Adversarial Attacks on Combinatorial Multi-Armed Bandits

Rishab Balasubramanian
Oregon State University

Jiawei Li
University of Illinois - Urbana-Champaign

Prasad Tadepalli
Oregon State University

Huazheng Wang
Oregon State University

Qingyun Wu
Pennsylvania State University

Haoyu Zhao
Princeton University

Main Contribution

We study adversarial attack of combinatorial multi-armed bandits (CMAB) and obtain the following results:

1. We define a new notion "polynomial attackability" that take the non-asymptotic behavior of attackability into account.
2. In the white-box adversarial attack setting, we define the notion of "gap", which fully characterize the attackabilty of the CMAB instance.
3. In the black-box adversarial attack setting, we show a hardness result that even if the CMAB instance is attackable in the white-box setting, it can still be unattackable in the black-box setting. This is the first result that shows a gap between the while-box and black-box attackability.
4. Experimental results also corroborate our theory.

Preliminary

1. ***Combinatorial Multi-Armed Bandits (CMAB):*** Select k base arms (= 1 super arm, $\mathcal{S}^{(t)}$) that give the cumulative maximal reward from a set of size m .

2. ***Probabilistically Triggered Arms:*** Subset of base arms $\tau_t \subseteq [m]$ that are triggered when super arm $\mathcal{S}^{(t)}$ is selected.

3. ***Reward:*** *The base arm get the realization* $\boldsymbol{X}^{(t)}$ The player observes the realization of base arms $\tau_t \subseteq [m]$ and receives reward $R(\mathcal{S}^{(t)}, X^{(t)}, \tau_t)$

4. ***Reward Poisoning Attack:*** Adversary modifies the base arm realizations to $\tilde{X}^{(t)}$, aiming to fool the algorithm to pull a spacific target arm . The total cost of the attack is computed as $C(T) = \sum_{t=1}^{T} ||\tilde{X}^{(t)} - X^{(t)}||$

	Cascading Bandit	Online Shortest Path / Minimum Spanning Tree	Influence Maximization
Base Arm	Individual Items	Each edge in the graph	Each node in the graph
Super Arm	Permutation of a subset of items	A set of edges that form a valid path/ spanning tree	Initially activated set of seed nodes
Reward poisoning - Attacker Modifies:	Reward feedback of the base arms played	Length/Weight of edges in the graph	Weight (Triggering probability) of edge between nodes

Polynomial Attackability

Attackability: A bandit instance is attackable if there is a strategy to manipulate the realization of arms, using $o(T)$ cost and fool the bandit algorithm to pull the target arm for $1 - o(T)$ times.

- The attackability only considers the asymptotic behavior
- For MAB, there exist efficient attack strategy
- For CMAB, because the number of super arm can be exponential, using naïve attack strategy will lead to the cost exponentially large (although still satisfy the condition)

Polynomial Attackability: A CMAB instance is **polynomially attackable** with respect to a set of super arms $\mathcal{M}$ if for any learning algorithm with polynomial regret there exists an attack method that uses at most $T^{1-\gamma'}$ attack cost and fool the algorithm to pull $\mathcal{S} \in \mathcal{M}$ for $T - T^{1-\gamma'}$ times with high probability for any $T \geq T^*$, where $\gamma, \gamma' > 0$, and T^* polynomially depends on $m, 1/p^*, T, K$.

Attackability of CMAB

Gap: For each super arm $\mathcal{S}$, we define the gap as:

$$\Delta_{\mathcal{S}} := r_{\mathcal{S}}(\mu) - \max_{\mathcal{S}' \neq \mathcal{S}} r_{\mathcal{S}'}(\mu \odot \mathcal{S})$$

where $\mu \in \mathcal{R}^m$ denote the mean of base arms, and $\mathcal{S} \in \{0,1\}^m$ is the all the arms activated by super arm $\mathcal{S}$. $\odot$ denote element-wise product.

Condition for attackability (main result): Given a CMAB instance and a target set of super arms $\mathcal{M}$. If $\Delta_{\mathcal{M}} > 0$ the instance is polynomially attackable, and when $\Delta_{\mathcal{M}} < 0$, the instance is polynomially unattackable.

```
Algorithm 1 Attack algorithm for CMAB instance
Require: Target arm set $\mathcal{M}$ such that $\Delta_{\mathcal{M}} > 0$, CMAB
    algorithm Alg.
1: Find a super arm $\mathcal{S} \in \mathcal{M}$ such that $\Delta_{\mathcal{S}} > 0$.
2: for $t = 1, 2, \ldots, T$ do
3:     Alg returns super arm $\mathcal{S}^{(t)}$.
4:     Adversary returns $\tilde{X}_i^t = 0$ for all $i \in \tau^{(t)} \setminus \mathcal{O}_{\mathcal{S}}$ and
       keep other outcome $X_i^t$ unchanged.
5: end for
```

Hardness Result in Black-box Setting

For MAB:

- Always attackable in white-box and black-box setting.

For linear bandit:

- If an instance is attackable in white-box setting, it is also attackable in black-box setting.

For CMAB:

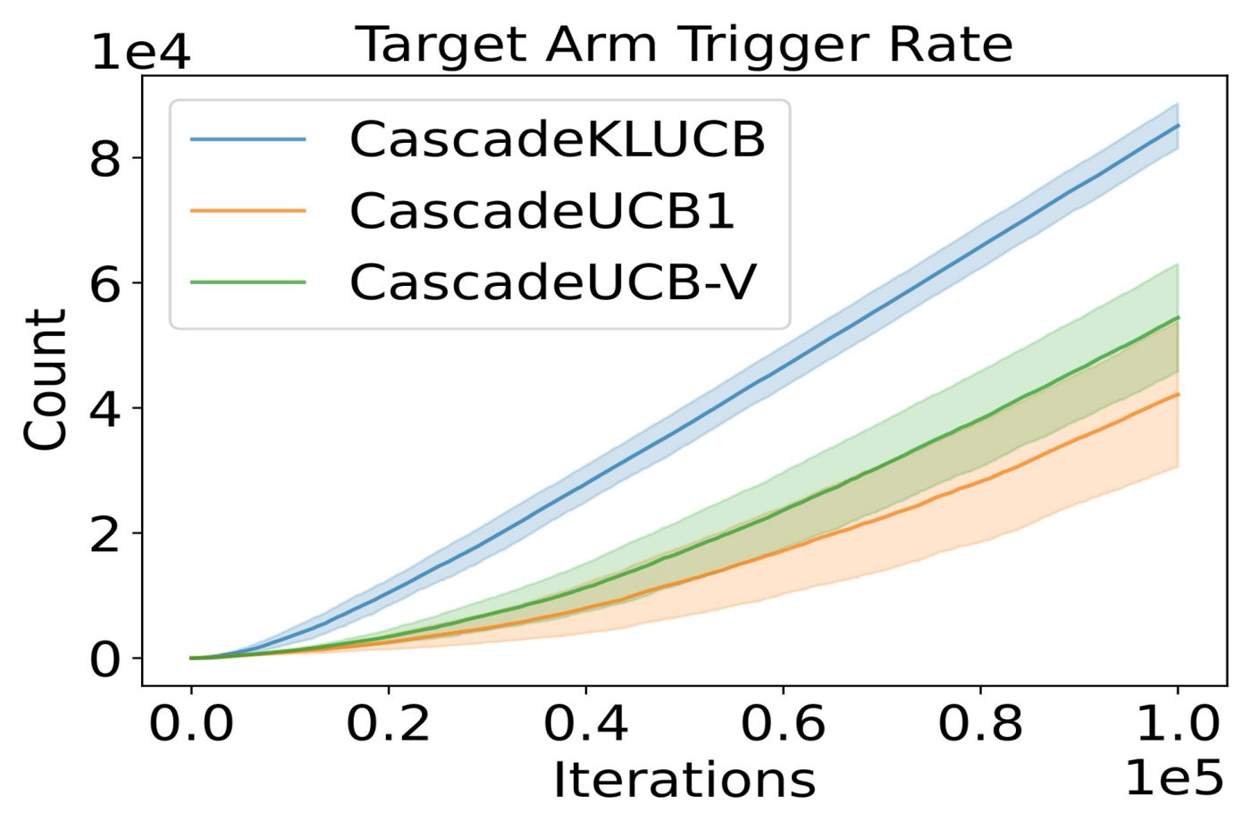
- There exist instances that are polynomially attackable in while-box setting, but not polynomially attackable in the black-box setting.
- First result indicating a gap between white-box and black-box settings.

Empirical Results

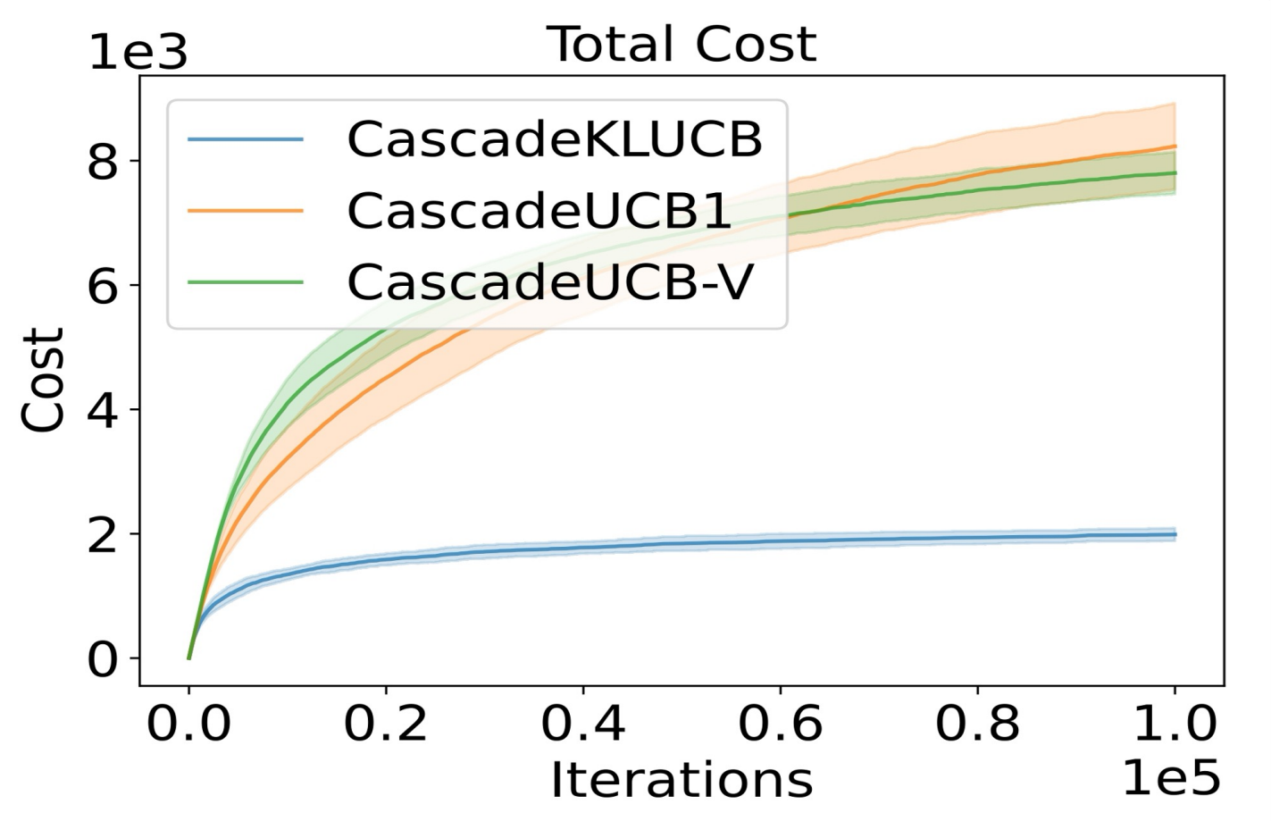
We test our attack on 4 different CMAB applications (Cascading Bandits, Probabilistic Max Coverage, Online Shortest Path, and Online Minimum Spanning Tree. Below we show the results for our attack on Cascading Bandits and Online Shortest Path.

- Cascading bandit: always attackable (as proved in the paper)
- Online shortest path: there exist unattackable target (specifically chosen target), whose attack cost grows linear in T

Cascading Bandits

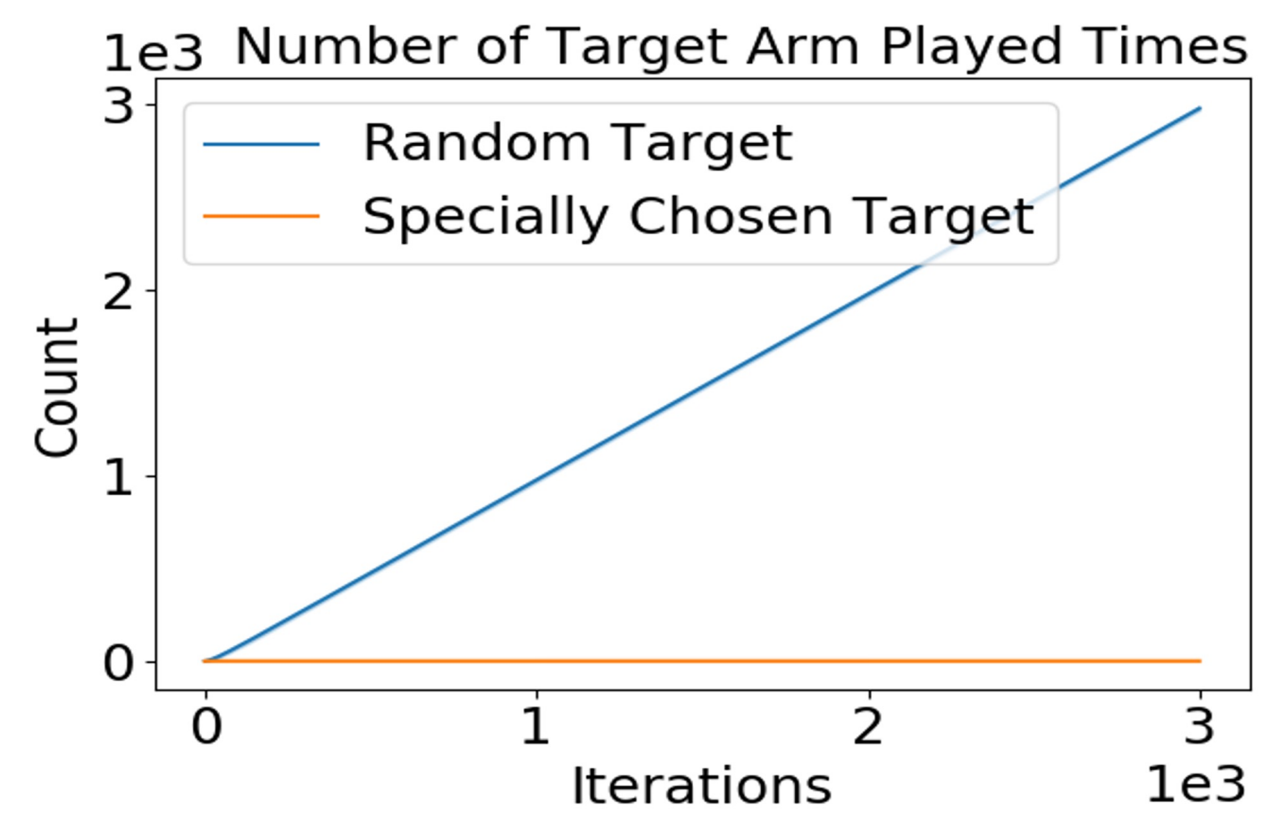


Target Arm Trigger Rate

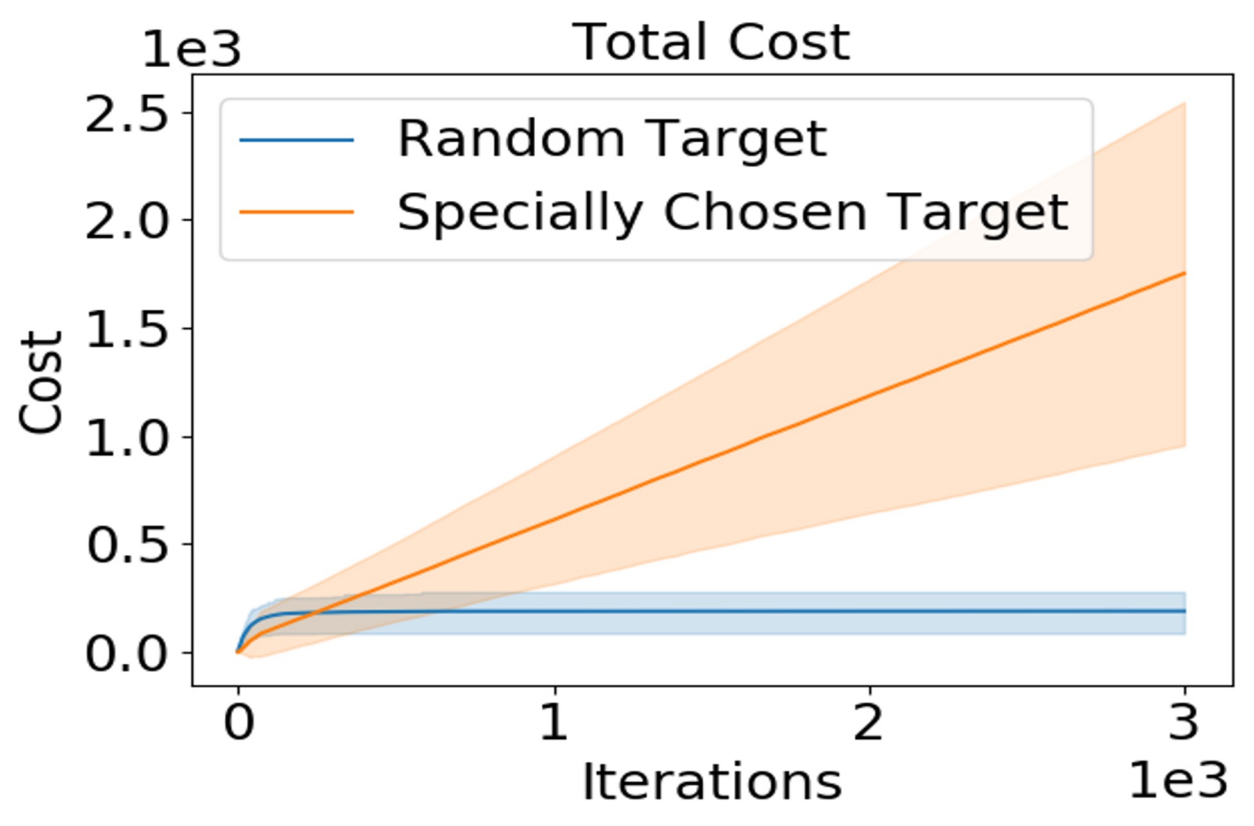


Total Attack Cost

Online Shortest Path



Target Arm Trigger Rate



Total Attack Cost