Applied Cryptography

Symmetric Cryptography, Assignment 1, Wednesday, February 16, 2022

Exercises with answers and grading.

1. **(10 points)** This question is about the non-tightness of the equation of lecture 2 slide 12. In other words, it is about the existence of a MAC function that is unforgeable but not PRF-secure. Suppose we are given a pseudorandom function $F : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$. Consider MAC function

$$\mathsf{MAC}_K(M) = F_K(M) \parallel F_K(M)\,.$$

(a) Prove that MAC is unforgeable up to bound $q_v/2^n$, i.e., that

$$\mathbf{Adv}^{\mathrm{unf}}_{\mathsf{MAC}}(q_m, q_v) \leq \frac{q_v}{2^n} + \mathbf{Adv}^{\mathrm{prf}}_{F}(q_m + q_v)\,.$$

You do *not* have to *explicitly* write a reduction from the unforgeability of MAC to the PRF-security of F.

(b) For PRF-security, we consider the setup of a distinguisher that has access to either $\mathsf{MAC}_K : M \mapsto T$ or to a random oracle $\mathsf{RO} : M \mapsto T$. Consider the following distinguisher $\mathcal{D}$:

- Fix an arbitrary M and query the oracle on M to receive a tag T;
- If the left and right half of T are equal, return 1. If the left and right half of T are unequal, return 0.

Determine the exact PRF-advantage of this particular distinguisher $\mathcal{D}$, $\mathbf{Adv}^{\mathrm{prf}}_{\mathsf{MAC}}(\mathcal{D})$.

Begin Secret Info:...

(a) A first step replaces F_K by a random function f. As F_K is evaluated for q_m+q_v different inputs, this step comes at a cost of $\mathbf{Adv}^{\mathrm{prf}}_{F}(q_m + q_v)$. (This step is very comparable to the first step of the security of CTR-mode.)

Now, consider the MAC function $f(M) \parallel f(M)$. Any forgery attempt has a tag of the form $T = X\|Y$. If $X \neq Y$, the forgery succeeds with probability 0. If $X = Y$, the forgery succeeds with probability $1/2^n$. As the adversary can make q_v forgery attempts, its success probability is at most $q_v/2^n$.

(b) The PRF-advantage of $\mathcal{D}$ is defined as

$$\mathbf{Adv}^{\mathrm{prf}}_{\mathsf{MAC}}(\mathcal{D}) = \left|\mathbf{Pr}\left(\mathcal{D}^{\mathsf{MAC}_K} = 1\right) - \mathbf{Pr}\left(\mathcal{D}^{\mathsf{RO}} = 1\right)\right|\,. \tag{1}$$

We define $\mathcal{D}$ to return 1 iff the left and right half of T are equal. Thus, if $\mathcal{D}$ is conversing with the real world MAC_K, it always outputs 1:

$$\mathbf{Pr}\left(\mathcal{D}^{\mathsf{MAC}_K} = 1\right) = 1\,.$$

On the other hand, if it is conversing with the ideal world RO, it outputs 1 with probability $1/2^n$:

$$\mathbf{Pr}\left(\mathcal{D}^{\mathsf{RO}} = 1\right) = 1/2^n\,.$$

We conclude for (1):

$$\mathbf{Adv}^{\mathrm{prf}}_{\mathsf{MAC}}(\mathcal{D}) = 1 - 1/2^n\,.$$

End Secret Info ....

2. **(5 points)** Consider the function $H : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ defined as $H_L(M) = L \otimes M$, i.e., defined as finite-field multiplication over $\mathrm{GF}(2^n)$.

   (a) Prove that this function is 2^{-n}-XOR-universal.

   (b) If plugged into the Wegman-Carter MAC function of lecture 2 slide 14, we obtain

   $$\mathbf{Adv}^{\mathrm{unf}}_{\mathsf{WC}}(q_m, q_v) \le q_v/2^n + \mathbf{Adv}^{\mathrm{prf}}_{F}(q_m + q_v)\,,$$

   provided that the adversary does not query WC_K for repeated nonces. Assume you can evaluate this function for repeated nonces. Mount a forgery attack in $q_m = 3$ MAC queries and $q_v = 1$ VFY query.

Begin Secret Info: ....

(a) For any $M \ne M'$ and T:

$$\mathbf{Pr}_L\left(H_L(M) \oplus H_L(M') = T\right) = \mathbf{Pr}_L\left(L \otimes (M \oplus M') = T\right) \quad = \mathbf{Pr}_L\left(L = T \otimes (M \oplus M')^{-1}\right) = 1/2^n\,.$$

(b) Make the following three MAC queries, for arbitrary M, M', N, N':

$$\begin{aligned}(N, M) &\mapsto T\\ (N, M') &\mapsto T'\\ (N', M) &\mapsto T''\,.\end{aligned}$$

Then, we know that $(N', M', T \oplus T' \oplus T'')$, so this is our forgery attempt:

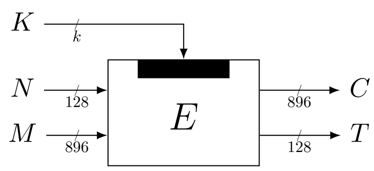
$$\begin{aligned}\mathsf{WC}_K(N', M') &= F_K(N') \oplus H_L(M')\\ &= \Big(T'' \oplus H_L(M)\Big) \oplus H_L(M')\\ &= T'' \oplus H_L(M) \oplus H_L(M')\\ &= T'' \oplus \Big(T \oplus F_K(N)\Big) \oplus \Big(T' \oplus F_K(N)\Big)\\ &= T \oplus T' \oplus T''\,.\end{aligned}$$

End Secret Info ....

3. **(10 points)** Suppose we are given a block cipher $E : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ for large n, in this case $n = 1024$. Consider the following authenticated encryption scheme

$$\begin{aligned}\mathsf{AE}\colon \{0,1\}^k \times \{0,1\}^{128} \times \{0,1\}^{896} &\to \{0,1\}^{896} \times \{0,1\}^{128}\,,\\ (K, N, M) &\mapsto (C, T)\,,\end{aligned}$$

defined as follows:



We will consider the nonce-misuse-resistance of this scheme. In other words, we consider security of this construction in the model of lecture 3 slide 4, $\mathbf{Adv}^{\text{ae}}_{\mathsf{AE}}(q_e, q_v)$, *with the difference that $\mathcal{D}$ may repeat nonces.* Here, q_e and q_v denote the total number of encryption and decryption queries, respectively.

(a) Describe how the authenticated decryption function AE_K^{-1} operates.

(b) The first step in the security proof of AE will be to replace the keyed block cipher E_K by a random permutation p. Apply the triangle inequality to do so, with explicitly mentioning the loss incurred by this triangle inequality:

$$\Delta_{\mathcal{D}}\left(\mathsf{AE}_K, \mathsf{AE}_K^{-1} \ ; \ \$, \perp\right) \leq \Delta_{\mathcal{D}}\left(\mathsf{AE}[p], \mathsf{AE}[p]^{-1} \ ; \ \$, \perp\right) + \dots$$

Explain your answer in words.

(c) We are left with the task of bounding $\Delta_{\mathcal{D}}\left(\mathsf{AE}[p], \mathsf{AE}[p]^{-1} \ ; \ \$, \perp\right)$. We will perform another triangle inequality:

$$\Delta_{\mathcal{D}}\left(\mathsf{AE}[p], \mathsf{AE}[p]^{-1} \ ; \ \$, \perp\right) \leq \Delta_{\mathcal{D}}\left(\mathsf{AE}[p], \mathsf{AE}[p]^{-1} \ ; \ \mathsf{AE}[p], \perp\right) + \Delta_{\mathcal{D}}\left(\mathsf{AE}[p], \perp \ ; \ \$, \perp\right) . \tag{2}$$

The first distance of (2) is a bit peculiar and will be ignored. Derive a bound on the second distance of (2), $\Delta_{\mathcal{D}}\left(\mathsf{AE}[p], \perp \ ; \ \$, \perp\right)$.

Begin Secret Info:...

(a) AE_K^{-1} *gets as input a tuple (N, C, T). It evaluates $E_K^{-1}(C, T)$ and parses the outcome as $M \| N^*$. If $N = N^*$ it outputs M, otherwise it outputs $\perp$.*

(b) As in the lectures:

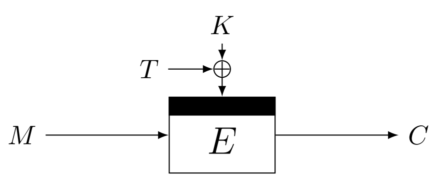
$$\begin{aligned}\Delta_{\mathcal{D}}\left(\mathsf{AE}_K, \mathsf{AE}_K^{-1} \ ; \ \$, \perp\right) &\leq \Delta_{\mathcal{D}}\left(\mathsf{AE}[p], \mathsf{AE}[p]^{-1} \ ; \ \$, \perp\right) + \Delta_{\mathcal{D}}\left(\mathsf{AE}_K, \mathsf{AE}_K^{-1} \ ; \ \mathsf{AE}[p], \mathsf{AE}[p]^{-1}\right) \\ &\leq \Delta_{\mathcal{D}}\left(\mathsf{AE}[p], \mathsf{AE}[p]^{-1} \ ; \ \$, \perp\right) + \mathbf{Adv}^{\text{sprp}}_{E}(q_e + q_v) .\end{aligned}$$

Here, it is important to note that we take SPRP security and not PRP security as the adversary can technically trigger inverse evaluations of E.

(c) Note that the decryption oracle is redundant, and we have to basically consider the PRF-security of $\mathsf{AE}[p]$ under q_e encryption queries. First apply the RP-to-RF-switch (i.e., replace p by f) at the cost of $\binom{q_e}{2}/2^{1024}$. Then, any response is uniformly randomly distributed from $\{0,1\}^{1024}$ and the worlds are indistinguishable.

End Secret Info ....

4. **(5 points)** Consider a tweakable block cipher $\widetilde{E} : \{0,1\}^k \times \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$, i.e., with k-bit key and tweak and n-bit data path, built from a block cipher $E : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ as follows:



It is possible to recover the secret key K with high probability in $2^{k/2}$ evaluations of $\widetilde{E}_K$ and $2^{k/2}$ offline evaluations of E. Describe the attack. You can assume that $k \ll n$, i.e., there is no need to make additional queries to eliminate false positives.

Begin Secret Info:...

Let $q = 2^{k/2}$. The attacker makes the following queries:

- *q construction queries $(T_i, 0) \mapsto C_i = \widetilde{E}_K(T_i, 0)$ for varying T_i;*
- *q primitive queries $(L_j, 0) \mapsto Y_j = E_{L_j}(0)$ for varying L_j;*
- *If there exist i, j such that $C_i = Y_j$, then the key satisfies $K = T_i \oplus L_j$.*

As $k \ll n$, the probability that the collision $C_i = Y_j$ happens even though $K \neq T_i \oplus L_j$ is negligible and can be discarded. If $k \geq n$, one must make a verification query to eliminate false positives.

End Secret Info ....

5. **(10 points)** Let $n = 128$, take $E : \{0,1\}^{128} \times \{0,1\}^{128} \to \{0,1\}^{128}$ to be your favorite block cipher, and consider the XEX construction XEX_K of lecture 3 slide 24. As this question is particularly about the masking, we will have to explicitly define what multiplication means in this context. To any string $a = a_{127}a_{126}\ldots a_0 \in \{0,1\}^{128}$, we associate its polynomial $a(X) = a_{127}X^{127} + a_{126}X^{126} + \cdots + a_0$. Addition of bit strings is defined as the bitwise XOR, as usual. Multiplication of two bit strings is defined as the multiplication of the two polynomials in GF(2^{128}) modulo $q(X) = X^{128} + X^7 + X^2 + X + 1$.

   (a) The masking is of the form $2^\alpha 3^\beta 7^\gamma \cdot E_K(N)$. Give the polynomials associated with "2", "3", and "7".

   (b) Suppose that for a certain value of N, $E_K(N) = \underbrace{0\ldots0}_{123}10101$. Compute $2^3 \cdot E_K(N)$ and $2^3 3 \cdot E_K(N)$.

   (c) Suppose that for a certain value of N, $E_K(N) = 1\underbrace{0\ldots0}_{127}$. Compute $2 \cdot E_K(N)$.

   (d) It is rather weird that XEX_K uses $2, 3, 7$ as masks and not $2, 3, 5$. Try to find out why. (Hint: admissible domains.)

Begin Secret Info:...

(a) *These are*

$$2(X) = X\,,$$
$$3(X) = X + 1\,,$$
$$7(X) = X^2 + X + 1\,.$$

(b) *Multiplication by 2^3 is a left-shift of 3, multiplication by 3 is a left-shift of 1 plus a single XOR:*

$$2^3 \cdot E_K(N) = \underbrace{0\ldots0}_{120}10101000\,,$$
$$2^3 3 \cdot E_K(N) = \underbrace{0\ldots0}_{119}111111000\,.$$

(c) *Now, the output is modulated:*

$$2 \cdot E_K(N) = \underbrace{0\ldots0}_{120}10000111\,.$$

(d) *It happens to be the case that* $3^2 = 5$. *This means that you can easily come up with different tweaks* $(\alpha, \beta, \gamma), (\alpha', \beta', \gamma')$ *such that the masks collide. For example, if* $(\alpha', \beta', \gamma') = (\alpha, \beta + 2, \gamma - 1)$,

$$2^{\alpha'} 3^{\beta'} 5^{\gamma'} = 2^{\alpha} 3^{\beta+2} 5^{\gamma-1} = 2^{\alpha} 3^{\beta} 5^{\gamma} .$$

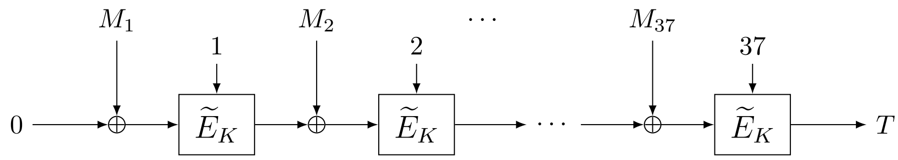
This is undesirable: one prefers the "admissible domains" for α, β, γ *non-trivial. This happens to be the case if you take* $(2, 3, 7)$.

End Secret Info . . . .

6. **(10 points)** Let $\widetilde{E} : \{0,1\}^k \times [1, 37] \times \{0,1\}^n \to \{0,1\}^n$ be a tweakable block cipher. Consider the PRF construction

$$F : \{0,1\}^k \times (\{0,1\}^n)^{37} \to \{0,1\}^n$$

that operates by first splitting the $37n$-bit message into 37 n-bit chunks $M_1 \| \cdots \| M_{37}$ and then processing this message as follows:



For this assignment, it is important to note that F generates authentication tags for messages that are of size *EXACTLY* $37n$ bits.

(a) We will consider the PRF security of F against any distinguisher that can make q construction queries of $37n$ bits. Prove that F is a secure PRF up to the following bound:

$$\mathbf{Adv}_F^{\mathrm{prf}}(q) \le 2 \cdot 37 \binom{q}{2} / 2^n + \mathbf{Adv}_{\widetilde{E}}^{\mathrm{tprp}}(37q) .$$

We have seen proofs in earlier assignments, but this one is a little bit harder. Therefore, we will give you some hints:

- It is easier to reason about the construction if the underlying primitives behave as random functions. The first two steps will move you from above construction to a construction based on random functions.
- Then, note that if for two *different* queries (i.e., with $M^{(i_1)} \neq M^{(i_2)}$) the input to the last random function never collides, we are fine as the output tags are independently generated using a random function.
- So, the big question is to upper bound a *non-trivial* (i.e., with $M^{(i_1)} \neq M^{(i_2)}$) collision at the last random function, and here you will have to apply induction.
- There is no page limit, but as a reference: in the solutions of this assignment the proof takes around 1 page including two figures.
- Remark: it is possible to derive a slightly stronger bound. In particular, if you would opt for the so-called "H-coefficient technique", this is possible and you will get a slightly tighter bound, but the analysis is a bit more cumbersome.

Good luck!

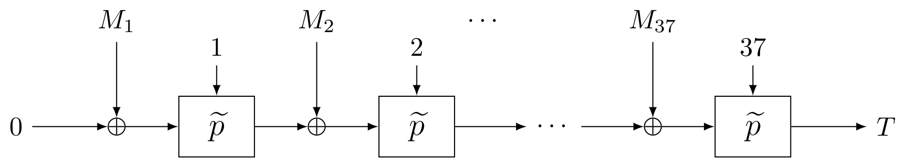
(b) Suppose we would stretch the usage of F and allow it for all messages of size a positive multiple of n bits, up to $37n$ bits. In other words, for an n-bit message M_1, one generates tag $T = \widetilde{E}_K(1, M_1)$, for a $2n$-bit message $M_1 \| M_2$ one generates tag $T = \widetilde{E}_K(2, \widetilde{E}_K(1, M_1) \oplus M_2)$, etc. Then, the scheme is vulnerable to a trivial distinguishing attack. Describe the attack. You do not have to derive a success probability.

Begin Secret Info:. . . . .
A sharp eye will recognize a tweakable block cipher based version of CBC-MAC in the function F. Because it is tweakable block cipher based, proving security (question (a)) will become easier but finding an attack (question (b)) a bit harder.

(a) *Consider any distinguisher $\mathcal{D}$ making q queries. The first step consists of replacing the tweakable block cipher $\widetilde{E}$ by a random tweakable permutation $\widetilde{p}$. This comes at a cost*

$$\mathbf{Adv}^{\mathrm{tprp}}_{\widetilde{E}}(37q)\,,$$

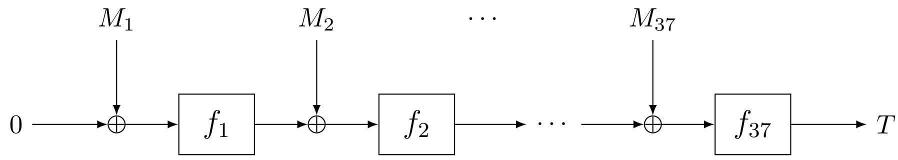
and leaves us with the following scheme:



Note that the permutations $\widetilde{p}(1,\cdot)$, $\widetilde{p}(2,\cdot)$, $\ldots$, $\widetilde{p}(37,\cdot)$ are independent random permutations, and using 37 *RP-to-RF switches (see also the solution to question 3), we can replace them by random functions f_1, f_2, $\ldots$, f_{37} at a cost of*

$$37\binom{q}{2}/2^n\,.$$

We are left with



So, to recap, we obtained that

$$\begin{aligned}\mathbf{Adv}^{\mathrm{prf}}_{F}(\mathcal{D}) &= \Delta_{\mathcal{D}}\left(F_K\ ;\ \mathcal{RO}\right)\\ &\le \Delta_{\mathcal{D}}\left(F[f_1,\ldots,f_{37}]\ ;\ \mathcal{RO}\right) + 37\binom{q}{2}/2^n + \mathbf{Adv}^{\mathrm{tprp}}_{\widetilde{E}}(37q)\,.\end{aligned}$$

Define the inputs to f_j by X_j for $j = 1,\ldots,37$, and define

$$\begin{aligned}\mathsf{bad}_j :=\ &\textit{there exist two queries } i_1, i_2 \textit{ such that } M_1^{(i_1)}\|\cdots\|M_j^{(i_1)} \neq M_1^{(i_2)}\|\cdots\|M_j^{(i_2)}\\ &\textit{and } X_j^{(i_1)} = X_j^{(i_2)}\,.\end{aligned}$$

Clearly, under the assumption that bad_{37} *does not hold, $F[f_1,\ldots,f_{37}]$ is perfectly indistinguishable from $\mathcal{RO}$. Thus,*

$$\Delta_{\mathcal{D}}\left(F[f_1,\ldots,f_{37}]\ ;\ \mathcal{RO}\right) \le \mathbf{Pr}\left(\mathsf{bad}_{37}\right)\,.$$

Unfortunately, this event can only be analyzed under the assumption that bad_{36} *does not happen:*

$$\begin{aligned}\mathbf{Pr}\left(\mathsf{bad}_{37}\right) &\le \mathbf{Pr}\left(\mathsf{bad}_{37} \mid \neg\mathsf{bad}_{36}\right) + \mathbf{Pr}\left(\mathsf{bad}_{36}\right)\\ &\le \binom{q}{2}/2^n + \mathbf{Pr}\left(\mathsf{bad}_{36}\right)\,.\end{aligned}$$

Now, we can induct with the observation that $\mathbf{Pr}\left(\mathsf{bad}_1\right) = 0$, *and obtain that*

$$\mathbf{Pr}\left(\mathsf{bad}_{37}\right) \leq 37 \cdot \binom{q}{2} / 2^n \,.$$

This completes the proof.

(b) The distinguisher fixes any $M_1 \neq M_1'$ *and and* M_2, *and makes the following queries:*

- $F_K(M_1) = T$;
- $F_K(M_1') = T'$;
- $F_K(M_1 \| (M_2 \oplus T)) = T''$;
- $F_K(M_1' \| (M_2 \oplus T')) = T'''$.

By construction, $T'' = T'''$ *(verify this!), which is unlikely to happen for the random oracle* $\mathcal{RO}$.

End Secret Info ....