HW 1: Symmetric Cryptography – Due date: February 21, 2023 before tutorial (corrected version)

Exercise 1. *PRF implies PRG*

Let $F : \{0,1\}^s \times \{0,1\}^n \to \{0,1\}^m$ be a secure Pseudo-Random Function (PRF). We define the following PRGs $G_d : \{0,1\}^s \to \{0,1\}^{md}$, for $d \leq \text{poly}(m)$ such that:

$$\forall k \in \{0,1\}^s, G_d(k) = F(k,\bar{0})||F(k,\bar{1})||\dots||F(k,\overline{d-1}),$$

where $||$ denotes the concatenation operator and $\bar{i}$ denotes the binary decomposition of i, written over m bits.

1. Prove that G_d is a secure PRG.

   ☞ Assuming that there exists some d and adversary $\mathcal{A}$, with non-negligible advantage ε of distinguishing between $G(U(\{0,1\}^s))$ and $U(\{0,1\}^{md})$, we build $\mathcal{A}'$ an adversary against the security game of F that does the following in Exp_b, $b \in \{0,1\}$:

$\mathcal{C}$	$\mathcal{A}'$	$\mathcal{A}$
Set $F' = bU(\{0,1\}^n \to \{0,2\}^m) + (1-b)F(U(\{0,1\}^s,\cdot))$	Ask $\mathcal{C}$ for $F'(\bar{i}), \forall i < d$	
Return $F'(\bar{i}), \forall i < d$	Compute and send to $\mathcal{A}$ $y = F'(0)\|\|F'(1)\|\|\dots\|\|F'(d-1)$	Return a bit b'
	Output the returned bit b'.	

   In the case of Exp_0, we have that $y \hookleftarrow G(U(\{0,1\}^s))$ but in Exp_1, $y \hookleftarrow U(\{0,1\}^{md})$. Then $\mathcal{A}$ distinguishes between the two experiments with advantage ε, and that is also the case for $\mathcal{A}'$.

Exercise 2. *PRG implies PRF*

Let $G : \{0,1\}^s \to \{0,1\}^{2s}$ be a secure length-doubling PRG. We have aldready how to get such a PRG from any PRG in the previous tutorials. The Goldreich-Goldwasser-Micali construction shows how to build a secure Pseudo-Random Function for any input size from G.

1. Let us denote $G(k) =: G_0(k)||G_1(k)$ for any $k \in \{0,1\}^s$ where $G_0, G_1 : \{0,1\}^s \to \{0,1\}^s$. Define $F_0 : \{0,1\}^s \times \{0,1\} \to \{0,1\}^s$ such that:

   $$\forall k \in \{0,1\}^s, \forall b \in \{0,1\}, F_0(k,b) := G_b(k).$$

   Prove that F_0 is a secure PRF.

   ☞ Let $\mathcal{A}$ be an attacker against the security of the PRF. We build $\mathcal{B}$ against the security of the PRG that does the following: on input $y = G(k)$ or $y \hookleftarrow U(\{0,1\}^{2s})$, the adversary calls $\mathcal{A}$ and answers its queries using y.

   - If y is uniform then the answers to the queries are too and $\mathcal{A}$'s view is that of a uniform function.
   - If $y = G(k)$, then $\mathcal{A}$'s view is that of F_0.

   Then $\mathcal{A}$ behaves as if it was in the security game of the PRF. Setting $\mathcal{B}$ to output the same bit as $\mathcal{A}$, they both have the same advatange. As such, the advantage of $\mathcal{A}$ cannot be non-negligible.

We now expand our construction to arbitrary input size n. Define the iterated PRF $F_n : \{0,1\}^s \times \{0,1\}^n \to \{0,1\}^s$ that does the following: on inputs k and $x = x_0x_1\dots x_{n-1}$, define $k_0 := k$ and compute recursively $k_i := G_{x_{i-1}}(k_{i-1})$ for $i = 1$ to n. Finally output k_n.
Remark: This can be seen as going down a binary tree.

2. Before proving the security of F_n, we prove that the distribution $(G(k_1), G(k_2), \dots, G(k_Q))$, where $k_i \hookleftarrow U(\{0,1\}^s)$ is indistinguishable from $U(\{0,1\}^{2Qs})$ for any $Q = \text{poly}(s)$, under the security of G.

   We use the hybrid argument by defining the following hybrid distributions:

   $$\forall i \in [0,Q], D_i := (G(k_1), \dots, G(k_i), U(\{0,1\}^{2s(Q-i)}) \text{ where } k_j \hookleftarrow U(\{0,1\}^s) \forall j \leq i.$$

Notice that D_0 and D_Q correspond to the distributions defined previously.

Prove that D_0 and D_Q are indistinguishable under the security of G. Estimate the security loss.

☞ Assume that D_0 and D_Q are distinguishable with non-negligible advantage. This implies that there exists $0 < i \leq Q$ such that there exists some adversary $\mathcal{A}$ which distinguishes with non-negligible probability between D_i and D_{i-1}.

We build a distinguisher $\mathcal{B}$ against the security of the PRG that does the following:

$\mathcal{C}$	$\mathcal{B}$	$\mathcal{A}$
Sample $b \hookleftarrow U(\{0,1\})$ Send to $\mathcal{B}$ $y \hookleftarrow bG(U(\{0,1\}^s)) + (1-b)U(\{0,1\}^{2s})$	Sample $k_j \hookleftarrow U(\{0,1\}^s), \forall j \leq i-1$ and $y_j \hookleftarrow U(\{0,1\}^{2s}), \forall j > i$ Send to $\mathcal{A}$: $Y := (G(k_1), \dots, G(k_{i-1}), y, y_{i+1}, \dots, y_Q)$ Return b'	Return a bit b'

We see that Y follows the distribution $bD_{i-1} + (1-b)D_i$: it is distributed following D_i if y is sampled according to G and it is distributed following D_{i-1} if y is uniform. As such,the guess b' of $\mathcal{A}$ is right with non-negligible probability and $\mathcal{B}$ breaks the security of G.

This proves that D_0 and D_Q are indistinguishable.

Moreover if an adversary has advantage at most ε in the distinguishing game of G then any adversary distinguishes between D_0 and D_Q with advantage at most $Q\varepsilon$.

We move on to the proof that F_n is secure.

3. To do so, we use the hybrid argument by introducing the following hybrid experiments. Let us first define

$$F_{n,i}^{(R_i)} : (x_0, \dots, x_{n-1}) \mapsto G_{x_{n-1}}(\dots(G_{x_i}(R_i(x_0, \dots x_{i-1})))),$$

where $R_i : \{0,1\}^i \to \{0,1\}^s$ is a map.

(a) Prove that $F_{n,0}^{(U(\{\varepsilon\} \to \{0,1\}^s))}(\cdot)$ is actually the distribution $F_n(U(\{0,1\}^s), \cdot)$.

☞ Since the map is constant, there are $\{0,1\}^s$ possible maps, which corresponds to choosing the constant. Moreover, the map is called after n iterations of G, so the constant we uniformly choose is actually the key of F_n.

(b) Prove that $F_{n,n}^{(U(\{0,1\}^n \to \{0,1\}^s))}$ is actually the distribution $U(\{0,1\}^n \to \{0,1\}^s)$.

☞ Notice that for any $c \in \{0,1\}^n$, we have $F_{n,n}^{(R_n)}(x) = R_n(x)$. Thus, if R_n is sampled uniformly at random, the map $F_{n,n}^{(R_n)}(\cdot)$ is also uniformly sampled over $\{0,1\}^n \to \{0,1\}^s$.

(c) We define the hybrid experiment Exp_i for $i \in [1,n]$ as: the challenger flips a coin b and samples R uniformly over $\{0,1\}^{i-b} \to \{0,1\}^n$. The adversary is then given access to an oracle, which on query $x \in \{0,1\}^n$ answers with $F_{n,i-b}^{(R)}(x)$. Eventually, the adversary outputs a guess b' and wins if and only if $b = b'$.

Prove that the PRF F_n is secure under the security of the PRG G and estimate the advantage loss.

☞ Assuming that the PRF is not secure, there exists some $i \geq 1$ such that there exists an adversary $\mathcal{A}$ with non-negligible advantage in Exp_i. We denote by Q the maximal number of queries that $\mathcal{A}$ does: it is at most $Q = \mathrm{poly}(s)$ as this adversary is ppt.

We now build $\mathcal{B}$ an adversary that distinguishes between $(G(k_1), G(k_2), \dots, G(k_Q))$ for uniform keys $\{k_i\}_{i=1}^Q \hookleftarrow U(\{0,1\}^{sQ})$ and $U(\{0,1\}^{2sQ})$.

On input $y_1, \dots y_Q$ following either of these two distributions, $\mathcal{B}$ calls $\mathcal{A}$, creates an empty list L and sets $\ell := 1$. When $\mathcal{A}$ queries $x \in \{0,1\}^n$, the adversary $\mathcal{B}$ does the following:

- If there does not exist some m such that $(m, x_0 \dots x_{i-2}) \in L$, it adds $(\ell, x_0 \dots x_{i-2})$ to L and increments ℓ.
- It computes and return $y := G_{x_{n-1}}(\dots(G_{x_i}(y_m[x_{i-1}]))\dots)$, where $y_m[x_{i-1}]$ denotes the s first (resp. last) bit of y_m if $x_{i-1} = 0$ (resp. $x_{i-1} = 1$).

When $\mathcal{A}$ eventually outputs a bit b', adversary $\mathcal{B}$ outputs the same.

In the case where $y_j = G(k_j)$, then for $x_0 \dots x_{i-2}$ such that $(j, x_0 \dots x_{i-2}) \in L$, let $R^{(i-2)}(x_0 \dots x_{i-2}) := k_j$, which is uniformly sampled. In this case, algorithm $\mathcal{A}$'s view is identical to that of case $b = 1$ in Exp_i.

In the case where $y_j \hookleftarrow U(\{0,1\}^{2s})$, let $x_0 \dots x_{i-2}$ such that $(j, x_0 \dots x_{i-2}) \in L$ and define $R^{(i)}(x_0 \dots x_{i-1}) := y_m[x_{i-1}]$ for any $x_{i-1} \in \{0,1\}$. Then $R^{(i)}$ is also uniformly sampled (or at least, what $\mathcal{A}$ sees of $R^{(i)}$ is uniformly sampled): in this case, algorithm $\mathcal{A}$'s view is that of the case $b = 0$ in Exp_i.

As such, it holds that the advantage of $\mathcal{B}$ is equal to the advantage of $\mathcal{A}$, which is non-negligible.

This proves that F_n is secure under the security of the PRG G. The advantage loss is the same as in the previous reduction.

Exercise 3. *Encrypting with a PRF*

Let F be a PRF function from $\{0,1\}^s \times \{0,1\}^n \to \{0,1\}^m$, we define the following encryption scheme: To encrypt a message $M \in \{0,1\}^m$ with a key $k \in \{0,1\}^s$, choose r uniformly in $\{0,1\}^n$ and return $c = (r \| F(k,r) \oplus M)$.

Show that this scheme is secure. More precisely, show if that there exists a PPT adversary $\mathcal{A}$ against the encryption scheme, then there exists a PPT adversary $\mathcal{B}$ against the PRF function F such that:

$$\mathrm{Adv}^{CPA}_{\mathcal{A}}(Enc) \leq 2\mathrm{Adv}^{PRF}_{\mathcal{B}}(F) + Q^2/2^n,$$

where Q is the number of encryptions queried by $\mathcal{A}$.

☞

Consider an ideal scheme where $F(k,.)$ is replaced by a truly random function: $(r, f(r) \oplus M)$. What is the maximum advantage in that case?

Exp'_b

$\mathcal{C}$		$\mathcal{A}$
	queries $\longleftarrow$ m_{i_0}, m_{i_1}	
choose r uniformly in $\{0,1\}^n$	$\longrightarrow$ $r_i, f(r_i) \oplus m_{i_b}$	
		$b' = b$?

Let repeat be the event where there are two identical r_i in Q experiments.

$$\begin{aligned}
\mathrm{Adv}_{\mathcal{A}} &= |\Pr[\mathcal{A}^{exp'_0} \to 1] - \Pr[\mathcal{A}^{exp'_1} \to 1]| \\
&\leq |\Pr[\mathcal{A}^{exp'_0} \to 1 | \mathtt{norepeat}] \Pr(\mathtt{norepeat}) - \Pr[\mathcal{A}^{exp'_1} \to 1 | \mathtt{norepeat}] \Pr(\mathtt{norepeat})| \\
&+ |\Pr[\mathcal{A}^{exp'_0} \to 1 | \mathtt{repeat}] \Pr(\mathtt{repeat}) - \Pr[\mathcal{A}^{exp'_1} \to 1 | \mathtt{repeat}] \Pr(\mathtt{repeat})| \\
&\leq \Pr(\mathtt{norepeat}) * 0 + \Pr(\mathtt{repeat}) \\
&\leq Q^2/2^n.
\end{aligned}$$

Indeed, if there is no repeat then the advantage of the adversary is 0 because f is a truly random function.

$$\begin{aligned}
\Pr(repeat) &= \Pr(\text{exists } i < j \leq Q : r_i = r_j) \\
&\leqslant \sum_{i<j\leq Q} \sum_r \Pr(r_i = r \text{ and } r_j = r) \\
&\leqslant \frac{Q(Q-1)}{2} \sum_r \Pr(r_i = r) \Pr(r_j = r) \\
&\leqslant \frac{Q(Q-1)}{2} 2^n 2^{-n} 2^{-n} \\
&\leqslant \frac{Q^2}{2^n}.
\end{aligned}$$

We now define the experiment Exp_b associated to the security of the scheme.

Exp_b

$\mathcal{C}$		$\mathcal{A}$
choose $k \in U(\{0,1\}^t)$		
	$\longleftarrow$ m_{i_0}, m_{i_1}	
choose r uniformly in $\{0,1\}^n$	$\longrightarrow$ $r_i, F(k,r_i) \oplus m_{i_b}$	
		$b' = b$?

Finally we do the following experiment: $\mathcal{C}$ outputs either $f(r_i)$, either $F(k,r_i)$.

$\mathcal{C}$		$\mathcal{B}$		$\mathcal{A}$
choose $k \in U(\{0,1\}^t)$				
			$\longleftarrow$ m_{i_0}, m_{i_1}	
		choose (r,b) uniformly in $\{0,1\}^n \times \{0,1\}$		
	$\longleftarrow$ r_i			
	$\longrightarrow$ y_i			
			$\longrightarrow$ $r_i, y_i \oplus m_{i_b}$	
			$\longleftarrow$ b'	
		$b' = b$?		

Distinguisher $\mathcal{B}$ for $F(k,.)$: uses $F(k,.)$ or $f(.)$ (uniform) for encrypting. For $b \in \{0,1\}$, it chooses Exp_b with probability $\frac{1}{2}$. To answer the queries of $\mathcal{A}$, $\mathcal{B}$ queries $F(k,.)$ or $f(.)$. Choice of experiment:

- Case 1: we are in Exp_0 / Exp_1, with probability 1/2 each
- Case 2: we are in Exp'_0 / Exp'_1, with probability 1/2 each.

If $\mathcal{A}$ guesses correctly, i.e. $b = b'$, we return 1 (which corresponds to "Exp"); else, we return 0.

$$\begin{aligned}
\mathrm{Adv}_{\mathcal{B}}^{PRF}(F) &= |\Pr[\mathcal{B}(F)=1] - \Pr[\mathcal{B}(f)=1]| \\
&= \left|\frac{1}{2}(\Pr[\mathcal{A}^{Exp_0} \to 0] + \Pr[\mathcal{A}^{Exp_1} \to 1] - \frac{1}{2}(\Pr[\mathcal{A}^{Exp'_0} \to 0] + \Pr[\mathcal{A}^{Exp'_1} \to 1])\right| \\
&= \left|\frac{1}{2}(1 - \Pr[\mathcal{A}^{Exp_0} \to 1] + \Pr[\mathcal{A}^{Exp_1} \to 1]) - \frac{1}{2}(1 - \Pr[\mathcal{A}^{Exp'_0} \to 1] + \Pr[\mathcal{A}^{Exp'_1} \to 1])\right| \\
&\geq \frac{1}{2}\left(\mathrm{Adv}_{\mathcal{A}}^{CPA}(Enc) - Q^2/2^n\right)
\end{aligned}$$

Exercise 4. *IND-CCA secure symmetric encryption*

Consider the following construction of symmetric encryption, where $\Pi = (\mathsf{Gen}, \mathsf{Mac}, \mathsf{Verify})$ is a MAC.

Gen(1^λ)**:** Choose a random key $K_1 \leftarrow \mathsf{Gen}'(1^\lambda)$ for an IND-CPA secure symmetric encryption scheme $(\mathsf{Gen}', \mathsf{Enc}', \mathsf{Dec}')$. Choose a random key $K_0 \leftarrow \Pi.\mathsf{Gen}(1^\lambda)$ for the MAC Π. The secret key is $K = (K_0, K_1)$.

Enc(K, M)**:** To encrypt M, do the following.

1. Compute $c = \mathsf{Enc}'(K_1, M)$.
2. Compute $t = \Pi.\mathsf{Mac}(K_0, c)$.

Return $C = (t, c)$.

Dec(K, C)**:** Return $\perp$ if $\Pi.\mathsf{Verify}(K_0, c, t) = 0$. Otherwise, return $M = \mathsf{Dec}'(K_1, c)$.

1. Assume that the MAC is weakly unforgeable. Assume however that there exists an algorithm $\mathcal{F}$, which on input a valid message for the MAC and a tag (M, t), outputs a forgery (M, t') such that $t \neq t'$. In particular, the MAC is not strongly unforgeable. Show that the scheme is not IND-CCA secure.

☞ Any CCA adversary can call the forger for $(c^\star, t^\star)$, to obtain $(c^\star, t') \neq (c^\star, t^\star)$, which it can hand to its decryption oracle.

2. We assume that: (i) $(\mathsf{Gen}', \mathsf{Enc}', \mathsf{Dec}')$ is IND-CPA-secure; (ii) Π is strongly unforgeable under chosen-message attacks. We will prove in this question the IND-CCA security of the new encryption scheme under these assumptions. Let $\mathcal{A}$ be an adversary against the IND-CCA security of the scheme.

(a) Define the event Valid as the event where $\mathcal{A}$ makes a valid (i.e. accepted by the MAC) decryption query for (c, t) where the ciphertext c was not encrypted by the encryption oracle nor is (c, t) the challenge ciphertext. Prove that if $\Pr(\mathsf{Valid})$ is non-negligible then there exists an adversary with non-negligible advantage against the strong unforgeability of the MAC.

☞ Katz-Lindell (p 145).

It appears that $\Pr(\mathsf{Valid})$ is negligible as this would otherwise lead to an attack against the strong unforgeability of the MAC: let $\mathcal{A}_0$ against the strong unforgeability of the MAC. It samples a uniform key K_1 for the encryption scheme and calls $\mathcal{A}$. With the key and its oracle access to $\Pi.\mathsf{Mac}$, it can answer any encryption query of $\mathcal{A}$. The challenge ciphertext is computed as any encryption query, except that the adversary first samples a uniform bit b and encrypts the message b. Then for any decryption query (c, t), either c is an encryption from a query (and the tag is valid) and it answers the corresponding message, or this is not the case but the tag is still valid, and it stops the experiment and outputs (c, t) as a forgery. Otherwise it answers $\perp$ to $\mathcal{A}$.

We see that $\mathcal{A}_0$ outputs a forgery if and only if the event Valid is realized. As such, it must have negligible probability.

The intuition is that since this event has negligible probability, the decryption oracle is useless to an attacker $\mathcal{A}$.

(b) Prove that if $|\Pr(\mathcal{A} \text{ wins} \wedge \overline{\mathsf{Valid}}) - 1/2|$ is non-negligible, then there exists an efficient adversary against the IND-CPA security of the encryption scheme $(\mathsf{Gen}, \mathsf{Enc}', \mathsf{Dec}')$.

☞ Let us build the following adversary $\mathcal{A}_1$ against the CPA security of the encryption scheme. It starts by sampling a MAC key K_0 and calls $\mathcal{A}$. It can answer all of its encryption queries thanks to the MAC key and its encryption oracle. For the challenge, on query (m_0, m_1) from $\mathcal{A}$, it also chooses (m_0, m_1) as challenge, and uses the MAC to complete the ciphertext for $\mathcal{A}$. For any decryption query (c, t), if t is a valid tag and c was encrypted by the encryption oracle of $\mathcal{A}_1$, it outputs the corresponding message. In any other case, it returns $\perp$. When $\mathcal{A}$ outputs a bit, it outputs the same bit.

We see that $\mathcal{A}_1$ wins if $\mathcal{A}$ wins and Valid does not occur. As such, $\Pr(\mathcal{A} \text{ wins} \wedge \overline{\mathsf{Valid}})$ is negligible under the CPA-security of the encryption scheme.

(c) Conclude.

☞ The advantage of $\mathcal{A}$ is $\leq \Pr(\mathsf{Valid}) + |\Pr(\mathcal{A} \text{ wins} \wedge \overline{\mathsf{Valid}}) - 1/2|$.

This concludes the proof.