Private Stochastic Multi-Arm Bandit

Nikita Mishra (U Chicago), Abhradeep Thakurta (Yahoo! Labs)

Arms $A$: $a_1$, $a_2$, $a_k$

Receives a reward $r_{a_2}(t) \in [0,1]$ at time $t \in [T]$

In general $r: A \times T \rightarrow [0,1]$
This paper:

1. Differentially private algorithm for stochastic multi-arm bandits
   a) Based on Upper Confidence Bound sampling and Thompson sampling
   b) Guarantee: Nearly optimal regret up to $\text{poly log } T$

2. Algorithmic extension to contextual multi-arm bandits

3. Experimental evaluation on simulated and real-world data sets