Privacy-Preserving Bandits

Mohammad Malekzadeh[1]**, Dimitrios Athanasakis**[2]**, Hamed Haddadi**[2,3]**, Benjamin Livshits**[2,3]

[1]Queen Mary University of London, [2]Brave Software, [3]Imperial College London

m.malekzadeh@qmul.ac.uk, {dathanasakis,hamed,ben}@brave.com

Problem

- Online agents process our private data to provide personalization.
- To protect privacy, we can run agents locally (on users' devices).
- Local agents require longer to produce useful recommendations.

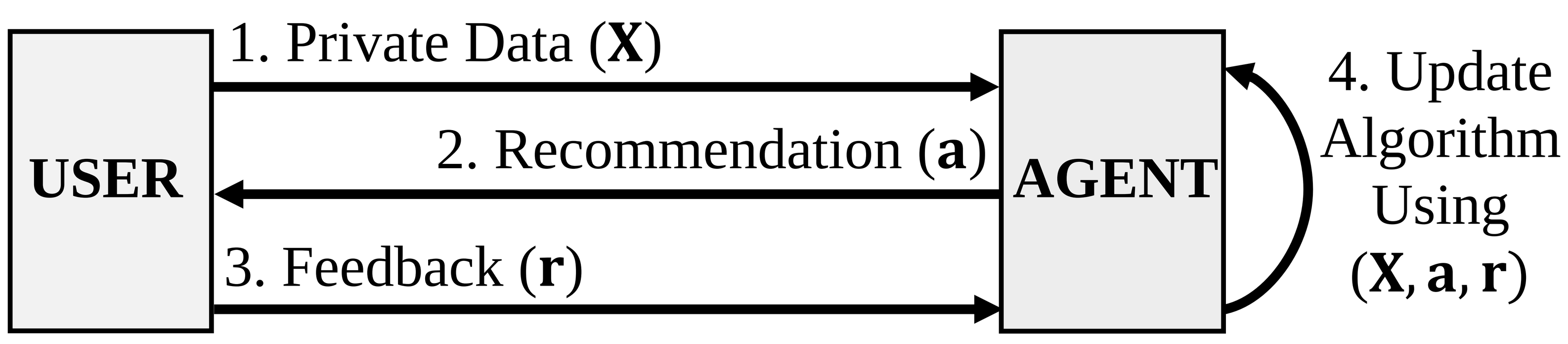

Figure 1: Overview of a Contextual Bandit Algorithm for Online Recommendation.

Contribution

$*$ P2B, a system for updating local agents, by collecting feedback from other agents, in a differentially-private manner.

$*$ We show that P2B can result in a small $\epsilon < 1$ value for differential privacy; a concrete and desirable privacy guarantee.

$*$ P2B is competitive in terms of predictive utility with approaches that provide no privacy protections. At the same time, it substantially outperforms local cold-start agents that do not share data.

Architecture

1. Every user runs their own *local agent*.
2. With probability p, each agent sends an encoded data to the *Shuffler*.
3. Shuffler periodically sends a refined batch of data to the *server*.
4. Upon receiving a new batch, the server updates the *global model*.
5. The global model is used as a *warm-start* by new local agents.

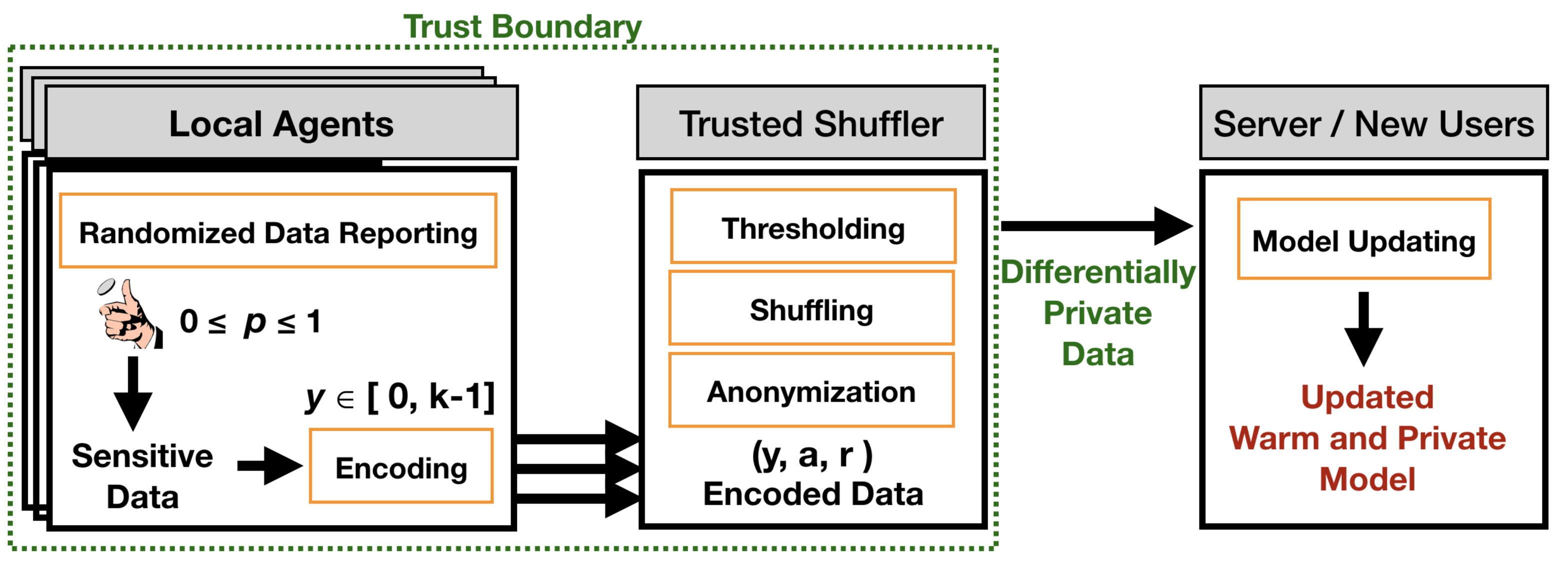

Figure 2: System Architecture for Privacy-Preserving Model Updating in P2B.

Methodology

- Randomized Data Reporting: probability p.
- Encoder: $\mathbf{x} \rightarrow y \in \{1, 2, \ldots, k\}$.
- Shuffler: An instance of ESA (PROCHLO) architecture [1].
  - o *Anonymization*: eliminating all the meta-data.
  - o *Shuffling*: shuffling the order of received data.
  - o *Thresholding*: to ensure every data blends in a crowd [2].

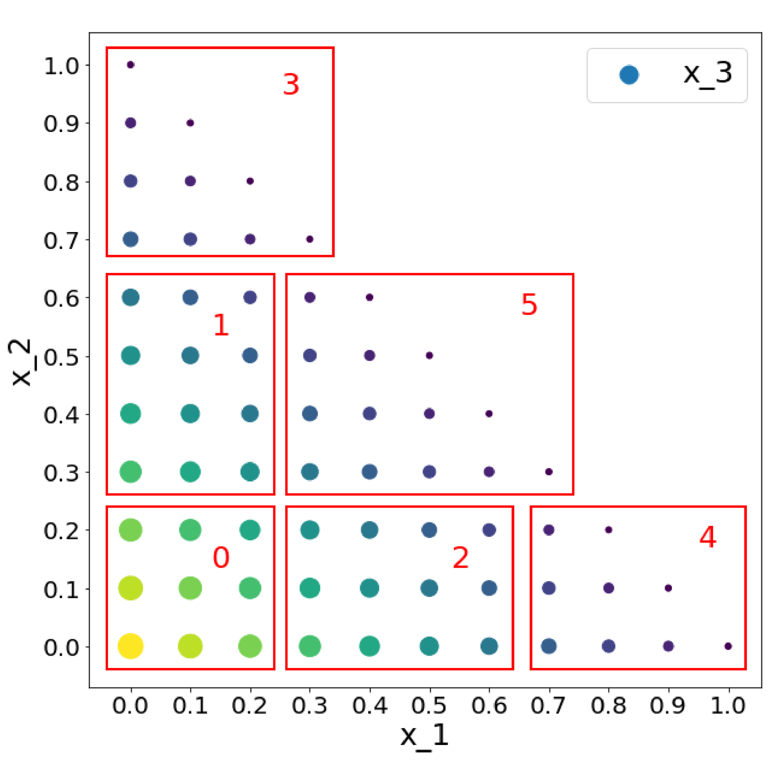

Figure 3: System architecture for P2B.

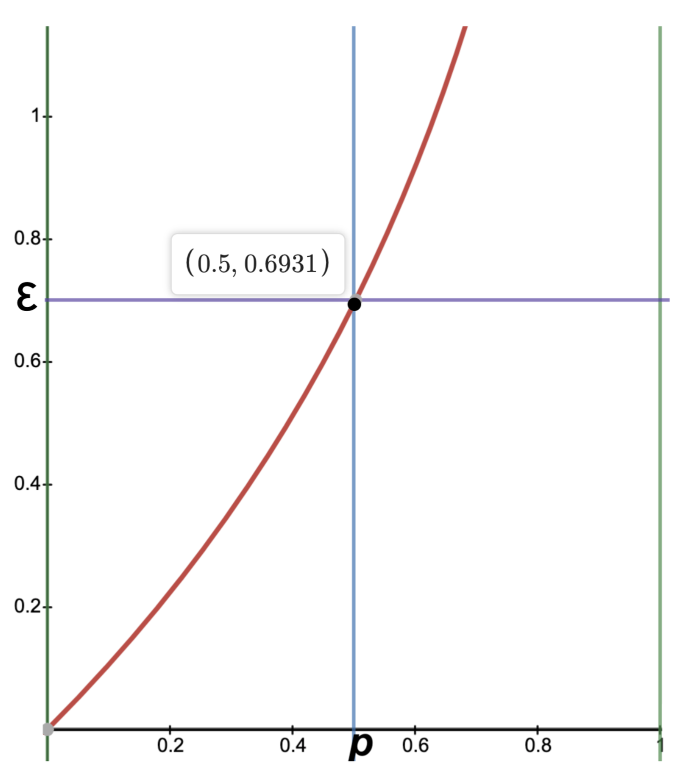

Figure 4: ϵ as a result of the p.

Privacy

$*$ **Pre-Sampling**: only a p fraction of users send data to the server.

$*$ **Crowd-Blending [2]**: each user hides in a crowd of size $> l = U/k$.

$*$ **Differential Privacy**: The combination of:

1. pre-sampling with probability p, and
2. $(l, \bar{\epsilon} = 0)$−crowd-blending.

$$\epsilon = \ln\left(p \cdot \left(\frac{2-p}{1-p}e^{\bar{\epsilon}}\right) + (1-p)\right) \quad \text{and} \quad \delta = e^{-\Omega(l \cdot (1-p)^2)}.$$

Evaluation

- Contextual Linear Upper Confidence Bound algorithm [3].
  - o **Cold**: no communication to the server at any point.
  - o **Warm and Non-Private**: sending data in its original form: $\mathbf{x}$.
  - o **Warm and Private**: sending data via P2B in the encoded form: y.

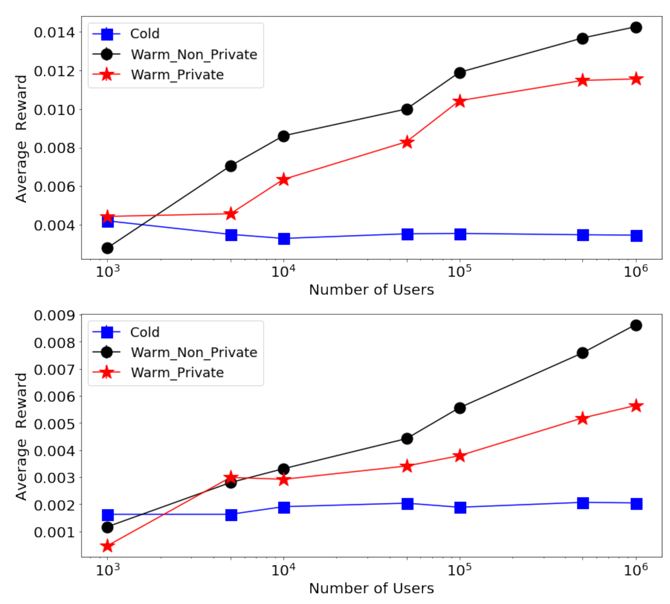

Figure 5: Synthetic Benchmarks: (Top) $A = 20$ and (Bottom) $A = 50$. For all: $d = 10$ and $T = 10$. The expected reward in this setting has a strong dependence on number of arms as agents will spend considerable time exploring alternative actions.

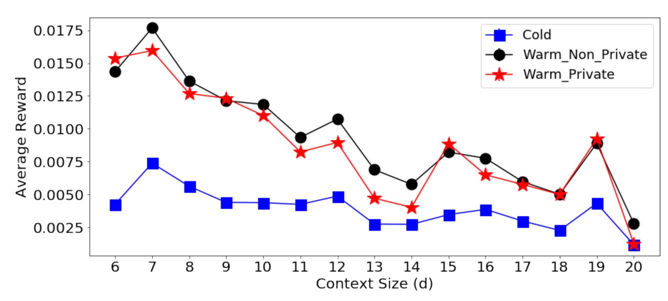

Figure 6: Synthetic Benchmarks: $U = 20000$, $A = 20$, $T = 20$, and $d = \{6, 7, \ldots, 20\}$. As the dimensionality of the context increases the average reward for this settings is reduced as agents spend more time trying to explore their environment.

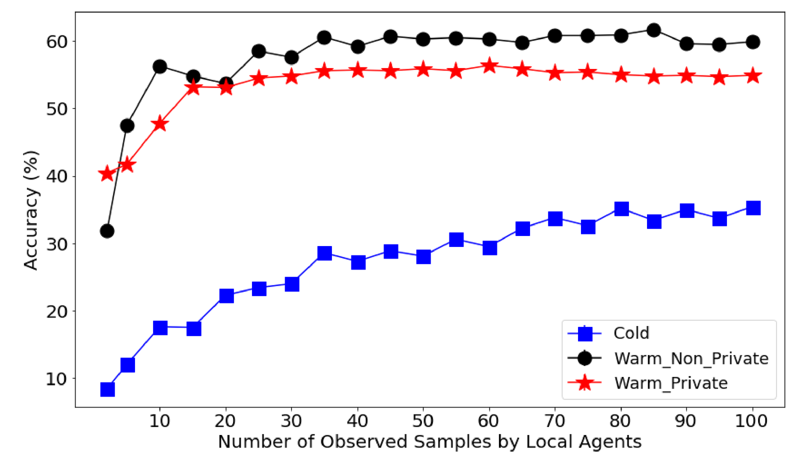

Figure 7: Multi-Label Dataset accuracy: Text-Mining with $d = 20$ and $A = 20$. As local agents observe more interactions they obtain better accuracy. This has a multiplicative effect in the distributed settings where agents reach to the plateau much faster.

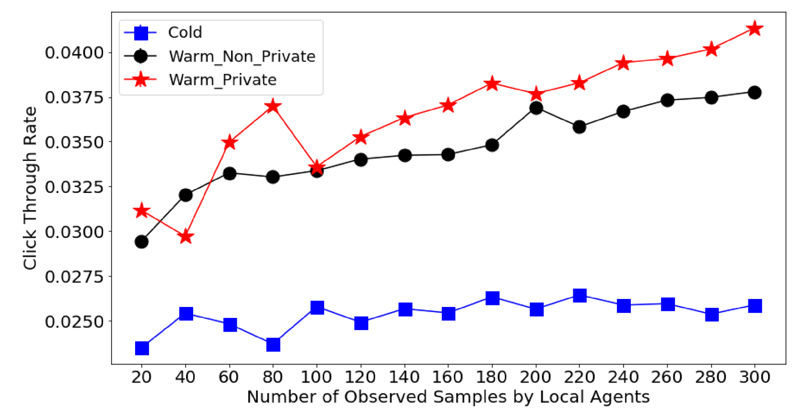

Figure 8: Criteo results. $d = 10$, $A = 40$, $k = 2^7$. The private and non-private agents obtain similar performances for low numbers of local interactions. As the number of interactions increase the private agents perform better than their non-private counterparts.

Conclusions

$*$ P2B: the intersection of differentially-private data collection and contextual bandit algorithms for privacy-preserving personalization.

$*$ A particularly viable option for settings where large user populations participate in a large amount of local interactions, where the performance penalty for privacy is vanishingly small.

$*$ As future work, we aim to study the behavior of more encoding approaches as well as their interplay with alternative contextual bandit algorithms.

References

[1] Andrea Bittau, Ifar Erlingsson, Petros Maniatis, Ilya Mironov, Ananth Raghunathan, David Lie, Mitch Rudominer, Usharsee Kode, Julien Tinnes, and Bernhard Seefeld. Prochlo: Strong Privacy for Analytics in the Crowd. 2017.

[2] Johannes Gehrke, Michael Hay, Edward Lui, and Rafael Pass. Crowd-blending privacy. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012.

[3] Wei Chu, Lihong Li, Lev Reyzin, and Robert E. Schapire. Contextual bandits with linear Payoff functions. In *Journal of Machine Learning Research*, 2011.