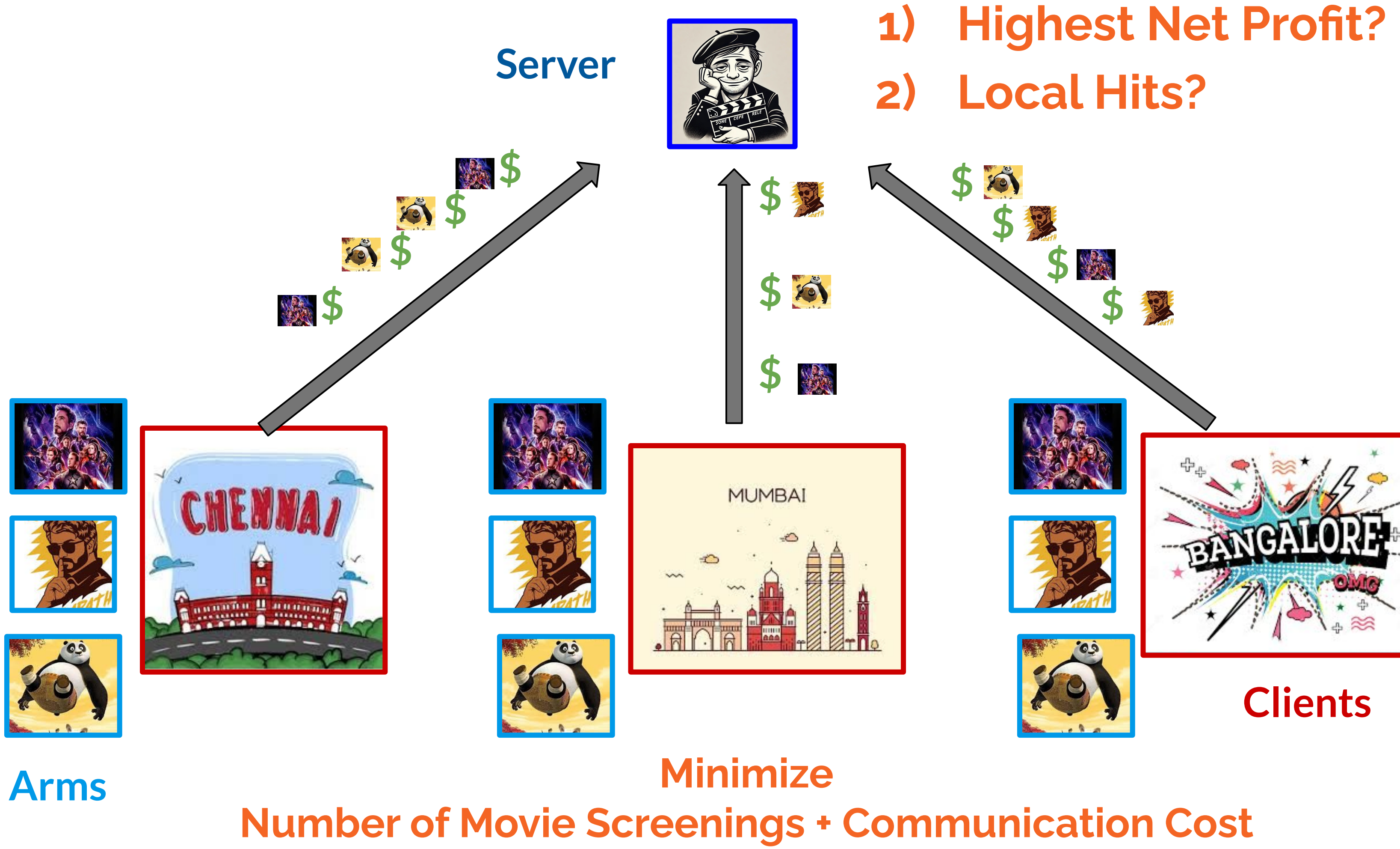
1) Highest Net Profit?

2) Local Hits?

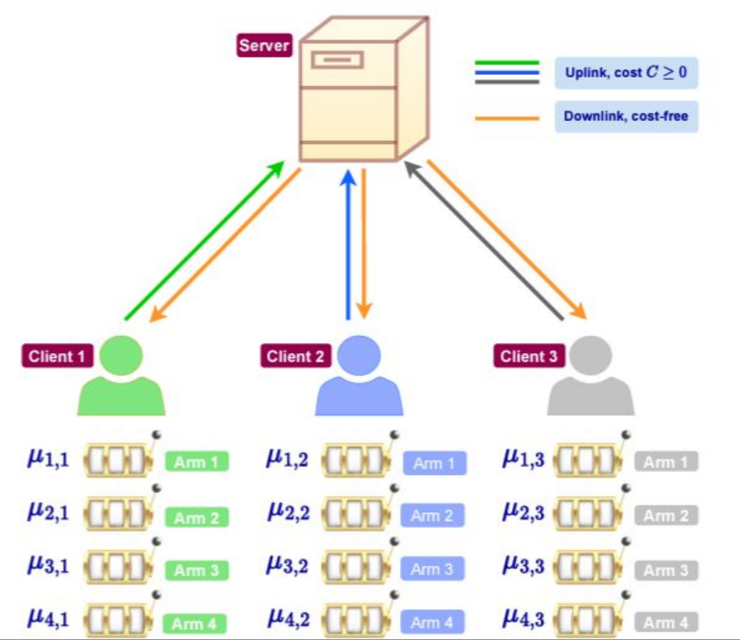


Minimize

Number of Movie Screenings + Communication Cost


1) Highest Net Profit?
2) Local Hits?
Server
CHENNAI
MUMBAI
BANGALORE
OMG
Arms
Clients
Minimize
Number of Movie Screenings + Communication Cost

Problem Definition:



GOAL:
Identify with HIGH confidence

$$S(\mu) := (k_1^*, k_2^*, \ldots, k_M^*, k^*) \in [K]^{m+1}$$

- Central Server
- M Clients: [M] = {1, 2, 3, … , M}
- K arms : [K] = {1,2,3, … , K}
- Rewards: N(mean, 1)
- Uplink Cost: C per usage per link
- Problem Instance:

$$\mu = [\mu_{k,m} : k \in [K], m \in [M]] \in \mathbb{R}^{K \times M}$$

- **Local Best Arm:**

$$k_m^* := \arg\max_k \mu_{k,m}, \quad \text{with mean} \quad \mu_m^* := \mu_{k_m^*,m} = \max_k \mu_{k,m}$$

- Average Arm Mean across Clients:

$$\mu_k := \frac{1}{M} \sum_{m=1}^{M} \mu_{k,m}$$

- **Global Best Arm:**

$$k^* := \arg\max_k \mu_k, \quad \text{with mean} \quad \mu^* := \mu_{k^*} = \max_k \mu_k$$

Successive Elimination of Arms (SEA)

- Initialize Candidate set S = [K]
- Pull every arm in S
- Eliminate every arm satisfying

$$\hat{\mu}_k(n) < \max_{j \in S} \hat{\mu}_j(n) - 2\alpha(n)$$

- Repeat
- Terminate if only 1 element left
- Report Element as your answer

Initialization

Selection Rule

Termination Rule

Declaration Rule

FEDerated learning successive ELIMination algorithm (FEDELIM)

Initialization: Local Active arms: $S_{l,m}$, Global Active arms S_g

Selection Rule: At each client **m**, pull $S_{l,m} \cup S_g$

Eliminate Local arms if $\hat{\mu}_{k,m}(n) < \max_{j \in S_{l,m}} \hat{\mu}_{j,m}(n) - 2\alpha_l(n)$

Eliminate Global arms if $\hat{\mu}_k(n) < \max_{j \in S_g} \hat{\mu}_j(n) - 2\alpha_g(n)$

Declaration Rule: Size of $S_{l,m}$ = 1; Report as Local Best Arm of m

Size of S_g = 1; Report as Global Best Arm

Termination Rule: If all local and global contenders are found

FEDerated learning successive ELIMination algorithm (FEDELIM)

Communication Rule:

- All remaining candidates for Global Best Arm communicated, Every 2^t timesteps (Exponential)
- Exponentially Sparse Communication - Almost Cost Free

Baseline : FEDELIM0

- Every Step Communication assuming 0 uplink Cost

Algorithm 1: Federated Learning Successive Elimination Algorithm (FEDELIM)

Input: $K \in \mathbb{N}$, $M \in \mathbb{N}$, $\delta \in (0,1)$

Output: $(\hat{k}_1^*, \ldots, \hat{k}_M^*, \hat{k}^*) \in [K]^{M+1}$

Initialize: $n = 0$, $\hat{\mu}_{k,m}(n) = 0$ and $\mathcal{S}_{l,m} = [K]$ for all k, m, $\hat{\mu}_k(n) = 0$ and $\mathcal{S}_g = [K]$ for all k, $\texttt{run} = \texttt{true}$

while $\texttt{run} = \texttt{true}$ **do**
- $n \leftarrow n + 1$
- **for** $m = 1 : M$ **do**
  - $\mathcal{S}_m \leftarrow \mathcal{S}_{l,m} \cup \mathcal{S}_g$ // Arms client m selects
  - **if** $|\mathcal{S}_m| > 1$ **then** // Selection rule
    - **for** $k \in \mathcal{S}_m$ **do**
      - pull arm k of client m and update its empirical mean $\hat{\mu}_{k,m}(n)$
  - **if** $|\mathcal{S}_{l,m}| > 1$ **then**
    - Set $\hat{\mu}_{*,m}(n) = \max_{k \in \mathcal{S}_{l,m}} \hat{\mu}_{k,m}(n)$
    - **for** $k \in \mathcal{S}_{l,m}$ *such that* $\hat{\mu}_{*,m}(n) - \hat{\mu}_{k,m}(n) \geq 2\alpha_l(n)$ **do**
      - // Inactive local arms elimination
      - $\mathcal{S}_{l,m} \leftarrow \mathcal{S}_{l,m} \backslash \{k\}$
  - **if** $|\mathcal{S}_{l,m}| = 1$ **then** // Declaration rule
    - Output $\hat{k}_m^* \in \mathcal{S}_{l,m}$
    - $\mathcal{S}_{l,m} \leftarrow \emptyset$
- **if** $|\mathcal{S}_g| > 1$ *and* $n = 2^t$ *for some* $t \in \mathbb{N}_0$ **then** // Communication rule
  - **for** $k \in \mathcal{S}_g$ **do**
    - For each $m \in [M]$, client m sends $\hat{\mu}_{k,m}(n)$ to the server.
    - Set $\hat{\mu}_k(n) = \sum_{m=1}^{M} \hat{\mu}_{k,m}(n)/M$
  - Set $\hat{\mu}_*(n) = \max_{k \in \mathcal{S}_g} \hat{\mu}_k(n)$
  - **for** $k \in \mathcal{S}_g$ *such that* $\hat{\mu}_*(n) - \hat{\mu}_k(n) \geq 2\alpha_g(n)$ **do** // Inactive global arms elimination
    - $\mathcal{S}_g \leftarrow \mathcal{S}_g \backslash \{k\}$
- **if** $|\mathcal{S}_g| = 1$ **then** // Declaration rule
  - Output $\hat{k}^* \in \mathcal{S}_g$
  - $\mathcal{S}_g \leftarrow \emptyset$
- **if** $|\mathcal{S}_m| = 0$ *for all* $m \in [M]$ **then** // Termination rule
  - $\texttt{run} = \texttt{false}$

Initialization

Selection and Local Arm Elimination

Local Arm Declaration

Communication Rule

Global Arm Elimination

Global Arm Declaration

Termination

Contrast with Prior Work:

- Shi, Shen, and Yang(2021):
  - Same Federated Bandits setup - Federated Learning MAB (FLMAB)
  - **Their Goal**: Regret Minimization. **Our Goal**: Pure Exploration

- Mitra, Hassani, and Pappas (2021)
  - Same Federated Bandits setup
  - **Them:** Global Best arm = Highest Mean among Clients
  - **Us:** Global Best arm = Highest Average Mean among Clients
  - **Them:** Use periodic communication (Used as baseline)

- Hillel et al. (2013) and Tao, Zhang, and Zhou (2019)
  - Same reward distribution across all clients.
  - **Goal:** Collaborative Learning

Key Theoretical Results:

Good Event:

$$\mathcal{E} := \bigcap_{n\in\mathbb{N}, k\in[K], m\in[M]} \left\{ \begin{array}{l} |\hat{\mu}_k(n) - \mu_k| \leq \alpha_g(n), \\ |\hat{\mu}_{k,m}(n) - \mu_{k,m}| \leq \alpha_1(n) \end{array} \right\}$$

- Event occurs with confidence 1 - delta
- If Event Happens, Algorithm gives the correct answer

Key Theoretical Results:

Performance Bound:

For a given problem instance, uplink cost C > 0, and delta (0,1) such that $C\ln T_k <= T_k$ for all k in [K], if

$T^C_{FedElim}$ - total number of arm selections

$C^{comm}_{FedElim}$ - the communication cost

$C^{total}_{FedElim}$ - total cost of FEDELIM to identify the local and global best arms

Under **Good Event:**

$$T^C_{\text{FedElim}} \leq \sum_{k=1}^{K}\sum_{m=1}^{M} \max\{T_{k,m}, 2T_k\} \leq 2T$$

$$C^{\text{comm}}_{\text{FedElim}} \leq C \cdot M \cdot \sum_{k=1}^{K} \left\lceil \frac{\ln T_k}{\ln 2} \right\rceil$$

$$C^{\text{total}}_{\text{FedElim}} = T^C_{\text{FedElim}} + C^{\text{comm}}_{\text{FedElim}} \leq 3T$$

T = Max number of arm pulls under FEDELIM0

$$T_{\text{FedElim0}} \leq T := \sum_{k=1}^{K}\sum_{m=1}^{M} \max\{T_{k,m}, T_k\}$$

$$T_{k,m} := 102 \cdot \frac{\ln\left(\frac{64\sqrt{\frac{8KM}{\delta}}}{\Delta^2_{k,m}}\right)}{\Delta^2_{k,m}} + 1, \quad T_k := 102 \cdot \frac{\ln\left(\frac{64\sqrt{\frac{8K}{\delta}}}{M\Delta^2_k}\right)}{M\Delta^2_k} + 1$$

Key Theoretical Results:

- **Intuition behind factor of 2:**
  - Info at $t = T$ conveyed utmost by $t = 2T$
- **C doesn't occur explicitly**
  - Exponentially sparse communication
- **Constant Interval Communication**
  - Higher Communication Cost
  - Bounds depend on Cost. Factor $(1 + C/H)$
- **Super Exponential Communication 2^{2^t}**
  - Higher Number of pulls $O(T^2)$

Simulations: (Recreated Results)

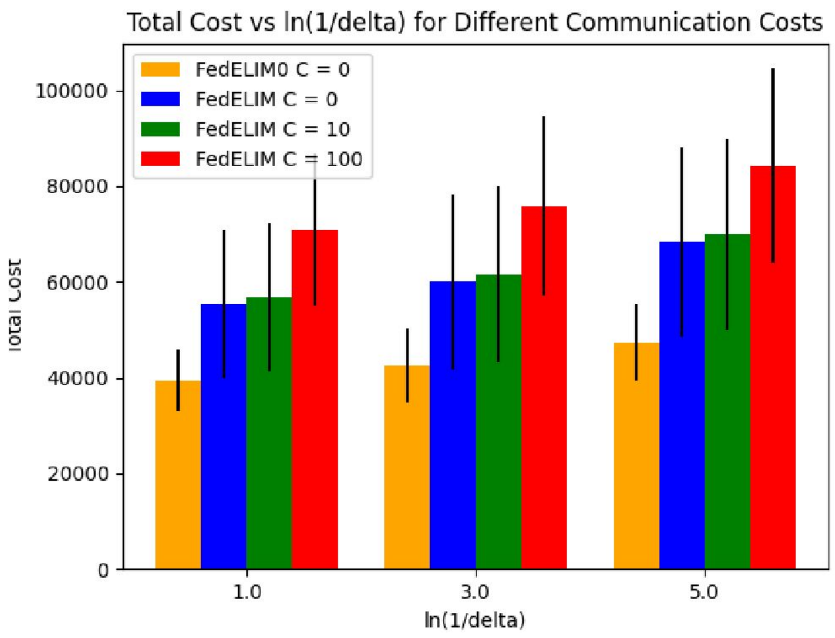
Synthetic Dataset:

$$\mu = \begin{bmatrix} 0.9 & 0.1 & 0.1 \\ 0.1 & 0.9 & 0.1 \\ 0.1 & 0.1 & 0.9 \\ 0.5 & 0.5 & 0.5 \end{bmatrix} \in \mathbb{R}^{4\times 3}$$

- 4 Arms
- 3 Clients
- Gaussian Rewards

Test 1:

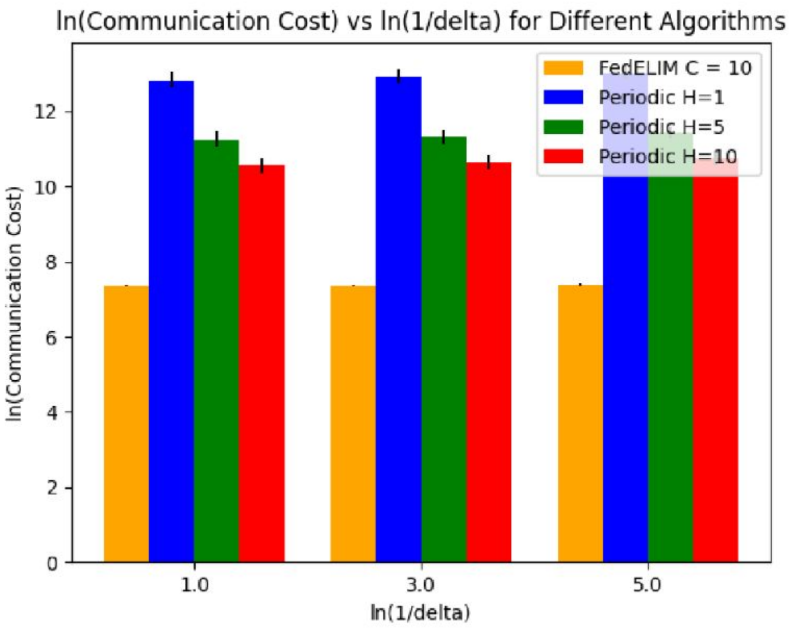
Comparison of Total Cost with FEDELIM0



- FEDELIM0 **always** has lower cost than FEDELIM when C=0
- Total cost of FEDELIM is within theoretical multiplicative bounds

Test 2:

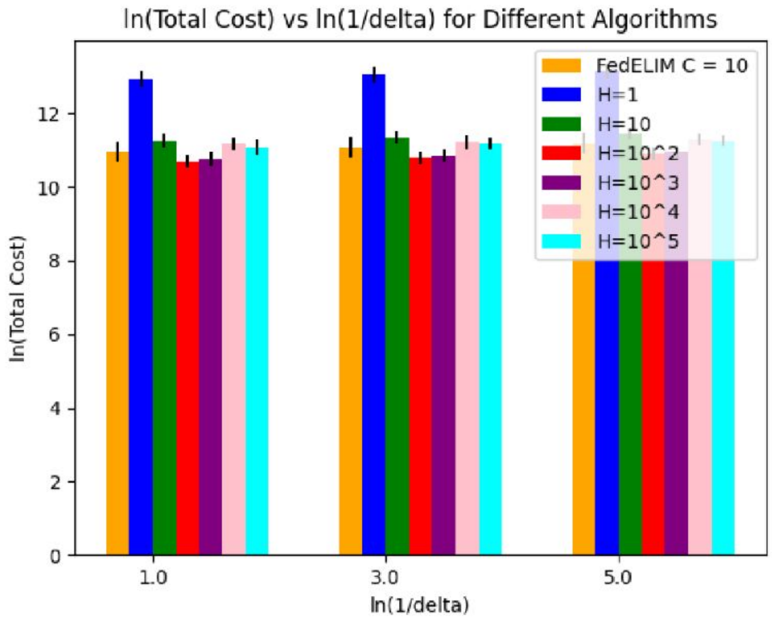
Communication Cost Comparison with Periodic Variant



- Communication cost in FEDELIM is **orders of magnitude** lesser than periodic variant

Test 3:

Total Cost Comparison with Periodic Variant



- Periodic variant is very sensitive to **cost-dependent parameter** H.
- FEDELIM performs consistently well without any tuning

Future Directions:

Hanna, Yang, and Fragouli (2022)

- **Communication Cost:** Function of number of bits

Solve Federated BAI with fewest number of bits