Ranked Prioritization of Groups in Combinatorial Bandit Allocation

Lily Xu[1], Arpita Biswas[1], Fei Fang[2], and Milind Tambe[1]

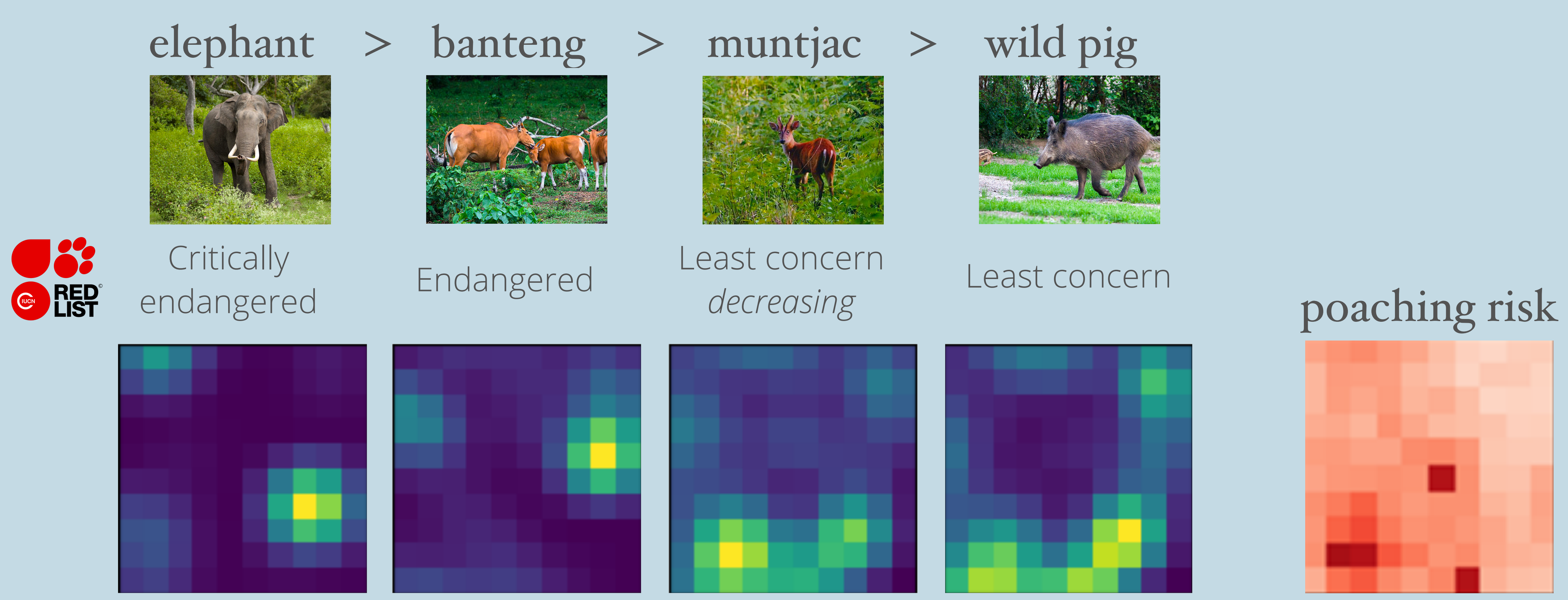
[1]Harvard University
[2]Carnegie Mellon University

RankedCUCB

- **Novel bandit objective** for prioritization in ranked settings
- **No-regret analysis** for weighted objective
- **Empirical results** on real-world data

Motivation

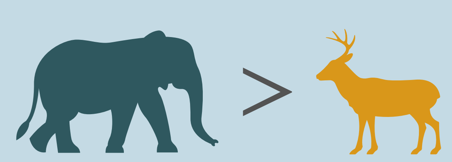
In online resource allocation, actions may have disparate impacts on different groups.



Actions with high reward may not be the same as actions that do most for vulnerable groups.

Problem statement

Allocate resources in an online fashion across ***groups with ranked priority***



Challenges

- Combinatorial allocation
- How to measure “prioritization” with rankings
- Rewards unknown *a priori*

Model

- N locations
- G groups of interest
- d_{gi} density in location i
- Action: Effort $\overrightarrow{\beta}$ subject to budget B
- Reward $\mu_i(\beta_i)$ from effort β_i at location i

Measuring ranked priority

Kendall tau

$$\frac{(\text{\# concordant pairs}) - (\text{\# discordant pairs})}{\binom{G}{2}}$$

$$\text{benefit}(g) = \sum_{i=1}^{N} d_{gi}\mu_i(\beta_i)$$

Prioritization metric

$$\mathscr{P}(\overrightarrow{\beta}) = \frac{\sum_{g,h\in[G]} \mathbf{1}(g < h) \cdot \big(\text{benefit}(g) - \text{benefit}(h)\big)}{\binom{G}{2}}$$

Approach

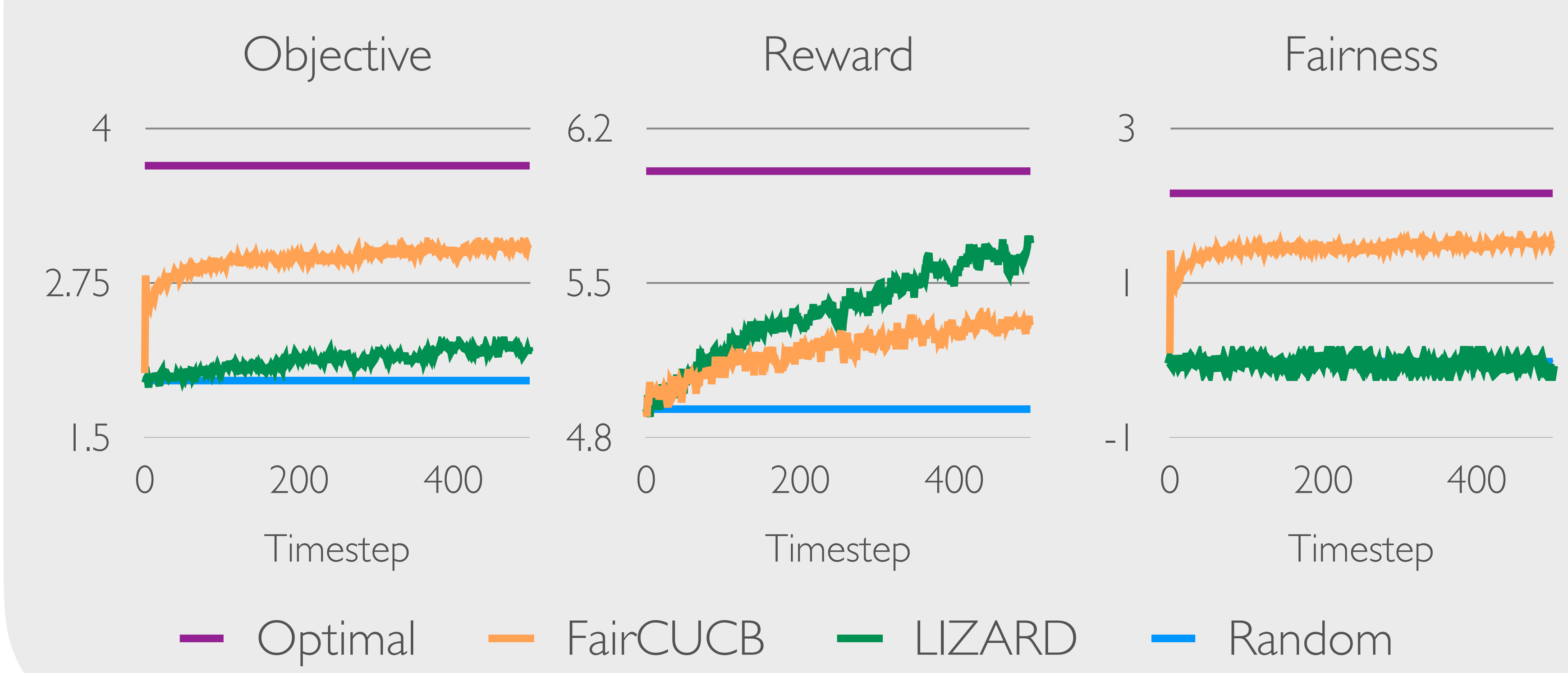
$$\text{obj}(\overrightarrow{\beta}) = \lambda \mu(\overrightarrow{\beta}) + (1-\lambda)\, \mathscr{P}(\overrightarrow{\beta})$$

reward prioritization

$$\text{obj}(\overrightarrow{\beta}) = \sum_{i=1}^{N} \mu_i(\beta_i)\Gamma_i \qquad \Gamma_i = \lambda + (1-\lambda)\cdot \frac{\sum_{g=1}^{G-1}\sum_{h=g+1}^{G}(d_{gi} - d_{hi})}{\binom{G}{2}}$$

Cumulative regret: $O\left(\frac{J \ln T}{N} + NJ\right)$

Experiments



I'm happy to chat! lily_xu@g.harvard.edu