Stochastic Bandits Robust to Adversarial Attacks

Xuchuang Wang[1], Maoli Liu[2], Jinhang Zuo[3],

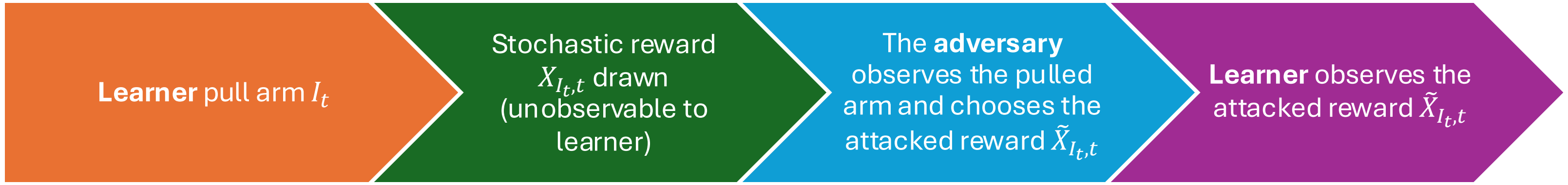
Xutong Liu[4], John C.S. Lui[2], Mohammad Hajiesmaili[1]

[1] University of Massachusetts, Amherst, [2] Chinese University of Hong Kong

[3] City University of Hong Kong, [4] Carnegie Mellon University

Model: Bandits with Adversarial Attacks

- K arms, each with a stochastic reward X_k with unknown mean μ_k
  - $\Delta_k := \mu_{k^*} - \mu_k$ where $k^* = \arg\max_k \mu_k$
- T decision rounds. Regret: $R_T := T\mu_{k^*} - \sum_{t=1}^{T} \mu_{I_t}$
- Total attack budget: $C := \sum_{t=1}^{T} \left| X_{k,t} - \tilde{X}_{k,t} \right|$



Attack Procedure

Corruption Procedure

See our paper for a detailed separation of these two models.

Algorithms for Known Attack Budget

Known Attack Budget

SE-WR		SE-WR-Stop
$O\left(\sum_{k\neq k^*} \frac{\log T}{\Delta_k} + KC\right)$	Stop Cond. ①	$O(\sqrt{KT}(\log T + C))$

<u>SE−WR</u>: Successive Elimination with Wide Confidence Radius

Standard CR: $\mu_k \in \left(\hat{\mu}_k - \sqrt{\frac{\log\frac{2KT}{\delta}}{N_k}}, \hat{\mu}_k + \sqrt{\frac{\log\frac{2KT}{\delta}}{N_k}}\right)$

Wide CR: $\mu_k \in \left(\hat{\mu}_k - \sqrt{\frac{\log\frac{2KT}{\delta}}{N_k}} - \frac{C}{N_k}, \hat{\mu}_k + \sqrt{\frac{\log\frac{2KT}{\delta}}{N_k}} + \frac{C}{N_k}\right)$

Improved regret analysis upon Lykouris et al. (2018, Theorem 1)

Stop Cond. ①

<u>SE−WR−Stop: SE−WE with Stop Condition</u>

Option a: Stop when $N_k \leq \frac{T}{K} + C\sqrt{\frac{T}{K\log\frac{KT}{\delta}}}$

$$R_T \leq O\left(\sqrt{KT\log KT} + KC\right)$$

Option b: Step when $N_k \leq \frac{T}{K}$

$$R_T \leq O\left(\sqrt{KT}(\log KT + C)\right)$$

Algorithms for Unknown Attack Budget

Known Attack Budget

SE−WR

$O\left(\sum_{k \neq k^*} \frac{\log T}{\Delta_k} + KC\right)$

Stop Cond.

①

SE−WR−Stop

$O(\sqrt{KT(\log T + C)})$

② Multi-Phase

③ Model Selection

PE−WR

$\tilde{O}\left(\sqrt{KT} + KC^2\right)$

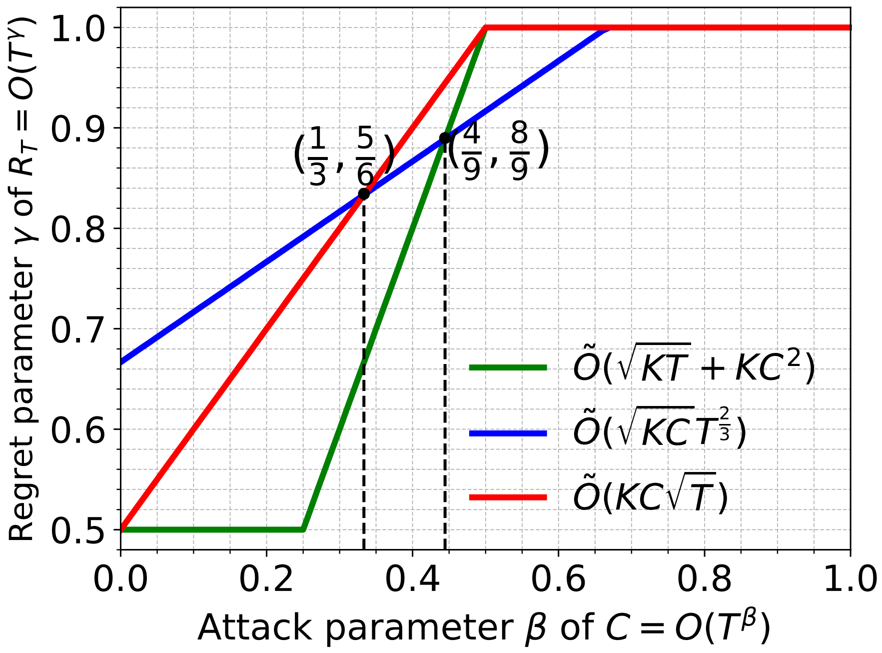
MS−SE−WR

$\tilde{O}(\sqrt{KC}T^{\frac{2}{3}})$, $\tilde{O}(KC\sqrt{T})$

Unknown Attack Budget

② An additive regret bound is obtained by phased-based elimination with a wide confidence interval (PE−WR).

③ Two multiplicative bounds are obtained via a model selection technique from Pacchiano et al. (2020, Theorem 5.3).



Additive vs. Multiplicative Bounds

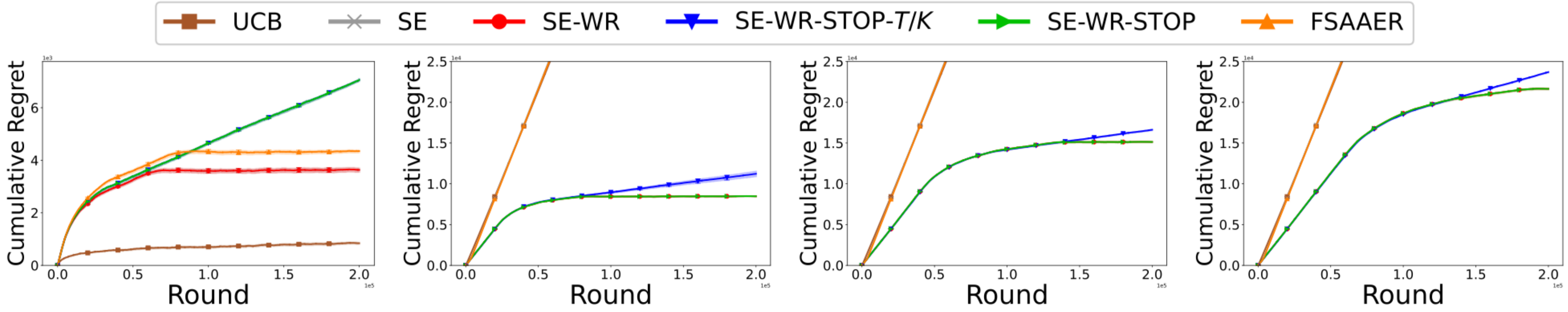
Results: Regret Lower Bounds

- For known attack budget C
  - General: $\Omega(KC)$
  - Gap-dependent: $\Omega\left(\sum_k \frac{\log T}{\Delta_k} + KC\right)$
  - Gap-independent: $\Omega\left(\sqrt{KT} + KC\right)$

- For unknown attack budget C
  - Additive bound: $\Omega\left(T^{\alpha} + C^{\frac{1}{\alpha}}\right)$
  - Multiplicative bound: $\Omega\left(C^{\frac{1}{\alpha}-1}T^{\alpha}\right)$

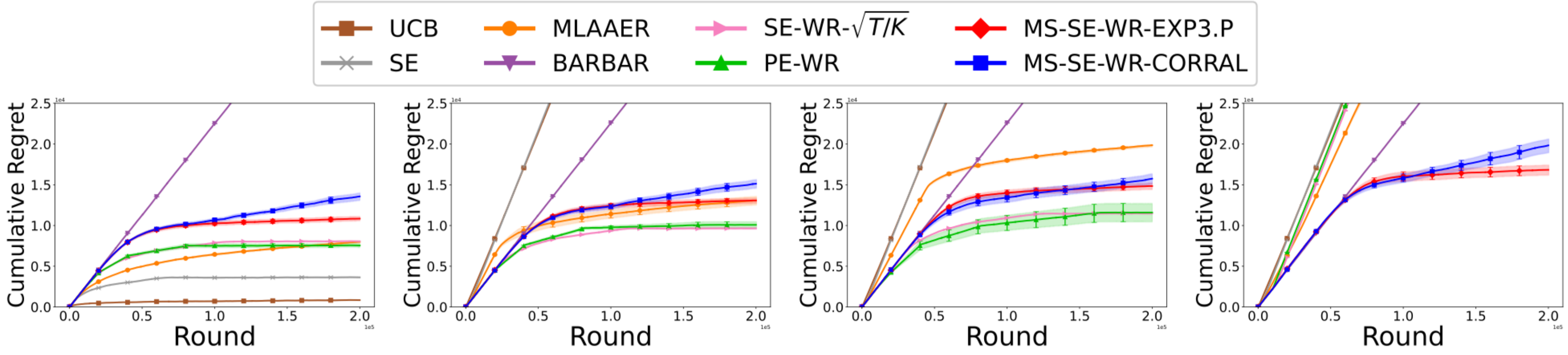
All upper bounds are tight to some logarithmic factors.

Simulations



(a) $C = 0$ (b) $C = 3000$ (c) $C = 6000$ (d) $C = 9000$

Figure 3: Regret comparison of algorithms with *known* attack budgets when varying budget C



(a) $C = 0$ (b) $C = 3000$ (c) $C = 6000$ (d) $C = 9000$

Figure 4: Regret comparison of algorithms with *unknown* attack budgets when varying budget C