ICML 2014

Optimal PAC Multiple Arm Identification with Applications to Crowdsourcing

Jian Li

Institute for Interdisciplinary Information Sciences

Tsinghua University

joint work with Yuan Zhou (CMU) and Xi Chen (Berkeley)

The Stochastic Multi-armed Bandit

- Stochastic Multi-armed Bandit
  - Set of n arms
  - Each arm is associated with an unknown reward distribution supported on $[0,1]$ with mean θ_i
  - Each time, sample an arm and receive the reward independently drawn from the reward distribution




The Stochastic Multi-armed Bandit

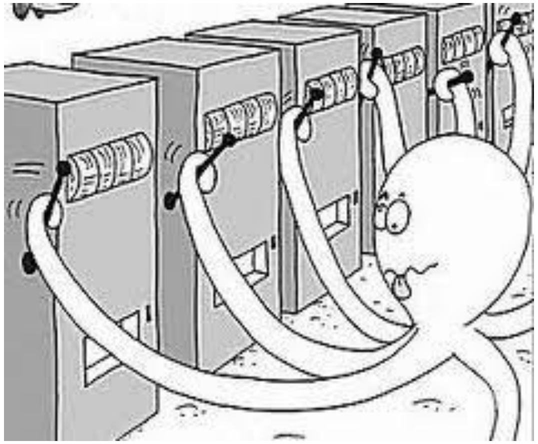
- Top-K Arm identification problem

  You can take N samples

  -A sample: Choose an arm, play it once, and observe the reward

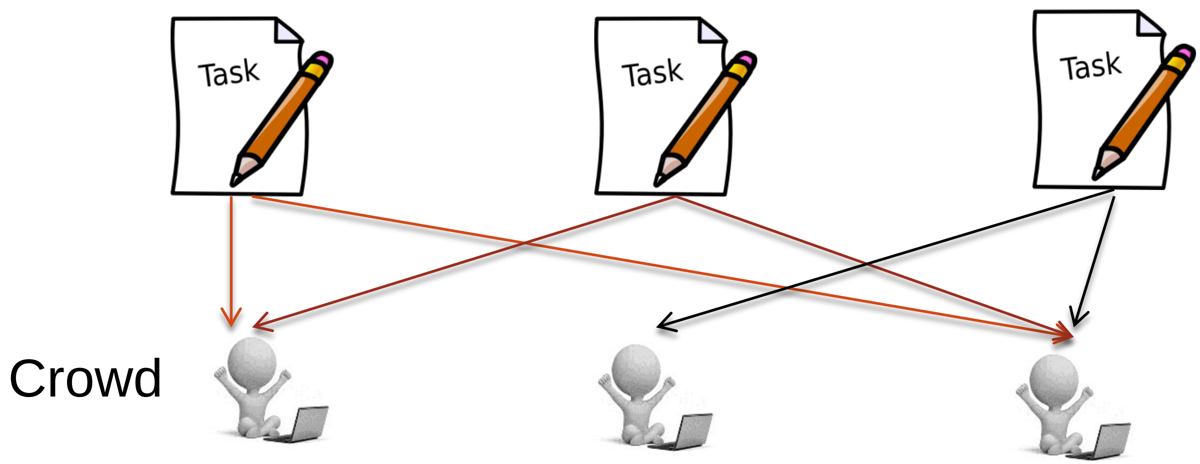
  Goal: (Approximately) Identify the best K arms (arms with largest means)

  Use as few samples as possible (i.e., minimize N)



Motivating Applications

- Wide Applications:
  - Industrial Engineering (Koenig & Law, 85), Evolutionary Computing (Schmidt, 06), Simulation Optimization (Chen, Fu, Shi 08)
- Motivating Application: Crowdsourcing



Motivating Applications

- Workers are noisy

0.95

0.99

0.5

- How to identify reliable workers and exclude unreliable workers ?
- Test workers by golden tasks (i.e., tasks with known answers)

❖ Each test costs money. How to identify the best K workers with minimum amount of money?

Top-K Arm Identification	
Worker	Bernoulli arm with mean θ_i (θ_i: i-th worker's reliability)
Test with golden task	Obtain a binary-valued sample (correct/wrong)

Evaluation Metric

- Sorted means $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_n$
- Goal: find a set of K arms T to minimize the aggregate regret

$$L_T = \frac{1}{K}\left(\sum_{i=1}^{K}\theta_i - \sum_{i\in T}\theta_i\right)$$

- Given any $\epsilon, \delta,$ the algorithm outputs a set T of K arms such that $L_T \leq \epsilon$, with probability at least $1-\delta$ (PAC learning)

- For $K=1,$ i.e., find $\hat{\imath}$: $\theta_1 - \theta_{\hat{\imath}} \leq \epsilon$ w.p. $1-\delta$
  - [Evan-Dar, Mannor and Mansour, 06]
  - [Mannor, Tsitsiklis, 04]
- This Talk: For general K

Simplification

- Assume Bernoulli distributions from now on
- Think of a collection of biased coins
- Try to (approximately) find K coins with largest bias (towards head)

0.5 0.55 0.6 0.45 0.8

Why aggregate regret?

- Misidentification Probability (Bubeck et. al., 13):

$$\Pr(T \neq \{1,2,\ldots,K\})$$

- Consider the case: (K=1)

1

0.99999

Distinguish such two coins with high confidence
requires approx 10^5 samples
(#samples depends on the gap $\theta_1 - \theta_2$)

Using regret (say with $\epsilon = 0.01$), we may choose either of them

Why aggregate regret?

- Explore-K (Kalyanakrishnan et al., 12, 13)
  - Select a set of K arms T: $\forall i \in T, \theta_i > \theta_K - \epsilon$ w.h.p. (θ_K: K-th largest mean)
  - Example: $\theta_1 \geq \cdots \geq \theta_{K-1} \gg \theta_K$ and $\theta_{i+K} > \theta_K - \epsilon$ for $i = 1, \ldots, K$
  - Set $T = \{K+1, K+2 \ldots, 2K\}$ satisfies the requirement

Naïve Solution

Uniform Sampling

Sample each coin M times

Pick the K coins with the largest empirical means

empirical mean: #heads/M

How large M needs to be (in order to achieve ϵ-regret)??

$$M = O(\frac{1}{\epsilon^2}\left(\log\frac{n}{K} + \frac{1}{K}\log\frac{1}{\delta}\right)) = O(\log n)$$

So the total number of samples is O(nlogn)

Naïve Solution

Uniform Sampling

- With M=O(logn), we can get an estimate θ_i' for θ_i such that $|\theta_i - \theta_i'| \leq \epsilon$ with very high probability (say $1 - \frac{1}{n^2}$)
  - This can be proved easily using Chernoff Bound (Concentration bound).
- What if we use M=O(1) (let us say M=10)
  - E.g., consider the following example (K=1):
    - 0.9, 0.5, 0.5, ..., 0.5 (a million coins with mean 0.5)
    - Consider a coin with mean 0.5,

      Pr[All samples from this coin are head]=(1/2)^10
    - With const prob, there are more than 500 coins whose samples are all heads

Uniform Sampling

- In fact, we can show a matching lower bound

$$M = \Theta(\frac{1}{\epsilon^2}\left(\log\frac{n}{K} + \frac{1}{K}\log\frac{1}{\delta}\right)) = \Theta(\log n)$$

One observation: if $K = \Theta(n), M = O(1)$.

Can we do better??

- Consider the following example:
  - 0.9, 0.5, 0.5, ..., 0.5 (a million coins with mean 0.5)
  - Uniform sampling spends too many samples on bad coins.
  - Should spend more samples on good coins
    - However, we do not know which one is good and which is bad……
  - Sample each coin $M=O(1)$ times.
    - If the empirical mean of a coin is large, we DO NOT know whether it is good or bad
    - But if the empirical mean of a coin is very small, we DO know it is bad (with high probability)

Optimal Multiple Arm Identification (OptMAI)

- Input: n (no. of arms), K (top-K arms), Q (total no. of samples/budget)
- Initialization: Active set of arms $S_0 = \{1,2,\ldots,n\}$, Set of top arms $T_0 = \emptyset$
  Iteration Index $r = 0$, Parameter $\beta \in (0.75, 1)$
- While $|T_r| < K$ and $|S_r| > 0$ do
  - If $|S_r| > 4K$ then
    - S_{r+1}=Quartile-Elimination$(S_r, \beta^r(1-\beta)Q)$
  - Else ($|S_r| \leq 4K$)
    - Identify the best K arms for at most 4K arms, using uniform sampling
  - $r = r + 1$
- Output: set of selected K arms T_r

Eliminate one quarter arms with lowest empirical means

Quartile-Elimination

- Idea: uniformly sample each arm in the active set S and discard the worst quarter of arms (with the lowest empirical mean)

- Input: S (active arms), Q (budget)
- Sample each arm $i \in S$ for $Q/|S|$ times & let $\hat{\theta}_i$ be the empirical mean
- Find the lower quartile of the empirical mean $\hat{q}$: $|\{i: \hat{\theta}_i < \hat{q}\}| = |S|/4$
- Output: $S' = S \setminus \{i: \hat{\theta}_i < \hat{q}\}$

Sample Complexity

- Sample complexity Q:

Outputs K arms s.t. $L_T = \frac{1}{K}(\sum_{i=1}^{K} \theta_i - \sum_{i \in T} \theta_i) \le \epsilon$, w.p. $1 - \delta$.

- $K \le \frac{n}{2}$: $Q = O\left(\frac{n}{\epsilon^2}\left(1 + \frac{\ln\left(\frac{1}{\delta}\right)}{K}\right)\right)$ (this is linear!)

- $K \ge \frac{n}{2}$: $Q = O\left(\frac{n-K}{K}\frac{n}{\epsilon^2}\left(\frac{n-K}{K} + \frac{\ln\left(\frac{1}{\delta}\right)}{K}\right)\right)$ (which can be sublinear!)

  - Apply our algorithm to identify the worst $(n-K)$ arms.

Sample Complexity

- Sample complexity Q:

Outputs K arms s.t. $L_T = \frac{1}{K}\left(\sum_{i=1}^{K}\theta_i - \sum_{i\in T}\theta_i\right) \le \epsilon$, w.p. $1-\delta$.

- $K \le \frac{n}{2}$: $Q = O\left(\frac{n}{\epsilon^2}\left(1 + \frac{\ln\left(\frac{1}{\delta}\right)}{K}\right)\right)$ (this is linear!)

Better bound if K is larger!

- $K \ge \frac{n}{2}$: $Q = O\left(\frac{n-K}{K}\frac{n}{\epsilon^2}\left(\frac{n-K}{K} + \frac{\ln\left(\frac{1}{\delta}\right)}{K}\right)\right)$ (which can be sublinear!)
  - Reduce to the $K \le \frac{n}{2}$ case by identifying the worst $(n-K)$ arms.

Sample Complexity

- $K \leq \frac{n}{2}$: $Q = O\left(\frac{n}{\epsilon^2}\left(1 + \frac{\ln\left(\frac{1}{\delta}\right)}{K}\right)\right)$
- $K = 1,\ Q = O\left(\frac{n}{\epsilon^2}\ln\left(\frac{1}{\delta}\right)\right)$ [Even-Dar et. al., 06]

- For larger $K,$ the sample complexity is smaller: identify K arms is simpler !
- Why? Example: $\theta_1 = \frac{1}{2} + 2\epsilon, \theta_2 = \theta_3 = \cdots \theta_n = \frac{1}{2}$.
  - Identify the first arm ($K = 1$) is hard ! Cannot pick the wrong arm.
  - Since $L_T \leq \frac{2\epsilon}{K}$, for $K \geq 2$, any set is fine.

- Naïve Uniform Sampling: $Q = \Omega(n\log(n))$, $\log(n)$ factor worse

Matching Lower Bounds

- $K \leq \frac{n}{2}$: there is an underlying $\{\theta_i\}$ such that for any randomized algorithm, to identify a set T with $L_T \leq \epsilon$ w.p. at least $1 - \delta$,

$$E[Q] = \Omega\left(\frac{n}{\epsilon^2}\left(1 + \frac{\ln\left(\frac{1}{\delta}\right)}{K}\right)\right)$$

- $K > \frac{n}{2}$: $E[Q] = \Omega\left(\frac{n-K}{K}\frac{n}{\epsilon^2}\left(\frac{n-K}{K} + \frac{\ln\left(\frac{1}{\delta}\right)}{K}\right)\right)$

Our algorithm is optimal for every value of n, K, ϵ, δ!

Matching Lower Bounds

- First Lower bound: $K \leq \frac{n}{2},\ Q \geq \Omega\left(\frac{n}{\epsilon^2}\right)$
  - Reduction to distinguishing two Bernoulli arms with means $\frac{1}{2}$ and $\frac{1}{2} + \epsilon$ with probability > 0.51, which requires at least $\Omega\left(\frac{1}{\epsilon^2}\right)$ samples [Chernoff, 72]

    (anti-concentration)

- Second Lower bound: $K \leq \frac{n}{2},\ Q \geq \Omega\left(\frac{n}{\epsilon^2}\left(\frac{\ln\left(\frac{1}{\delta}\right)}{K}\right)\right)$
  - A standard technique in statistical decision theory

Experiments

OptMAI	$\beta = 0.8, \quad \beta = 0.9$
SAR	Bubeck et. al., 13
LUCB	Kalyanakrishnan et. al., 12
Uniform	Naïve Uniform Sampling

Simulated Experiments:

No. of Arms: $n = 1000$
Total Budget: $Q = 20n, Q = 50n, Q = 100n$
Top-K Arms: $K = 10, 20, \ldots, 500$
Report average result over 100 independent runs

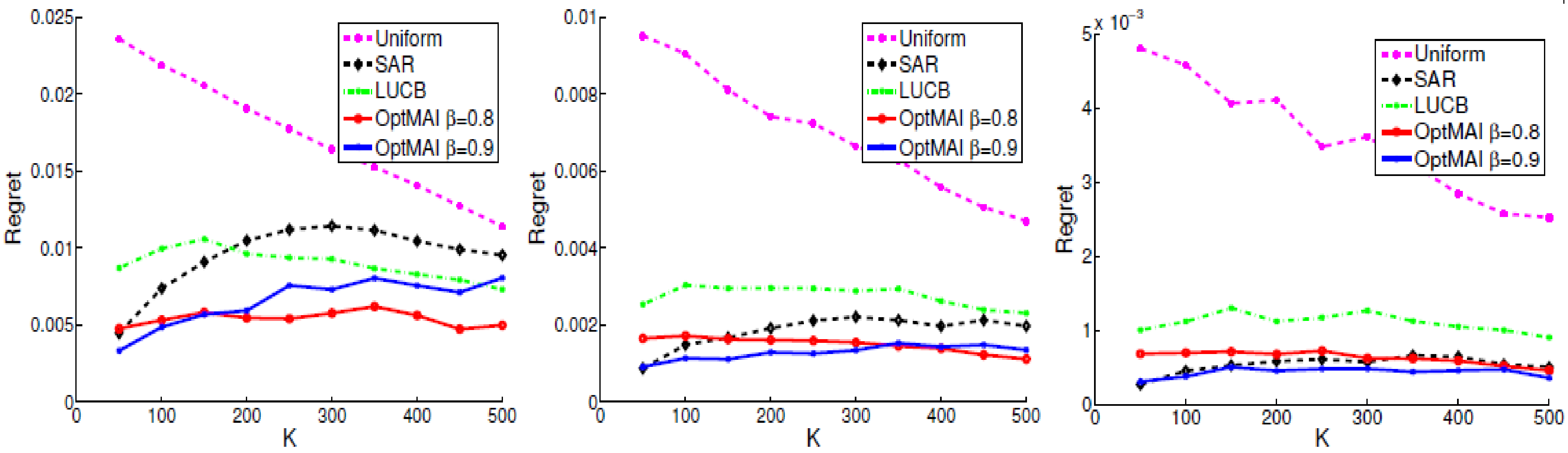
Underlying distributions:

(1) $\theta_i \sim Uniform[0,1]$
(2) $\theta_i = 0.6$ for $i = 1, \ldots, K$, $\theta_i = 0.5$ for $i = K+1, \ldots, n$

Metric: regret L_T

Simulated Experiment

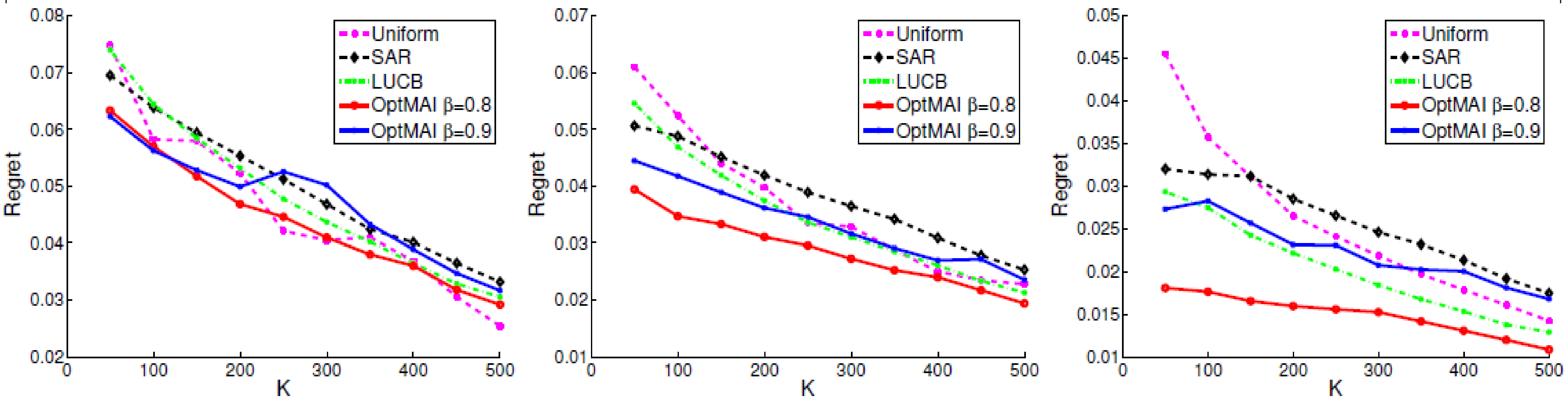
- $\theta_i \sim Uniform[0,1]$



$\theta \sim \mathrm{Unif}[0,1],\ Q = 20 \cdot n$ $\theta \sim \mathrm{Unif}[0,1],\ Q = 50 \cdot n$ $\theta \sim \mathrm{Unif}[0,1],\ Q = 100 \cdot n$

Simulated Data

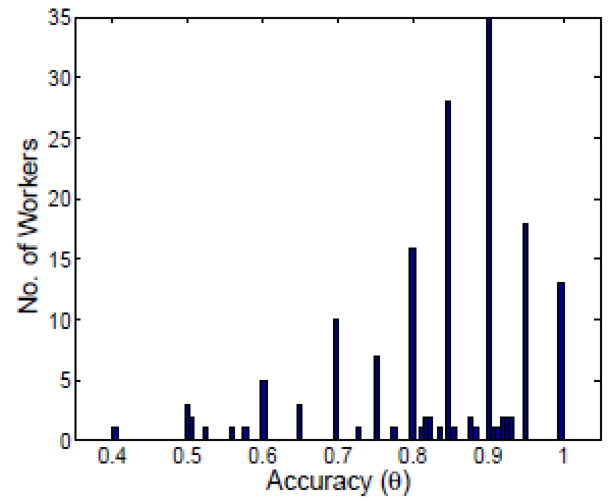
- $\theta_i = 0.6$ for $i = 1, \dots, K, \theta_i = 0.5$ for $i = K + 1, \dots, n$



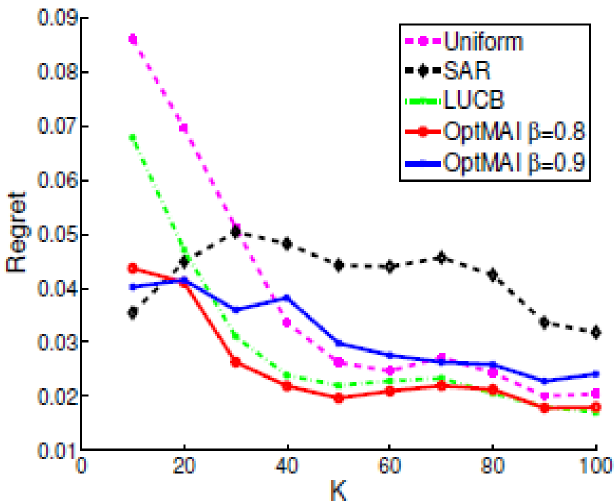
$\theta = 0.6/0.5,\ Q = 20 \cdot n$ $\theta = 0.6/0.5,\ Q = 50 \cdot n$ $\theta = 0.6/0.5,\ Q = 100 \cdot n$

Real Data

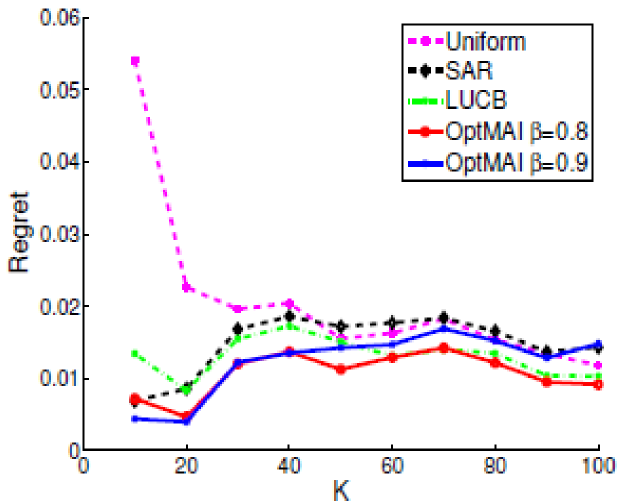
- RTE data for textual entailment (Snow et. al., 08)
- 800 binary labeling tasks with true labels
- 164 workers



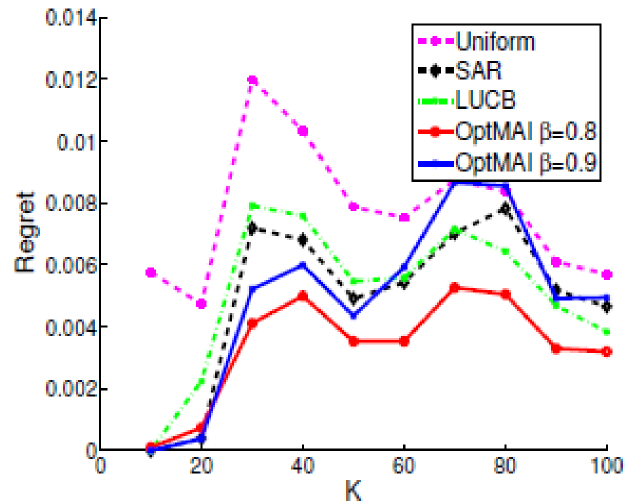
Histogram of θ_i



Regret $(Q = 10 \cdot n)$



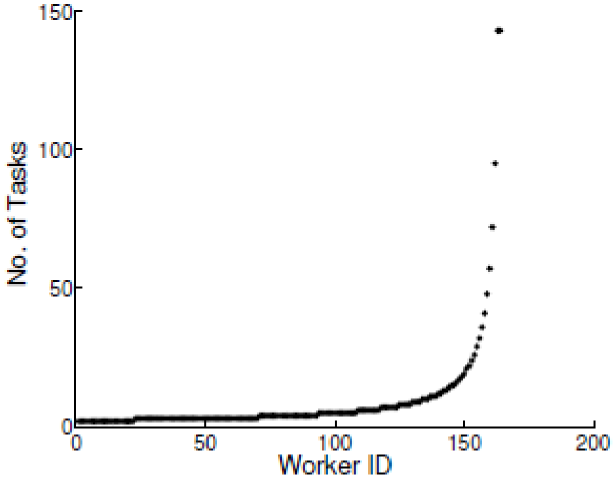
Regret $(Q = 20 \cdot n)$



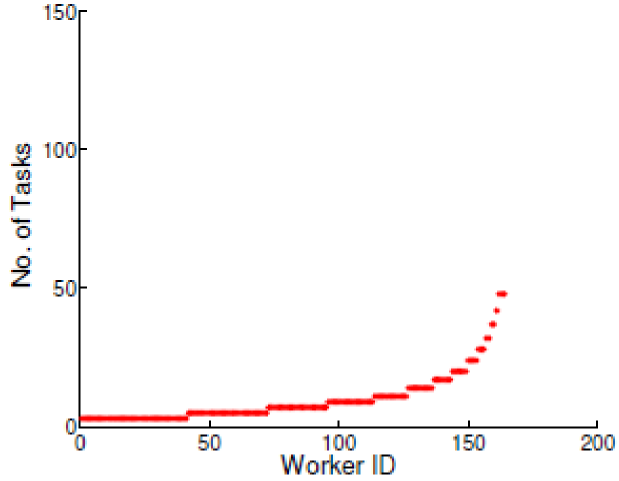
Regret $(Q = 50 \cdot n)$

Real Data

- Empirical distribution of the number tasks assigned to a worker
  $(\beta = 0.9, K = 10, Q = 20n)$



No. of Tasks (SAR)



No. of Tasks (OptMAI)

A worker receives at most 143 tasks

SAR queries an arm $\Omega\left(\frac{Q}{\log(n)}\right)$ times

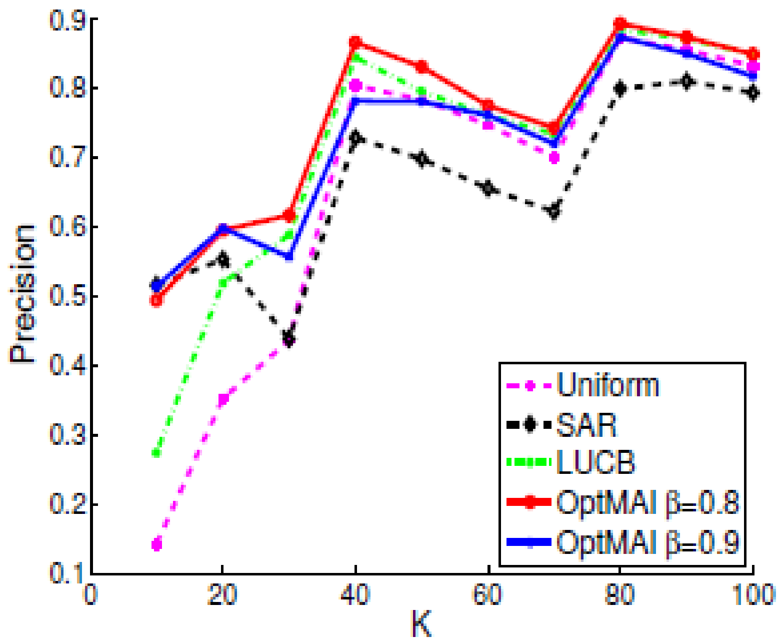
A worker receives at most 48 tasks

OptMAI queries an arm $O\left(\frac{Q}{n^{\Omega(1)}}\right)$ times

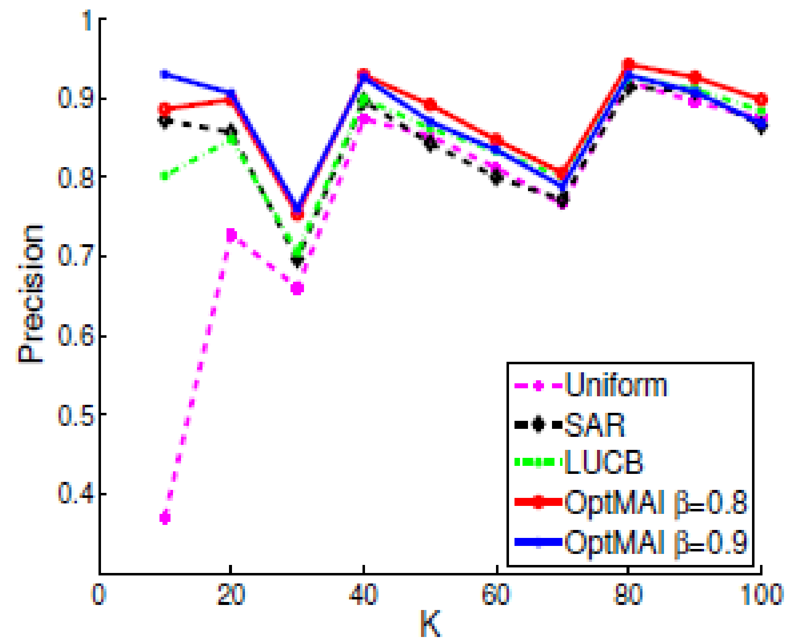
Crowdsourcing: Impossible to assign too many tasks to a single worker

Real Data

- Precision $= \frac{|T \cap \{1,\ldots,K\}|}{K}$: no. of arms in T belongs to the top K arms



Precision $(Q = 10 \cdot n)$



Precision $(Q = 20 \cdot n)$

Conclusion

- Top-k arm identification
- Application in crowdsourcing
- (Worse case) Optimal upper and lower bounds

- Further direction: some instances are "easier", i.e., 0.9,0.1,0.1,0.1,……. Can we get better upper bounds for these instance??

Thanks.

lapordge@gmail.com