The Price of Differential Privacy for Online Learning

Naman Agarwal, Karan Singh
Computer Science, Princeton University

PRINCETON UNIVERSITY

Framework: Private Online Learning

Online Learning is a framework for sequential decision making that offers distribution-free learning guarantees. Consequently, it is well suited to dynamic and adversarial environments where real-time learning from changing data is crucial.

Formal Setup: On each round $t = 1, 2, \dots T$

- The learner predicts $x_t \in \mathcal{X} \subseteq \mathbb{R}^N$ (convex) .
- The adversary chooses a loss vector $l_t \in \mathcal{Y}$.
- The learner suffers $\langle l_t, x_t \rangle$ and observes l_t in the **full-information** setting (and, in contrast, only $\langle l_t, x_t \rangle$ under **bandit** feedback).

$$\text{Regret} = \mathbb{E}\Bigg[\underbrace{\sum_{t=1}^{T} \langle l_t, x_t \rangle}_{\text{Loss of the learner}} - \underbrace{\min_{x \in \mathcal{X}} \sum_{t=1}^{T} \langle l_t, x \rangle}_{\text{Loss of the best fixed decision}} \Bigg]$$

$$O(\sqrt{T}) \text{ Regret} \implies O\left(\frac{1}{\varepsilon^2}\right) \text{ Sample Complexity}$$

Privacy Guarantee: A randomized online learning algorithm $\mathcal{A}$ is **ε-differentially private** if whenever

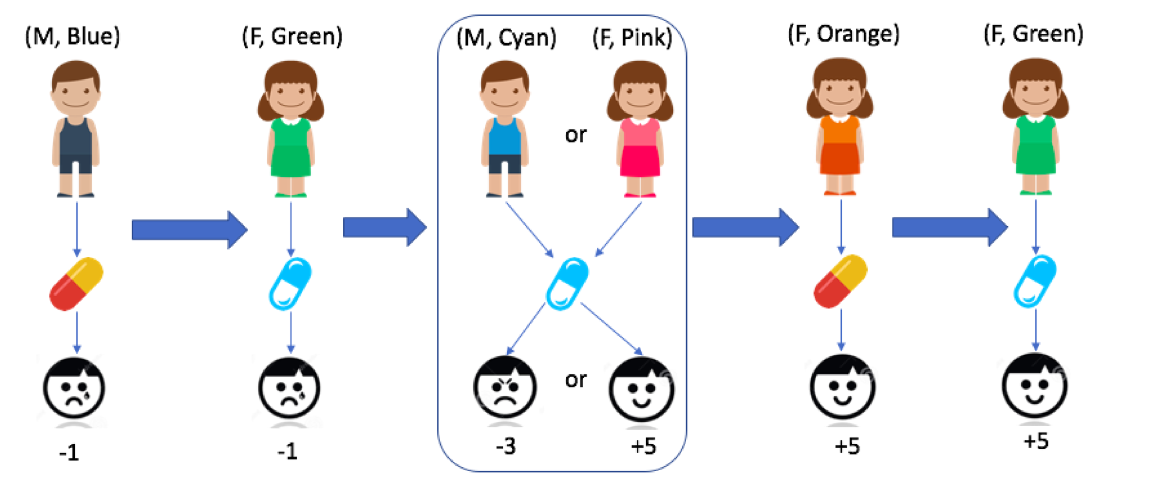
$$L = (l_1, \dots, l_t, \dots l_T) \xrightarrow{\mathcal{A}} (x_1, \dots x_T)$$
$$L' = \underbrace{(l_1, \dots, l'_t, \dots l_T)}_{\text{Input}} \xrightarrow{\mathcal{A}} \underbrace{(x'_1, \dots x'_T)}_{\text{Output}},$$

for any possible set $S \subseteq \mathcal{X}^T$ of output sequences

$$\mathbb{P}(\underbrace{(x_1, \dots x_T)}_{\text{Output of } \mathcal{A} \text{ on } L} \in S) \leq e^{\varepsilon} \mathbb{P}(\underbrace{(x'_1, \dots x'_T)}_{\text{Output of } \mathcal{A} \text{ on } L'} \in S).$$

$$\text{Price} = \lim_{T \to \infty} \left(\frac{\varepsilon\text{-DP Regret}(T)}{\text{Non-private Regret}(T)} - 1 \right)$$

Illustration: The Promise of Privacy



Loss Vector $_{\text{(in OL)}}$ ≡ Feature Vector + Reward $_{\text{(in Sup L)}}$

Our Contributions

Full-Information Setting

Meta-Theorem: Any regularization-based low-regret algorithm can be adapted to achieve

$$\text{Regret}_{\varepsilon\text{-DP}} = \text{Regret}_{\text{Non-private}} + O\left(\frac{\log^2 T}{\varepsilon}\right)$$

while ensuring ε-differential privacy.

- Privacy is Free! as long as $\varepsilon \geq \frac{1}{\sqrt{T}}$.
- Previous best[JKT12,ST13] scale as $O\left(\frac{\sqrt{T}}{\varepsilon}\right)$.
- Adapts to the **Geometry** of the problem.
  - Optimal dependence on N.

Bandit Feedback

Meta-Theorem: Any low-regret bandit algorithm can be adapted to achieve

$$\text{Regret}_{\varepsilon\text{-DP}} = O\left(\frac{\text{Regret}_{\text{Non-private}}}{\varepsilon}\right)$$

while ensuring ε-differential privacy.

- Optimal $O(\sqrt{T})$ dependence on #rounds.
- Previous best[ST13] scale as $O(T^{\frac{2}{3}})$.
- Works for general convex sets.

Full-Information Algorithm

FTRL Template

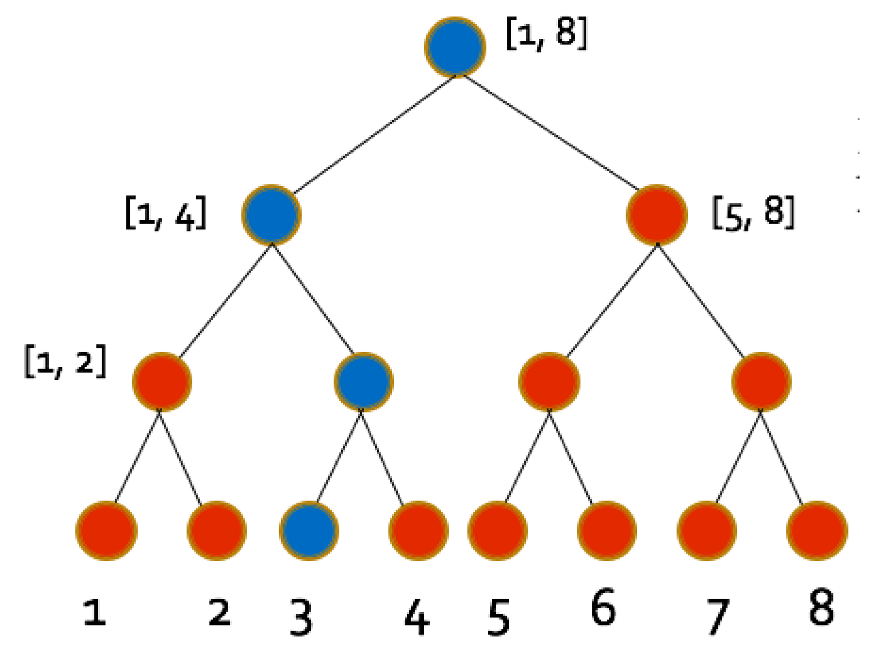
Initialize an empty binary tree B to compute differentially private estimates of $\sum_{i=1}^{t} l_i$.
for $t = 1$ *to* T **do**
  $x_t = argmin_{x \in \mathcal{X}} (\eta \langle x, \tilde{L}_{t-1} \rangle + R(x))$.
  $\tilde{L}_{t-1} \approx \sum_{i=1}^{t} l_i$+noise
  Observe l_t, and suffer a loss of $\langle l_t, x_t \rangle$.
  $(\tilde{L}_t, B) \leftarrow$ `TreeBasedAgg`(l_t, B).

Tree-based Aggregation

Input: A sequence of vectors $(l_1, \dots l_T)$.

Output: ε-DP estimates $\tilde{L}_t$ of sums $\left(\sum_{i=1}^{t} l_i\right)$.

Utility: $|\tilde{L}_t - \sum_{i=1}^{T} l_i| \approx \frac{\log^2 T}{\varepsilon}$. [DNPR10, JKT12]



Summary of Our Results

Full-Information Setting

	Previous Best	Our Regret Bound	Non-private
Expert Advice	$\tilde{O}\left(\frac{\sqrt{T}\log N}{\varepsilon}\right)$ [DR14]	$O\left(\sqrt{T \log N} + \frac{N \log N \log^2 T}{\varepsilon}\right)$	$O(\sqrt{T \log N})$
Sphere	$\tilde{O}\left(\frac{\sqrt{NT}}{\varepsilon}\right)$ [ST13]	$O\left(\sqrt{T} + \frac{N \log^2 T}{\varepsilon}\right)$	$O(\sqrt{T})$
Cube	$\tilde{O}\left(\frac{\sqrt{NT}}{\varepsilon}\right)$ [ST13]	$O\left(\sqrt{NT} + \frac{N \log^2 T}{\varepsilon}\right)$	$O(\sqrt{NT})$
General OLO*	$\tilde{O}\left(\frac{\sqrt{T}}{\varepsilon}\right)$ [ST13]	$O\left(\sqrt{T} + \frac{\log^2 T}{\varepsilon}\right)$	$O(\sqrt{T})$

Bandit Feedback

	Previous Best	Our Regret Bound	Non-private
Multi-armed Bandits	$\tilde{O}\left(\frac{NT^{\frac{2}{3}}}{\varepsilon}\right)$ [ST13]	$\tilde{O}\left(\frac{\sqrt{TN \log N}}{\varepsilon}\right)$	$O(\sqrt{NT})$
Bandit Linear Optimization*	$\tilde{O}\left(\frac{T^{\frac{2}{3}}}{\varepsilon}\right)$ [ST13]	$\tilde{O}\left(\frac{\sqrt{T}}{\varepsilon}\right)$	$O(\sqrt{T})$

Bandit Algorithm

Reduction to Non-private Setting

Require: Bandit Algorithm $\mathcal{A}$.
for $t = 1$ *to* T **do**
  Receive x_t from $\mathcal{A}$ and output x_t.
  Receive a loss value $\langle l_t, x_t \rangle$ from the adversary.
  Sample $Z_t \sim Lap\left(\frac{1}{\varepsilon}\right)$.
  Forward $\langle l_t, x_t \rangle + \langle Z_t, x_t \rangle$ as input to $\mathcal{A}$.

Key Points: Regret Analysis

Full-Information Setting

- The Tree-based Aggregation scheme adds $\approx \frac{\log^2 T}{\varepsilon}$ noise on the true cumulative sums.
- Treating these perturbations as worst-case loss vectors leads to $O\left(\frac{\sqrt{T}\log^2 T}{\varepsilon}\right)$ regret.
- **(FTPL Analysis)** Once these perturbations are made identical in distribution, the regret of the proposed algorithm is the same as that of FTRL algorithm injecting all noise at $t = 0$.

Bandit Feedback

- Since bandit algorithms utilize importance sampling, adding a perturbation of Z_t drastically reduces the stability.
  - A careful analysis leads to $O(T^{\frac{2}{3}})$ regret.
- A perturbation of $\langle Z_t, x_t \rangle$ permits one to *pretend* the magnitude of loss vector is $\approx \frac{1}{\varepsilon}$.

See ↑ for more details.

References

[DNPR10] Dwork, Cynthia, Naor, Moni, Pitassi, Toniann, and Rothblum, Guy N. Differential privacy under continual observation. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pp. 715–724. ACM, 2010.

[JKT12] Jain, Prateek, Kothari, Pravesh, and Thakurta, Abhradeep. Differentially private online learning. In *COLT*, volume 23, pp. 24–1, 2012.

[ST13] Smith, Adam and Thakurta, Abhradeep Guha. (nearly) optimal algorithms for private online learning in full-information and bandit settings. In *Advances in Neural Information Processing Systems*, pp. 2733–2741, 2013.

[DR14] Dwork, Cynthia, Roth, Aaron. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 211–407, 2014..