Verifying Snapping Mechanism

October 28, 2019

1 Formalization

Definition 1 ($\mathsf{Snap}(\mu, a) : \mathsf{Distr}(U) \to A \to \mathsf{Distr}(B)$)
The ideal Snapping mechanism $\mathsf{Snap}(\mu, a)$ is defined as:

$$u \xleftarrow{\$} \mu; y = \ln(u); s \xleftarrow{\$} \{-1, 1\}; y' = s * y; z = \frac{y'}{\epsilon}; x = f(a); w = x + z; w' = \lfloor w \rfloor_\Lambda; r = \mathsf{clamp}_B(w')$$

where f is the query function over input $a \in A$, ϵ is the privacy budget and S sampled from $\{-1, +1\}$ with $\mathsf{Bernoulli}(0.5)$.

Definition 2
Let $\epsilon \leq 0$. The ϵ-DP divergence $\Delta_\epsilon(\mu_1, \mu_2)$ between two sub-distributions $\mu_1 \in \mathsf{Distr}(U)$, $\mu_2 \in \mathsf{Distr}(U)$is defined as:

$$\sup_{E \in U} \Big(\Pr_{x \leftarrow \mu_1}[x \in E] - \exp(\epsilon) \Pr_{x \leftarrow \mu_2}[x \in \cdot E]] \Big)$$

Definition 3 (ϵ **- dilation**)
.

Let $\epsilon \geq 0$. The ϵ-dilation $D_\epsilon(\mu_1, \mu_2)$ between two sub-distributions $\mu_1 \in \mathsf{Distr}(U)$, $\mu_2 \in \mathsf{Distr}(U)$is defined as:

$$\sup_{E \in U} \Big(\Pr_{x \leftarrow \mu_1}[x \in E] - \exp(\epsilon) \Pr_{x \leftarrow \mu_2}[x \in \exp(-\epsilon) \cdot E]] \Big)$$

Proposition 1 (**(ϵ, δ)-differential privacy**)
For every pair of sub-distributions $\mu_1 \in \mathsf{Distr}(U)$, $\mu_2 \in \mathsf{Distr}(U)$, s.t.

$$D_\epsilon(\mu_1, \mu_2) \leq \delta,$$

The snapping mechanism $\mathsf{Snap}(\mu, a) : \mathsf{Distr}(U) \to A \to \mathsf{Distr}(B)$ is (ϵ, δ) - differentially private w.r.t. an adjacency relation Φ for every two adjacent inputs a, a' and μ_1, μ_2

Proof. Followed directly by unfolding the Snap mechanism.

$$\begin{aligned}
\Pr_{x \leftarrow \mathsf{Snap}(\mu_1, a)}[x = e] &= \Pr_{u \leftarrow \mu_1}[\lfloor f(a) + \tfrac{S \cdot \log(u)}{\epsilon} \rfloor_\Lambda = e] \\
&= \Pr_{u \leftarrow \mu_1}[u \in [\tfrac{\exp((e - \frac{\Lambda}{2} - f(a))\epsilon)}{S}, \tfrac{\exp((e + \frac{\Lambda}{2} - f(a))\epsilon)}{S})] \\
&\leq \exp(\epsilon) \Pr_{u \leftarrow \mu_2}[u \in \exp(-\epsilon)[\tfrac{\exp((e - \frac{\Lambda}{2} - f(a))\epsilon)}{S}, \tfrac{\exp((e + \frac{\Lambda}{2} - f(a))\epsilon)}{S})] \\
&= \exp(\epsilon) \Pr_{u \leftarrow \mu_2}[\lfloor f(a') + \tfrac{S \cdot \log(u)}{\epsilon} \rfloor_\Lambda = e] \\
&= \exp(\epsilon) \Pr_{x \leftarrow \mathsf{Snap}(\mu_2, a')}[x = e]
\end{aligned}$$

□

Definition 4 ((ϵ, δ) **- Dilation lifting**)
Two sub-distributions $\mu_1 \in \mathsf{Distr}(U_1)$, $\mu_2 \in \mathsf{Distr}(U_2)$are related by the (ϵ, δ) - dilation lifting of $\Psi \subseteq U_1 \times U_2$, written $\mu_1 \Psi^{d(\epsilon,\delta)} \mu_2$, if there exist two witness sub-distributions $\mu_L \in \mathsf{Distr}(U_1 \times U_2)$ and $\mu_R \in \mathsf{Distr}(U_1, U_2)$ s.t.:

1. $\pi_1(\mu_L) = \mu_1$ and $\pi_2(\mu_R) = \mu_2$;
2. $\mathsf{supp}(\mu_L) \subseteq \Psi$ and $\mathsf{supp}(\mu_R) \subseteq \Psi$; and
3. $D_\epsilon(\mu_L, \mu_R) \leq \delta$.

It is easy to see that two sub-distributions μ_1 and μ_2 are related by $=^{d(\epsilon,\delta)}$ iff $D_\epsilon(\mu_1, \mu_2) \leq \delta$. Therefore, the Snap mechanism $\mathsf{Distr}(U) \to A \to \mathsf{Distr}(B)$ is $(\epsilon, \delta)-$ differentially private w.r.t. and adjacency relation Φ and μ_1, μ_2 iff:

$$\mu_1 =^{d(\epsilon,\delta)} \mu_2$$

for every two adjacent inputs a and a'.

Definition 5 **((ϵ, δ) - lifting [1])**
Two sub-distributions $\mu_1 \in \mathsf{Distr}(U_1)$, $\mu_2 \in \mathsf{Distr}(U_2)$are related by the (ϵ, δ) - dilation lifting of $\Psi \subseteq U_1 \times U_2$, written $\mu_1 \Psi^{\#(\epsilon,\delta)} \mu_2$, if there exist two witness sub-distributions $\mu_L \in \mathsf{Distr}(U_1 \times U_2)$ and $\mu_R \in \mathsf{Distr}(U_1, U_2)$ s.t.:

1. $\pi_1(\mu_L) = \mu_1$ and $\pi_2(\mu_R) = \mu_2$;
2. $\mathsf{supp}(\mu_L) \subseteq \Psi$ and $\mathsf{supp}(\mu_R) \subseteq \Psi$; and
3. $\Delta_\epsilon(\mu_L, \mu_R) \leq \delta$.

Theorem 2
if $\mu_1 \Psi^{d(\epsilon,\delta)} \mu_2$, then we can prove $\mathsf{Snap}(\mu_1, a) \Psi^{\#(\epsilon,\delta)} \mathsf{Snap}(\mu_2, a')$, w.r.t. an adjacent relation Φ for every $a \Phi a'$

Theorem 3
The coupling of two Snap mechanisms: $\mathsf{Snap}(\mu_1, a_1)$, $\mathsf{Snap}(\mu_2, a_2)$.

$$\frac{}{u_1 \xleftarrow{\$} \mu \sim_{\epsilon,0} u_2 \xleftarrow{\$} \mu : T \Rightarrow u_1 = (e^{\epsilon}) u_2}$$

$$\frac{}{y_1 = \ln(u_1) \sim_{0,0} y_2 = \ln(u_2) : u_1 = e^{\epsilon} u_2 \Rightarrow y_1 = \epsilon + y_2}$$

$$\frac{}{s_1 \xleftarrow{\$} \mu \sim_{0,0} s_1 \xleftarrow{\$} \mu : T \Rightarrow s_1 = s_2}$$

$$\frac{}{y_1' = s_1 * y_1 \sim_{0,0} y_2' = s_2 * y_2 : s_1 = s_2 \wedge y_1 = e^{\epsilon} u_2 \Rightarrow y_1' = \epsilon + y_2'}$$

$$\frac{}{z_1 = \dfrac{y_1'}{\epsilon} \sim_{0,0} z_2 = \dfrac{y_2'}{\epsilon} : y_1' = \epsilon + y_2' \Rightarrow z_1 = z_2 + 1}$$

$$\frac{}{x_1 = f(a_1) \sim_{0,0} x_2 = f(a_2) : x_1 = x_2 + 1 \Rightarrow a_1 = a_2 + 1}$$

$$\frac{}{w_1 = x_1 + z_1 \sim_{0,0} w_2 = x_2 + z_2 : x_1 + 1 = x_2 \wedge z_1 = z_2 + 1 \Rightarrow w_1 = w_2}$$

$$\frac{}{w_1' = \lfloor w_1 \rceil_{\Lambda} \sim_{0,0} w_2' = \lfloor w_2 \rceil_{\Lambda} : w_1 = w_2 \Rightarrow w_1' = w_2'}$$

$$\frac{}{r_1 = \mathsf{clamp}_B(w_1') \sim_{0,0} r_2 = \mathsf{clamp}_B(w_2') : w_1' = w_2' \Rightarrow r_1 = r_2}$$

Figure 1: coupling of two Snap mechanisms: $\mathsf{Snap}(\mu_1, a_1)$, $\mathsf{Snap}(\mu_2, a_2)$

References

[1] Gilles Barthe, Marco Gaboardi, Benjamin Grégoire, Justin Hsu, and Pierre-Yves Strub. Proving differential privacy via probabilistic couplings. In *LICS 2016*.