



Multi-Agent Stochastic Bandits Robust to Adversarial Corruptions

Fatemeh Ghaffari¹, Xuchuang Wang¹, Jinhang Zuo², Mohammad Hajiesmaili¹

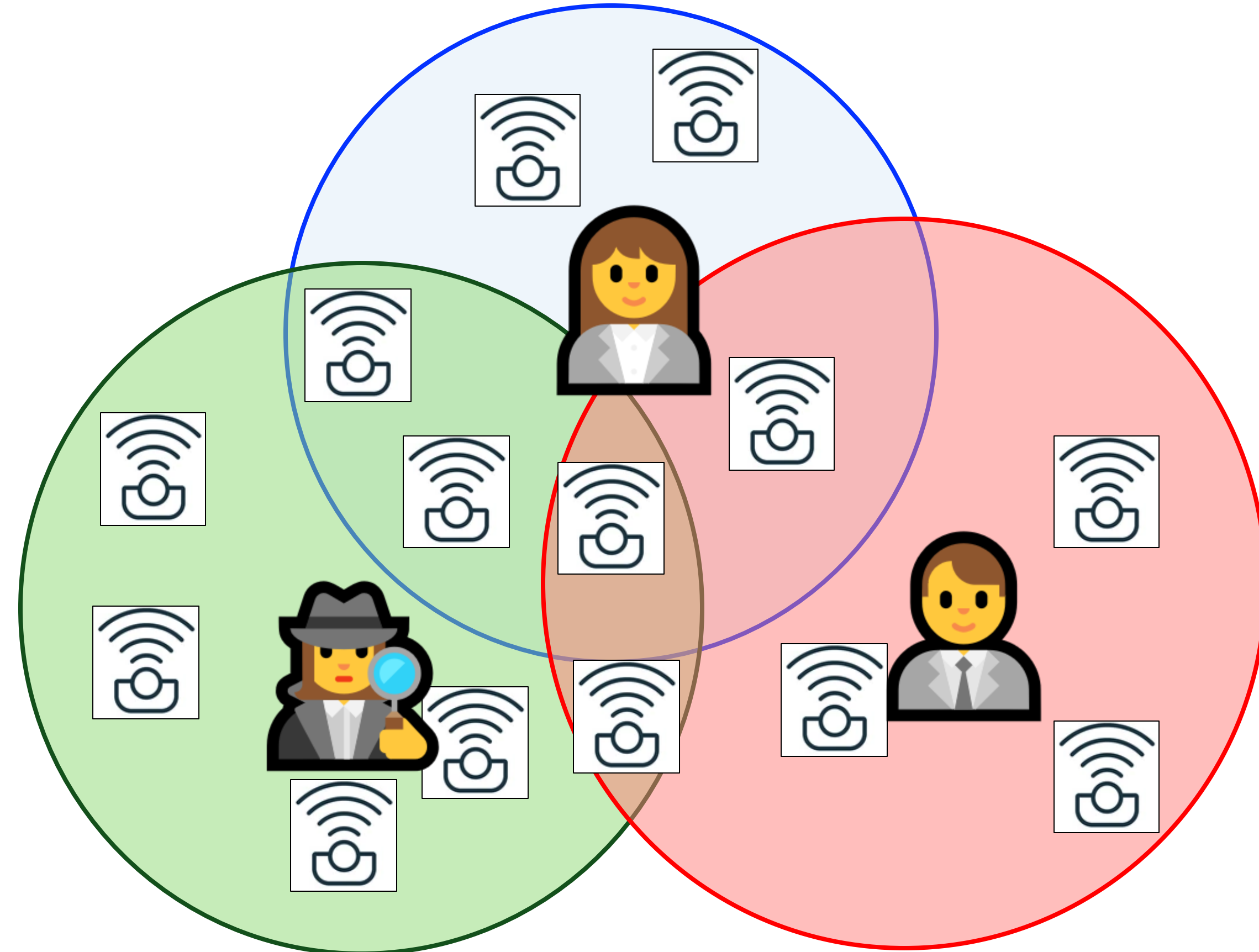


University of
Massachusetts
Amherst

香港城市大學
City University of Hong Kong



Heterogeneous Multi-Agent Bandits



- K arms, each w. reward X_k , **unknown** mean μ_k
- $\Delta_k := \mu_{k^*} - \mu_k$ where $k^* := \arg \max_k \mu_k$
- L agents, with subset arm access
- T sequential decision rounds.
- Regret:

$$R_T := \sum_{\ell \in [L]} T \mu_{k_\ell^*} - \sum_{t=1}^T \mu_{I_{\ell,t}}$$

Adversarial Corruption Procedure

Corruption Budget: $C := \sum_{t=1}^T |X_{k,t} - \tilde{X}_{k,t}|$

Learner pull arm I_t

Stochastic reward $X_{k,t}$ drawn from all arms

The **adversary** chooses the corrupted reward $\tilde{X}_{k,t}$ for all arms

Learner observes the corrupted reward $\tilde{X}_{I_t,t}$

DRAA Algorithm Design

For each epoch m (with $O(4^m)$ rounds) & For each agent $\ell \in [L]$

P1: Construct Active Arm Set and Calculate Probabilities

Active arms \mathcal{A}_ℓ^m : potentially good.

Pulling Prob. $p_{k,\ell}^m \leftarrow \begin{cases} \text{small} & \text{for } k \notin \mathcal{A}_\ell^m \\ \text{uniform} & \text{for } k \in \mathcal{A}_\ell^m \end{cases}$

P2: Pull Arm & Comm.

Pick arms to pull according to $p_{k,\ell}^m$.

P3: Reward Estimate

$(p_{k,\ell}^m)^{-1}$ -weighted normalization for low-bias estimate

Theoretical Analysis

Hetero MA2B: **DRAA** achieves $R_T \leq \frac{L}{L_{\min}} C + \frac{K}{\Delta_{\min}} \log T \cdot \log \log T$

Homo MA2B:

Active Arm Elim. [Liu et. al., 2021]

$$R_T \leq LC + \frac{K}{\Delta_{\min}} \log T \cdot \log \log T$$

DRAA for Homo MA2B

$$R_T \leq C + \frac{K}{\Delta_{\min}} \log T \cdot \log \log T$$

Basic MAB:

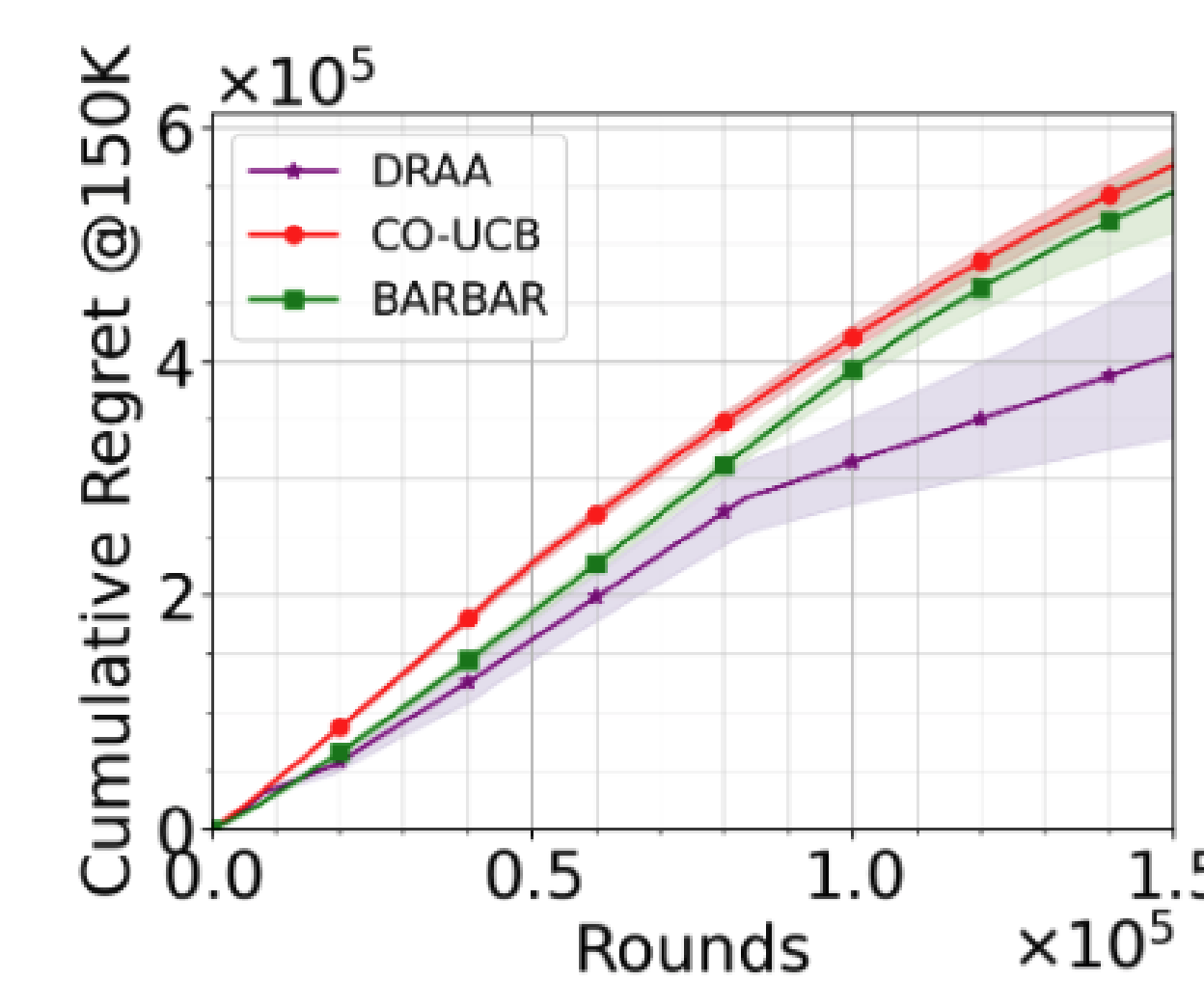
BARBAR [Gupta et. al., 2019]

$$R_T \leq KC + \frac{K}{\Delta_{\min}} \log T \cdot \log \log T$$

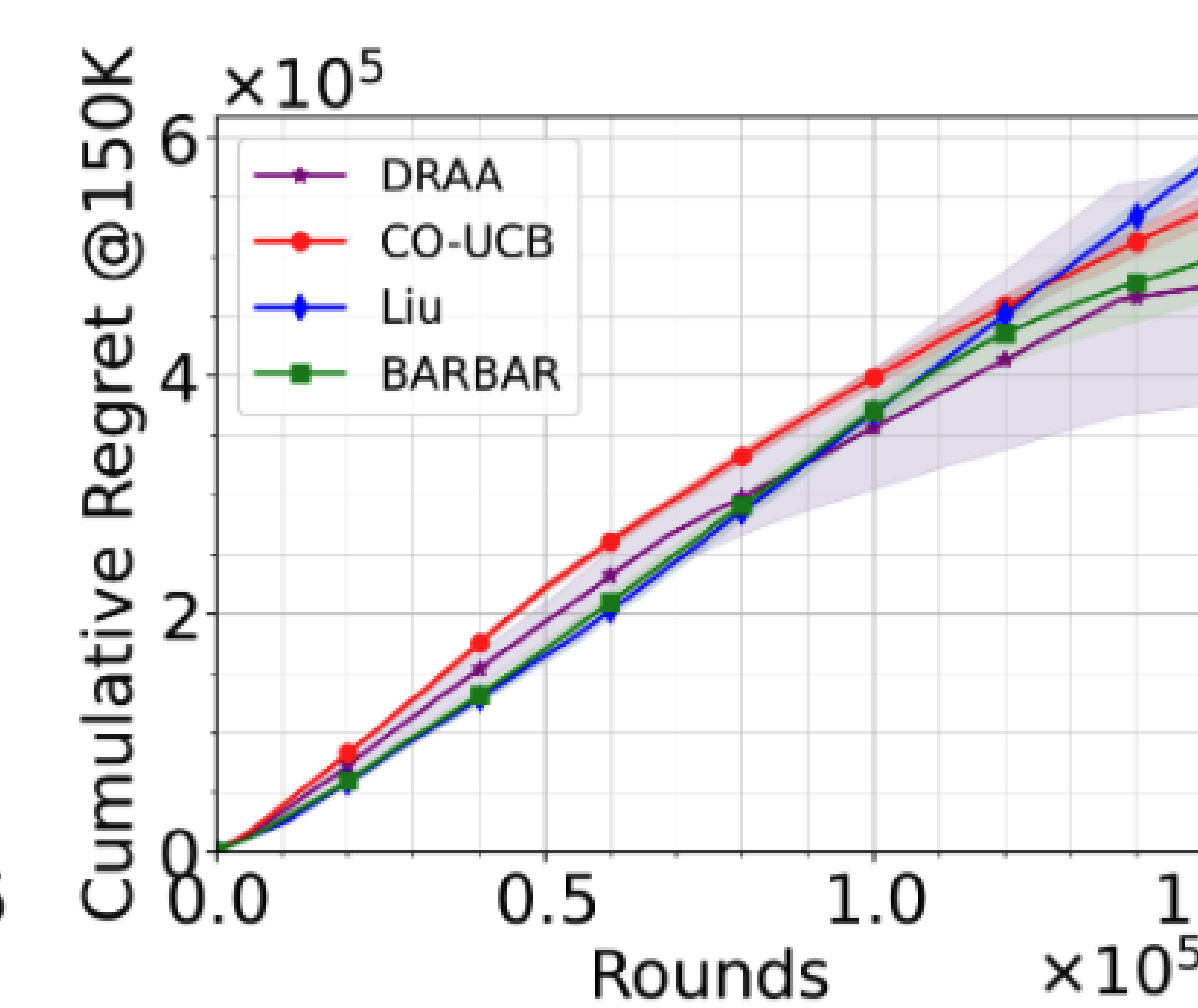
DRAA for Stochastic Bandits

$$R_T \leq C + \frac{K}{\Delta_{\min}} \log T \cdot \log \log T$$

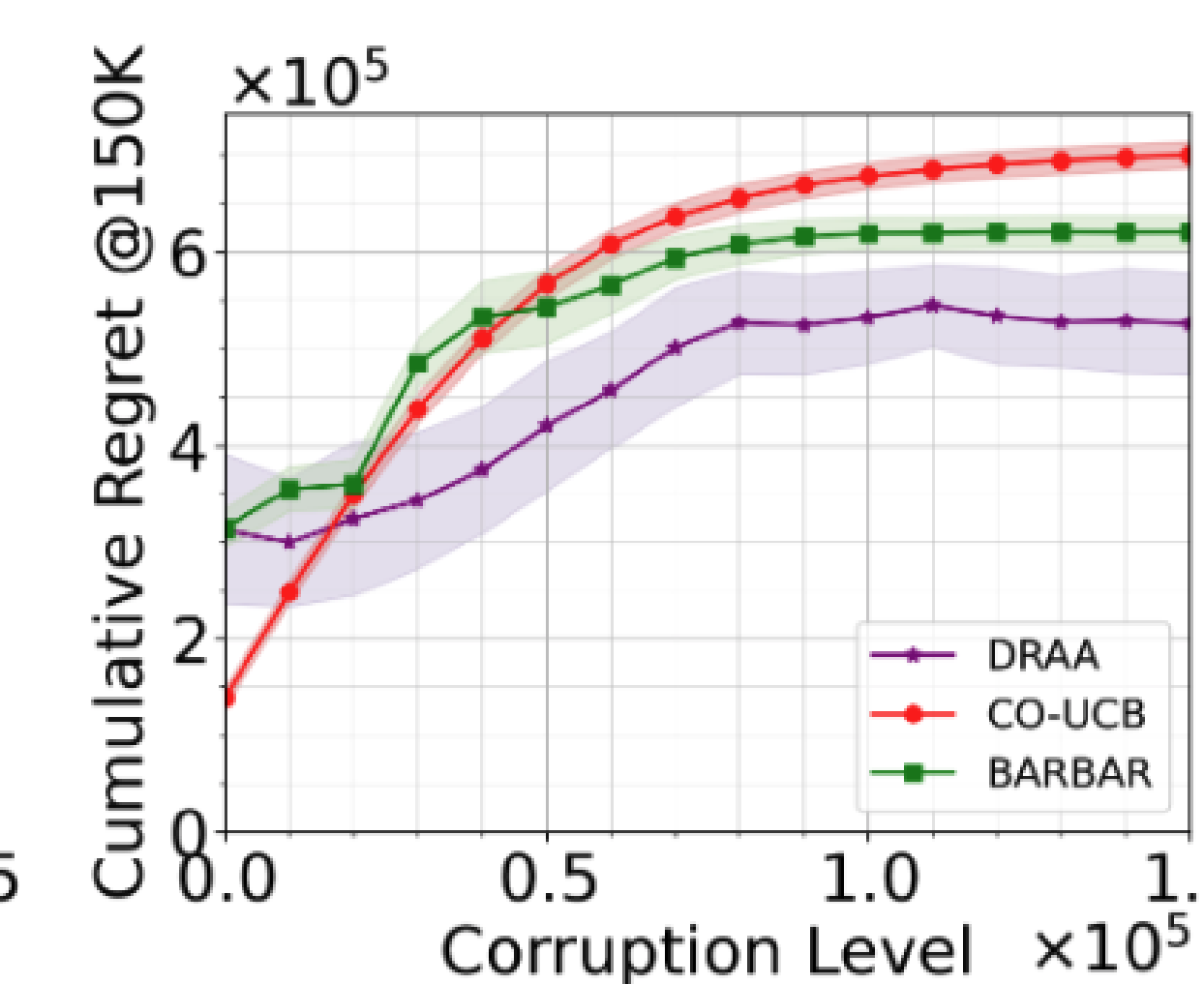
Empirical Regret Comparison



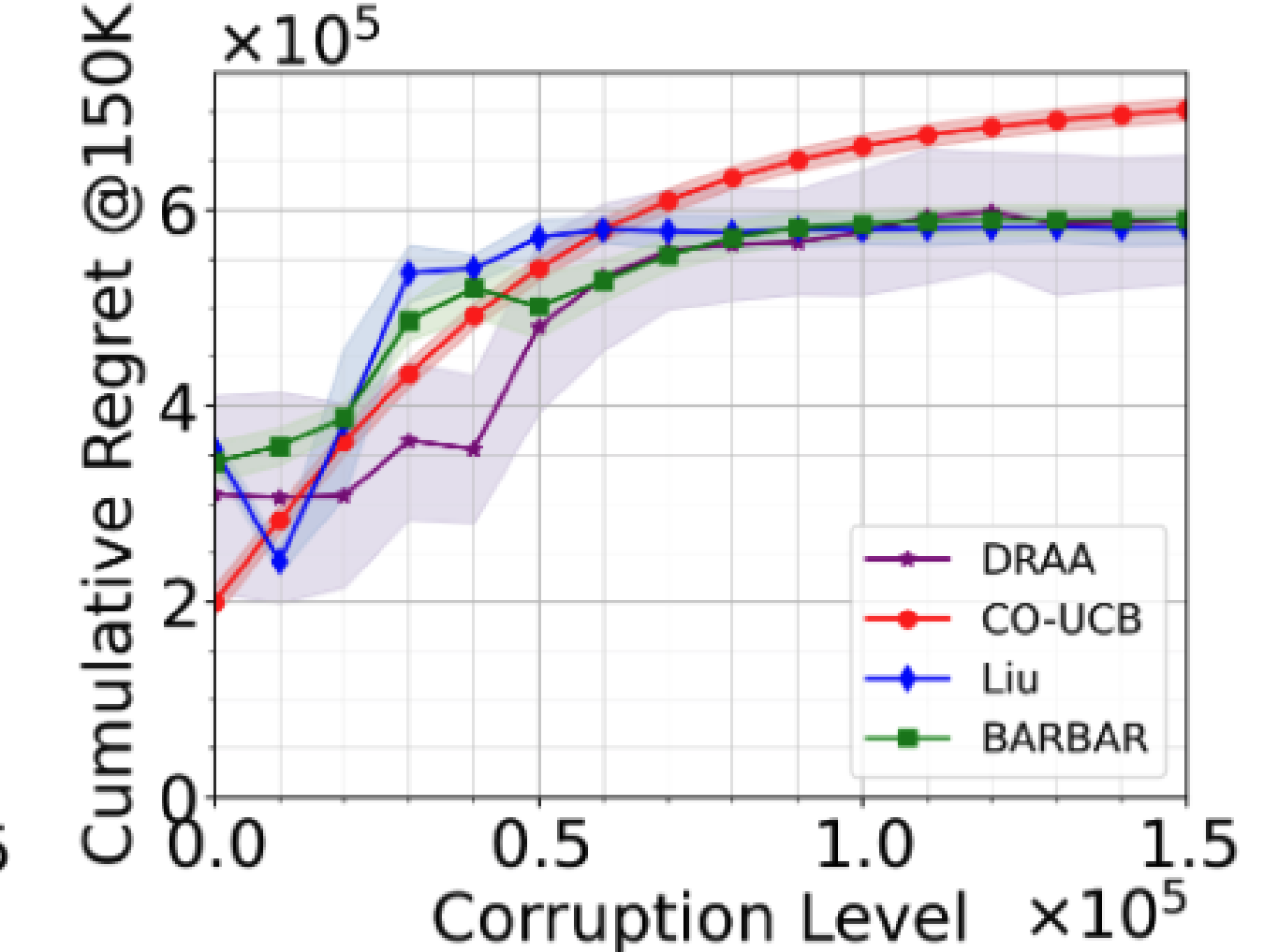
(a) Regret vs. Rounds
Heterogeneous



(b) Regret vs. Rounds
Homogeneous



(c) Vary Corruption Rate
Heterogeneous



(d) Vary Corruption Rate
Homogeneous