Differentially Private Best-Arm Identification

Achraf Azize, **Marc Jourdan**, Aymen Al Marjani, Debabrota Basu

July 1, 2024

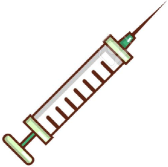
Inria
CRIStAL Centre de Recherche en Informatique, Signal et Automatique de Lille
Université de Lille
cnrs

Phase III clinical trials

μ_1

μ_2

μ_3



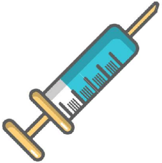
μ_4

Goal: Identify a treatment with a high efficiency.

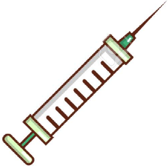
Phase III clinical trials

μ_1

μ_2



μ_3



μ_4

Goal: Identify a treatment with a high efficiency.

Setting: Pure exploration for stochastic multi-armed bandits.

☞ Sequential hypothesis testing with adaptive data collection.

Sequential decision making under uncertainty

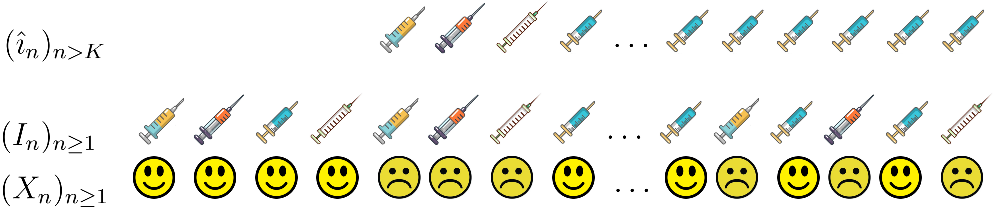
After treating $n-1$ patients, the physician has

- ☞ a guessed answer for a good treatment $\hat{\imath}_n \in [K]$.

As the n-th patient enters, the physician selects

- ☞ a treatment $I_n \in [K]$ for administration.

Then, it observes a realization $X_n \sim \nu_{I_n}$ with $\nu_i = \mathcal{B}(\mu_i)$.



Best-Arm Identification (BAI)

K arms: arm $i \in [K]$ with $\nu_i = \mathcal{B}(\mu_i) \in \mathcal{D}$ where $\mu_i \in (0,1)$.

Goal: identify the unique **best arm** $i^\star = \arg\max_{i \in [K]} \mu_i$.

Best-Arm Identification (BAI)

K arms: arm $i \in [K]$ with $\nu_i = \mathcal{B}(\mu_i) \in \mathcal{D}$ where $\mu_i \in (0,1)$.

Goal: identify the unique **best arm** $i^\star = \arg\max_{i \in [K]} \mu_i$.

Algorithm: at time n ,

- *Recommendation rule*: recommend a candidate answer $\hat{\imath}_n$.
- *Stopping rule*: dictate when to stop sampling .
- **Sampling rule**: pull an arm I_n and observe $X_n \sim \nu_{I_n}$.

Best-Arm Identification (BAI)

K arms: arm $i \in [K]$ with $\nu_i = \mathcal{B}(\mu_i) \in \mathcal{D}$ where $\mu_i \in (0,1)$.

Goal: identify the unique **best arm** $i^\star = \arg\max_{i \in [K]} \mu_i$.

Algorithm: at time n ,

- *Recommendation rule*: recommend a candidate answer $\hat{\imath}_n$.
- *Stopping rule*: dictate when to stop sampling .
- **Sampling rule**: pull an arm I_n and observe $X_n \sim \nu_{I_n}$.

Fixed-confidence: given a confidence pair δ , define a δ-correct stopping time τ_δ , i.e. $\mathbb{P}_\nu(\tau_\delta < +\infty, \hat{\imath}_{\tau_\delta} \neq i^\star) \leq \delta$.

☞ Minimize the **expected sample complexity** $\mathbb{E}_\nu[\tau_\delta]$.

Lower bound on the expected sample complexity

(Garivier and Kaufmann, 2016) For all δ-correct algorithm,

$$\forall \nu \in \mathcal{D}^K, \quad \liminf_{\delta \to 0} \frac{\mathbb{E}_\nu[\tau_\delta]}{\log(1/\delta)} \geq T^\star_{\mathrm{KL}}(\nu) ,$$

Lower bound on the expected sample complexity

(Garivier and Kaufmann, 2016) For all δ-correct algorithm,

$$\forall \nu \in \mathcal{D}^K, \quad \liminf_{\delta \to 0} \frac{\mathbb{E}_\nu[\tau_\delta]}{\log(1/\delta)} \geq T^\star_{\mathrm{KL}}(\nu) \, ,$$

where the inverse of the **characteristic time** is

$$T^\star_{\mathrm{KL}}(\nu)^{-1} = \max_{w \in \triangle_K} \min_{j \neq i^\star} C_{\mathrm{KL}}(i^\star, j; \nu, w) \, ,$$

with $$C_{\mathrm{KL}}(i, j; \nu, w) \approx \mathbb{1}\left(\mu_i > \mu_j\right) \frac{2(\mu_i - \mu_j)^2}{1/w_i + 1/w_j} \, .$$

Lower bound on the expected sample complexity

(Garivier and Kaufmann, 2016) For all δ-correct algorithm,

$$\forall \nu \in \mathcal{D}^K, \quad \liminf_{\delta \to 0} \frac{\mathbb{E}_\nu[\tau_\delta]}{\log(1/\delta)} \geq T^\star_{\mathrm{KL}}(\nu) ,$$

where the inverse of the **characteristic time** is

$$T^\star_{\mathrm{KL}}(\nu)^{-1} = \max_{w \in \triangle_K} \min_{j \neq i^\star} C_{\mathrm{KL}}(i^\star, j; \nu, w) ,$$

with $\quad C_{\mathrm{KL}}(i, j; \nu, w) \approx \mathbb{1}\left(\mu_i > \mu_j\right) \frac{2(\mu_i - \mu_j)^2}{1/w_i + 1/w_j} .$

Algorithms: Track-and-Stop, online optimization, **Top Two**.

TTUCB (Jourdan and Degenne, 2023)

- Recommend the empirical best arm $\hat{\imath}_n = \arg\max_{i \in [K]} \mu_{n,i}$.

TTUCB (Jourdan and Degenne, 2023)

- Recommend the empirical best arm $\hat{\imath}_n = \arg\max_{i \in [K]} \mu_{n,i}$.
- Generalized likelihood ratio (**GLR**) stopping rule

$$\tau_\delta = \inf\{n \in \mathbb{N} \mid \min_{j \neq \hat{\imath}_n} C_{\mathrm{KL},n}(\hat{\imath}_n, j) > c(n-1, \delta)\} ,$$

with $C_{\mathrm{KL},n}(i,j) = C_{\mathrm{KL}}(i,j;\nu_n,N_n)$ and $c(n,\delta) \approx \log(1/\delta) + \mathcal{O}(\log n)$.

TTUCB (Jourdan and Degenne, 2023)

- Recommend the empirical best arm $\hat{\imath}_n = \arg\max_{i \in [K]} \mu_{n,i}$.
- Generalized likelihood ratio (**GLR**) stopping rule

$$\tau_\delta = \inf\{n \in \mathbb{N} \mid \min_{j \neq \hat{\imath}_n} C_{\mathrm{KL},n}(\hat{\imath}_n, j) > c(n-1, \delta)\} ,$$

with $C_{\mathrm{KL},n}(i,j) = C_{\mathrm{KL}}(i,j;\nu_n, N_n)$ and $c(n,\delta) \approx \log(1/\delta) + \mathcal{O}(\log n)$.

- Sample $I_n \in \{B_n, C_n\}$ uniformly at random where

UCB leader: $B_n = \arg\max_{i \in [K]} \left\{ \mu_{n,i} + \sqrt{\log(n)/N_{n,i}} \right\}$,

TC challenger: $C_n = \arg\min_{j \neq B_n} C_{\mathrm{KL},n}(B_n, j)$.

Differential privacy

⚠ Rewards may reveal sensitive information about individuals !

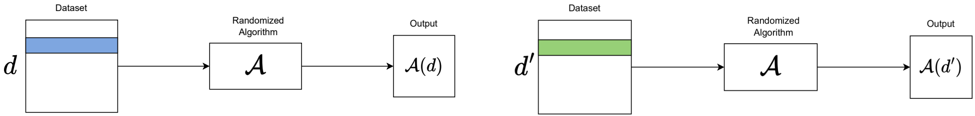
Differential privacy

⚠ Rewards may reveal sensitive information about individuals !

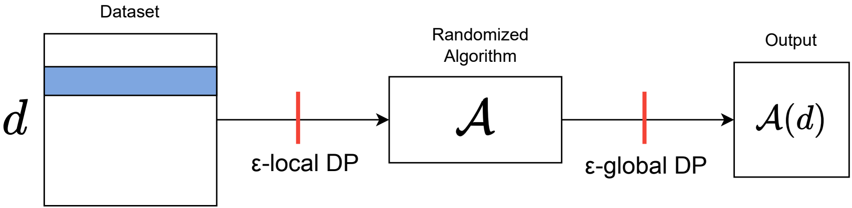
Definition (Dwork and Roth, 2014)

A randomised algorithm $\mathcal{A}$ satisfies ε-DP if for any two neighbouring datasets d and d' that differ only in one row and for all sets of output $\mathcal{O}$,

$$\mathbb{P}(\mathcal{A}(d) \in \mathcal{O}) \leq \exp(\varepsilon)\mathbb{P}\left(\mathcal{A}\left(d'\right) \in \mathcal{O}\right) .$$



Trust models for differentially private BAI



ε**-local** differential privacy:

☞ $\mathcal{A}$ has only access to private rewards.

ε**-global** differential privacy:

☞ $\mathcal{A}$ has access to the true rewards, but its output is private.

Local Differentially Private Best Arm Identification

Theorem

For all δ-correct ε-local DP algorithm and all instance ν ,

$$\mathbb{E}_{\nu}[\tau_\delta] \geq \max\left\{T^\star_{\mathrm{KL}}(\nu), c(\varepsilon)^{-1}T^\star_{\mathrm{TV}^2}(\nu)\right\} \log \frac{1}{2.4\delta} ,$$

with $c(\varepsilon) = \min\{4, e^{2\varepsilon}\}(e^{\varepsilon} - 1)^2$ *and* $T^\star_{\mathrm{TV}^2}(\nu) = T^\star_{\mathrm{KL}}(\nu_G)/2$ *.*

Local Differentially Private Best Arm Identification

Theorem

For all δ-correct ε-local DP algorithm and all instance ν ,

$$\mathbb{E}_\nu[\tau_\delta] \geq \max\left\{T^\star_{\mathrm{KL}}(\nu), c(\varepsilon)^{-1} T^\star_{\mathrm{TV}^2}(\nu)\right\} \log \frac{1}{2.4\delta} ,$$

with $c(\varepsilon) = \min\{4, e^{2\varepsilon}\}(e^\varepsilon - 1)^2$ *and* $T^\star_{\mathrm{TV}^2}(\nu) = T^\star_{\mathrm{KL}}(\nu_G)/2$ *.*

Two hardness regimes depending on ε and the environment ν .

- ☞ *Low-privacy*: $c(\varepsilon) > \frac{T^\star_{\mathrm{TV}^2}(\nu)}{T^\star_{\mathrm{KL}}(\nu)}$. **Privacy is for "free"**
- ☞ *High-privacy:* $c(\varepsilon) > \frac{T^\star_{\mathrm{TV}^2}(\nu)}{T^\star_{\mathrm{KL}}(\nu)}$. **Privacy scales the cost by $1/\varepsilon^2$**

CTB-TT: ε-local DP version of TTUCB

❶ Private estimator $\widetilde{\mu}_n$ based on randomised response:

☞ Observe private rewards $\widetilde{X}_n \sim \mathcal{B}\left(\frac{X_n(e^{\varepsilon}-1)+1}{e^{\varepsilon}+1}\right)$ instead of X_n .

❷ **Plug** $\widetilde{\mu}_n$ in TTUCB.

CTB-TT: ε-local DP version of TTUCB

❶ Private estimator $\widetilde{\mu}_n$ based on randomised response:

☞ Observe private rewards $\widetilde{X}_n \sim \mathcal{B}\left(\frac{X_n(e^\varepsilon - 1)+1}{e^\varepsilon + 1}\right)$ instead of X_n .

❷ **Plug** $\widetilde{\mu}_n$ in TTUCB.

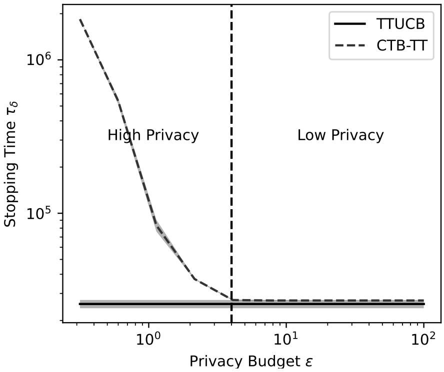
Theorem

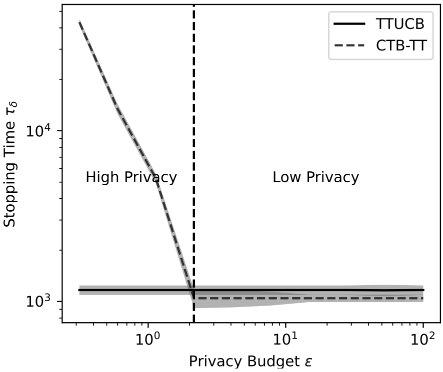
CTB-TT is ε-local DP, δ-correct and satisfies

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_\nu[\tau_\delta]}{\log(1/\delta)} \leq \left(1 + \frac{2}{e^\varepsilon - 1}\right)^2 T^\star_{\mathrm{TV}^2}(\nu) .$$

Empirical stopping time ($\delta = 0.01$)

(left) $\mu_1 = (0.95, 0.9, 0.9, 0.9, 0.5)$ and (right) $\mu_2 = (0.75, 0.7, 0.7, 0.7, 0.7)$.





Global Differentially Private Best Arm Identification

Theorem

For all δ-correct ε-global DP algorithm and all instance ν ,

$$\mathbb{E}_\nu[\tau_\delta] \geq \max\{T^\star_{\mathrm{KL}}(\nu), T^\star_{\mathrm{TV}}(\nu)/(6\varepsilon)\} \log \frac{1}{2.4\delta} ,$$

where $T^\star_{\mathrm{KL}}(\nu) \approx \sum_{i \neq i^\star} (\mu_{i^\star} - \mu_i)^{-2}$ *and* $T^\star_{\mathrm{TV}}(\nu) \approx \sum_{i \neq i^\star} (\mu_{i^\star} - \mu_i)^{-1}$ *.*

Global Differentially Private Best Arm Identification

Theorem

For all δ-correct ε-global DP algorithm and all instance ν ,

$$\mathbb{E}_\nu[\tau_\delta] \geq \max\{T^\star_{\mathrm{KL}}(\nu), T^\star_{\mathrm{TV}}(\nu)/(6\varepsilon)\} \log \frac{1}{2.4\delta} ,$$

where $T^\star_{\mathrm{KL}}(\nu) \approx \sum_{i \neq i^\star}(\mu_{i^\star} - \mu_i)^{-2}$ and $T^\star_{\mathrm{TV}}(\nu) \approx \sum_{i \neq i^\star}(\mu_{i^\star} - \mu_i)^{-1}$.

Two hardness regimes depending on ε and the environment ν .

- ☞ *Low-privacy regime*: $6\varepsilon > \frac{T^\star_{\mathrm{TV}}(\nu)}{T^\star_{\mathrm{KL}}(\nu)}$. **Privacy is for "free"**.
- ☞ *High-privacy regime:* $6\varepsilon < \frac{T^\star_{\mathrm{TV}}(\nu)}{T^\star_{\mathrm{KL}}(\nu)}$. **Privacy is "dominating"**.

AdaP-TT: ε-global DP version of TTUCB

❶ Private estimator with Laplace noise: $\widetilde{\mu}_n = \mu_n + \text{Lap}\left(\frac{1}{\varepsilon N_n}\right)$.

☞ **Doubling and forgetting**, i.e. phases per arm.

❷ **Plug** $\widetilde{\mu}_n$ in TTUCB.

☞ Private stopping threshold: $c(n, \delta) \approx \log(1/\delta) + \frac{1}{n\varepsilon^2} \log(1/\delta)^2$.

AdaP-TT: ε-global DP version of TTUCB

❶ Private estimator with Laplace noise: $\widetilde{\mu}_n = \mu_n + \text{Lap}\left(\frac{1}{\varepsilon N_n}\right)$.

☞ **Doubling and forgetting**, i.e. phases per arm.

❷ **Plug** $\widetilde{\mu}_n$ in TTUCB.

☞ Private stopping threshold: $c(n, \delta) \approx \log(1/\delta) + \frac{1}{n\varepsilon^2} \log(1/\delta)^2$.

Theorem

AdaP-TT is ε-global DP, δ-correct and satisfies

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_\nu[\tau_\delta]}{\log(1/\delta)} \leq 4T^\star_{\mathrm{KL},\beta}(\nu) \left(1 + \sqrt{1 + (\Delta_{\max}/\varepsilon)^2}\right) ,$$

which is $\mathcal{O}\left(\max\left\{T^\star_{\mathrm{KL}}(\nu), T^\star_{\mathrm{TV}}(\nu)/\varepsilon\right\}\right)$ for most instances.

AdaP-TT: ε-global DP version of TTUCB

❶ Private estimator with Laplace noise: $\widetilde{\mu}_n = \mu_n + \text{Lap}\left(\frac{1}{\varepsilon N_n}\right)$.

☞ **Doubling and forgetting**, i.e. phases per arm.

❷ **Plug** $\widetilde{\mu}_n$ in TTUCB.

☞ Private stopping threshold: $c(n, \delta) \approx \log(1/\delta) + \frac{1}{n\varepsilon^2} \log(1/\delta)^2$.

Theorem

AdaP-TT is ε-global DP, δ-correct and satisfies

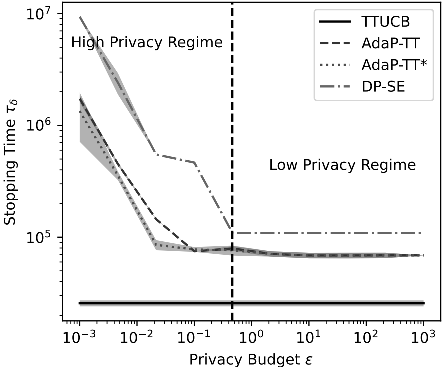
$$\limsup_{\delta \to 0} \frac{\mathbb{E}_\nu[\tau_\delta]}{\log(1/\delta)} \leq 4T^\star_{\text{KL},\beta}(\nu) \left(1 + \sqrt{1 + (\Delta_{\max}/\varepsilon)^2}\right),$$

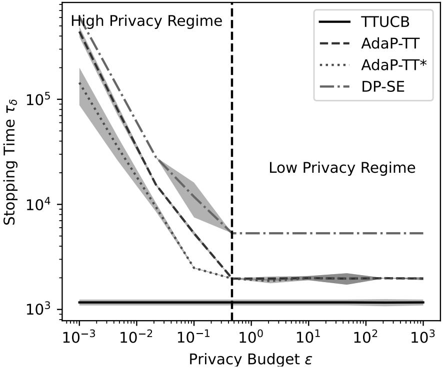
which is $\mathcal{O}\left(\max\left\{T^\star_{\text{KL}}(\nu), T^\star_{\text{TV}}(\nu)/\varepsilon\right\}\right)$ *for most instances.*

AdaP-TT* algorithm: modified private transportation costs.

Empirical stopping time ($\delta = 0.01$)

(left) $\mu_1 = (0.95, 0.9, 0.9, 0.9, 0.5)$ and (right) $\mu_2 = (0.75, 0.7, 0.7, 0.7, 0.7)$.





Conclusion

Differentially Private Best Arm Identification:

- ☞ ε-local and ε-global trust models,
- ☞ lower bounds on the expected sample complexity,
- ☞ matching upper bounds for modified TTUCB.

Perspectives:

- other trust models, e.g. shuffle DP,
- other DP settings, e.g. (ε, δ)-DP or Rény-DP.

References

Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407.

Garivier, A. and Kaufmann, E. (2016). Optimal best arm identification with fixed confidence. In *Proceedings of the 29th Conference On Learning Theory*.

Jourdan, M. and Degenne, R. (2023). Non-asymptotic analysis of a ucb-based top two algorithm. *Thirty-Seventh Conference on Neural Information Processing Systems*.