A Simple Proof of a Distinguishing Bound of Iterated Uniform Random Permutation

Mridul Nandi

Applied Statistics Unit
Indian Statistical Institute, Kolkata
mridul.nandi@gmail.com,

Abstract. Let $\mathbf{P}$ be chosen uniformly from the set $P := \mathsf{Perm}(S)$, the set of all permutations over a set S of size N. In Crypto 2015, Minaud and Seurin proved that for any unbounded time adversary A, making at most q queries, the distinguishing advantage between $\mathbf{P}^r$ (after sampling $\mathbf{P}$, compose it for r times) and $\mathbf{P}$, denoted $\Delta_A(\mathbf{P}^r\ ;\ \mathbf{P})$, is at most $(2r+1)q/N$. In this paper we provide an alternative simple proof of this result for an upper bound $2q(r+1)^2/N$ by using well known coefficient H-technique.

Keywords: iterated random permutation, blockcipher, cascade encryption.

1 Introduction

Let S be a set of size $N > 0$. Let $P := \mathsf{Perm}(S)$ denote the set of all permutations over the set S. For any finite set X, we denote the uniform random variable over X as $\mathbf{X}$. In this notation, $\mathbf{P}$, called **uniform random permutation**, is the uniform random variable chosen from the set of all permutations over S. Let $\mathbf{P}^r := \mathbf{P} \circ \cdots \circ \mathbf{P}$ (r times) define the transformed random variable which is obtained by composing $\mathbf{P}$ for r times. It is also known as iterated random permutation. Similarly, $\mathbf{P}^{-1}$ is the random variable applying inverse after we sample $\mathbf{P}$.

Given a block cipher E over the plaintext space S, the cascade encryption of length r encrypts a message x as $E_{k_r} \circ \cdots \circ E_{k_1}(x)$. It has been extensively studied in the setting where the keys $k_1, \ldots, k_r$ are chosen uniformly and independently. However, before [1, 2], virtually nothing is known regarding the security of cascade encryption when the keys are identical. Analyzing iterated random permutation is relevant for the cascaded (or multiple) encryption schemes under same key. In Crypto 2015, Minaud and Seurin [2] proved that for any unbounded time adversary A, making at most q forward an inverse queries, the distinguishing advantage between $\mathbf{P}^r$ and $\mathbf{P}$, denoted $\Delta_A(\mathbf{P}^r\ ;\ \mathbf{P})$, is at most $(2r+1)q/N$. By using the standard reduction, one can claim that $E_k^r(\cdot)$ is strong pseudorandom permutation as long as $q = o(N/r)$ and E is also a strong pseudorandom permutation with the same number of queries. The main result of [2] is to show that $\Delta_A(\mathbf{P}^r\ ;\ \mathbf{P}) \leq (2r+1)q/N$.

1.1 Proof Ideas

We first describe the proof idea of [1]. Minaud and Seurin use the game playing technique as a language of the proof. In between the games corresponding to the oracles $\mathbf{P}$ and $\mathbf{P}^r$, they considered intermediate games $\mathbf{C}$ and $\mathbf{C}^r$ where $\mathbf{C}$ is a permutation chosen uniformly from the set C of all cycles over S. The result will follow by bounding the two terms (1) $\Delta_A(\mathbf{P}\ ;\ \mathbf{C})$ and (2) $\Delta_A(\mathbf{C}\ ;\ \mathbf{C}^r)$. The bound of the first term follows from the observation that the response of $\mathbf{C}$ (and $\mathbf{P}$) on the i^{th} query is uniformly distributed from a set of size $N-i$ (and $N-i+1$ respectively). The $\mathbf{C}^r$ is nothing but union of d cycles of same size where $d = (N, r)$ (g.c.d. of N and r). To bound the

second term, the authors showed that $\mathbf{C}^r$ and $\mathbf{C}$ are identical except for queries belong to a secretly chosen set of size r.

In our proof, we chose the intermediate game in a way so that we do not need to consider the second term. In particular, we consider an intermediate oracle $\mathbf{C}_{\ell,1}$, a permutation chosen uniformly from the set of all permutations union of ℓ self loops and a cycle. If we choose ℓ such that $(\ell, r) = 1$ then $\mathbf{C}^r_{\ell,1} = \mathbf{C}_{\ell,1}$. So we do not need to bound the second term. The bounding first term, however, would be similar.

2 Preliminaries

We quickly recall the basic definition, notation and known results which would be used in the paper.

2.1 Partial functions and Associated Graphs

Any set of pairs $v \subseteq S^2$ such that $(x, y_1), (x, y_2) \in F \Rightarrow y_1 = y_2$ is called partial function relation. Given such a set v we associate a partial function, abusing notation, $v : S \to S$ such that $(x, y) \in v$ if and only if $v(x) = y$. In this paper the partial function and the partial function relation will be identified. We write $Dom(v) = \{x \in S : \exists y \in S,\ (x, y) \in v\}$. Given any partial function v, we associate a directed graph $G^v := (S, E^v)$ where $E^v = \{(a, v(a)) : a \in Dom(v)\}$. In this paper, unless mentioned otherwise, we assume only injective partial functions (subset of a permutation relation). Some simple observations are as follows.

1. The graph G^π for a permutation π is union of cycles (allowing self-loops, cycles of size 1).
2. For any partial (injective) function v with domain size q, the graph G^v is a union of $a \geq 1$ straight line graphs and b cycles (removing the isolated nodes) with total number of nodes $q + a$.

2.2 Coefficient H-technique to bound Distinguishing Advantage

Interpolation Probability. Patarin's coefficient-H technique [3] (see also [4]) is a tool for showing an upper bound for the distinguishing advantage. It mainly requires to compute interpolation probabilities as defined below.

Definition 1. *For any $P' \subseteq P$, and a partial injective function $v = \{(x_1, y_1), \ldots, (x_q, y_q)\}$ with (domain) size q, the interpolation probability is defined as*

$$\mathcal{I}_{P'}(v) = Pr[\mathbf{P}'(x_1) = y_1, \ldots, \mathbf{P}'(x_q) = y_q] = \frac{\#\{\pi \in P' : \pi(x_i) = y_i, 1 \leq i \leq q\}}{\#P'}.$$

If $\pi(x_i) = y_i$ for all $1 \leq i \leq q$ then we also say that π is an extension of v. In this case, $v \subseteq \pi$ (viewing as a function relation) and G^v is a subgraph of G^π. For any partial injective function v of size q, the number of permutations extending v is exactly $(N-q)!$ as we can only choose a bijection function over a set of size $(N - q)$ freely. This proves the following lemma for the interpolation probability of the uniform random permutation.

Lemma 1. *For any partial injective function v with domain size $q \leq N$, $\mathcal{I}_{P'}(v) = 1/P(N, q)$ where $P(a, b) = a(a-1)\cdots(a-b+1)$, $1 \leq b \leq a$.*

Let $C_1 = \mathsf{Cycl}(S)$, the set of all cycles over S. We fix an element $x_0 \in S$. A cycle can map x_0 to any element x_1 from $S \setminus \{x_0\}$. Similarly, image of x_1 can be any element from $S \setminus \{x_0, x_1\}$ and so on. So $|C_1| = (N-1)!$. Now given a partial injective function v of size $q < N$ with no cycle present in G^v (if there is a cycle of size smaller than N then clearly we can not extend it) then we now show that the number of cycles extending v is exactly $(N-q-1)!$.

Lemma 2. *Let $v = \{(x_1, y_1), \ldots, (x_q, y_q)\}$ be a partial injective function of size $q < N$ with no cycle. Then,*

$$\#\{\pi \in C_1 : \pi(x_1) = y_1, \ldots, \pi(x_q) = y_q\} = (N-q-1)!.$$

Proof. Suppose G^v is decomposed into a straight line graphs $L_0, L_1, \ldots, L_{a-1}$. We denote s_i and t_i to represent the staring and end node of L_i, $0 \le i \le a$. Let $\pi \in C_1$ be any extension of v and so G^v is a subgraph of the cyclic graph G^π. So these a straight lines actually produce a gaps in the cycle G^π. Now, the L_i's can be in any order and there are $(a-1)!$ such orders. For each fixed order, we can choose a non-negative integers $x_1, \ldots x_a$, representing the number of nodes present in the gap. Note that the number of non-isolated nodes in G^v is exactly $q+a$. Thus, we can choose any non-negative x_i's such that $x_1 + \cdots + x_a = N' := N - q - a$. We know that there are $\binom{N'+a-1}{a-1}$ such solutions. If we fix any such solution, the N' isolated nodes in G^v can appear in any order in G^π. So we can order them in $N'!$ ways. Thus, the number of cycles extending v is exactly

$$(a-1)! \times N'! \times \binom{N'+a-1}{a-1} = (N-q-1)!. \qquad \square$$

Corollary 1. *For any partial injective function v with domain size $q \le N$ and no cycle, $\mathcal{I}_{C_1}(v) = 1/P(N-1, q)$.*

Another interesting class of permutations is $C_{\ell,1}$ which is union of exactly ℓ self-loops and a cycle with the rest of the nodes. Here we assume that $0 \le \ell \le N-2$. Clearly, $|C_{\ell,1}| = \binom{N}{\ell} \times (N-\ell-1)!$. Now given a partial injective function q of size q with no cycle, the number of permutations from $C_{\ell,1}$ extending v is at least $\binom{N-2q}{\ell} \times (N-\ell-q-1)!$. This is because, we first choose a set A of ℓ nodes from the isolated nodes of G^v and then we construct a cycle in $S \setminus A$ extending v. So

$$\mathcal{I}_{C_{\ell,1}}(v) \ge \frac{\binom{N-2q}{\ell} \times (N-\ell-q-1)!}{\binom{N}{\ell} \times (N-\ell-1)!}.$$

The right hand side of the above equation further can be lower bounded by $(1 - \frac{q\ell}{N-q-\ell}) \times \frac{1}{P(N,q)}$. So we have the following result.

Lemma 3. *For any partial injective function v with domain size $q \le N$ and no cycle,*

$$\mathcal{I}_{C_{\ell,1}}(v) \ge (1 - \frac{q\ell}{N-q-\ell}) \times \frac{1}{P(N,q)} = (1 - \frac{q\ell}{N-q-\ell}) \times \mathcal{I}_P(v).$$

Distinguishing Advantage. Let $P_1, P_2 \subseteq P$ and A be an oracle algorithm which makes at most q queries (both forward and backward or inverse). We define

$$\Delta_A^{\pm}(\mathbf{P}_1 \ ; \ \mathbf{P}_2) := |\mathsf{Pr}[A^{\mathbf{P}_1, \mathbf{P}_1^{-1}} = 1] - \mathsf{Pr}[A^{\mathbf{P}_1, \mathbf{P}_1^{-1}} = 1]|.$$

Let $\Delta_q^{\pm}(\mathbf{P}_1 \ ; \ \mathbf{P}_2) := \max_A \Delta_A^{\pm}(P_1 \ ; \ P_2)$ where maximum is taken over all adversaries making at most q forward and backward queries. It is easy to see that Δ_A and Δ_q satisfy the triangle inequality. Now we state a known result which can be proved by using standard reduction argument.

Fact 1. $\Delta_q^{\pm}(\mathbf{P}_1^r \ ; \ \mathbf{P}_2^r) \leq \Delta_{rq}^{\pm}(\mathbf{P}_1 \ ; \ \mathbf{P}_2)$.

Now we describe the Coefficient H-technique (expressed in our notation) which would be used to prove the main result. Let F_q denote the set of all injective partial functions of size q.

Theorem 1 (Patarin [4]). *Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be two oracle algorithms over S. Suppose there exist a set of partial functions $\mathcal{V}_{bad} \subseteq F_q$ and $\varepsilon > 0$ such that the following conditions hold:*

1. *For all $\{(x_1, y_1), \ldots, (x_q, y_q)\} \notin \mathcal{V}_{bad}$,*

$$\Pr[\mathcal{O}_1(x_1) = y_1, \ldots, \mathcal{O}_1(x_q) = y_q] \geq (1 - \varepsilon) \Pr[\mathcal{O}_2(x_1) = y_1, \ldots, \mathcal{O}_2(x_q) = y_q]$$

   (the above probabilities are defined as interpolation probabilities).
2. *For all A making at most q queries to $\mathcal{O}_2$, $\Pr[\mathrm{Trans}(A^{\mathcal{O}_2}) \in \mathcal{V}_{bad}] \leq \delta$ where $\mathrm{Trans}(A^{\mathcal{O}_2}) = \{(x_1, y_1), \ldots, (x_q, y_q)\} \in F_q$, x_i and y_i denote the i^{th} query and response of A to $\mathcal{O}_2$.*

Then,

$$\Delta_q(\mathcal{O}_1 \ ; \ \mathcal{O}_2) \leq \varepsilon + \delta.$$

The above result can be applied for more than one oracle, in particular a permutation and its inverse chosen uniformly from a subset P' of P. If we have an oracle $\mathcal{O}$ and its inverse $\mathcal{O}^{-1}$ then the interpolation probability for both $\mathcal{O}$ and $\mathcal{O}^{-1}$ can be simply expressed through the interpolation probability of $\mathcal{O}$ only. For example, if an adversary makes a query y to $\mathcal{O}^{-1}$ and obtains the response x, we can write $\mathcal{O}(x) = y$. Therefore, under the conditions of Theorem 1 we also have $\Delta_q^{\pm}(\mathcal{O}; \mathcal{O}') \leq \varepsilon + \delta$.

3 Iterated Random Permutation is SPRP Indistinguishable

Theorem 2 (Minaud and Seurin [2]). *For any positive integers $q, r \leq N$,*

$$\Delta_q^{\pm}(\mathbf{P}^r, \mathbf{P}) \leq \frac{q(2r+1)}{N}.$$

In this paper we provide a simpler proof for a little bit larger bound in terms of the order of r. However, when r is trated as constant the both bound matched in order which is $\Theta(q/N)$.

Theorem 3. *For any positive integers $q, r \leq N$,*

$$\Delta_q^{\pm}(\mathbf{P}^r, \mathbf{P}) \leq \frac{2q(r+1)^2}{N}.$$

Proof. There exists $0 \leq \ell < r$ such that $(N - \ell, r) = 1$. In fact, if $rk < N \leq (k+1)N$ for some integer k then we define $N - \ell = rk + 1$ and clearly, $(rk + 1, rk) = 1$. So we take $\ell = N - rk - 1$. Now for all $\pi \in C_{\ell,1}$, $\pi^r = \pi$. In other words, $\mathbf{C}_{\ell,1}^r = \mathbf{C}_{\ell,1}$ as a random variable. Let

$$\mathcal{V}_{bad} = \{v : G^v \text{ has a loop for a partial function} v \text{ over } S\}.$$

We have already seen that for all injective $v \in F_q \setminus \mathcal{V}_{bad}$ with $|Dom(v)| = q$,

$$\mathcal{I}_{C_{\ell,1}}(v) \geq (1 - \ell q/(N - q - \ell)) \times \mathcal{I}_P(v).$$

Now we show the following claim.

Claim 1. For all A making at most q queries $\mathsf{Pr}[\mathrm{Trans}(A^{\mathbf{P},\mathbf{P}^{-1}}) \in \mathcal{V}_{bad}] \leq q/(N - q + 1)$.

Proof of the Claim. Suppose the cycle occurs first time on i^{th} query. Let v_{i-1} denote the transcript up to $i - 1$ queries. Let us assume that the i^{th} query be x_i which is a $\mathbf{P}$-query. Then it can make a cycle only if the response takes the value x where $x = x_i$ if x_i is isolated in $G^{v_{i-1}}$, else it is the starting node of the straight line graph for which x_i is the end node. So this can happen with probability exactly $1/(N - i + 1) \leq 1/(N - q + 1)$. Summing over all i, the claim is proved. □

Using coefficient H-technique as described in Sect. 2, we have

$$\Delta_q(\mathbf{C}_{\ell,1} \ ; \ \mathbf{P}) \leq \frac{q\ell}{N - q - \ell} + \frac{q}{N - q + 1} \leq \frac{2q(r + 1)}{N}.$$

Here we assume that $q + r \leq N/2$. Now, $\Delta_q(\mathbf{P}^r \ ; \ \mathbf{P}) \leq \Delta_q(\mathbf{P}^r \ ; \ \mathbf{C}^r_{\ell,1}) + \Delta_q(\mathbf{C}_{\ell,1} \ ; \ \mathbf{P}) \leq \Delta_{rq}(\mathbf{P} \ ; \ \mathbf{C}_{\ell,1}) + \Delta_q(\mathbf{C}_{\ell,1} \ ; \ \mathbf{P})$. Using the above bound (replacing q by rq for bounding the first term) we have

$$\Delta_q(\mathbf{P}^r, \mathbf{P}) \leq \frac{2q(r + 1)^2}{N}. \quad \square$$

References

1. Brice Minaud and Yannick Seurin. The iterated random permutation problem with applications to cascade encryption. Technical report, Cryptology ePrint Archive, Report 2015/504, 2014. http://eprint. iacr. org.
2. Brice Minaud and Yannick Seurin. The iterated random permutation problem with applications to cascade encryption. In *To be appeared in Advances in Cryptology, CRYPT0-15*, volume xxxx, pages xxx–xxx. Springer, 2015.
3. J. Patarin. Etude des Générateurs de Permutations Basés sur le Schéma du D.E.S. Phd Thèsis de Doctorat de l'Université de Paris 6, 1991.
4. Jacques Patarin. The coefficients h technique. In *Selected Areas in Cryptography*, pages 328–345. Springer, 2009.