King Abdullah University
of Science and Technology

Homework 3

Konstantin Burlachenko

Assignment in Contemporary Topics in Computer Security.
Last updated on November 15, 2021

Contents

Task 1

Privacy loss is defined as $l_{D,D'}(y) := \ln\left(\frac{p_{A(D)}(y)}{p_{A(D')}(y)}\right)$ where $y \sim A(D)$.

The Laplacian mechanism for the case when $f : X^n \to \mathbb{R}$ working as follows:

$$A(X) = f(X) + Z, Z \sim \frac{1}{2b}\exp\left(-\frac{|z|}{b}\right), b = \frac{GS_{f,1}}{\varepsilon}$$

Lemma If W has p.d.f $p_w(w)$, then $W + c, \forall c \in \mathbb{R}$ has p.d.f $p_w(w - c)$:

$$\mathbf{P}(c + W < t) = \mathbf{P}(W < t - c) = \int_{-\infty}^{t-c} p_w(w)dw = |u - c = w| = \int_{-\infty}^{t} p_w(u - c)du.$$

$$l_{D,D'}(y) := \ln\left(\frac{p_{A(D)}(y)}{p_{A(D')}(y)}\right) = \ln\frac{\frac{1}{2b}\exp\left(-\frac{|y-f(D)|}{b}\right)}{\frac{1}{2b}\exp\left(-\frac{|y-f(D')|}{b}\right)} = \ln\frac{\exp\left(-\frac{|y-f(D)|}{b}\right)}{\exp\left(-\frac{|y-f(D')|}{b}\right)} =$$
$$\ln\exp\left(\frac{|y-f(D')| - |y-f(D)|}{b}\right) = \varepsilon\left(\frac{|y-f(D')| - |y-f(D)|}{GS_{f,1}}\right) =$$
$$\varepsilon\left(\frac{|f(D)+Z-f(D')| - |f(D)+Z-f(D)|}{GS_{f,1}}\right) =$$
$$\varepsilon\left(\frac{|f(D)-f(D')+Z| - |Z|}{GS_{f,1}}\right)$$

Know we use assumptions from the task:

1. $GS_{f,1} = 11$
2. $f(D) = 0$
3. $f(D') = 1$

With using this concrete values we can obtain:

$$l_{D,D'}(y) = \varepsilon\left(|Z - 1| - |Z|\right)$$

Where $y = f(D) + Z, Z \sim \frac{1}{2b}\exp\left(-\frac{|z|}{b}\right), b = \frac{1}{\varepsilon}$.

So $l_{D,D'}$ is a random variable.

The p.d.f. os Laplacian distribution allow integrate analytically. Below is a derivation probability that $Z \in [a, b], a \geq 0, b \geq 0$.

$$\mathbf{P}(Z \in [a,b] | a \geq 0, b \geq 0) = \varepsilon/2\int_a^b \exp(-\varepsilon z)dz = -1/2\cdot\exp(-\varepsilon z)|_a^b = 1/2(\exp(-\varepsilon a) - \exp(-\varepsilon b))$$

We can state that $l_{D,D'}$ is a random variable with CDF:

$$
\begin{aligned}
F_{l_{D,D'}}(y) = \int_{\varepsilon(|z-1|-|z|)<y} \varepsilon/2\exp(-\varepsilon z)dz = \varepsilon/2 \int_{|z-1|-|z|<y/\varepsilon} \exp(-\varepsilon|z|)dz = \\
\varepsilon/2 \int_{|z-1|-|z|<y/\varepsilon \cap \{z\in[-\infty,0] \cup z\in[0,1] \cup z\in[1,+\infty]\}} \exp(-\varepsilon|z|)dz = \\
\varepsilon/2 \int_{(1-(-z))-(-z)<y/\varepsilon \cap \{z\in[-\infty,0]\}} \exp(-\varepsilon|z|)dz + \\
\varepsilon/2 \int_{(1-z)-(z)<y/\varepsilon \cap \{z\in[0,1]\}} \exp(-\varepsilon|z|)dz + \\
\varepsilon/2 \int_{(z-1)-(z)<y/\varepsilon \cap \{z\in[1,+\infty]\}} \exp(-\varepsilon|z|)dz = \\
\varepsilon/2 \int_{(1+2z)<y/\varepsilon \cap \{z\in[-\infty,0]\}} \exp(-\varepsilon|z|)dz + \\
\varepsilon/2 \int_{(1-2z)<y/\varepsilon \cap \{z\in[0,1]\}} \exp(-\varepsilon|z|)dz + \\
\varepsilon/2 \int_{\{-1<y\}/\varepsilon \cap \{z\in[1,+\infty]\}} \exp(-\varepsilon|z|)dz = \\
\varepsilon/2(\int_{(1+2z)<y/\varepsilon \cap \{z\in[-\infty,0]\}} \exp(\varepsilon z)dz + \\
\int_{(1-2z)<y/\varepsilon \cap \{z\in[0,1]\}} \exp(-\varepsilon z)dz + \\
\int_{\{-1<y\}/\varepsilon \cap \{z\in[1,+\infty]\}} \exp(-\varepsilon z)dz) = \\
1/2(\int_{\{z<(y/\varepsilon-1)/2\} \cap \{z\in[-\infty,0]\}} d(\exp(\varepsilon z)) - \int_{\{z>-(y/\varepsilon-1)/2\} \cap \{z\in[0,1]\}} d(\exp(-\varepsilon z)) - \\
\int_{\{-1<y\}/\varepsilon \cap \{z\in[1,+\infty]\}} d(\exp(-\varepsilon z)
\end{aligned}
$$

Task 2

For Top-k selection we repeat exponential mechanism just k times without replacement. Firstly we consider apply several algorithm with sampling elements with Exponential mechanism and plugging result into next round. The mechanism $A(D)$ that is applied after first, second, etc. selection take extra arguments about previously selected items, and even formally algorithm $A(D)$ defined on input dataset D in fact it exclude all previous items from consideration. The Composition Theorems allows to consider such scenario when algorithms are worryingly in dependent way.

Exponential Mechanism is ε-DP.
By basic composition theorem via repeating this mechanism it will be $k \cdot \varepsilon$-DP. So given a target budget ε' it's enough to perform each maximum selection with ε'/k privacy budget.

By advanced composition theorem via repeating this mechanism it will be:
$(\varepsilon\sqrt{2k \cdot \ln(\frac{1}{\delta'})} + k\varepsilon\frac{\exp(\varepsilon)-1}{\exp(\varepsilon)+1}, \delta' k)$-approximate DP.

For small ε we have $\exp(\varepsilon) \approx 1 + \varepsilon$ and $\exp(\varepsilon) + 1 \approx 2$ and so the last privacy description of the algorithm bringing that in such circumstances:

$(\varepsilon\sqrt{2k \cdot \ln(\frac{1}{\delta'})} + k\frac{\varepsilon^2}{2}, \delta' k)$-approximate DP.

Under another reasonable assumption mentioned in the lecture that we want to have $\varepsilon \cdot k < 1, \varepsilon < 1 \implies \varepsilon^2 k < 1 \implies \sqrt{k}\varepsilon > \varepsilon^2 k$. And under that assumptions mechanism will be: $(2\varepsilon\sqrt{2k \cdot \ln(\frac{2}{\delta'})}, \delta' k)$-approximate DP.

So given a target budget ε', δ' it's enough to perform each maximum selection with $(\frac{\varepsilon'}{(2\sqrt{2k \cdot \ln(\frac{2}{\delta'})}}, \delta'/k)$ privacy budget. Because underlying mechanism (Exponential Mechanism) in not configurable by δ it's enough that we will require to be it only $\frac{\varepsilon'}{(2\sqrt{2k \cdot \ln(\frac{2}{\delta'})}}$-DP. As we see advanced mechanism allow us to have more privacy budget for each operation. $\propto 1/\sqrt{k}$, but with the cost that result mechanism will be only approximate DP.

Now we move to analyze expected accuracy. Let $q_k(D)$ be the score function of a best (more higher value of function q) k-th item.

Let's look into value of $\mathbf{P}(q_k(D) - \min_{j \in S} q(j, D) \geq h)$ for value $h \in \mathbb{R}$.

Let's consider another event: $\cup_{j \in S}(q_k(D) - q(j, D) \geq h)$

From one side:
$q_k(D) - \min_{j \in S} q(j, D) \geq h \implies \exists j' \in S : q_k(D) - q(j', D) \geq h \implies \cup_{j \in S}(q_k(D) - q(j, D) \geq h)$

From another side:
$\cup_{j \in S}(q_k(D) - q(j, D) \geq h) \implies \exists j' \in S : q_k(D) - q(j', D) \geq h$, but $q_k(D) - \min_{j \in S} q(j, D) \geq q_k(D) - q(j', D) \geq h \implies q_k(D) - \min_{j \in S} q(j, D) \geq h$

And with using union bound we can obtain:
$\mathbf{P}(q_k(D) - \min_{j \in S} q(j, D) \geq h) = \mathbf{P}(\cup_{j \in S}(q_k(D) - q(j, D) \geq h)) \leq \sum_{j \in S} \mathbf{P}(q_k(D) - q(j, D) \geq h)$

Theorem 5.7 from the lecture 5 provides probability of the following event and upper bound for it:

$$\mathbf{P}\left(q(y, D) \leq q_{max} - \frac{2\Delta(\ln(d) + t)}{\varepsilon}\right) \leq \exp(-t)$$

In that inequality q_{max} is most aggressive bound, but in fact we can decrease it and substitute $q_k(D)$. For any sequence of top-k pulled elements, the obtained event with replacing $q_{max}(D)$ into $q_k(D)$ will be a subset of original event considered in equation, and it's probability will be less then original event. Nevertheless we can use bound $\exp(-t)$ for it.

$$\mathbf{P}\left(q(y, D) \leq q_k - \frac{2\Delta(\ln(d) + t)}{\varepsilon}\right) \leq \mathbf{P}\left(q(y, D) \leq q_{max} - \frac{2\Delta(\ln(d) + t)}{\varepsilon}\right) \leq \exp(-t)$$

This bound is valid bound for sampling once $q(y, D)$, but $q(j, D)$ is sample number j. One way is consider $q(j, D)$ is sampling strategy with maximum element from dataset $\hat{D}$ obtained from original subset D via removing sampled elements from the previous $j - 1$ rounds. We don't know this previous samples, but whatever they are the only sensibleness quality for dataset change is sensitivity bound Δ. For our derivations we fix it as global value. And finally:

$$\mathbf{P}\left(q(j,D) \leq q_k - \frac{2\Delta(\ln(d)+t)}{\varepsilon}\right) = \mathbf{P}\left(\frac{2\Delta(\ln(d)+t)}{\varepsilon} \leq q_k - q(j,D)\right) \leq \exp(-t)$$

With using that bound and substitute into original:
$\mathbf{P}\left(q_k(D) - \min_{j\in S} q(j,D) \geq \dfrac{2\Delta(\ln(d)+t)}{\varepsilon}\right) \leq |S|\exp(-t)$

The r.v. $q_k(D) - \min_{j\in S} q(j,D)$ is non-negative, and for any non negative r.v.: $\mathbf{E}[Z] = \int_{z=0}^{z=+\infty} \mathbf{P}(Z > z)dz$, in general: $\mathbf{E}[Z] = \int_0^{+\infty} \mathbf{P}(Z > z)dz + \int_{-\infty}^{0} \mathbf{P}(Z < z)dz$.
Let's define: $Z = \varepsilon/2\Delta \cdot (q_k(D) - \min_{j\in S} q(j,D))$

$$\mathbf{E}[Z] = \int_0^{+\infty} \mathbf{P}(Z > z)dz = \int_{-\ln(d)}^{+\infty} \mathbf{P}(Z > \ln(d)+t)dt \leq \ln(d)+|S|\int_0^{+\infty} \exp(-t)dt = \ln(d)+|S|$$

So: $\mathbf{E}[q_k(D) - \min_{j\in S} q(j,D) = 2\Delta/\varepsilon \cdot \mathbf{E}[Z] \leq 2\Delta/\varepsilon(\ln(d)+k)$, where d is the dimension of the set from which we sample.

1. With basic composition theorem and having DP privacy budget (ε') we use $\varepsilon = \frac{\varepsilon'}{k}$.
   And:
   $\mathbf{E}[q_k(D) - \min_{j\in S} q(j,D) \leq 2\Delta/\varepsilon'(k\ln(d)+k^2)$
2. With advanced composition theorem and having approximate-DP privacy budget (ε',δ') we use $\varepsilon = \frac{\varepsilon'}{(2\sqrt{2k\cdot\ln(\frac{2}{\delta'})}}$.
   And:
   $\mathbf{E}[q_k(D) - \min_{j\in S} q(j,D) \leq 4\Delta/\varepsilon'(\sqrt{2\cdot\ln(\frac{2}{\delta'})}\cdot(\sqrt{k}\ln(d)+k\sqrt{k})$.

Task 3

Lemma about p.d.f. for dependent r.v. If $z = f(w)$, where f is not decreasing function, and w is r.v C.d.f. for Z will have the following form: $F_z(z) = P(Z < z) = P(f(w) < z) = P(w < f^{-1}(z)) = F_w(f^{-1}(z))$ And p.d.f. for z will have the following form: $f_z(z) = F_z'(z) = f_w(w)|_{w=f^{-1}(z)} \cdot (f^{-1}(z))_z' = f_w(w)|_{w=f^{-1}(z)} \cdot |(f^{-1}(z))_z'|$
If $z = f(w)$, where f is not increasing function, and w is r.v C.d.f. for Z will have the following form: $F_z(z) = P(Z < z) = P(f(w) < z) = P(w > f^{-1}(z)) = 1 - F_w(f^{-1}(z))$ And p.d.f. for z will have the following form: $f_z(z) = F_z'(z) = -f_w(w)|_{w=f^{-1}(z)} \cdot (f^{-1}(z))_z' = f_w(w)|_{w=f^{-1}(z)} \cdot |(f^{-1}(z))_z'|$

Lemma about p.d.f. for shifted and normal r.v. $z = au + b, u \sim N(m,\sigma^2)$ With using the previous lemma we have $f^{-1}(z) = z - b/a$, $|(f^{-1})'| = |1/a|$.

$$f_z = 1/|a| \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{((y-b)/a-m)^2}{2\sigma^2}\right) =$$
$$\frac{1}{\sqrt{2\pi(|a|\sigma)^2}} \cdot \exp\left(-\frac{(y/a-b/a-ma/a)^2}{2\sigma^2}\right) =$$
$$\frac{1}{\sqrt{2\pi(|a|\sigma)^2}} \cdot \exp\left(-\frac{(y-(ma+b))^2}{2(|a|\sigma)^2}\right) \implies z \sim N(am+b,(a\sigma)^2)$$

If $U \sim N(0, \sigma^2)$, then $F(U) \sim N(b, a^2\sigma^2) + N(0, \rho^2) \sim N(b, a^2\sigma^2 + \rho^2)$.

Here we used previous lemma, and also the fact that for two i.r.v. with Gaussian distribution, the p.d.f. is also Gaussian with additive mean and variance from both distribution. It's possible to show this from convolution formula for two r.v.

If $V \sim N(\Delta, \sigma^2)$, then $F(V) \sim N(a\Delta + b, a^2\sigma^2) + N(0, \rho^2) \sim N(a\Delta + b, a^2\sigma^2 + \rho^2)$.

Here g.s. $\Delta := \max_{D\sim D'} \|f(D) - f(D')\|$. Finally we need to show that there exist some a, b, ρ such that:
$F(U) \sim f(D) + N(0, \sigma^2) \iff N(b, a^2\sigma^2 + \rho^2) \sim N(f(D), \sigma^2)$.
$F(V) \sim f(D') + N(0, \sigma^2) \iff N(a\Delta + b, a^2\sigma^2 + \rho^2) \sim N(f(D'), \sigma^2)$.

The p.d.f of Gaussian r.v. completely defined by it's parameters, and two r.v. will have the same distribution if and only if parameters of distribution are exactly the same.
From first condition on $F(U)$ we have: $b = f(D)$.
From second condition on $F(V)$ we have: $a = \dfrac{f(D') - b}{\Delta} = \dfrac{f(D') - f(D)}{\Delta}$.

Finally via considering variance part in $F(U), F(V)$ we have the following condition: $\rho^2 = \sigma^2(1 - a^2) = \sigma^2 \left(1 - \left(\dfrac{f(D') - f(D)}{\Delta}\right)^2\right)$.

For example we can take: $\rho = \sigma\sqrt{\left(1 - \left(\dfrac{f(D') - f(D)}{\Delta}\right)^2\right)}$.

Because g.s. defined as $\Delta := \max_{D\sim D'} \|f(D) - f(D')\|$ the last expression has sense in terms of usual real arithmetic.

Task 4

For purpose of that problem we will consider adaptive composition of k executions of the Gaussian Mechanism.

First of all it's enough to prove advanced composition theorem for U, V such that $U_i \sim N(0, \sigma^2)$ and $V_i \sim N(\Delta, \sigma^2)$. The property that allows to do it is the post-processing property that is valid for pure-DP and (ε, δ)-DP of differential privacy, so such prove about that DP r.v. will imply the same guarantee for $A(D)$ and $A(D')$.

So our goal is prove $U \approx_{\hat{\varepsilon}} V$. Let's compute privacy loss fo $z_j \sim N(0, \sigma^2)$.

$$l_{D,D'}(y) = \ln\left(\frac{p_U(z; D)}{p_V(z; D')}\right) = \sum_{j=1}^{k} \ln\left(\frac{p_{U_j}(z_j; D)}{p_{V_j}(z_j; D')}\right) = \sum_{j=1}^{k} \ln\left(\frac{\exp\left(-\dfrac{(z_j - 0)^2}{2(\sigma)^2}\right)}{\exp\left(-\dfrac{(z_j - \Delta_j)^2}{2(\sigma)^2}\right)}\right) =$$

$$\sum_{j=1}^{k}(-\frac{(z_j - 0)^2}{2(\sigma)^2}) - (-\frac{(z_j - \Delta)^2}{2(\sigma)^2}) = -\frac{1}{2\sigma^2} \cdot \sum_{j=1}^{k}(z_j^2 - (z_j - \Delta)^2) = \frac{1}{2\sigma^2} \cdot \sum_{j=1}^{k}((z_j - \Delta)^2 - z_j^2) =$$

$$\frac{1}{2\sigma^2} \cdot \sum_{j=1}^{k}(\Delta^2 - 2z_j\Delta) = \frac{\Delta^2}{2\sigma^2} - \frac{1}{\sigma^2}\sum_{j=1}^{K} z_j\Delta = \frac{\Delta^2}{2\sigma^2}K + \frac{\Delta}{\sigma}\sqrt{K} \cdot Z, Z \sim N(0, 1)$$

The last reduction happens due to that sum of Normally distributed i.r.v. z_j has normal distribution. And reduction to a single r.v. happens due to the following $E[\sum_j -z_j\Delta_j/\sigma^2] = 0$, and $Var[\sum_j -z_j\Delta_j/\sigma^2] = \sum_j \Delta_j^2/\sigma^4 Var[z_j] = K\Delta_j^2/\sigma^4 \cdot \sigma^2 = K\Delta_j^2/\sigma^2$.

Standard Gaussian tail bound: $\mathbf{P}(Z \geq v) \leq \exp(-v^2/2)$

$$\mathbf{P}(l_{D,D'}(y) \geq \varepsilon) = \mathbf{P}(\frac{\Delta^2}{2\sigma^2}K + \frac{\Delta}{\sigma}\sqrt{K} \cdot Z \geq \varepsilon) =$$

$$\mathbf{P}\left(Z \geq \frac{(\varepsilon - \dfrac{\Delta^2}{2\sigma^2}K)}{(\Delta/\sigma)\sqrt{K}}\right) \leq \exp\left(-\left(\frac{(\varepsilon - \dfrac{\Delta^2}{2\sigma^2}K)}{(\Delta/\sigma\sqrt{K})}\right)^2/2\right) = \delta$$

$$\implies \mathbf{P}(l_{D,D'}(y) \geq \varepsilon) \leq \sigma$$

Sow we have found for $\forall\delta > 0$ the mechanism if (ε, δ) - DP. Where:

$$-\left(\frac{(\varepsilon - \dfrac{\Delta^2}{2\sigma^2}K)}{(\Delta/\sigma)\sqrt{K}}\right)^2 = 2\ln(\delta) \iff$$

$$\left(\frac{(\varepsilon - \dfrac{\Delta^2}{2\sigma^2}K)}{(\Delta/\sigma)\sqrt{K}}\right)^2 = 2\ln(1/\delta) \iff$$

$$\frac{(\varepsilon - \dfrac{\Delta^2}{2\sigma^2}K)}{(\Delta/\sigma)\sqrt{K}} = \sqrt{2\ln(1/\delta)} \iff$$

$$\varepsilon = \frac{\Delta^2}{2\sigma^2}K + \Delta/\sigma\sqrt{K}\sqrt{2\ln(1/\delta)} = \frac{\Delta^2}{\sigma^2}(1/2 + \sqrt{2\ln(1/\delta)})$$

Let's assume $\sigma = \frac{\Delta\sqrt{K}}{\varepsilon}\sqrt{2\ln(1/\delta)} \cdot 1/t$, and $t \geq 0$ will be defined later. Next, ε should be thought of as a small constant. Anything between (say) 0.1 and 5 might be a reasonable level privacy guarantee:

1. Smaller corresponds to stronger privacy (but smaller accuracy)
2. Bigger corresponds to weaker privacy (but bigger accuracy)

$$\varepsilon = \frac{\varepsilon^2}{4 \cdot \ln(1/\delta)}t^2 + \varepsilon t \iff \frac{\varepsilon^2}{4 \cdot \ln(1/\delta)}t^2 + \varepsilon t - \varepsilon = 0$$

So answer for t is:

$$t = \frac{-\varepsilon \pm \sqrt{\varepsilon^2 - 4(\frac{\varepsilon^2}{4\cdot\ln(1/\delta)})(-\varepsilon)}}{2(\frac{\varepsilon^2}{4\cdot\ln(1/\delta)})} =$$

$$\frac{-1 \pm \sqrt{1 - (\frac{1}{\ln(1/\delta)})(-\varepsilon)}}{2(\frac{\varepsilon}{4\cdot\ln(1/\delta)})} = \frac{\sqrt{1 + (\frac{\varepsilon}{\ln(1/\delta)})} - 1}{(\frac{\varepsilon}{2\cdot\ln(1/\delta)})} \approx \frac{1 + (1/2\frac{\varepsilon}{\ln(1/\delta)}) - 1}{(\frac{\varepsilon}{2\cdot\ln(1/\delta)})} = 1$$

In last inequality we get rid of negative root for t, which impossible in our task assumes that $\varepsilon/\ln(1/\delta)$ is small enough.

So we have proved that: $\sigma = \frac{\Delta\sqrt{K}}{\varepsilon}\sqrt{2\ln(1/\delta)} \cdot 1/1 = \frac{\Delta\sqrt{K}}{\varepsilon}\sqrt{2\ln(1/\delta)}$