1 DiPA*

Definition 1.1. Fix parameters ϵ, N. Let C be the guard conditions $\{\texttt{true}, \texttt{insample} \geq \texttt{x}, \texttt{insample} < \texttt{x}, n \geq N\}$ A **DiP* automaton** (DiPA*) $\mathcal{A}$ is defined as the tuple $\mathcal{A} = (Q, \Sigma, \Gamma, q_0, X, P, \delta)$, where:

- Q = finite set of states; partitioned into input states Q_{in} and non-input states Q_{non}
- Σ is the input alphabet (taken to be $\mathbb{R}$)
- Γ is a finite output alphabet
- $q_0 \in Q$ is the starting state
- X= $\{\texttt{x}, \texttt{insample}, \texttt{insample'}, \texttt{n}\}$ is a set of variables. $\texttt{x}, \texttt{insample}, \texttt{insample'} \in \mathbb{R}$; $\texttt{n} \in \mathbb{N}$ and is initialized to 0.
- $P : Q \to \mathbb{Q}^{\geq 0} \times \mathbb{Q} \times \mathbb{Q}^{\geq 0} \times \mathbb{Q}$ describing the parameters for sampling from Laplace distributions at each state.
- $\delta : (Q \times C) \to Q \times (\Gamma \cup \{\texttt{insample}, \texttt{insample'}\} \cup \{\phi\}) \times \{\texttt{true}, \texttt{false}\} \times \{0, 1\}$ is the transition function (technically a relation) that defines what state to transition to, what symbol or real value to output, whether or not x is assigned to, and whether or not n is incremented based on the current state and transition guard.

There are certain conditions that δ must satisfy; these are almost all the same as the restrictions on transition functions of DiPA, but with some slight modifications and one major addition (marked in blue):

- **Determinism:** For any state $q \in Q$, if $\delta(q, \texttt{true})$ is defined, then $\delta(q, \texttt{insample} \geq \texttt{x})$, $\delta(q, \texttt{insample} < \texttt{x})$, and $\delta(q, n \geq N)$ are undefined. In addition, if two guard conditions are simultaneously true, then the automaton will follow the transition with guard $n \geq N$.
- **Output Distinction:** For any state $q \in Q$, if $\delta(q, \texttt{insample} \geq \texttt{x}) = (q_1, o_1, b_1, i_1)$ and $\delta(q, \texttt{insample} < \texttt{x}) = (q_2, o_2, b_2, i_2)$, then $o_1 \neq o_2$ and at least one of $o_1 \in \Gamma$ and $o_2 \in \Gamma$ is true. In addition, $o_1 \neq \phi$ and $o_2 \neq \phi$ and if $\delta(q, n \geq N) = (q', o', b', i')$, then $o' = \phi$, i.e., the ϕ output symbol is reserved for transitions with guard $n \geq N$, which must output ϕ.
- **Initialization:** The initial state q_0 has only one outgoing transition of the form $\delta(q_0, \texttt{true}) = (q, o, \texttt{true}, i)$.
- **Non-input transition:** From any $q \in Q_{non}$, if $\delta(q, c)$ is defined, then $c = \texttt{true}$.
- **Control Flow Separation:** Consider the underlying graph G of $\mathcal{A}$. For all states $q \in Q$, if $\delta(q, n \geq N) = (q', o, b, i)$, let G' be the graph of G with the edge corresponding to the transition with guard $n \geq N$ from q removed. Then q and q' must be in different connected components of G'.

Note that the **control flow separation** condition implies that no cycle in G can contain an edge that corresponds to a transition with guard $n \geq N$. In addition, determinism combined with control flow separation imply that no two transitions (i.e. transitions with different guards) can be from some state q to the same state q'.

1.1 TODO: path probabilities

tldr $n \geq N$ transitions $\rightarrow$ same probability as `true` if guard condition is met, and $\texttt{insample} \geq \texttt{x}, \texttt{insample} < \texttt{x}$ transitions also need to take this into account.

2 Violations of Differential Privacy

Definition 2.1. A **bounded** cycle C in a DiPA* $\mathcal{A}$ is a cycle in $\mathcal{A}$ where there exists at least one transition $(q', \sigma, t, 1)$ (i.e. `n` gets incremented) and there exists some $q \in Q$ ("exit state") in the cycle such that $f(q, n \geq N) = (q', \sigma, t, i)$ where q' is not in the cycle. Otherwise, the cycle is **unbounded**.

Definition 2.2. A cycle C with an exit state with transition $n \geq N$ is a **trivially exiting** cycle if, for *all* paths $\rho = q_0 q_1 ... q_m)$ from the start state to a state $q_m \in C$, at least N transitions $q_i \rightarrow q_{i+1}$ are increment transitions or some transition $q_i \rightarrow q_{i+1}$ has guard $n \geq N$.

Proposition 2.3. *If a DiPA* $\mathcal{A}$ *has a reachable unbounded leaking non-trivially exiting cycle, then it is not differentially private.*

Proof. Let $C = a_1 a_2 ... a_{n-1} a_n; a_1 = a_n$ be such a cycle in $\mathcal{A}$. We will reduce the analysis to an analogous DiPA.

Case 1: C does not have an exit state.

Consider an abstract path $\eta = q_0 \sigma_0 q_1 ... q_{m+n-1} \sigma_{m+n-1} q_m$ such that $a_1 ... a_n = q_m ... q_{m+n}$ (i.e. the last n states of the path are the cycle C).

For $\ell > 0$, let η_ℓ be the path $\eta_\ell = q_0 \sigma_0 q_1 \sigma_1 ... q_{m+\ell n-1} \sigma_{m+\ell n-1} q_{m+\ell n}$ such that $q_k = q_{k-n}$ and $\sigma_k = \sigma_{k-n}$ for all $m + n \leq k \leq m + \ell n$. This is the path η with the cycle C repeated ℓ times. Note that because C has no exit state, for all states $a_i \in C$, all transitions from a_i have a guard that is *not* $n \geq N$. This means that the path η_ℓ in $\mathcal{A}$ exists for all $\ell > 0$. Thus, the same input sequences α_ℓ and β_ℓ as described in Lemma 6 of [1] are witnesses to a violation of differential privacy. In particular, the same analysis holds because there is some fixed number f such that η_ℓ has at most f transitions with guard $n \geq N$, even as ℓ can vary arbitrarily.

Case 2: Suppose that C has no increment transition.

Because C is non-trivially exiting, there exists some path $\eta = q_0 \sigma_0 q_1 \cdots q_{m+n-1} \sigma_{m+n-1} q_m$ such that $a_1 \cdots a_m = q_m \cdots q_{m+n}$ and at $q_m = a_1$, $n < N$.

As in Case 1, for $\ell > 0$, consider $\eta_\ell = q_0 \sigma_0 q_1 \sigma_1 \cdots q_{m+\ell n-1} \sigma_{m+\ell n-1} q_{m+\ell n}$ such that $q_k = q_{k-n}$ and $\sigma_k = \sigma_{k-n}$ for all $m + n \leq k \leq m + \ell n$. Because there are no increment transitions in C, $\forall 0 \leq i \leq \ell n$, $n < N$ at state q_i. So for all states $a_i \in C$, a transition from a_i with guard

$n \geq N$ will never be taken by $\mathcal{A}$. As before, then, the path η_ℓ in $\mathcal{A}$ exists for all $\ell > 0$, so α_ℓ and β_ℓ from Lemma 6 of [1] are witnesses to a violation of differential privacy.

□

Definition 2.4. A bounded pair is a pair of cycles (C, C') such that at least one of C and C' is a bounded cycle. Similarly, a pair of cycles (C, C') is trivially exiting if at least one of C and C' are trivially exiting.

Proposition 2.5. *If a DiPA* $\mathcal{A}$ has an unbounded non-trivially exiting leaking pair of cycles (C, C') where C is reachable, then it is not differentially private.*

Proposition 2.6. *If a DiPA* $\mathcal{A}$ has a reachable unbounded non-trivially exiting disclosing cycle, then it is not differentially private.*

Definition 2.7. A unbounded and non-trivially exiting privacy violating path is a path ρ of length n in a DiPA* $\mathcal{A}$ such that one of the following hold:

- $\texttt{tail}(\rho)$ is an $\texttt{AG}$-path such that $\texttt{last}(\rho)$ is in a unbounded non-triviallt exiting $\texttt{G}$-cycle and the 0th transition is an assignment transition that outputs `insample` (or $\texttt{AL}, \texttt{L}$-cycle, respectively)
- ρ is an `AG`-path such that $\texttt{first}(\rho)$ is in a unbounded non-trivially exiting `G`-cycle and the 0th transition has guard `insample < x` and outputs `insample` (or similar with `L`-cycles)
- ρ is an `AG`-path such that $\texttt{first}(\rho)$ is in a unbounded non-trivially exiting `L`-cycle and the last transition has guard $\texttt{insample} \geq \texttt{x}$ and outputs `insample` (or similar)

Proposition 2.8. *If a DiPA* $\mathcal{A}$ has an unbounded non-trivially exiting privacy violating path, then it is not differentially private.*

If this text is still here, I didn't have time to write out the details, but propositions 2.5, 2.6, and 2.8 are all proved extremely simlarly to proposition 2.3.

Definition 2.9. A DiPA* $\mathcal{A}$ is well-formed if $\mathcal{A}$ has no unbounded non-trivially exiting leaking cycles, unbounded non-trivially exiting leaking pairs, unbounded non-trivially exiting disclosing cycles, or unbounded non-trivially exiting privacy violating paths.

Theorem 2.10. *If a DiPA* is not well-formed, then it is not differentially private.*

Proof. Follows from propositions 2.3, 2.5, 2.6, and 2.8. □

3 Proving Differential Privacy

Let $t = (p, c, q, o, b)$ be a transition in a DiPA* $\mathcal{A}$ with parameters ϵ, N, where the transition is from $p \to q$, c is the guard of the transition, o is the output of the transition, and b is whether or not it assigns to x. Let d, μ be the parameters for sampling `insample` and d', μ' be the parameters for sampling `insample'` at state p. t is a critical transition if it is not in a cycle in $\mathcal{A}$ or, if it is in a cycle C, that C trivially exits. t is an input transition if p is an input state.

We will assign a cost to each transition as follows:

$$cost(t) = \begin{cases} d & t \text{ is a critical non-input transition} \\ 2d & t \text{ is a critical input transition and } o \neq \texttt{insample'} \\ 2d+d' & t \text{ is a critical input transition and } o = \texttt{insample'} \\ dN & t \text{ is a non-input transition on a bounded cycle} \\ 2dN & t \text{ is a critical input transition on a bounded cycle and } o \neq \texttt{insample'} \\ (2d+d')N & t \text{ is a critical input transition on a bounded cycle and } o = \texttt{insample'} \\ 0 & \text{otherwise} \end{cases}$$

Note to self: this may not be necessary? if each path already accounts for this

Theorem 3.1. *If a DiPA* is well-formed, then it is differentially private.*

Proof. Let $\mathcal{A}^* = (Q, \Sigma, \Gamma, q_0, X^*, P^*, \delta^*)$ be a well-formed DiPA* with parameters ϵ and N.

Let $G = \{\texttt{true}, \texttt{insample} \geq \texttt{x}, \texttt{insample} < \texttt{x}\}$ be the set of guard conditions for DiPAs.

For $n \in \mathbb{N}$, let $[N] = \{0, 1, \cdots, N\}$. Construct the DiPA $\mathcal{A} = (Q \times [N], \Sigma, \Gamma \cup \{\phi\}, (q_0, 0), X, P, \delta)$ as follows:

For each state $q \in Q^*$:

For $g \in G$, if $\delta^*(q, g) = (q', \sigma, \mathbf{b}, x)$ is defined, then for all $k \in [N-1]$, define the transition

$$\delta((q, k), g) = ((q', k + x), \sigma, \mathbf{b})$$

If $\delta^*(q, n \geq N) = (q', \sigma, \mathbf{b}, x)$ is defined, then define the transition

$$\delta((q, N), \texttt{true}) = ((q', N), \sigma, \mathbf{b})$$

Otherwise if $\delta^*(q, g) = (q', \sigma, \mathbf{b}, x)$ is defined, then define the transition

$$\delta((q, N), g) = ((q', N), \sigma, \mathbf{b})$$

Note that we need to separate out the $k = N$ case to satisfy the condition of determinism.

Intuitively, at state (q, k) in $\mathcal{A}$, k will track the value of n in $\mathcal{A}^*$ (since everything above N is treated the same, we compress all of those values together).

First, **for every bounded cycle C^* in $\mathcal{A}^*$:** We will "unroll" C^* to loop for N iterations.

For each state $q^* \in C^*$:

For $0 \leq k \leq N$, $q^{*(k)} \in Q$; i.e. we create $N+1$ copies of C^*.

For guard conditions $g \in G$ and if $0 \leq k < N$, we define new transitions as follows:

If $\delta^*(q^*, g) = (q', \sigma, \mathbf{b}, x)$ is defined, then if $x = 0$,

$$\delta(q^{*(k)}, g) = \begin{cases} (q'^{(k)}, \sigma, \mathbf{b}) & \text{if } q' \text{ is in } \textit{any} \text{ bounded cycle} \\ (q', \sigma, \mathbf{b}) & \text{otherwise} \end{cases}$$

Otherwise if $x = 1$, then

$$\delta(q^{*(k)}, g) = \begin{cases} (q'^{(k+1)}, \sigma, \mathbf{b}) & \text{if } q' \text{ is in } \textit{any} \text{ bounded cycle} \\ (q', \sigma, \mathbf{b}) & \text{otherwise} \end{cases}$$

Now consider when $k = N$. If $\delta^*(q^*, n \geq N) = (q', \sigma, \mathbf{b}, x)$ is defined,

$$\delta(q^{*(N)}, \texttt{true}) = \begin{cases} (q'^{(N)}, \sigma, \mathbf{b}) & \text{if } q' \text{ is in } \textit{any} \text{ bounded cycle} \\ (q', \sigma, \mathbf{b}) & \text{otherwise} \end{cases}$$

Otherwise for $g \in G$, if $\delta^*(q^*, g) = (q', \sigma, \mathbf{b}, x)$ is defined,

$$\delta(q^{*(N)}, g) = \begin{cases} (q'^{(N)}, \sigma, \mathbf{b}) & \text{if } q' \text{ is in } \textit{any} \text{ bounded cycle} \\ (q', \sigma, \mathbf{b}) & \text{otherwise} \end{cases}$$

Additionally, **for every unbounded trivially exiting cycle C^* in $\mathcal{A}$:**

For each $q^* \in C^*$, $q^* \in Q$. For guard conditions $g \in G$, if $\delta^*(q^*, g) = (q', \sigma, \mathbf{b}, x)$, then $\delta(q^*, g) = (q', \sigma, \mathbf{b})$. Otherwise, if $\delta^*(q^*, n \geq N) = (q', \sigma, \mathbf{b}, x)$, then $\delta(q^*, \texttt{true}) = (q', \sigma, \mathbf{b})$. As before, also remove all non-true transitions from q^*.

If q^* is not in a bounded or trivially exiting cycle, then $q^* \in Q$.

Let a be the minimum number of increment transitions over *all* paths from the start state q_0^* to q^* in $\mathcal{A}^*$.

Then for each guard $g \in G$, if $\delta^*(q^*, g) = (q', \sigma, \mathbf{b}, 0)$, then

$$\delta(q^*, g) = \begin{cases} (q'^{(\min\{N,a\})}, \sigma, \mathbf{b}) & \text{if } q' \text{ is in a bounded cycle } C \\ (q', \sigma, \mathbf{b}) & \text{if } q' \text{ is not in a bounded cycle} \end{cases}$$

Otherwise if $\delta^*(q^*, g) = (q', \sigma, \mathbf{b}, 1)$, then

$$\delta(q^*, g) = \begin{cases} (q'^{((\min\{N,a+1\})}, \sigma, \mathbf{b}) & \text{if } q' \text{ is in a bounded cycle } C \\ (q', \sigma, \mathbf{b}) & \text{if } q' \text{ is not in a bounded cycle} \end{cases}$$

For each state $q \in Q$, let $P(q) = P(q^*)$ for the corresponding state $q^* \in Q^*$.

Lemma 3.2. *If $\mathcal{A}^*$ is well-formed, then $\mathcal{A}$ is well-formed.*

This follows from the following lemma[1]:

Lemma 3.3. *If there exists a cycle C in $\mathcal{A}$, then there exists an unbounded non-trivially exiting cycle C^* in $\mathcal{A}^*$.*

Proof. Let $C = a_0 a_1 \cdots a_{m-1} a_0$ be a cycle in $\mathcal{A}$. Let T be the set of transitions in $\mathcal{A}$ and let T^* be the set of transitions in $\mathcal{A}^*$ where each transition is described by a 5-tuple (p, c, q, o, b) as before.

Note that by the construction of $\mathcal{A}$, for each state $q \in Q$, there exists some corresponding state $q^* \in Q^*$. Let $f_q : Q \to Q^*$ be a function describing this correspondence. Similarly, note that each transition $t = (p, c, q, o, b)$ corresponds to a transition t^* in $\mathcal{A}^*$ and let $f_t : T \to T^*$ describe this correspondence (f_q and f_t are not necessarily injective nor surjective).

Let $a_i, a_{i+1} \in C$ (for convenience, assume that a_{i+1} is actually $a_{i+1 \bmod m}$). Let t_i be the transition (a_i, c, a_{i+1}, o, b) (recall that t_i is unique because of determinism and control flow separation). Note that $f_t(t_i) = (f_q(a_i), c', f_q(a_{i+1}), o', b')$ from the construction of $\mathcal{A}$.

Then the path $C^* = f_q(a_0) f_q(a_1) \cdots f_q(a_{m-2}) f_q(a_{m-1}) f_q(a_0)$ exists in $\mathcal{A}^*$ and is a cycle.

Suppose that C^* is a bounded cycle for the sake of contradiction. Then by construction of $\mathcal{A}$, all a_i must be of the form $a_i = f_q(a_i)^{(x)}$ for some x. In addition, note that there cannot[2] be any transitions $\mathcal{A}$ from a state $q_1^{(x)}$ to a state $q_2^{(y)}$ for $y < x$. Let $a_0 = f_q(a_0)^{(k)}$ for some fixed k. Then, $a_{m-1} = f_q(a_{m-1})^{(k)}$. However, this is impossible, since there must be some increment transition in C^* such that for some a_i, $a_i = f_q(a_i)^{(k+1)}$ because C^* is a bounded cycle. Thus, C^* cannot be a bounded cycle.

Now suppose that C^* is a trivially exiting cycle for the sake of contradiction. Let $f_q(a_i)$ be an exit state of C^*. Then $\delta(a_i, \texttt{true}) = (a_{i+1} = q', \sigma, \mathbf{b})$ for some q' outside of C^* by definition of an exit state and by construction of $\mathcal{A}$. However, it is then impossible for C to be a cycle because of control flow separation, since it would imply that, in the graph of $\mathcal{A}^*$ with the $f_q(a_i) \to f_q(q')$ transition removed, the component that contains $f_q(a_i)$ and the component that contains $f_q(q')$ are the same. Thus, C^* is non-trivially exiting. □

Lemma 3.4. *If $\mathcal{A}$ is $wt(\mathcal{A})$-DP, then $\mathcal{A}^*$ is $wt(\mathcal{A})$-DP.*

Proof. Let ρ^*, ρ'^* be two equivalent paths on adjacent inputs in $\mathcal{A}^*$. We will show that there exist paths ρ, ρ' in $\mathcal{A}$ such that $\frac{\mathbb{P}[\epsilon,\rho^*]}{\mathbb{P}[\epsilon,\rho'^*]} \leq \frac{\mathbb{P}[\epsilon,\rho]}{\mathbb{P}[\epsilon,\rho']}$.

sketch: for each state traversed in ρ^*, there is an equivalent state traversed in ρ, except for the fact that $\mathcal{A}$ may possibly enter bounded cycles "earlier' (i.e. with smaller n) than $\mathcal{A}^*$ does.

□

[1]TODO: check if this connection needs to be elaborated on

[2]Unsure if this is trivial or needs to be elaborated on

Lemmas 3.2 and 3.4 together prove the theorem. □

References

[1] Rohit Chadha, A. Prasad Sistla, and Mahesh Viswanathan. On Linear Time Decidability of Differential Privacy for Programs with Unbounded Inputs, April 2021.