Cryptographic Engineering

Important: This exam is composed of 4 parts:

Part 1: P. Karpman, **7.5 points**

Part 2: C. Pernet, **6.5 points**

Part 3: E. Peyre, **3 points**

Part 4: C. Ene, **6 points**

— Any paper document allowed. All electronic devices are forbidden.

— **Each of the 4 parts has to be answered on a separate answer sheet.**

— The grading over 23 points will not be scaled, hence it is not necessary to answer correctly all questions to get the maximum grade of 20.

— Your answers have to be short but clearly and cleanly argued or commented.

— You may assume the results of unanswered questions to proceed to the next ones.

Part 1: Symmetric Cryptography (P. Karpman)

Exercise 1.1 (7.5 pts.): Format-preserving block ciphers

We first briefly recall the following security definitions.

UP. Let F be an arbitrary keyed function $\{0,1\}^\kappa \times \mathcal{M} \to \mathcal{C}$. An adversary in the game $\textsc{Forge}^F$ is given oracle access to $\mathbb{O} = F(k,\cdot)$ for $k \twoheadleftarrow \{0,1\}^\kappa$; it wins iff. it returns a couple (x,y) s.t.:

1. x was not queried to $\mathbb{O}$
2. $F(k,x) = y$

One then defines:

$$\mathbf{InSec}_F^{\mathsf{UP}}(q,t) = \max_{A_{q,t}} \Pr[A_{q,t}^{\mathbb{O}}() \text{ wins } \textsc{Forge}^F]$$

where $A_{q,t}$ makes q queries to $\mathbb{O}$ and runs in time t.

PRP. Let F be an arbitrary keyed function $\{0,1\}^\kappa \times \mathcal{M} \to \mathcal{M}$, and $\mathrm{Perm}(\mathcal{M})$ denote the set of all permutations over $\mathcal{M}$. One defines:

$$\mathbf{Adv}_F^{\mathsf{PRP}}(q,t) = \max_{A_{q,t}} \left| \Pr[A_{q,t}^{\mathbb{O}}() = 1 : \mathbb{O} \twoheadleftarrow \mathrm{Perm}(\mathcal{M})] - \Pr[A_{q,t}^{\mathbb{O}}() = 1 : \mathbb{O} = F(k,\cdot), k \twoheadleftarrow \{0,1\}^\kappa] \right|$$

The goal of this exercise is to study a generic construction that reduces the message domain of an n-bit block cipher $E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ to some subset of $\{0,1\}^n$. (This subset may be arbitrary, and in particular is not guaranteed to possess a rich algebraic structure; for instance, it may be the subset of strings representing n-bit prime numbers, or valid Dutch *burgerservicenummer*.)

Given $\mathcal{S} \subset \{0,1\}^n$, the *cycle walking* construction generically allows to build from E a block cipher $\mathsf{CW}[E,\mathcal{S}] : \{0,1\}^\kappa \times \mathcal{S} \to \mathcal{S}$. It works as follows: to evaluate $\mathsf{CW}[E,\mathcal{S}](k,\cdot)$ on $x \in \mathcal{S}$, compute $x' := E(k, x)$; then if $x' \in \mathcal{S}$ return x'; otherwise iterate the process by computing $x'' = E(k, x')$ and test if it is in $\mathcal{S}$, etc., and return the first encountered $x'^{\cdots'}$ that is in $\mathcal{S}$.

This construction may also be applied to any fixed permutation P (rather than on a block cipher) in the obvious way, and we will admit that if $P \twoheadleftarrow \mathrm{Perm}(\{0,1\}^n)$ is a uniformly sampled permutation of domain $\{0,1\}^n$, then $\mathsf{CW}[P, \mathcal{S}]$ is a uniformly sampled permutation of domain $\mathcal{S}$. We will also make the (obviously wrong) simplifying hypothesis that for every $\mathcal{S} \subset \{0,1\}^n$, for every $x \in \{0,1\}^n$, for $c := \lceil 2^n / \#\mathcal{S} \rceil$, the probability (over the sampling of P) that none of the values $P(x), P \circ P(x), \cdots, P^c(x)$ is in $\mathcal{S}$ is equal to zero (where P^c denotes the c-time composition of P).

Q.1 (***correctness & efficiency***)

1. Informally state two necessary conditions on $\mathcal{S}$ for $\mathsf{CW}[E, \mathcal{S}]$ to be an "efficient" block cipher of message domain $\mathcal{S}$, when E is any "efficient" block cipher of message domain $\{0,1\}^n$.

2. Give an efficient[1] algorithm to compute the inverse cipher $\mathsf{CW}[E, \mathcal{S}]^{-1} : \{0,1\}^\kappa \times \mathcal{S} \to \mathcal{S}$ of $\mathsf{CW}[E, \mathcal{S}]$. (That is, given k, $y := \mathsf{CW}[E](k, x)$ and the knowledge of E and $\mathcal{S}$, this algorithm must return x.)

Q.2 (***PRP security***)

1. Show that under the above simplifying hypothesis and given $\mathcal{S}$ and $x \in \mathcal{S}$, a PRP adversary for E that cannot compute $\mathsf{CW}[\mathbb{O}, \mathcal{S}](x)$ with at most c queries to its oracle $\mathbb{O}$ is able to win the PRP game with advantage one. (That is, show that when the relevant probabilities are conditioned by this event, the PRP advantage is one.)

2. Show by an explicit reduction that under the above simplifying hypothesis one has:

$$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{CW}[E,\mathcal{S}]}(q, t) \leq \mathbf{Adv}^{\mathsf{PRP}}_{E}(cq, ct)$$

Be careful to justify your answer as much as possible.

Q.3 (***UP security***)

1. Show by an explicit reduction that under the above simplifying hypothesis, one has:

$$\mathbf{InSec}^{\mathsf{UP}}_{\mathsf{CW}[E,\mathcal{S}]}(q, t) \leq \mathbf{Adv}^{\mathsf{PRP}}_{E}(c(q+1), c(t+1)) + \frac{1}{\#\mathcal{S} - q}$$

2. Does the above reduction strategy also work to reduce the UP security of $\mathsf{CW}[E, \mathcal{S}]$ to the UP (and not PRP) security of E?

3. Could it be useful to reduce the UP security of $\mathsf{CW}[E, \mathcal{S}]$ to the UP security of E?

Q.4 (***application***)

1. Suppose that one wishes to use $\mathsf{CW}[E, \mathcal{S}]$ to implement an encryption scheme over $\mathcal{S}$ whose security will be quantified w.r.t. $\mathsf{IND\text{-}CPA}$ security. Which of the two above security definitions for $\mathsf{CW}[E, \mathcal{S}]$ is the most relevant for that?

2. Suppose that one wishes to design a MAC whose message domain is $\mathcal{S}$ and whose tag space may be arbitrary. Do you think that using $\mathsf{CW}[E, \mathcal{S}]$ as a basis is a good idea?

---

[1] As much as $\mathsf{CW}[E, \mathcal{S}]$.

Part 2: Asymetric Cryptography (C. Pernet)

Exercise 2.1 (6.5 pts.): McEliece

Recall that the Mc Eliece cryptosystem based on a code $\mathcal{C}$ over a field $\mathbb{K}$ is defined by:

— the private key is composed of a generator matrix $G \in \mathbb{K}^{k\times n}$ of a code with an efficient decoding algorithm up to t errors, an invertible matrix $S \in \mathbb{K}^{k\times k}$, a permutation matrix $P \in \mathbb{K}^{n\times n}$;

— the public key is $(\hat{G}, t)$ where $\hat{G} = SGP$

— the encryption function: $E : m \mapsto c = m\hat{G} + e$ where e is sampled uniformly with $w_H(e) = t$

1. (0.5 pts) Recall how the decryption algorithm works.
2. In order to ensure a sufficiently good resistance against know attacks, we are requested to use a linear code of length 1024 able to correct up to 50 errors.

   2.1 (1pt) If we choose to work over a Reed-Solomon code, what would be the parameters of the code (base field, length, dimension)? What would be the size in kilobytes of the public key ?

   2.2 (1pt) Same question if we choose to work over a binary Goppa code. We recall that a binary Goppa code $\mathcal{G}$ of length n and parameters (m, r) is obtained as $\mathbb{F}_2^n \bigcap \mathrm{GRS}_{2^m}(n, n-r)$, where $\mathrm{GRS}_q(n, k)$ is a generalized Reed-Solomon code over the field $\mathbb{F}_q$ of length n and dimension k. This construction ensures that the dimension of $\mathcal{G}$ is $\geq n - rm$ and its minimum distance is $\geq 2r + 1$.

   For an arbitrary field, suppose that a same message m is sent twice using McEliece cryptosystem. An attacker, has then access to two different ciphertexts $y^{(1)}$ and $y^{(2)}$ for the same message m.

3. (1.5pts) Given two vectors $e, f \in \mathbb{F}_q^n$ with t non zero coefficients each, sampled uniformly at random (both the positions and the values of the non-zero coefficients):

   (a) For a fixed index i, what are the probabilities $P[e_i = 0], P[f_i = 0]$ and $P[e_i + f_i = 0]$. (express them as functions of q, n and t)

   (b) What is the probability $P[e_i = 0 \mid e_i + f_i = 0]$?

4. (0.5pts) Consequently, explain why the attacker can deduce, k positions in $y^{(1)}$ at which the corresponding error $e^{(1)}$ is zero, with a high probability.
5. (1pts) Deduce that there is then a polynomial time algorithm (state its cost) to compute the clear text m without knowing the private key.
6. (0.5pts) Explain how does this attack generalizes for the *related plaintext attack*: when the ciphertexts c_1 and c_2 correspond to plain texts which difference is known to the attacker.
7. (0.5pts) Propose a countermeasure for these attacks.

Part 3: Elliptic curves (E. Peyre)

Exercise 3.1: (3 pts) Elliptic curves

Let E be the elliptic curve defined by the affine equation

$$Y^2 = X^3 + 3$$

over the field $\mathbf{F}_{11} = \mathbf{Z}/11\mathbf{Z}$.

1. List all the elements of $E(\mathbf{F}_{11})$.
2. Find all points of order 2 in $E(\mathbf{F}_{11})$.
3. How do we know that the group $E(\mathbf{F}_{11})$ is isomorphic to $\mathbf{Z}/12\mathbf{Z}$?

Part 4: Security Proofs (C. Ene)

Exercice 4.1 (4.0 pts.)

In this exercise, $\langle \_ , \_ \rangle$ represents concatenation, $[\ \_\ ]_\_$ represents a symmetric encryption scheme, $\{\ \_\ \}_\_$ an asymmetric encryption scheme, $pr(u)$ is the inverse secret key associated to $pk(u)$ and $\oplus$ denotes the usual bitwise xor over equal-length bitstrings, e.g. $0011 \oplus 1110 = 1101$. Consider the following protocol:

$$\begin{array}{llcll} 1. & A & \rightarrow & B & : \{\ \langle\langle B, A\rangle, N_a\rangle\ \}_{pk(B)} \\ 2. & B & \rightarrow & A & : \langle\{\ \langle K \oplus N_a, A\rangle\ \}_{pk(A)}, [\ N_a\ ]_K\rangle \\ 3. & A & \rightarrow & B & : \{\ \langle\langle A, B\rangle, K\rangle\ \}_{pk(B)} \end{array}$$

The goal of this protocol is to provide both secrecy and authentication: at the end of a session between two honest participants a and b, **"k" (the instantiation of the variable K in the specification of the protocol) should be a new shared secret value known only by a and b.** This target session between honest participants a and b may be part of a richer scenario containing other running sessions in parallel where the active adversary i can be involved.

1. Describe in details (as a list) A's and B's actions at receipt of messages 2 and 3 and what beliefs they have at that stage.

2. Show (**using the McAllester's Algorithm**) that k (the instantiation of the variable K in the specification of the protocol) remains secret in presence of a passive Dolev-Yao intruder.

3. What do you think about the correctness of the protocol in presence of an active Dolev-Yao intruder? If you think that the protocol is correct, then give a justification. Otherwise,

   — give an attack on the target session between honest participants a and b where the intruder i will learn k;

   — propose a correction of the protocol.

Exercice 4.2 (2.0 pts.)

In this exercise, $|\cdot|$ denotes the length of a bitstring, $\overline{x}$ is the bitwise complement of x (e.g. $\overline{1101} = 0010$) and $\oplus$ denotes the usual bitwise xor over equal-length bitstrings, e.g. $0011 \oplus 1110 = 1101$. A one-way function is a function that is easy to compute but hard to invert. Formally, $f : \{0,1\}^* \mapsto \{0,1\}^*$ is a one-way function, if for all probabilistic polynomial-time families of adversaries $\mathcal{A}$ the following probability:

$$p(k) \stackrel{def}{=} Pr_{b \stackrel{R}{\leftarrow} [x \stackrel{R}{\leftarrow} \{0,1\}^k;\ y \leftarrow f(x);\ x' \stackrel{R}{\leftarrow} \mathcal{A}(y)\ :\ \text{return } f(x')=y]}(b = true)$$

(simpler written $p(k) \stackrel{def}{=} Pr[f(x') = y \mid x \stackrel{R}{\leftarrow} \{0,1\}^k;\ y \leftarrow f(x);\ x' \stackrel{R}{\leftarrow} \mathcal{A}(y)]$)
is a negligible function in k. That is, the probability that a probabilistic polynomial-time algorithm $\mathcal{A}$ is able to find a preimage x' for a given image $y = f(x)$ of an uniformly sampled x is negligible. In this exercise, we assume the existence of at least one such one-way function denoted by f_0.

For each of the assertions below, prove or disprove that they are valid for arbitrary one-way functions f and g (we assume that $\forall x \in \{0,1\}^*$, $|f(x)| = |g(x)|$). That is, if the assertion is valid give a proof by reduction. If it is not, give a counterexample of one-way functions f and g such that the obtained function is not a one-way function.

1. Let $CXor(f) : \{0,1\}^* \mapsto \{0,1\}^*$ be the function defined by $CXor(f)(x) = \overline{f(x)}$, i.e. $CXor(f)$ is the function that applies the function f to the argument and then computes the bitwise complement of the result.
   If f is a one-way function then $CXor(f)$ is also a one-way function.

2. Let $BXor(f, g) : \{0,1\}^* \mapsto \{0,1\}^*$ be the function defined by $BXor(f, g) = f(x) \oplus \overline{g(x)}$.
   If f and g are one-way functions then $BXor(f, g)$ is also a one-way function.