CS6222 Cryptography

Topic: Construction of PRG from OWF
Lecturer: Wei-Kai Lin (TA: Arup Sarker)

Date: Oct 8, 2024
Scriber: Liran Li

---

1 Construction of PRG from OWF

Today, we continue and finish the proof of the following Theorem by Goldreich and Levin. Let's recall and continue.

Theorem 1. ***Construction of Hard-Core Predicate for Permutation***
Suppose f is an one way permutation, and we sample $x, r \leftarrow \{0,1\}^n$, and we construct:

$$f'(x,r) \coloneqq f(x) \| r$$
$$h(x,r) \coloneqq x \odot r = \sum_{i=1}^{n} x_i \cdot r_i \ (mod\ 2)$$

then f' is an one way permutation, while h is a hard-core predicate with respect to f'.

Proof. We first check f'. It is a permutation because $(f')^{-1}(a,b) = (f^{-1}(a), b)$. Then f' is OWF, it is straightforward that it is easy to compute. For the analysis of hardness to invert, it is similar to the proof of Theorem **??**, note that we substitute the sampling of a uniformly random bit $b \xleftarrow{\$} \{0,1\}$ to $b' \xleftarrow{\$} \{0,1\}^n$ and everything follows similarly. It remains to check that h is a hard-core predicate, we proceed by contradiction. Suppose h is not a hard core predicate, then by definition there exists an NUPPT A and a polynomial p such that for infinitely many $n \in \mathbb{N}$:

$$\Pr_{x,r}[A(f'(x,r)) = h(x,r)] \geq \frac{1}{2} + \frac{1}{p(n)}$$

We want to construct an adversary B that can invert f. Firstly we consider the perfect case, which is:

$$\Pr_{x,r}[A(f'(x,r)) = h(x,r)] = 1$$

In this case we observe that it means $\forall\ x, r$, we have $A(f(x)\|r) = x \odot r$. For example we have:

$$A(f(x)\|10\cdots0) = x_1 \cdots x_n \odot 10\cdots0 = x_1$$

thus in general, this applies to the case when for the strings $e_1, ..., e_n$, for $e_i = 0, ..., 1, ...0$, it means i^{th} bit is 1, while all other bits are zero. Then for B, we can just simply run $A(y\|e_i) = x_i$ for n times, and we can see that B will always successfully invert f, which is a contradiction. Then we relax the probability of a little bit, say:

$$\Pr_{x,r}[A(f'(x,r)) = h(x,r)] = \frac{3}{4} + \frac{1}{p(n)}$$

Then $B(y)$ will be :$\forall i \in \{1, ..., n\}$, we will get $z_i \leftarrow A(y\|e_i)$, and we repeat this procedure n times, we get $z_1 \cdots z_n$ and hopefully this will be x. However, the overall successful probability for this is low, thus this will not work as we wished. Now we replace:

$$z_i \leftarrow A(y\|r) \oplus A(y\|r \oplus e_i)$$

for uniformly random r, this is well-defined because one can show:

$$(x \odot r) \oplus (x \odot (r \oplus e_i)) = x \odot e_i = x_i$$

and we will succeed. Note that we will have the following probability:

$$\Pr_{x,r}[z_i = x_i] \geq \frac{1}{2} + 2\frac{1}{p(n)}$$

following from the constructions above and by union bound. However, we note that repeat this procedure for another n round is not satisfactory enough.
Now, we want to ask this question: can this procedure work for all pairs x and y? The answer is No, because this procedure depends on fixed y, we cannot randomly sample another y, making this probability highly dependent. We make a new claim: there are many "good" x: we define the set:

$$G \coloneqq \{x \in \{0,1\}^n : \Pr_r[A(f(x)\|r) = x \odot r] \geq \frac{3}{4} + \frac{1}{2p(n)}\}$$

and we claim that the cardinality of this set is $|G| \geq \frac{1}{2p(n)} \cdot 2^n$. We modify our procedure:
For $j = 1, ..., m$:

1. Randomly sample $r \leftarrow \{0,1\}^n$.
2. We get $z_{ij} \leftarrow A(y\|r) \oplus A(y\|r \oplus e_i)$. Note that:

   $$(x \odot r) \oplus (x \odot (r \oplus e_i)) = x \odot e_i = x_i$$

   and

   $$\Pr_{x,r}[z_{ij} = x_i] \geq \frac{1}{2} + \frac{1}{p(n)}$$

3. Note that z_{ij} is independent from $j \in [m]$, and we just take z_i to be the majority:

   $$z_i \leftarrow maj(z_{i1}, ..., z_{in})$$

We consider the event $\mathbb{X}_j \coloneqq$ the event $z_{ij} = x_i$, and the expectation is:

$$\mu = \mathbb{E}[\mathbb{X}_j = 1] = \frac{1}{2} + \frac{1}{p(n)}$$

and we have

$$\Pr[\sum_{j \in [m]} \leq \frac{m}{2}] = (1 - \delta) \cdot m \cdot \mu \leq e^{-\Omega(m \cdot \alpha^2)}$$

by Chernoff bound. Which indicates if we take $m = n \cdot \alpha^2$, everything will follow. Observe that in this reduction, we have two failing cases:

1. $x \notin G$.
2. x_i fails in the above procedure.

Thus our total failure probability will be:

$$\Pr[\text{failure}] \leq \Pr[x \notin G] + n \cdot e^{-\Omega(n)} = (1 - \frac{1}{2p(n)}) + negl$$

so our success probability should be

$$\Pr[B \text{ successfully invert}] \geq \frac{1}{2p(n)} - negl \geq \frac{1}{poly(n)}$$

which conclude this case.
Now, we face the next challenge: how can we push the probability from $\frac{3}{4}$ to our desired $\frac{1}{2} + \frac{1}{p(n)}$? This is not easy, if we take $\frac{3}{5}$ instead of $\frac{3}{4}$ in our procedure, we will have the probability

$$\Pr_{x,r}[z_{ij} = x_i] \geq \frac{1}{5} + \frac{1}{p(n)}$$

which indicates that the union bound fails the majority selection. Before pushing the probability, let's prove the claim regarding

$$G \coloneqq \{x \in \{0,1\}^n : \Pr_r[A(f(x)\|r) = x \odot r] \geq \frac{3}{4} + \frac{1}{2p(n)}\}$$

We proceed by contradiction, assume $|G| < \frac{1}{2p(n)} \cdot 2^n$, and we denote $A(f(x)\|r) = x \odot r$ by event T. Then we have

$$\begin{aligned}
\Pr_{x,r}[T] &= \Pr_{x,r}[x \in G] + \Pr_{x,r}[T \mid x \notin G] \cdot \Pr_x[x \notin G] \\
&\leq \frac{|G|}{2^n} + (\frac{3}{4} + \frac{1}{2p(n)}) \cdot 1 \\
&< \frac{1}{2p(n)} + \frac{3}{4} + \frac{1}{2p(n)} \\
&= \frac{3}{4} + \frac{1}{p(n)}
\end{aligned}$$

But we start by $\Pr_{x,r}[T] \geq \frac{3}{4} + \frac{1}{p(n)}$, which is a contradiction. Hence the claim is proved.
Now, we back to the main theorem, we hope the term $A(y\|r)$ in $z_{ij} \leftarrow A(y\|r) \oplus A(y\|r \oplus e_i)$ to be $x \odot r_j$, but we cannot guarantee this anymore. One can check that it is equivalent to just guessing random g_j, we hope $g_j = x \odot r_j$. From now on, we won't call $A(y\|r_i)$ anymore.
Now, we are ready to do the final and formal reduction $B(y)$:

1. We sample $u_1, ..., u_\ell \leftarrow \{0,1\}^n$ independently, and we get pairwise independent $r_1, ..., r_m$ by performing bitwise XOR on subsets of $(u_1, ..., u_\ell)$, with $\ell = \log(m)$. We also sample $b_1, ..., b_\ell \leftarrow \{0,1\}$ independently, and again we perform bitwise XOR to obtain pairwise independent $g_1, ..., g_m$. We hope this will hold:

   $$b_k = x \odot u_k \ \forall k \in [\ell]$$

   If this hold, we will have the following probability:

   $$\Pr[b_k = x \odot u_k \ \forall k \in [\ell]] = \frac{1}{2^\ell} = \frac{1}{m}$$

and then we will have $g_j = x \odot r_j$ for all $j \in [m]$, which indicates:

$$\bigoplus_{k \in \text{Set } j} b_k = x \odot \bigoplus_{k \in \text{Set } j} u_k$$

then the following probability hold:

$$\Pr[g_j = x \odot r_j \ \forall j \in [m]] = \frac{1}{m}$$

2. Then for each i, we compute the following:

$$z_{ij} \leftarrow A(y \| r_j \oplus e_i) \oplus g_j \ \forall j \in [m]$$

3. We pick the major: $z_i = maj(z_{i1}, ..., z_{im})$
4. Output $z_1, ..., z_n$.

By choosing $m = poly(n, \alpha)$, we claim that we can achieve the desired result. We skip the detailed probability analysis, we just provide the sketch here: Suppose $f^{-1}(y) \in G$, and suppose $b_k = x \odot u_k$ for all $k \in [\ell]$, then by the Chebyshev's inequality (Theorem 2 below), we will have

$$\Pr[z_i \neq x_i] \leq O(\frac{1}{m})$$

and by the union bound, we will have:

$$\Pr[\exists \ i \text{ s.t. } \ z_i \neq x_i] \leq O(\frac{n}{m})$$

From these calculations, we see that the event $f^{-1}(y) \in G$ fails with probability $1 - \frac{1}{2p(n)}$, while the event $b_k = x \odot u_k$ for all $k \in [\ell]$ fails with probability $1 - \frac{1}{m}$. Combining everything together, we conclude the proof, the result follows. □

Theorem 2. ***Chebyshev's Inequality***
For pairwise independent variables $X_1, X_2, ..., X_m \in [0, 1]$, we have

$$\Pr[|\sum_{j \in [m]} x_j - m \cdot \mu| \geq \delta m] \leq \frac{1 - \mu}{m - \delta^2}$$

where $\mu = \mathbb{E}[x_j]$ for all j.

Acknowledgement

References