Crypto Engineering
Symmetric encryption (mostly)

2023-W39

Exercise 1: A random sequence *(M1 Exam '18)*

Let $\mathcal{S}$ be a set of size N; let $(u_n)_{n\in\mathbb{N}}$ be a sequence whose elements are drawn independently and uniformly at random from $\mathcal{S}$, i.e. for all i, $u_i \twoheadleftarrow \mathcal{S}$. Suppose that you do not initially know $\mathcal{S}$,[*] nor N.

1. Give an algorithm that takes as input a finite number of elements of (u_n) and that returns an approximation of N (you do not need to rigorously prove your approximation).

2. What is the time and memory complexity of your algorithm (be careful to specify the data structures you may use)?

Exercise 2: 18+

Recall that Ber_p denotes the probability distribution over $\{0,1\}$ s.t. the event '1' occurs with probability p. We also denote by Ber_p^Q the product (or vectorial) distribution $\mathrm{Ber}_p \otimes \cdots \otimes \mathrm{Ber}_p$ formed by Q independent distributions Ber_p.
Recall also that the total variation distance between two distributions $\mathfrak{D}$ and $\mathfrak{D}'$ over a finite universe Ω is defined as $1/2 \sum_{\omega\in\Omega} |\mathfrak{D}(\omega) - \mathfrak{D}'(\omega)|$, or equivalently $\sum_{\{\omega:\mathfrak{D}(\omega)>\mathfrak{D}'(\omega)\}} \mathfrak{D}(\omega) - \mathfrak{D}'(\omega)$.

1. Suppose (w.l.o.g.) that $p = 1/2 + \varepsilon$, for $\varepsilon > 0$; specify a (tentatively optimal) distinguisher for Ber_p and $\mathrm{Ber}_{1/2}$ (i.e. the uniform distribution over $\{0,1\}$), and compute its advantage in function of ε.

2. Same question for Ber_p^2 and $\mathrm{Ber}_{1/2}^2$. What remark can you make about the growth of the advantage compared to previous question?

3. Compute the total variation distance between Ber_p^2 and $\mathrm{Ber}_{1/2}^2$. Was your previous distinguisher optimal?

4. Propose a general distinguishing strategy for an arbitrary Q.

Remark. It is possible to show that to distinguish Ber_p^Q and $\mathrm{Ber}_{1/2}^Q$ with constant advantage (in function of p), one needs $Q = \Omega(1/\varepsilon^2)$.

Exercise 3: MTP, CTR[RF]

For the sake of simplicity and without loss of generality, we assume in this exercise that all messages are of a fixed bitlength n.

We define the *many-time pad* encryption algorithm as follows: given a key $k \in \{0,1\}^n$ and a message $m \in \{0,1\}^n$, it computes $\mathsf{MTP}(k, m)$ as $m \oplus k$.

1. Give an efficient attack on MTP, w.r.t. the IND-CPA security definition

[*] Be careful that the elements of $\mathcal{S}$ need not be integers. For instance $\mathcal{S}$ could be equal to $\{$*martes martes*, *martes foina*, *martes zibellina*$\}$.

We now define the *Counter mode with random function* encryption algorithm $\mathsf{CTR}[\mathsf{RF}]$ as follows. Let $\mathsf{Func}(\{0,1\}^n)$ denote the set of all functions of $\{0,1\}^n \to \{0,1\}^n$. Further let $\mathtt{c}$ be a stateful "counter" in $[\![0, 2^n - 1]\!]$ initialised to zero, and $\mathtt{c++}$ the expression that evaluates to the n-bit string representing $\mathtt{c}$, and then has the side-effect of incrementing its value by one, modulo 2^n. Given a key $F \in \mathsf{Func}(\{0,1\}^n)$ and a message m, $\mathsf{CTR}[\mathsf{RF}](F, m)$ is then computed as $F(\mathtt{c++}) \oplus m$

2. Show that $\boldsymbol{Adv}^{\text{IND-CPA}}_{\mathsf{CTR}[\mathsf{RF}]}(< 2^n, \infty) = 0$

3. Show that $\boldsymbol{Adv}^{\text{IND-CPA}}_{\mathsf{CTR}[\mathsf{RF}]}(2^n, \infty) = 1$

Exercise 4: PRP $\Rightarrow$ UP

Consider a block cipher $E : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$.

1. Show that $\boldsymbol{Adv}^{\text{PRP}}_{E}(q+1, t+1) \geq \left|\boldsymbol{InSec}^{\text{UP}}_{E}(q,t) - \frac{1}{2^n - q}\right|$.

   Hint: Show that any UP adversary may be used as a black box to derive a PRP adversary.

2. Give a specification for E s.t.:

   (a) $\boldsymbol{InSec}^{\text{UP}}_{E}(q, \infty) = 1/(2^{n-1} - q)$;

   (b) $\boldsymbol{Adv}^{\text{PRP}}_{E}(q, \infty) \geq 1/2$.

3. Explain why one may be justified in saying that for block ciphers, "PRP security implies UP security, but the converse is false" (or equivalently, that "UP security reduces to PRP security, but not the converse").

4. Is the reduction of the previous question *tight*, considering that for a "good" E with $\mathcal{K} = \{0,1\}^n$, one expects $\boldsymbol{Adv}^{\text{PRP}}_{E}(1,1) \approx 1/2^{n/2}$ and $\boldsymbol{InSec}^{\text{UP}}_{E}(1,1) \approx 1/2^n$?

Exercise 5: RKA-UP

Consider a block cipher $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$, and Φ a set of functions of $\{0,1\}^n \to \{0,1\}^n$. We define the *related-key unpredictability* "RKA-UP[Φ]" of E with parameter Φ (and its associated insecurity function), similarly as the usual unpredictability, except that in addition to its oracle for $E(k, \cdot)$, the adversary may now additionally query *related-key* oracles $E(\phi(k), \cdot)$ for any $\phi \in \Phi$.

For $\Delta \in \{0,1\}^n$, let $\boxplus\Delta$ (resp. $\oplus\Delta$) denote the function: $x \mapsto x \boxplus \Delta$ (resp. $x \mapsto x \oplus \Delta$).[†]
We first consider $\Phi^{\boxplus,\oplus} := \{\boxplus\Delta : \Delta \in \{0,1\}^n\} \cup \{\oplus\Delta : \Delta \in \{0,1\}^n\}$.

1. Show that for some $q, t = \Theta(n)$, one has $\boldsymbol{InSec}^{\text{RKA-UP}[\Phi^{\boxplus,\oplus}]}_{E}(q,t) \approx 1$.

   Hint: Use the related-key oracles to recover k (with overwhelming probability, possibly up to equivalence) by exploiting the fact that the carry propagation rule implies that $x \boxplus 1 = x \oplus 1$ iff. x has its least-significant bit set to zero.

We now restrict $\Phi^{\oplus}$ to $\{\oplus\Delta : \Delta \in \{0,1\}^n\}$.

2. Show that for some $q, t = \Theta(2^{n/2})$, one has (under very mild assumptions on E) $\boldsymbol{InSec}^{\text{RKA-UP}[\Phi^{\oplus}]}_{E}(q,t) \approx 1$.

   Hint: Use the related-key oracles to recover k (with overwhelming probability, under very mild assumptions on E) by exploiting collisions between the sets $\{k' : k' \twoheadleftarrow \{0,1\}^n\}$ and $\{k \oplus \Delta' : \Delta' \twoheadleftarrow \{0,1\}^n\}$.

[†] Here '$\oplus$' is the usual notation for the bitwise XOR, and $\boxplus$ the (somewhat less) usual notation for addition modulo the string size (of the canonical embedding of the strings to the integers).

3. As far as you can tell, is RKA-UP$[\Phi^{\boxplus,\oplus}]$ a meaningful security notion? Same question for RKA-UP$[\Phi^{\oplus}]$.

Remark. One usually rather considers related-key *PRP* security (which is a bit lengthier to define), for which the attacks studied in this exercise naturally carry over (cf. for instance the previous exercise). In the case of $\Phi^{\oplus}$, one may show that for black-box block ciphers, the above attack is essentially optimal.

Exercise 6: Active ciphertext-only attack on raw CTR mode

We consider the following setup: two parties exchange messages of the form $m_0||\cdots||m_7$, where $m_0, \ldots, m_6$ are 8-bit strings representing 7-bit ASCII characters,[‡] and m_7 represents the sum modulo 256 of $m_0, \ldots, m_6$ interpreted as integers.

The messages are encrypted in CTR mode, instantiated with a 64-bit block cipher E which is assumed to be a "good" PRP. The ciphertexts of (one-block) messages are of the form $c||\,E(k,c) \oplus m$, where c is the (public) counter. The selection of the counter is assumed to be well-implemented, either through a stateful mechanism or from a high-quality random source.

Upon receiving a message, one decrypts it as $m'_0||\cdots||m'_7$ (using the counter provided in the first half of the ciphertext) and checks if m'_7 is indeed equal to the modular sum of $m'_0, \ldots, m'_6$; if not, it sends back an unencrypted error code "SENDAGN" asking to resend the message.

We then consider an *active* ciphertext-only adversary that can read every message exchanged between the two parties; intercept any of those messages (so that the receiver does not receive anything); inject arbitrary messages (that the receiver will treat in the same way as if they had been sent by the other party).

Finally, we say that the ordered pair (a, b) of two n-bit strings has a (one-bit) *signed difference* $+2^i$, $0 \le i < n$ if $a \boxplus 2^i = b$ (where $\boxplus$ is defined similarly as in the previous exercise). Equivalently, this means that a and b differ exactly in their i^{th} bit, and that this bit is equal to zero in a, or $i = n - 1$. We say symmetrically that (a, b) has signed difference -2^i if $a = b \boxplus 2^i$ (or, introducing the $\boxminus$ operator, $a \boxminus 2^i = b$).[§]

1. Let a, b, $c = a \boxplus b$ be 8-bit strings. Further let $a' = a \oplus 1$, $b' = b \oplus 1$, $c' = a' \boxplus b'$ (where 1 denotes here the string with only its least-significant bit set to 1). Show that $c = c'$ iff. the signed differences of (a, a') and (b, b') have different signs.

2. Show that the above is not true anymore for $a' = a \oplus 2^7$ and $b' = b \oplus 2^7$.

3. Design an attack that, given a ciphertext $c||\,E(k,c) \oplus m$, injects 7×7 ciphertexts, intercepts at most 7×7 error messages, and returns 2^7 possible candidates for m.

4. Implement a proof of concept of this attack.

5. Explain why a similar attack would not work if one had used a "XOR checksum" instead of a modular one (or more generally, any $\mathbb{F}_2$-linear error-detection mechanism).

[‡] This implies in particular that the most-significant of those 8 bits is always zero.

[§] These definitions also naturally extend to differences on more than one bit, but we will not need those in the present exercise. (When using multibit signed differences, one should be aware that there may be several possible differences for a given pair, since the signed representation of bit strings is redundant.)