Cryptography 101: Pseudorandom function

ComComX • 5 Mar 2026

Motivating examples

Identity authentication

Consider the scenario that Alice and Bob shared a short key before but now, instead of sending a secret message to Bob, Alice wants to prove her identity to Bob. What make this challenging is that Alice wants to it multiple times, and the communication channel is now controlled by an active adversary Mallory, who can read whatever Alice and Bob send to the channel.

The simple way is to use a password. E.g., Alice can use the shared key as a password and send it to Bob whenever she needs to prove her identity. But this is a very bad idea since Mallory can learn the password from the channel in the first interaction of Alice and Bob, and use it to impersonate Alice in future interactions. This is known as the *replay attack*. Can encryption with a PRG help? It can not, since PRG is a deterministic function and Mallory does not need to learn what is the password but only its valid encryption. We will see how pseudorandom function generator can help in this scenario.

Chosen plaintext attack

In this kind of attack, the adversary can choose any plaintext and view the corresponding ciphertext to gain information about the encryption scheme. Consider the security game between a PPT-adversary (Adv) and a PPT-challenger (Cha) on an encryption scheme $(\mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec})$ as follows.

IND-CPA(λ):

1. Cha chooses a random bit $b \leftarrow \{0,1\}$ and key $k \leftarrow \mathsf{KeyGen}(1^\lambda)$
2. For $i = 1, \ldots, t$ (with $t = \text{poly}(\lambda)$):
   - Adv chooses $\hat{m}_i \in \mathcal{M}$ and sends it to Cha
   - Cha sends back $\mathsf{Enc}(1^\lambda, k, \hat{m}_i)$
3. Adv chooses m_0, m_1 and sends to Cha

4. Cha sends back $\mathsf{Enc}(1^\lambda, k, m_b)$

5. Adv guesses a bit b'

6. Adv wins if $b' = b$.

Definition 1 (CPA-security)**.** *An encryption scheme is said to be IND-CPA-secure if for any PPT-adversary,*

$$\Pr[\textit{Adv wins IND-CPA}(\lambda)] = \frac{1}{2} + \text{negl}(\lambda).$$

Since the key is reused multiple time to encrypt $\hat{m}_i$, a PRG is insufficient for CPA-security.

Pseudorandom function

PRF is a mathematical object that behaves like an *indexable* (or *locally computable*) PRG.

Let F be a function family defined by $F = \{F_\lambda\}_{\lambda \in \mathbb{N}}$ where

$$F_\lambda = \{f_k : \{0,1\}^{n(\lambda)} \to \{0,1\}^{m(\lambda)} | k \in \mathcal{K}\}$$

where n and m are polynomials, the sampling of k and the computation of f_k are all efficient.

Let F_λ^{All} be the set of all functions $f : \{0,1\}^{n(\lambda)} \to \{0,1\}^{m(\lambda)}$. This means that sampling a random function from F_λ^{All} will give a truly random function.

Definition 2 (Informal)**.** *A PRF is secure if given the output from either a truly random function or the PRF, no PPT Adv can distingiush whether the output is produced by a truly random function or the by PRF.*

Consider the following security game:

<u>**FN-IND**</u>(λ):

1. Cha chooses a random bit $b \leftarrow \{0,1\}$

2. If $b = 1$: sample a random f from F_λ^{All} (truly random)

3. If $b = 0$: sample a random f from F_λ (pseudorandom)

4. For $i = 1, \ldots, t$ (with $t = \text{poly}(\lambda)$):

    ◦ Adv sends $x_i \in \{0,1\}^{n(\lambda)}$ to Cha

    ◦ Cha sends $f(x_i)$ to Adv

5. Adv guesses a bit b'

6. Adv wins if $b' = b$.

Definition 3 (PRF family)**.** *A function family F is a PRF if for any PPT Adv,*

$$\Pr[\textit{Adv wins FN-IND}(\lambda)] = \frac{1}{2} + \text{negl}(\lambda).$$

Alternative definition (Boaz Barak):

Definition 4 (PRF generator)**.** *Let $F : \{0,1\}^\lambda \times \{0,1\}^{n(\lambda)} \to \{0,1\}^{m(\lambda)}$ be an efficiently computable function that takes two inputs k and i and outputs $F(k,i)$. Then, F is a pseudorandom function generator if for any PPT adversary A outputting a single bit and any $i \in \{0,1\}^{n(\lambda)}$,*

$$\Pr_{k\leftarrow U_\lambda}[A(F(k,i),1^\lambda)=1] - \Pr_{s\leftarrow F^{\text{rand}}}[A(s,1^\lambda)=1] = \text{negl}(\lambda).$$

Similar to previous definition, we can think of F as a family or an ensemble of functions:

$$F = \{f_k : \{0,1\}^{n(\lambda)} \to \{0,1\}^{m(\lambda)} | k \in \{0,1\}^\lambda\}.$$

Intuitively, the adversary is given access to a function $f : \{0,1\}^{n(\lambda)} \to \{0,1\}^{m(\lambda)}$ and is asked to specify whether the function is sampled from F^{All} (a truly random function) or is sampled from F. The adversary can run as many queries as he wants. The function family F is said to be a PRF if the adversary can not do significantly better than random guessing.

Encryption using PRFs

An encryption is scheme with PRFs is defined as follows:

- $\mathsf{KeyGen}(1^\lambda)$: Sample a function $f_k \leftarrow F_\lambda$ (shared among parties)
- $\mathsf{Enc}(1^\lambda, k, m)$:
  1. Pick $i \leftarrow \{0,1\}^\lambda$
  2. $c = m \oplus f_k(i)$
  3. Output (i, c)
- $\mathsf{Dec}(1^\lambda, k, (i, m))$:
  1. $\hat{m} = c \oplus f_k(i)$
  2. Output $\hat{m}$.

Theorem 5. *The above encryption scheme is CPA-secure.*

Proof. Assume for contradiction that the encryption scheme is not CPA-secure. Specifically, there is a PPT adversary A that wins IND-CPA(λ) with a probability noticeably more than half:

$$\Pr[A \text{ wins IND-CPA}(\lambda)] \geq \frac{1}{2} + \frac{1}{p(\lambda)}$$

for some polynomial $p(\lambda)$. Under this assumption, we show that there is a PPT distinguisher D that distinguishes a PRF and a truly random function with non-negligible advantage.

Specifically, D is given access to a function $\hat{f} : \{0,1\}^\lambda \to \{0,1\}^\lambda$ and is to conclude $\hat{f}$ is a truly random function sampled from F^{All} or a pseudorandom function sampled from F_λ. The adversary D will use $\hat{f}$ while internally simulating A and the game IND-CPA as follows:

- Whenever A queries its challenger for the encryption of $\hat{m}_i$ (in the IND-CPA game):
  1. D samples $x_i \leftarrow U_\lambda$
  2. D outputs to A the string $(x_i, \hat{m}_i \oplus \hat{f}(x_i))$
- In the last encryption step when A gives two messages m_0 and m_1:
  1. D samples $b \leftarrow U_1$ and $r \leftarrow U_\lambda$
  2. D outputs to A the string $(r, m_b \oplus \hat{f}(r))$
- D outputs 1 if A guesses the bit b correctly. Otherwise, D outputs 0.

With the above D, we now have:

- If $\hat{f}$ is a pseudorandom function, then the probability that D gives a correct conclusion (outputting 1) equals the probability that A guesses the bit b correctly (winning the IND-CPA game), which is $\frac{1}{2} + \frac{1}{p(\lambda)}$ due to the assumption of A.
- On the other hand, if $\hat{f}$ is a truly random function, then the probability that D gives a correct conclusion (outputting 0) is negligibly more than half since this is the probability that A winning the game IND-CPA. This is because there are 2^λ possible values of r and A only observes poly(λ) values of $\hat{f}(r)$. Therefore, it is extremely unlikely (with probability of only $\text{poly}(\lambda)/2^\lambda$) that in the last interaction, the string r has already been used before. The winning probability of A is thus at most $\frac{1}{2} + \text{poly}(\lambda)/2^\lambda$, where the first term is the winning probability of random guessing and the second term is the probability that an r is reused.

This implies that

$$\begin{aligned}
&\Pr_{\hat{f}\leftarrow F_\lambda}[D(1^\lambda,\hat{f}(U_\lambda))=1]-\Pr_{\hat{f}\leftarrow F^{\text{All}}}[D(1^\lambda,\hat{f}(U_\lambda))=1]\\
&\quad=\Pr_{\hat{f}\leftarrow F_\lambda}[D(1^\lambda,\hat{f}(U_\lambda))=1]-(1-\Pr_{\hat{f}\leftarrow F^{\text{All}}}[D(1^\lambda,\hat{f}(U_\lambda))=0])\\
&\quad\geq\frac{1}{2}+\frac{1}{p(\lambda)}-1+\frac{1}{2}-\text{poly}(\lambda)/2^\lambda)\\
&\quad=\frac{1}{p(\lambda)}-\text{negl}(\lambda),
\end{aligned}$$

which is non-negligible. Here in the inequality, we have used the fact that the loosing probability of A is at least $1-(\frac{1}{2}+\text{poly}(\lambda)/2^\lambda)$. This contradicts the fact that $\{F_\lambda\}$ is a PRF. $\square$

Solution to Identity authentication: One time passwords

Suppose Alice and Bob have already shared a key k. Then the authentication protocol is as follows:

1. Alice requests Bob for authentication
2. Bob picks a random $i\leftarrow U_\lambda$ and sends to Alice
3. Alice sends $f_k(i)$ to Bob
4. Bob checks and accepts the session if the message sent by Alice is exactly $f_k(i)$, otherwise rejects.

Note: it is *not* crucial that i is random. What is crucial is that each i is used at most once to avoid the replay attack. Often in practice, i is computed as a function of time. This is what known as the one-time password.

Theorem 6 (Security of one-time password protocol)**.** *Suppose $\{f_k\}_{k\in\{0,1\}^\lambda}$ is a PRF generator. After observing $T=\text{poly}(\lambda)$ interactions by Alice and Bob, the probability that Mallory, with arbitrary efficient computation, succeeds in impersonating Alice is negligible (at most $2^{-\lambda}+\text{negl}(\lambda)$).*

Proof. Assume for contradiction that Mallory succeeds with non-negligible probability. We construct a PPT distinguisher D that distinguishes a PRF and a truly random function with non-negligible advantage.

Specifically, D is given an oracle $\hat{f}:\{0,1\}^\lambda\rightarrow\{0,1\}$ and is to conclude $\hat{f}$ is a truly random function or a pseudorandom function. Let D be as follows: D simulates the interactions of Alice, Bob and Mallory, and whenever f_k is called, D replaces f_k by the given oracle $\hat{f}$. Then, D outputs 1 (i.e., $\hat{f}$ is a PRF) if Bob accepts Mallory, and outputs 0 (i.e., $\hat{f}$ is a truly random function) otherwise.

With the above D, we now have:

- If $\hat{f}$ is a pseudorandom function, then the probability that D gives a correct conclusion is *non-negligible* due to assumption regarding the success probability of Mallory.
- On the other hand, if $\hat{f}$ is a truly random function, then the probability that D gives a correct conclusion is *negligible* since the success probability of Mallory is negligible. This is because there are 2^{λ} possible values of i and Mallory only observes poly(λ) interactions. Therefore, it is extremely unlikely that in the interaction with Mallory, Bob sends an i that has already been used before. Therefore, the success probability of Mallory is thus at most $\mathrm{poly}(\lambda)/2^{\lambda} + 2^{-\lambda}$, where the first term is the probability that an i is reused and the second term is the probability that Mallory guesses $\hat{f}(i)$ correctly.

This implies that the advantage of D in distinguishing a truly random function and a pseudorandom function is non-negative. □

Existence of PRFs

Theorem 7 (Existence of PRFs)**.** *Suppose that the PRG conjecture is true, i.e., there is a PRG that maps λ bits to $\lambda + 1$ bits. Then, there is a PRF generator $F = \{F_{\lambda}\}_{\lambda \in \mathbb{N}}$ where*

$$F_{\lambda} = \{ f_k : \{0,1\}^{\lambda} \to \{0,1\}^{\lambda} | k \in \{0,1\}^{\lambda} \}.$$

Proof idea. The idea is to use a length-doubling PRG G to construct F, and then use the hybrid argument to show that if there is a PPT adversary A that breaks the security of F then we can construct a PPT distinguisher D that breaks the security of G. □