Contents

1 Definitions

1.1 Probability

Definition 1.1. *The probability* $\Pr(\rho|X)$ *of a path* $\rho = q_0 \to q_1 \to \cdots \to q_m$ *given an input* $X = \langle a_1, \ldots, a_m \rangle$ *is defined recursively as the probability that all transitions in* ρ *are traversed in sequence given the input* X *starting at state* q_0*.*

Definition 1.2. *Let* $\mathcal{A}$ *be a DiPA, and* $s \in seg(\mathcal{A})$ *be a segment. The* ***privacy loss*** $\text{loss}(s)$ *of a segment* $s \in seg(\mathcal{A})$ *is defined as*

$$\text{loss}(s) = \sup_{\rho \in segF(s)} \sup_{X' \sim X} \left(\frac{Pr(\rho|X)}{P(\rho|X')} \right)$$

where X *and* X' *vary over all pairs of neighbouring datasets.*

Definition 1.3. *Let* $f_\varepsilon(x)$ *be the probability density function of a random variable* X *with* $X \sim Lap(0, 1/\varepsilon)$*.*

$$f_\varepsilon(x) = \frac{\varepsilon}{2} \exp(-\varepsilon|x|)$$

Definition 1.4. *Let* $F_\varepsilon(x)$ *be the cumulative distribution function of a random variable* X *with* $X \sim Lap(0, 1/\varepsilon)$*.*

$$F_\varepsilon(x) = P(X \leq x) = \begin{cases} \frac{1}{2}\exp(\varepsilon x) & x < 0 \\ 1 - \frac{1}{2}\exp(-\varepsilon x) & x \geq 0 \end{cases}$$

2 Global Optimization

2.1 An abstract description of the global optimization problem

Let $\mathcal{F}$ be a finite family of segment sequences. The global optimization problem is to find

$$\max_{s \in \mathcal{F}} \max_{\boldsymbol{\Delta} \in \{-1,0,1\}^{|s|}} \left(\min_{\boldsymbol{\gamma} \in [-1,1]^{|s|}} \quad c_{s,\boldsymbol{\Delta}}^T \cdot \boldsymbol{\gamma} \quad \text{subject to} \quad A_{s,\boldsymbol{\Delta}} \cdot \boldsymbol{\gamma} \geq 0 \right)$$

Here, the inner minimization problem is a linear program, where $A_{s,\boldsymbol{\Delta}}$ is a matrix of constraints that depends on s and $\boldsymbol{\Delta}$. Similarly, $c_{s,\boldsymbol{\Delta}}$ is a vector of costs that depends on s and $\boldsymbol{\Delta}$.

2.2 Specifying the inner minimization problem

Here, we specify how $A_{s,\boldsymbol{\Delta}}$ and $c_{s,\boldsymbol{\Delta}}$ are defined. Let $s = s_1 \hookrightarrow \cdots \hookrightarrow s_n$ be a sequence of segments with total number of transitions m. Let $\boldsymbol{\Delta} \in \{-1,0,1\}^m$ be a vector of input perturbations.

We define some notation as follows:

- Let t_j^i denote the jth transition in segment i, and ε_j^i denote the noise added to the input before that transition.
- Let Δ_j^i denote the entry of $\boldsymbol{\Delta}$ that corresponds to the input perturbation for the jth transition in segment i.
- Let γ_j^i denote the entry of $\boldsymbol{\gamma} \in [-1,1]^m$ that corresponds to the coupling shift for the jth transition in segment i – this is to be determined by the inner minimization problem.

Then, the minimization problem over $\boldsymbol{\gamma}$ is as follows:

$$\min_{\boldsymbol{\gamma} \in [-1,1]^m} \quad \sum_{i=1}^{n} \sum_{j=1}^{|s_i|} |\gamma_j^i - \Delta_j^i| \varepsilon_i$$

$$\begin{aligned}
\text{subject to} \quad & \gamma_k^i \leq \gamma_0^i & & \text{if } t_k^i \text{ has guard } < \\
& \gamma_k^i \geq \gamma_0^i & & \text{if } t_k^i \text{ has guard } \geq \\
& \gamma_0^i \leq \gamma_0^k & & \text{if } s_k \hookrightarrow s_i \text{ and guard}(s_i) \text{ is } < \\
& \gamma_0^i \geq \gamma_0^k & & \text{if } s_k \hookrightarrow s_i \text{ and guard}(s_i) \text{ is } \geq \\
& \gamma_k^i = 0 & & \text{if } t_k^i \text{ outputs } \texttt{insample} \\
& \gamma_k^i = \Delta_k^i & & \text{if } t_k^i \text{ belongs to a cycle}
\end{aligned}$$

This can be rewritten as a linear program using standard techniques, producing a constraint matrix $A_{s,\boldsymbol{\Delta}}$ and a cost vector $c_{s,\boldsymbol{\Delta}}$.

2.3 Suspected ways to simplify the problem

- We might be able to determine the maximizing $\boldsymbol{\Delta}$ in the second minimization problem in linear time given $s \in \mathcal{F}$.
  - I suspect this is true since I see that the maximizing $\boldsymbol{\Delta}$ always has $\Delta_j^i = -1$ if t_j^i has guard $\geq$, and $\Delta_j^i = 1$ if t_j^i has guard $<$. In the case that t_j^i has guard *true* and is an assignment transition, the value of Δ_j^i seems to depend on the costs ε_j^i in the segment s_i.
  - If this is true, we need not check exponentially many $\boldsymbol{\Delta}$ in the second maximization problem.
- It might be possible to solve a local minimization problem over segments instead of segment sequences, and then use the results to solve a global constraint system that is much smaller than the one described above.

2.4 Constraints for all coupling strategies on a segment

In this section, we will try to understand the constraints that all valid coupling strategies on a segment must satisfy. For the purpose of this section, we will assume that our DiPA consists of a single segment. This assumption will be relaxed in later sections.

Here are some assumptions that are made throughout this document.

1. The noise added to inputs on each state q_i is the same (ε).
2. Since we know tight coupling strategies for segments with cycles, we are restricting our attention to segments with no cycles.
3. We will only consider one segment at a time.

Some notation:

1. Let N be the number of transitions in the segment.
2. The raw input received on the ith transition is denoted by a_i.
3. If we are considering two datasets $X\langle 1\rangle$ and $X\langle 2\rangle$, we will use $a_i\langle 1\rangle$ to denote the value of a_i in the first dataset, and vice versa for $a_i\langle 2\rangle$.
4. Similarly, we use $x_i\langle 1\rangle$ to denote the random variable representing the value of `insample` before the ith transition when A receives the input $X\langle 1\rangle$, and vice versa for $x_i\langle 2\rangle$.

A coupling strategy is a choice of values $\gamma_0, \ldots, \gamma_N$ such that $\gamma_i \in \Gamma$ for all i.

3 Tightness of the strategies S^L, S^J

Last Updated: Wednesday, June 28th, 2023

The relevant definitions and lemmata for proofs in this section are in the appendix. It is also assumed, for now, that all transition outputs are in Γ.

3.1 S^L is tight when there is an L-cycle

Theorem 1. *(S^L is tight for segments with L-cycles) Consider a segment $s \in seg(\mathcal{A})$ corresponding to the sequence of states $q_0 \to q_1 \to \cdots \to q_m$. If s contains an L-cycle, then the L-cost of the segment gives a tight upper bound on the privacy loss of the segment. That is,*

$$\mathrm{loss}(s) = \exp\left(2\varepsilon_0 + \sum_{i>0:guard(a_i)=\texttt{insample}\geq\texttt{x}} 2\varepsilon_i\right)$$

given that state q_i draws from the distribution $Lap(0, 1/\varepsilon_i)$ to noise `insample`.

Proof. We will prove the result for when $\varepsilon_i = \varepsilon$ for all $i \geq 0$. The proof for the general case goes through in the same fashion. Let f, F be the probability density function and cumulative distribution function of a random variable X with $X \sim Lap(0, 1/\varepsilon)$ as defined in the appendix.

Since s has an L-cycle, there exists a sequence of paths ρ_i for $i \in \mathbb{N}$ each with l_i number of L-transitions such that $\lim_{i\to\infty} l_i = \infty$. Let m be the number of G-transitions in ρ_i. We will assume that this number is the same across all ρ_i.[1]

For each ρ_i, construct the adjacent pair of inputs X_i, X_i' as follows. Let $X_i[j] = 0$ for all $j \in \{1, \ldots, |\rho_i|\}$, where $|\rho_i|$ is the number of transitions in ρ. Define $X_i[j]$ as follows:

$$X_i[j] = \begin{cases} 1 & \text{if } \rho_i[j] \to \rho_i[j+1] \text{ is an assignment transition or has guard } \texttt{insample}\geq\texttt{x} \\ -1 & \text{otherwise, in which case } \rho_i[j] \to \rho_i[j+1] \text{ has guard } \texttt{insample<x} \end{cases}$$

[1] Otherwise, s has a G-cycle, and $\mathcal{A}$ is not differentially private. The privacy loss through s is ∞, which matches the L-cost.

Let $\tilde{a_j}$ be the random variable representing the value of `insample` before the jth transition in ρ on input X_i. Let $\tilde{b_j}$ be the random variable representing the value of `insample` before the jth transition in ρ on input X_i'. Further, let $\Gamma_L = \{j : \rho_i[j] \to \rho_i[j+1]$ has guard `insample<x`$\}$, and $\Gamma_G = \{j : \rho_i[j] \to \rho_i[j+1]$ has guard `insample≥x`$\}$.

Notice that $\tilde{a_j} = \tilde{b_j} + 1$ for $j \in \Gamma_L$, and $\tilde{a_j} + 1 = \tilde{b_j}$ for $j \in \{0\} \cup \Gamma_G$. Since $\tilde{a_j}$ is distributed as $Lap(X_i[j], 1/\varepsilon)$, we can write its probability density function as $f(x - X_i[j])$, and its cumulative distribution function as $F(x - X_i[j])$. A similar statement holds for $\tilde{b_j}$.

We may now compute and compare $Pr(\rho_i|X_i')$ and $\Pr(\rho_i|X_i)$ as follows.

$$\begin{aligned}
\Pr(\rho_i|X_i') &= \int_{-\infty}^{\infty} \Pr(\tilde{b_0} = x) \prod_{j\in\Gamma_L} \Pr(\tilde{b_j} < x) \prod_{j\in\Gamma_G} \Pr(\tilde{b_j} \geq x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \Pr(\tilde{b_0} = x) \prod_{j\in\Gamma_L} \Pr(\tilde{b_j} < x) \prod_{j\in\Gamma_G} \Pr(\tilde{b_j} \geq x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} f_\varepsilon(x - X_i[0]) \prod_{j\in\Gamma_L} F_\varepsilon(x - X_i[j]) \prod_{j\in\Gamma_G} (1 - F_\varepsilon(x - X_i[j]))\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} f(x-1)F(x+1)^{\ell_i}(1 - F(x-1))^m \\
&= \int_{-\infty}^{\infty} f(x)F(x+2)^{\ell_i}(1 - F(x))^m \\
&= \exp(2\varepsilon(m+1)) \left(\int_{(-\infty,-2)\cup(2,\infty)} f(x)F(x)^{\ell_i}(1-F(x))^m\,\mathrm{d}x + g(\ell_i) \int_{-2}^{2} f(x)F(x+2)^{\ell_i}(1-F(x))^m \right)
\end{aligned}$$

with $g(\ell_i) \to 1$ as $\ell_i \to \infty$. As we take $\ell_i \to \infty$, we see that

$$h(\ell_i) := \frac{\left(\int_{(-\infty,-2)\cup(2,\infty)} f(x)F(x)^{\ell_i}(1-F(x))^m\,\mathrm{d}x + g(\ell_i) \int_{-2}^{2} f(x)F(x+2)^{\ell_i}(1-F(x))^m \right)}{\Pr(\rho_i|X_i)} \to 1$$

and so as we take the supremum over ρ_i below, we get:

$$\begin{aligned}
\mathrm{loss}(s) \geq \sup_{\rho_i} \frac{\Pr(\rho_i|X_i')}{\Pr(\rho_i|X_i)} &= \exp(2\varepsilon(m+1)) \sup_{\rho_i} \{h(l_i)\} \\
&= \exp(2\varepsilon(m+1))
\end{aligned}$$

We know that S^L is tight, and gives the bound $\exp(2\varepsilon(m+1))$. Thus, we have shown that $\mathrm{loss}(s) = \exp(2\varepsilon(m+1))$, as desired. □

3.2 An alternative coupling strategy: S^J

Definition 3.1. *S^J is a coupling strategy in which we do not couple the noised threshold, but couple the results of all other transitions with twice the cost. [TODO: Describe in more detail]*

Theorem 2. *Let $s = q_0 \to \cdots \to q_m$ be a segment with only L-transitions. If S^J is the least-cost coupling strategy on s, then it provides a tight bound on* $\mathrm{loss}(s)$ *given by*

$$\mathrm{loss}(s) = \sum_{i=1}^{m} 2\varepsilon_i$$

Proof. I have a proof for this, but I will add it into this document soon. [TODO]

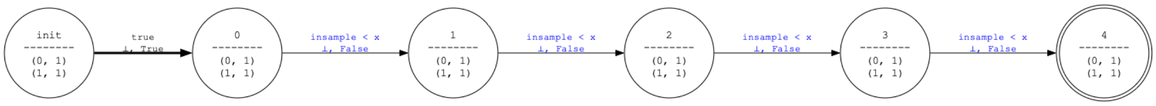

Figure 1: A segment s with only L-transitions.

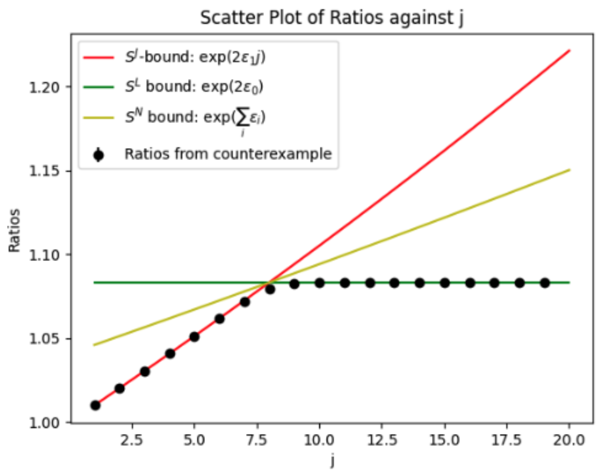

Figure 2:

$\square$

Hypothesis 3.1. *For segments which contain only L-transitions and for which the J-cost exceeds the L-cost, S^L is tight.*

Proof. I think this is true from the graph above, but I need to prove it.

Note June 28 2023: I think this is not true for segments that contain both L-transitions and G-transitions. $\square$

A Lemmata

A.1 Properties of f_ε and F_ε

Lemma 3. *For $x \leq 0$, we have*

$$F_\varepsilon(x) = \exp(2\varepsilon)F_\varepsilon(x-2)$$

and equivalently for $x \leq -2$, we have

$$F_\varepsilon(x+2) = \exp(2\varepsilon)F_\varepsilon(x)$$

Lemma 4. *For $x \geq 0$, we have*

$$1 - F_\varepsilon(x) = \exp(2\varepsilon)(1 - F_\varepsilon(x+2))$$

Lemma 5. *For $x \geq 0$, we have*

$$f_\varepsilon(x) = \exp(2\varepsilon)f_\varepsilon(x+2)$$

A.2 For the proof of Theorem 1

Lemma 6.

$$\int_{-\infty}^{-2} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x = \exp(2\varepsilon\ell)\int_{-\infty}^{-2} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x$$

Proof. From Lemma 3, we have that

$$\begin{aligned}\int_{-\infty}^{-2} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x &= \int_{-\infty}^{-2} f_\varepsilon(x)(\exp(2\varepsilon)F_\varepsilon(x))^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x \\ &= \exp(2\varepsilon\ell)\int_{-\infty}^{-2} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x\end{aligned}$$

□

Lemma 7.

$$\int_{0}^{\infty} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x = \exp(2\varepsilon m)\int_{2}^{\infty} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x$$

Proof. From Lemma 4 and 5, we have that

$$\begin{aligned}\int_{0}^{\infty} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x &= \int_{0}^{\infty} \exp(2\varepsilon)f_\varepsilon(x+2)F_\varepsilon(x+2)^\ell(\exp(2\varepsilon)(1-F_\varepsilon(x+2)))^m \,\mathrm{d}x \\ &= \exp(2\varepsilon m)\int_{0}^{\infty} f_\varepsilon(x+2)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x+2))^m \,\mathrm{d}x \\ &= \exp(2\varepsilon(m+1))\int_{2}^{\infty} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x\end{aligned}$$

□

Lemma 8. *There exists a function $g : \mathbb{N} \to \mathbb{R}$ such that*

$$\int_{-2}^{0} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x = g(\ell)\exp(2\varepsilon(m+1))\int_{-2}^{2} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m \,\mathrm{d}x$$

with $g(\ell) \to 1$ as $\ell \to \infty$.

Proof. I'm not sure yet how to prove this, although I strongly suspect that the $(m+1)$ term comes from the fact that $f_\varepsilon(x)$ is the derivative of $-(1-F_\varepsilon(x))$, and it is taken to the mth power. Its integral should behave like a polynomial of degree $m+1$ evaluated at 2, which corresponds to $\exp(2\varepsilon(m+1))$. □