Stochastically Constrained Best Arm Identification with Thompson Sampling

Le Yang
cooperated with Siyang Gao, Cheng Li, Yi Wang

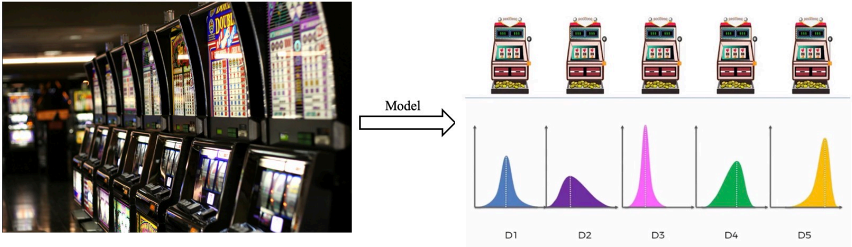
Email: le_yang@nus.edu.sg

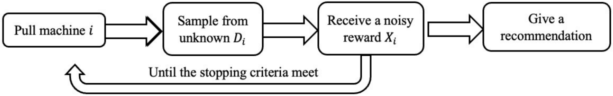
Date: September 16, 2024

Contents

Introduction

Best Arm Identification (BAI) Problem





Thompson Sampling (TS) Algorithm

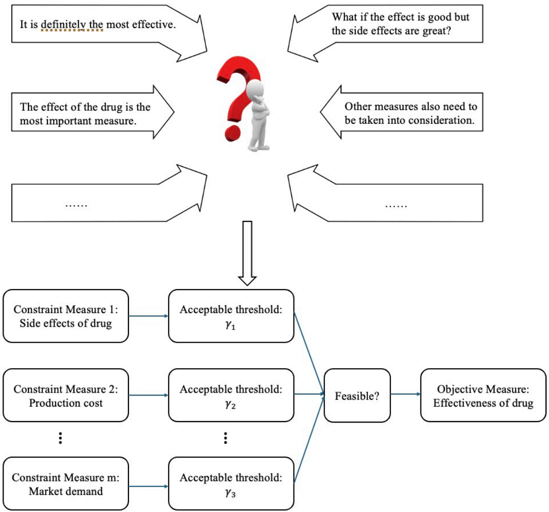
TTTS is an asymptotically optimal algorithm for BAI problems Russo [2020].

TS was originally proposed by Thompson in 1933 [Thompson, 1933] for the multi-armed bandit (MAB) problem.

TS has been extended to tackle a wide range of variant MAB problems.

- Combinatorial bandits [Sankararaman and Slivkins, 2018]
- Contextual bandits [Agrawal and Goyal, 2013]
- Online problems [Gopalan, Mannor, and Mansour, 2014]

Best Feasible Arm Identification (BFAI) Problem



Related Literature

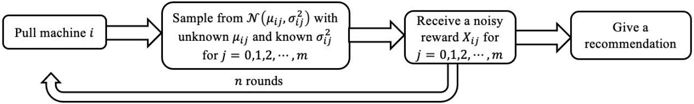
Communities	Related Papers	Differences from our work
Machine Learning	• Multi-Objective Problem: Drugan and Nowe (2013) Auer et al. (2016) Tekin and Turgay (2017) • Feasible Arm Identification: Katz-Samuels and Scott (2018)	Formulate into different problem
	• Top Feasible Arm Identification: Katz-Samuels and Scott (2019)	Fixed confidence not fixed budget
Simulation	Lee et al. (2012) Hunter and Pasupathy (2013) Pasupathy et al. (2014)	Focus on finding the sample allocations that approximately maximize PCS

Formulation

Best Feasible Arm Identification (BFAI) Problem

How to choose the most effective drug with acceptable side effects?

- Problem:

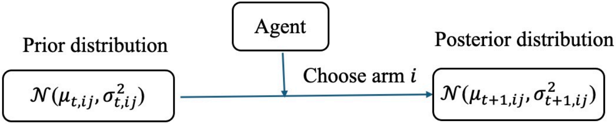


- Goal: Identify the best arm: $I^* = \{i: argmax_i \mu_{i0} \; s.t. \mu_{ij} \le \gamma_j, \forall\, j = 1,2,\cdots,m\}$.

- Assumption: The best feasible arm is unique. Let $I^* = 1$.

Bayesian Framework

Normal prior distributions:



$$\mu_{t+1,ij} = \begin{cases} \dfrac{\sigma_{t,ij}^{-2}\mu_{t,ij} + \sigma_{ij}^{-2} X_{t,ij}}{\sigma_{t,ij}^{-2} + \sigma_{ij}^{-2}} & \text{if } I_t = i, \\ \mu_{t,ij} & \text{if } I_t \neq i, \end{cases} \qquad \sigma^2_{t+1,ij} = \begin{cases} \dfrac{1}{\sigma_{t,ij}^{-2} + \sigma_{ij}^{-2}} & \text{if } I_t = i, \\ \sigma^2_{t,ij} & \text{if } I_t \neq i. \end{cases}$$

Posterior distribution:

$$\begin{aligned} \Pi_t =& \mathcal{N}(\mu_{t,10}, \sigma^2_{t,10}) \otimes \ldots \otimes \mathcal{N}(\mu_{t,1m}, \sigma^2_{t,1m}) \otimes \ldots \\ & \otimes \mathcal{N}(\mu_{t,k0}, \sigma^2_{t,k0}) \otimes \ldots \otimes \mathcal{N}(\mu_{t,km}, \sigma^2_{t,km}). \end{aligned}$$

Asymptotic Optimal

Asymptotic Optimal: (Rate Optimal) $\max_{\text{sampling rules}}$ convergence rate $= \max_{\text{sampling rules}} \liminf_{n\to\infty} -\frac{1}{n}\log(1-PCS).$

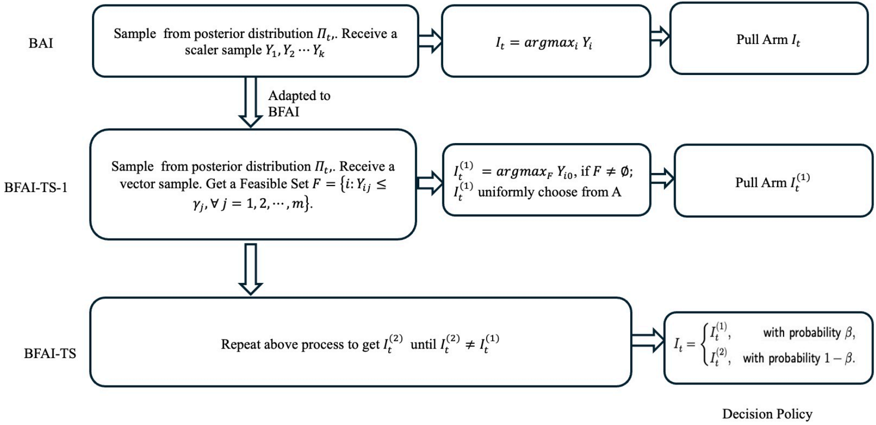
$$P_{t,i} \triangleq \mathbb{P}_{\theta\sim\Pi_t}\left(\bigcap_{i'\neq i}\left((\theta_{i0}<\theta_{i'0})\cap\bigcap_{j=1}^{m}(\theta_{i'j}\leq\gamma_j)\right)^c \cap \bigcap_{j=1}^{m}(\theta_{ij}\leq\gamma_j)\right)$$

Other arms are identified as the best feasible arm

Capture the feasibility of arm ***i***

BFAI-TS Algorithm

Process of Extending



Algorithm Description

Algorithm 4: BFAI-TS Algorithm

Input: $k \geq 2, \beta \in (0, 1), n$

Collect n_0 samples for each arm i;
while $t \leq n$ **do**
    Sample $\theta \sim \Pi_t$;
    Get the feasible set $F \triangleq \{i : \theta_{ij} \leq \gamma_j \text{ for } j = 1, 2, \ldots, m\}$;
    if $F \neq \emptyset$ **then**
        Set $I_t^{(1)} \leftarrow \text{argmax}\, \theta_{i0}$ for $i \in F$;
    else
        Choose $I_t^{(1)}$ uniformly from $\{1, 2, \ldots, k\}$
    Sample $B \sim$Bernoulli(β);
    if $B = 1$ **then**
        Play $I_t^{(1)}$;
    else
        repeat
            Sample $\theta \sim \Pi_t$;
            Get the feasible set F;
            if $F \neq \emptyset$ **then**
                Set $I_t^{(2)} \leftarrow \text{argmax}\, \theta_{i0}$ for $i \in F$;
            else
                Choose $I_t^{(2)}$ uniformly from $\{1, 2, \ldots, k\}$
        until $I_t^{(2)} \neq I_t^{(1)}$;
        Play $I_t^{(2)}$;
    Update posterior Π_{t+1};

Output: I^*

Algorithm Analysis

Define $\phi_{t,i} \triangleq \mathbb{P}(I_t = i|\mathcal{F}_{t-1})$ and $\bar{\phi}_{t,i} \triangleq \frac{\sum_{l=2}^{t} \phi_{l,i}}{t}$.

The probability of pulling arm i in round t:

$$\begin{aligned}\phi_{t,i} =& \frac{c_t}{k} + (1-\beta)P_{t,i}\sum_{i' \neq i}\left(\frac{P_{t,i'}}{1-P_{t,i'}}(1-c_t) + \frac{c_t}{k-1}\right) \\ &+ P_{t,i}\beta(1-c_t),\end{aligned}$$

where c_t is the probability that samples from all the arms are infeasible in round t.

$P_{t,1} \to 1$, $\phi_{t,1} \to \beta$, $c_t \to 0$

$\frac{P_{t,i}}{1-P_{t,1}} \to \frac{\phi_{t,i}}{1-\phi_{t,1}}$, $i \neq 1$.

Theoretical Results

Notations

- $\mathcal{F}$ the set of feasible arms, i.e., $\mathcal{F} \triangleq \{i : \mu_{ij} \leq \gamma_j, \forall j \in \{1, 2, ..., m\}\}$;
- I^* the best feasible arm, i.e., $I^* \triangleq \text{argmax}_{i \in \mathcal{F}} \mu_{i0}$;
- $\mathcal{F}_w$ the set of feasible but suboptimal arms, i.e., $\mathcal{F}_w \triangleq \{i : i \in \mathcal{F}$ and $i \neq I^*\}$;
- $\mathcal{I}_b$ the set of infeasible arms with objective performance no worse than I^*, i.e., $\mathcal{I}_b \triangleq \{i : \mu_{I^*0} \leq \mu_{i0}$ and $\exists j \in \{1, 2, \dots, m\}$ such that $\mu_{ij} > \gamma_j\}$;
- $\mathcal{I}_w$ the set of infeasible arms with objective performance worse than I^*, i.e., $\mathcal{I}_w \triangleq \{i : \mu_{I^*0} > \mu_{i0}$ and $\exists j \in \{1, 2, \dots, m\}$ such that $\mu_{ij} > \gamma_j\}$;
- $\mathcal{M}_F^i$ the set of constraints estimated as satisfied by arm i in round t, i.e. $\mathcal{M}_F^i \triangleq \{j : \mu_{ij} \leq \gamma_j$ for $j \in \{1, 2, \dots, m\}\}$;
- $\mathcal{M}_I^i$ the set of constraints estimated as violated by arm i in round t, i.e. $\mathcal{M}_I^i \triangleq \{j : \mu_{ij} > \gamma_j$ for $j \in \{1, 2, \dots, m\}\}$.

Definitions

$(\alpha_2^\beta, ..., \alpha_k^\beta)$:
the optimal sampling rates of the remaining $k-1$ arms, which satisfy the following optimality condition

$$\sum_{i=2}^{k} \alpha_i^\beta = 1-\beta, \text{ and } \mathcal{R}_i = \mathcal{R}_{i'} \text{ for any } i \neq i' \neq 1,$$

where
$\mathcal{R}_i = \frac{(\mu_{i0}-\mu_{10})^2}{(\sigma_{i0}^2/\alpha_i^\beta+\sigma_{10}^2/\beta)}\mathbf{1}\{i \in \mathcal{F}_w \cup \mathcal{I}_w\} + \alpha_i^\beta \sum_{j\in\mathcal{M}_I^i} \frac{(\mu_{ij}-\gamma_j)^2}{\sigma_{ij}^2}\mathbf{1}\{i \in \mathcal{I}_b \cup \mathcal{I}_w\}.$

Γ_β: the optimal sampling rates given $\beta \in (0,1)$, where

$$\Gamma_\beta = \min_{i\neq 1}\Bigg(\frac{(\mu_{i0}-\mu_{10})^2}{2(\sigma_{i0}^2/\alpha_i^\beta + \sigma_{10}^2/\beta)}\mathbf{1}\{i \in \mathcal{F}_w \cup \mathcal{I}_w\} + \alpha_i^\beta \sum_{j\in\mathcal{M}_I^i} \frac{(\mu_{ij}-\gamma_j)^2}{2\sigma_{ij}^2}\mathbf{1}\{i \in \mathcal{I}_b \cup \mathcal{I}_w\}, \min_{j\in\mathcal{M}_F^1} \beta\frac{(\mu_{1j}-\gamma_j)^2}{2\sigma_{1j}^2}\Bigg).$$

Theorem

For the BFAI-TS Algorithm, $\mathbb{E}[N_\beta^\epsilon] < \infty$ for any $\epsilon > 0$, where

$$N_\beta^\epsilon \triangleq \inf\{t \in \mathbb{N} : |\mu_{n,ij} - \mu_{ij}| \leq \epsilon \text{ and } |N_{n,i}/n - \alpha_i^\beta| \leq \epsilon, \forall i \in A \text{ and } n \geq t\}$$

given $\beta \in (0,1)$. The sample allocations of the algorithm is asymptotically optimal in the sense that

$$\lim_{n\to\infty} \frac{N_{n,i}}{n} \xrightarrow{p} \alpha_i^\beta \quad \forall i \in A,$$

Theorem

The following properties hold with probability 1:
1. For any $\beta \in (0,1)$, Γ_β shows the fastest rate of posterior convergence that any algorithm allocating β proportion of the total samples to the best feasible arm can possibly achieve

$$\limsup_{n\to\infty} -\frac{1}{n}\log(1-P_{n,1}) \leq \Gamma_\beta \tag{1}$$

and the BFAI-TS Algorithm achieves this rate with

$$\lim_{n\to\infty} -\frac{1}{n}\log(1-P_{n,1}) = \Gamma_\beta. \tag{2}$$

2. The term Γ_{β^*} shows the fastest rate of posterior convergence that any BFAI algorithm can possibly achieve

$$\limsup_{n\to\infty} -\frac{1}{n}\log(1-P_{n,1}) \leq \Gamma_{\beta^*} \tag{3}$$

and when the β of the BFAI-TS Algorithm is set to β^*, the algorithm achieves the optimal rate with

$$\lim_{n\to\infty} -\frac{1}{n}\log(1-P_{n,1}) = \Gamma_{\beta^*}. \tag{4}$$

Numerical Results

Probabilities of false selection for the tested algorithms.

Experiments / Algorithms \ Sample size	Experiment 1			Experiment 2			Experiment 3			Experiment 4			Experiment 5			Dose-finding		
	2100	2500	3400	1000	2500	3500	2000	2400	3700	2600	3300	3700	200	400	800	3500	6000	8000
BFAI-TS-1	0.20	0.17	0.13	0.38	0.14	0.10	0.24	0.19	0.12	0.13	0.06	0.04	0.37	0.25	0.22	0.18	0.09	0.06
OCBA-CO	0.13	0.10	0.04	0.25	0.14	0.05	0.14	0.11	0.06	0.09	0.06	0.05	0.31	0.24	0.20	0.20	0.12	0.10
Uniform	0.61	0.51	0.46	0.71	0.59	0.50	0.46	0.48	0.39	0.54	0.49	0.51	0.37	0.41	0.34	0.27	0.19	0.13
TF-LUCB	0.22	0.11	0.05	0.51	0.15	0.06	0.37	0.24	0.11	0.26	0.13	0.09	0.27	0.19	0.21	0.16	0.10	0.09
MD-UCBE	0.87	0.87	0.87	0.82	0.85	0.83	0.79	0.79	0.77	0.79	0.77	0.77	0.50	0.50	0.50	0.21	0.19	0.19
BFAI-TS ($\beta = 0.5$)	0.11	0.06	0.02	0.24	0.09	0.03	0.12	0.07	0.05	0.04	0.02	0.02	0.25	0.18	0.15	0.13	0.07	0.05
BFAI-TS ($\beta = \beta^*$)	0.07	0.02	0.01	0.19	0.02	0.01	0.09	0.05	0.01	0.02	0.01	0.00	0.11	0.03	0.00	0.11	0.05	0.01

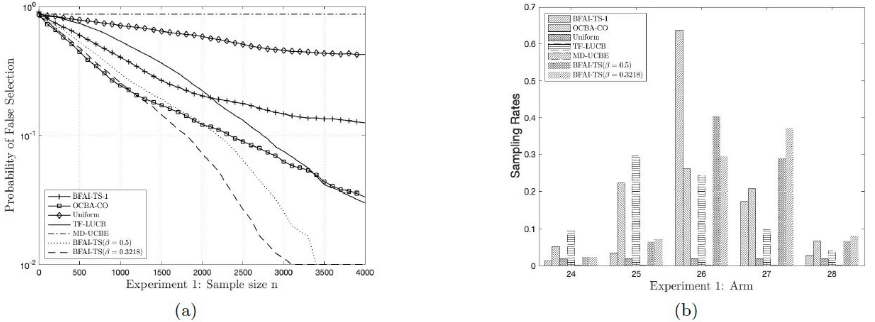


Fig. 1. PFS and their sampling rates on selected arms (Experiment 1)

Conclusions

Demonstrate the extensibility of the TS algorithm by proposing the BFAI-TS algorithm.

Solve a significant and common class of constrained optimization problem, the BFAI problem.

The BFAI-TS algorithm is asymptotically optimal and exhibits impressive numerical performance.

References

S. Agrawal and N. Goyal. Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning*, pages 127–135, 2013.

A. Gopalan, S. Mannor, and Y. Mansour. Thompson sampling for complex online problems. In *International Conference on Machine Learning*, pages 100–108. PMLR, 2014.

D. Russo. Simple Bayesian algorithms for best arm identification. *Operations Research*, 68(6):1625–1647, 2020.

K. A. Sankararaman and A. Slivkins. Combinatorial semi-bandits with knapsacks. In *International Conference on Artificial Intelligence and Statistics*, pages 1760–1770. PMLR, 2018.

W. R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3-4):285–294, 1933.

Thank you!