MUTUAL-INFORMATION-PRIVATE ONLINE GRADIENT DESCENT ALGORITHM

Ruochi Zhang, Parv Venkitasubramaniam

Electrical and Computer Engineering
Lehigh University, USA
{ruz614, parv.v}@lehigh.edu

ABSTRACT

A user implemented privacy preservation mechanism is proposed for the online gradient descent (OGD) algorithm. Privacy is measured through the information leakage as quantified by the mutual information between the users outputs and learners inputs. The input perturbation mechanism proposed can be implemented by individual users with a space and time complexity that is independent of the horizon T. For the proposed mechanism, the information leakage is shown to be bounded by the Gaussian channel capacity in the full information setting. The regret bound of the privacy preserving learning mechanism is identical to the non private OGD with only differing in constant factors.

Index Terms— Online machine learning, privacy, information theory, convex optimization, gradient descent.

1. INTRODUCTION

In this paper, we study privacy leakage in an emerging field of study: online machine learning. The goal of online learning is to make a sequence of accurate predictions given knowledge of the correct answer to previous prediction tasks and possibly additional available information [1]. Big data applications such as targeted advertising and online ranking has burgeoned the interest in developing efficient online learning algorithms. It has also been shown that online learning techniques can be used to obtain results for online posted price mechanisms and online auctions [2].

In practice, the learner is usually the service provider who gathers large amounts of personal information about the users of the service, and would in all likelihood contain individual data that is private or sensitive. Although individuals are willing to share their data, they are not expecting the disclosure of identities [3]. For example, a navigation app user wants to get current traffic condition model without his/her position being perfectly shared. In fact, the goal of learning is to uncover "relationships" or "trends" from historical data, which might be possible to be separated from the information of individual identities [3]. The purpose of this work is to propose a user driven privatization mechanism that allows the learner to infer the desired trends and patterns without compromising an individual users privacy.

It has been shown that many online learning problems and algorithms can be analyzed based on the online convex optimization model. In online convex optimization (OCO), the hypothesis set and the loss functions are forced to be convex to obtain stronger learning bounds. The model is based on a convex set and a sequence of convex cost functions.

Due to the popularity of OCO, considerable efforts have been devoted to designing differential private algorithms for the OCO problem [4, 5, 6]. Most of these works consider adversary who observes the output of the online learning algorithm and extracts sensitive data from users, and a differentially private algorithm guarantees that the statistics from observation will not deviate too much if at most one individual alters his input data [4].

In this work, we consider the problem from another perspective — we consider the situation where the adversary can observe the input data of online learning system, and we use mutual information as a measure of the information leaked from the user to the adversary [7]. We note that this approach, even in the absence of an explicit adversary, protects the users sensitive data from the service provider running the learning algorithm.

In this work, we modify the online gradient descent algorithm and provide a mutual-information-private version where we introduce an encryption layer to encrypt data through an additive noise mechanism. The scheme we propose can be implemented by each individual user with a space and time complexity that is independent of the learning horizon. We propose solutions for both the full information setting and the bandit setting.

Related work

There are several approaches that ensure privacy in online learning [4, 5, 6]. In [4], the authors provide a differentially private OCO method. This algorithm is generated by adding i.i.d. Gaussian noise into the output of non-private OCO algorithm, and taking a projection in case the output is out of

The authors would like to gratefully acknowledge the support of the National Science Foundation through the grants CCF-1149495 and CCF-1617889.

the convex set. The authors convert two popular OCO algorithms (Implicit Gradient Descent and Generalized Infinitesimal Gradient Ascent) into corresponding differentially private algorithms while guaranteeing $O(\sqrt{T})$ regret bounds. In [5], the authors consider online prediction from expert advice in a situation where each expert observes its own loss at each time while the loss cannot be disclosed to others. The regret bound (performance) of their proposed exponential weighting scheme is the same or almost the same as the well-known exponential weighting scheme in the full information model. In [6], the authors consider online learning with distributed data sources. The autonomous learners update local parameters based on local data sources and periodically exchange information with a small subset of neighbors in a communication network. This approach limits the power of the learner and the regret bound for strongly convex functions is still sublinear. A gradient descent based additive noise mechanism has been studied in the context of deep learning [8].

However, these encryption methods are implemented from the learner's side — the data is encrypted inside or after the online learning algorithm. Although we can guarantee differential privacy if the adversary has only the access to the output of learning system, we cannot claim privacy if the adversary has the access to the input of learning system, or even the learner itself is malicious. The method we propose is to encrypt the data from users' side, which is a very conservative way to protect sensitive information.

2. MUTUAL-INFORMATION-PRIVATE ONLINE GRADIENT DESCENT ALGORITHM

We will first describe briefly the online convex optimization (OCO) model. This model has potential to be highly applicable since many machine learning optimization problems are indeed convex [9]. In this model, the data from user is represented by a convex cost function and the learner's objective is to minimize the total cost by predicting appropriate parameters.

Definition 1 (Online Convex Optimization)**.** *Consider an online learning system that receives a stream of functions $(f_1, f_2, \cdots, f_T)$ and each $f_t : \mathcal{S} \to \mathbb{R}$ is a convex cost function representing data from one individual. The system is required to output a sequence of parameter estimates $(w_1, w_2, \cdots, w_T)$ with $w_t \in \mathcal{S} \subset \mathbb{R}^d$ that minimizes the total errors $\sum_{t=1}^{T} f_t(w_t)$. Due to causality, for every t, the algorithm computes w_t based only on $(f_1, f_2, \cdots, f_{t-1})$. We seek an algorithm $\mathcal{A}$ that minimize the* ***regret*** *defined by*

$$\mathrm{Regret}_T(\mathcal{A}) = \sum_{t=1}^{T} f_t(w_t) - \min_{w \in \mathcal{S}} f_t(w)$$

We consider situations where 1) the input functions $(f_1, f_2, \cdots, f_{t-1})$ are $L-$Lipschitz continuous, and 2) the hypothesis space $\mathcal{S}$ is bounded w.r.t. l^2-norm. Under these restrictions, the OCO problem can be solved by the online gradient descent (OGD) algorithm [10].

The OGD algorithm only takes the sub-gradient $z_t \in \partial f_t(w_t)$ as input. It is noticeable that user individual information may be inferred if an accurate sub-gradient is provided to the learner. We use linear regression as an example, where the learner wants to predict y_t by a feature vector x_t with a linear function $\langle w_t, x_t \rangle$, and the loss function has the form $f_t(w_t) = |\langle w_t, x_t \rangle - y_t|$. In this case, the sub-gradient will be $z_t = \pm x_t$ which reveals the feature vector x_t almost completely.

In light of this, we propose a privacy-preserving encryption in which each individual user adds independent noise to the sub-gradient and sends the modified sub-gradient value $\widetilde{z}_t$ to the learner. Or equivalently, the user can generate a modified cost function $\widetilde{f}_t : \mathcal{S} \to \mathbb{R}$ based on the perturbed sub-gradient value. The algorithm is shown in Algorithm 1 and Fig. 1.

Algorithm 1 Mutual-information-private OGD

1: **Encryption layer:**
2: Receive w_t from the learner
3: Pick a sub-gradient $z_t \in \partial f_t(w_t)$
4: Output $\widetilde{z}_t = z_t + v_t$ to the learner, where $v_t \sim \mathcal{N}(0, \sigma^2 I)$ is independent generated Gaussian noise
5: **Learner:**
6: Receive $\widetilde{z}_t$ from the encryption layer
7: Update $\theta_{t+1} = \theta_t - \widetilde{z}_t$, (initialize $\theta_1 = 0$)
8: Predict $w_{t+1} = \arg\min_{w \in \mathcal{S}} \|w - \eta\theta_{t+1}\|$

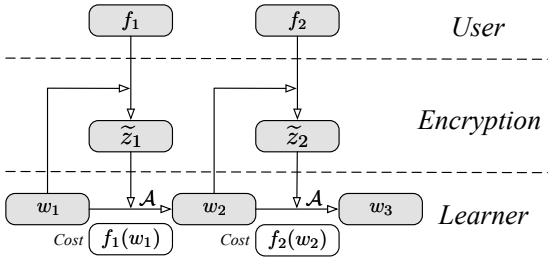

Fig. 1. Mutual-information-private OGD algorithm.

By simply adding noise into the sub-gradient before sending the cost function to the learner, we guarantee that the information leakage of individual user is upper-bounded when measured by mutual information. In fact, the additive noise results in a Gaussian channel between user and learner, which restricts the information flow by its channel capacity.

Theorem 1 (Privacy Guarantee)**.** *The noise adding mechanism in Algorithm 1 is $C-$mutual-information private. i.e., $I(f_k; \widetilde{z}_k) < C$ for every k, where $C = \frac{d}{2}\log(1 + \frac{L^2}{d\sigma^2})$*

Proof. Since $f_t \to z_t \to \widetilde{z}_t$ forms a Markov chain, by the data processing theorem, it suffices to show that $I(z_t; \widetilde{z}_t) < C$. Since f_t is L-Lipschitz w.r.t. norm $\|\cdot\|_2$, we have $\|z_t\|_2 \leq L$, that is, the power of z_t is bounded. Since $\mathbb{E}\|z_t\|_2^2 \leq L^2$ and $z_t \to \widetilde{z}_t$ is a Gaussian channel, we have the result. □

Now we analyze the learning performance of this algorithm given the learner is running a OGD algorithm. Since randomness is introduced in the algorithm, it is reasonable to use the average regret as a measure of performance (see Definition 2). It turns out that OGD algorithm is resistant to independent zero-mean noise-adding, we can show that the following regret is still sub-linear to the time horizon T. In fact, the regret is bounded by $B\sqrt{(L^2 + d\sigma^2)T}$ while the non-private OGD algorithm has a regret bound of $BL\sqrt{T}$ [1]. Therefore, this private algorithm has a similar regret bound $O(\sqrt{T})$ as the non-private version with an increment on constant factors.

Definition 2. *In a randomized algorithm $\mathcal{A}$, the regret is defined as*

$$\text{Regret}(\mathcal{A}) = \max_{u \in \mathcal{S}} \mathbb{E}\left\{\sum_{t=1}^{T}[f_t(w_t) - f_t(u)]\right\} \quad (1)$$

where $u \in \mathcal{S}$ is any fixed parameter in hypothesis space $\mathcal{S}$ and the expectation is taking with respect to the randomness of $\{v_k\}_{k=1}^T$.

Theorem 2 (Regret Guarantee)**.** *Let $\mathcal{A}_1$ be Algorithm 1, f_t is L-Lipschitz continuous for every t and $\max_{u\in\mathcal{S}}\|u\|_2 \leq B$. Then $\text{Regret}(\mathcal{A}_1)$ is sub-linear to T. Specifically,*

$$\text{Regret}(\mathcal{A}_1) \leq \frac{B^2}{2\eta} + \frac{\eta}{2}T(L^2 + d\sigma^2)$$

In particular, by setting $\eta = B/\sqrt{(L^2 + d\sigma^2)T}$ we obtain the bound $\text{Regret}(\mathcal{A}_1) \leq B\sqrt{(L^2 + d\sigma^2)T}$.

Proof. Starting from the definition of sub-gradient, since $z_t \in \partial f_t(w_t)$, for every $u \in \mathcal{S}$ we have

$$f_t(w_t) - f_t(u) \leq \langle w_t - u, z_t\rangle \quad (2)$$

Note that the noise v_t is zero-mean and independently generated, by the law of total probability we have

$$\mathbb{E}\left[\langle w_t - u, z_t\rangle\right] = \mathbb{E}\left[\langle w_t - u, \widetilde{z}_t\rangle\right] \quad (3)$$

Combining (2), (3) and summing over time we get

$$\mathbb{E}\left\{\sum_{t=1}^{T}[f_t(w_t) - f_t(u)]\right\} \leq \mathbb{E}\left[\sum_{t=1}^{T}\langle w_t - u, \widetilde{z}_t\rangle\right] \quad (4)$$

The sequence $\{w_t\}_{t=1}^T$ is the output of OGD algorithm running on the input sequence $\{\widetilde{z}_t\}_{t=1}^T$. Therefore (see Eq. 2.15 in [1]),

$$\sum_{t=1}^{T}\langle w_t - u, \widetilde{z}_t\rangle \leq \frac{1}{2\eta}\|u\|_2^2 + \frac{\eta}{2}\sum_{t=1}^{T}\|\widetilde{z}_t\|_2^2 \quad (5)$$

Since $\widetilde{z}_t = z_t + v_t$ where z_t is a sub-gradient of $L-$Lipschitz function, v_t is independent of z_t and $\mathbb{E}\|v_t\|_2^2 \leq d\sigma^2$, the expectation of last term is bounded by

$$\begin{aligned}\mathbb{E}\left[\frac{\eta}{2}\sum_{t=1}^{T}\|\widetilde{z}_t\|_2^2\right] &\leq \frac{\eta}{2}\sum_{t=1}^{T}\left[\mathbb{E}\|z_t\|_2^2 + \mathbb{E}\|v_t\|_2^2\right] \\ &\leq \frac{\eta T}{2}(L^2 + d\sigma^2) \quad (6)\end{aligned}$$

By (1), (4)-(6) and $\|u\|_2 \leq B$ the proof is complete. □

3. EXTENSION TO BANDIT SETTING

The bandit setting is useful in the situation where the learner (or even the user) only knows the value of the loss function but he doesn't know the value of the loss function at other points. We can easily extend the additive noise method in Algorithm 1 to bandit setting, where the encryption layer consists only of the value of $f_t(\cdot)$ at each step t.

We adapt our approach in Section 2 to the Bandit Online Gradient Descent (OGD) algorithm [11]. In our approach, the encryption layer computes an estimate of the gradient and adds Gaussian noise at each step. Note that the encryption layer can be implemented at the user's side since no input from other individual is required in this algorithm. We present our approach in Algorithm 2. As will be shown in Theorems 3 and 4, this algorithm has a comparable privacy and regret guarantee as Algorithm 1 (the full information setting). Furthermore, the regret bound achievable by the algorithm $O(T^{3/4})$ is also identical to the non-private version with only constant factors differing in the proof.

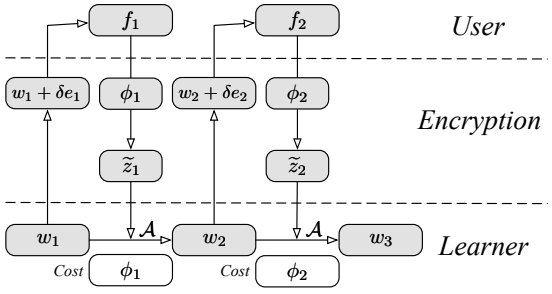

Fig. 2. Mutual-information-private OGD algorithm — Bandit setting.

Theorem 3 (Privacy Guarantee)**.** *Let $F = \max_{u\in\mathcal{S}, t\geq 1} f_t(u)$. If $F < \infty$, the noise adding mechanism in Algorithm 2 is $C-$mutual information private. i.e., $I(f_t; \widetilde{z}_t) < C$ for every t, where $C = \frac{d}{2}\log(1 + \frac{d(F/\delta + L)^2}{\sigma^2})$*

Proof. Since $f_t \to z_t \to \widetilde{z}_t$ forms a Markov chain, by the data processing theorem, it suffices to show that $I(z_t; \widetilde{z}_t) <$

Algorithm 2 Mutual-information-private OGD: Bandit setting

1: **Encryption layer:**
2: Receive w_t from the learner
3: Pick $e_t \sim U_{sp}$, where U_{sp} is the uniform distribution over the unit sphere $\{u : \|u\|_2^2 = 1\}$.
4: Send $w_t + \delta e_t$ to the user
5: Receive cost value $\phi_t = f_t(w_t + \delta e_t)$ from the user
6: $z_t = \frac{d}{\delta}\phi_t e_t$
7: Output $\widetilde{z}_t = z_t + v_t$ to the learner, where $v_t \sim \mathcal{N}(0, \sigma^2 I)$ is independent generated Gaussian noise
8: **Learner:**
9: Receive $\widetilde{z}_t$ from the encryption layer
10: Update $\theta_{t+1} = \theta_t - \widetilde{z}_t$, (initialize $\theta_1 = 0$)
11: Predict $w_{t+1} = \arg\min_{w \in \mathcal{S}} \|w - \eta\theta_{t+1}\|$

C. Since f_t is L-Lipschitz w.r.t. norm $\|\cdot\|_2$, we have

$$\begin{aligned}\mathbb{E}\|z_t\|_2^2 &= \frac{d^2}{\delta^2}\mathbb{E}\|f_t(w_t + \delta e_t)\|_2^2 \le \frac{d^2}{\delta^2}\mathbb{E}\|f_t(w_t) + L\delta\|_2^2 \\ &= \frac{d^2}{\delta^2}(F + L\delta)^2 = d^2(F/\delta + L)^2 \end{aligned} \tag{7}$$

That is the power of z_t is bounded by $d^2(F/\delta + L)^2$. Since $z_t \to \widetilde{z}_t$ is a Gaussian channel, we have the result. □

Theorem 4 (Regret Guarantee)**.** *Let $\mathcal{A}_2$ be Algorithm 2, f_t is L-Lipschitz continuous for every t, $F = \max_{u\in\mathcal{S}, t\ge 1} f_t(u) < \infty$ and $\max_{u\in\mathcal{S}}\|u\|_2 \le B$. Then* $\mathrm{Regret}(\mathcal{A}_2)$ *is sub-linear to T. Specifically,*

$$\mathrm{Regret}(\mathcal{A}_2) \le \frac{B^2}{2\eta} + \frac{\eta}{2}T(d^2(F/\eta + L)^2 + d\sigma^2) + 3TL\delta$$

In particular, if we set $\eta \sim T^{-3/4}$ and $\delta \sim T^{-1/4}$, the regret is bounded by $O(T^{3/4})$.

Proof. Define $\hat{f}_t(w) = \mathbb{E}_{e_t\sim U_{sp}}[f_t(w + \delta e_t)]$. By the $L-$Lipschitzness of f_t, two function f_t and $\hat{f}_t(w)$ are "close".

$$|\hat{f}_t(w) - f_t(w)| \le \max_e |f_t(w + \delta e) - f_t(w)| \le L\delta$$

$$\begin{aligned}\mathrm{Regret}(\mathcal{A}_2) &= \mathbb{E}\left\{\sum_{t=1}^T [f_t(w_t + \delta e_t) - f_t(u)]\right\} \\ &\le \mathbb{E}\left\{\sum_{t=1}^T [\hat{f}_t(w_t) - \hat{f}_t(u) + 3L\delta]\right\} \end{aligned} \tag{8}$$

Moreover, $\hat{f}_t$ is differentiable and z_t is the gradient of $\hat{f}_t$ at point w_t (see [11] Lemma 1). Therefore

$$\hat{f}_t(w_t) - \hat{f}_t(u) \le \langle w_t - u, z_t\rangle$$

Follow the same procedure in the proof of Theorem 2, we have

$$\mathbb{E}\left\{\sum_{t=1}^T [\hat{f}_t(w_t) - \hat{f}_t(u)]\right\} \le \frac{1}{2\eta}\|u\|_2^2 + \frac{\eta}{2}\sum_{t=1}^T \mathbb{E}\|\widetilde{z}_t\|_2^2 \tag{9}$$

Since $\widetilde{z}_t = z_t + v_t$, the variance of z_t is bounded by (7), v_t is independent of z_t and $\mathbb{E}\|v_t\|_2^2 \le d\sigma^2$, the last term is bounded by

$$\frac{\eta}{2}\sum_{t=1}^T \mathbb{E}\|\widetilde{z}_t\|_2^2 \le \frac{\eta}{2}T(d^2(F/\eta + L)^2 + d\sigma^2) \tag{10}$$

By combining (8)-(10) and $\|u\|_2 \le B$, the proof is complete. □

4. NUMERICAL RESULTS

We present numerical results to illustrate the learning performance of mutual-information-private OGD algorithms for a simple linear regression problem. Fig. 3 plots the regret of algorithm 1 and its non-private version ($\sigma = 0$). Here we set $d = L = B = \sigma^2 = 1$ and the encrypted OGD is $\frac{1}{2}\log 2$-mutual-information-private. The performance-privacy trade-off for full information setting is illustrated in Fig. 4, where we set $d = B = 1$ and $T = 25$, which provides the ROC curve of Algorithm 1.

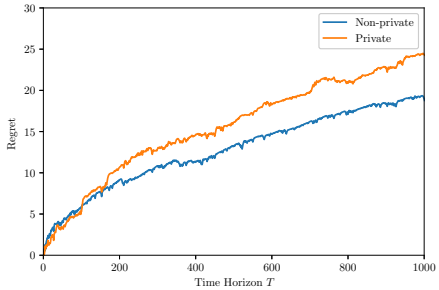

Fig. 3. Performance-privacy trade-off for Algorithm 1.

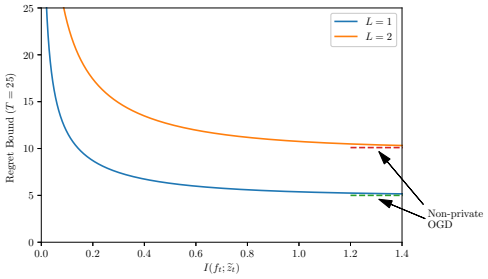

Fig. 4. Performance-privacy trade-off for Algorithm 1.

5. CONCLUSION

In this work, we proposed an additive noise mechanism for the online gradient descent (OGD) algorithm for both full-information setting and bandit setting. It is shown that our private preserving OGD provides a conservative way to protect users' data. The user's leaked information is bounded by the channel capacity of Gaussian channel while the regret of the learning system is sub-linear to the time horizon T.

6. REFERENCES

[1] Shai Shalev-Shwartz et al., "Online learning and online convex optimization," *Foundations and Trends® in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2012.

[2] Avrim Blum, Vijay Kumar, Atri Rudra, and Felix Wu, "Online learning in online auctions," *Theoretical Computer Science*, vol. 324, no. 2-3, pp. 137–146, 2004.

[3] Anand D Sarwate and Kamalika Chaudhuri, "Signal processing and machine learning with differential privacy: Algorithms and challenges for continuous data," *IEEE signal processing magazine*, vol. 30, no. 5, pp. 86–94, 2013.

[4] Prateek Jain, Pravesh Kothari, and Abhradeep Thakurta, "Differentially private online learning," in *Conference on Learning Theory*, 2012, pp. 24–1.

[5] Jun Sakuma and Hiromi Arai, "Online prediction with privacy," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 935–942.

[6] Feng Yan, Shreyas Sundaram, SVN Vishwanathan, and Yuan Qi, "Distributed autonomous online learning: Regrets and intrinsic privacy-preserving properties," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 11, pp. 2483–2493, 2013.

[7] Thomas M Cover and Joy A Thomas, *Elements of information theory*, John Wiley & Sons, 2012.

[8] Reza Shokri and Vitaly Shmatikov, "Privacy-preserving deep learning," in *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*. ACM, 2015, pp. 1310–1321.

[9] Martin Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, 2003, pp. 928–936.

[10] William C Davidon, "New least-square algorithms," *Journal of Optimization Theory and Applications*, vol. 18, no. 2, pp. 187–197, 1976.

[11] Abraham D Flaxman, Adam Tauman Kalai, and H Brendan McMahan, "Online convex optimization in the bandit setting: gradient descent without a gradient," in *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 2005, pp. 385–394.