Advanced Cryptography — Midterm Exam
Solution

Serge Vaudenay

13.5.2014

- duration: 3h00
- documents are allowed
- a pocket calculator is allowed
- communication devices are not allowed
- the exam invigilators will *not* answer any technical question during the exam
- readability and style of writing will be part of the grade

The exam grade follows a linear scale in which each question has the same weight.

1 Cryptosystem based on Matrices

We define a new cryptosystem. Let p be a large prime number. Let $a \in \mathbf{Z}_p$ and $b \in \mathbf{Z}_p^*$ be arbitrary such that $a^2 + b^2 \in \mathbf{Z}_p^*$. Let G be the matrix $G = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. The public parameters are given by (p, a, b). We let $x \in \mathbf{Z}$ be a secret key and let $Y = G^x$ be a public key. To encrypt $m \in \mathbf{Z}_p$, we pick a random integer r, compute $U = G^r$, $V = Y^r$, and $w = V_{1,1} + m \bmod p$ (where $V_{1,1}$ is the upper left coefficient of V). The ciphertext is the pair (U, w).

Q.1 Explain how to decrypt.

We note that $U^x = V$. So, $m = w - (U^x)_{1,1} \bmod p$.

Q.2 Let $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Show that the following properties are equivalent.

- $p \bmod 4 = 1$;
- there is an invertible 2×2 matrix P with coefficients in $\mathbf{Z}_p$ such that $P^{-1}JP$ is a diagonal matrix;
- there is an invertible 2×2 matrix P with coefficients in $\mathbf{Z}_p$ such that $P^{-1}GP$ is a diagonal matrix.

HINT: we recall that a 2×2 matrix is diagonalizable if and only if it has two different eigenvalues or it is already diagonalized.

We first explain the hint. In the linear algebra class, there was a theorem saying that whenever a $n \times n$ matrix has n pairwise different eigenvalues, then it is diagonalizable. If a 2×2 matrix is diagonalizable and its two eigenvalues are equal, then the matrix must be diagonal. So, non-diagonal and diagonalizable is equivalent to having two different eigenvalues for 2×2 matrices.
Let $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Since we can write $G = aI + bJ$ or $J = \frac{1}{b}(G - aI)$ and since $P^{-1}IP = I$, the facts that J or G is equivalent to a diagonal matrix are equivalent: if $J = PDP^{-1}$ with D diagonal, then $G = P(aI+bD)P^{-1}$ where $aI+bD$ is diagonal; and if $G = PDP^{-1}$ with D diagonal, then $J = P\frac{1}{b}(D - aI)P^{-1}$ where $\frac{1}{b}(D - aI)$ is diagonal.
Now, if J is equivalent to a diagonal matrix, then its characteristic polynomial must have roots in the field. This polynomial is $\begin{vmatrix} -x & 1 \\ -1 & -x \end{vmatrix} = x^2+1$. So, if J is equivalent to a diagonal matrix, then -1 must have a square root. This is equivalent to $(-1)^{\frac{p-1}{2}} = +1$ which is itself equivalent to $p \bmod 4 = 1$.
Conversely, if $p \bmod 4 = 1$, then $x^2 + 1$ has two square roots which are different. Since the characteristic polynomial has two different roots, the 2×2 matrix J is equivalent to a diagonal matrix.

Q.3 For $p \bmod 4 = 1$, show that the key recovery problem reduces to the discrete logarithm problem in $\mathbf{Z}_p^*$.

To solve $G^x = Y$, we first extract a square root θ of -1 then find eigenvectors of J to form a matrix P such that $P^{-1}JP$ is a diagonal matrix. Then, we have to solve $(P^{-1}GP)^x = P^{-1}YP$. By observing that $P^{-1}GP$ and $P^{-1}YP$ are diagonal matrices, this reduces to solving two equations of form $g^x = y$: we can solve $\lambda_1^{x_1} = y_1$ and $\lambda_2^{x_2} = y_2$ where λ_1 and λ_2 are the eigenvalues of G and y_1 and y_2 are the ones of Y. If λ_1 is a generator of $\mathbf{Z}_p^*$, then x_1 is unique modulo $p - 1$ so it must be the solution x. If λ_2 is a generator of $\mathbf{Z}_p^*$, then x_2 is unique modulo $p - 1$ so it must be the solution x. But if neither λ_1 and λ_2 are generators, then we must compute their order n_1 and n_2, respectively, by solving $\lambda_i = g^{\frac{p-1}{n_i}}$, then solve $x \equiv x_1 \pmod{n_1}$ and $x \equiv x_2 \pmod{n_2}$ by standard CRT tricks.

Q.4 For $p \bmod 4 = 3$, we define $\mathbf{K} = \mathbf{Z}_p[x]/(x^2 + 1)$, the field of $\mathbf{Z}_p$ extended with a root θ of $x^2 + 1$. By working with matrices with coefficients in $\mathbf{K}$, show that the key recovery problem reduces to the discrete logarithm problem in $\mathbf{K}^*$.
HINT: show that J is diagonalizable as a matrix with coefficients in $\mathbf{K}$.

Although -1 has no square root in $\mathbf{Z}_p$, it has the square roots θ and $-\theta$ in $\mathbf{K}$. By taking $P = \begin{pmatrix} 1 & 1 \\ \theta & -\theta \end{pmatrix}$ we have $P^{-1}JP = \begin{pmatrix} \theta & 0 \\ 0 & -\theta \end{pmatrix}$. Again, we observe that $P^{-1}GP = \begin{pmatrix} a+b\theta & 0 \\ 0 & a-b\theta \end{pmatrix}$ and $P^{-1}YP = \begin{pmatrix} a'+b'\theta & 0 \\ 0 & a'-b'\theta \end{pmatrix}$ for some a' and b'. So, we can just solve $(a + b\theta)^x = a + b'\theta$ and $(a - b\theta)^x = a - b'\theta$. This is a discrete logarithm problem in $\mathbf{K}^*$.

Q.5 In general, give a positive integer q such that G^q is the identity matrix.

As we have seen, the discrete logarithm problem (G, Y) reduces to a discrete logarithm problem in either $\mathbf{Z}_p^*$ or in $\mathbf{K}^*$. They have order $p-1$ and p^2-1 respectively. Since $p-1$ divides p^2-1, in general, $q = p^2-1$ is a multiple of the order of G.

2 Predicate Encryption

We define a new cryptographic primitive called *predicate encryption*. We consider a predicate P. A predicate encryption for P is defined by four algorithms:

$\mathsf{Setup}(1^\lambda) \to (\mathsf{pp}, \mathsf{msk})$**:** (probabilistic) given a security parameter λ, it generates a key pair where msk is the master key (secret) of the authority and pp is the public parameter, which is distributed to all participants.

$\mathsf{Keygen}(\mathsf{msk}, k) \to \mathsf{sk}$**:** (probabilistic) given a key k for the predicate $P(k, .)$, the authority generates a secret key sk for a participant Bob.

$\mathsf{Enc}(\mathsf{pp}, \mathsf{ind}, m) \to c$**:** (probabilistic) given a value ind called *index* and a message m, Alice generates a ciphertext c. Note that this is independent of k and Bob.

$\mathsf{Dec}(\mathsf{sk}, c)$**:** (deterministic) given the ciphertext and the secret sk, this decryption algorithm yields m if $P(k, \mathsf{ind})$ is true and $\perp$ otherwise. (I.e., this is the correctness property of the primitive.)

The security of this primitive specifies that from c, Bob (holding sk) does not learn m if $P(k, \mathsf{ind})$ is false.

Q.1 In *identity-based encryption* (IBE), there is an authority holding a master key, distributing some public parameters to everyone, and giving a secret key to each user. We want that if, e.g., Alice is offline and cannot connect to retrieve the public key of Bob to any public directory, she can still encrypt a message which can only be decrypted by Bob (and the authority). More precisely, we have
$\mathsf{IBE.Setup}(1^\lambda) \to (\mathsf{pp}, \mathsf{msk})$**:** generate the public parameters and the master key.
$\mathsf{IBE.Keygen}(\mathsf{msk}, \mathsf{id}) \to \mathsf{sk}$**:** given the identity of Bob, generate his secret key.
$\mathsf{IBE.Enc}(\mathsf{pp}, \mathsf{id}, m) \to c$**:** encrypt a message m for a given identity.
$\mathsf{IBE.Dec}(\mathsf{sk}, c) \to m$**:** decrypt the message given the correct secret key.
We want that a user who does not hold the correct sk learns nothing about m.
By well choosing a predicate, construct one IBE scheme with the above syntax based on predicate encryption.
Give an argument for the security.

An index and a key are just an identity: $\mathsf{ind} = k = \mathsf{id}$. *We set* $P(k, \mathsf{ind})$ *to* true *if and only if* $k = \mathsf{ind}$. *The algorithms for IBE match the ones of predicate encryption:* $\mathsf{IBE.Setup} = \mathsf{Setup}$, $\mathsf{IBE.Keygen} = \mathsf{Keygen}$, $\mathsf{IBE.Enc} = \mathsf{Enc}$, *and* $\mathsf{IBE.Dec} = \mathsf{Dec}$. *Clearly, the correctness property is satisfied.*
If sk *is not the correct key, since* Dec *returns* $\perp$*, the holder does not learn more than* $\perp$*. So, the confidentiality of* m *is preserved.*

Q.2 In *ciphertext-policy attribute-based encryption* (CP-ABE), there is an authority holding a master key, distributing some public parameters to everyone, and giving a secret key to each user. Each user has list $z = (z_1, \ldots, z_n)$ of attributes associated to a semantic. (E.g., if member of EPFL or not, if MSc student or faculty member or admin staff, if registered to the Advanced Cryptography class, etc.) Each message is encrypted with a formula φ. It can only be decrypted for holders of attributes satisfying the formula φ. (E.g., expressing that the message can only be decrypted by MSc students or admin staff of EPFL who registered to Advanced Cryptography.) More precisely, we have

$\mathsf{CPABE.Setup}(1^\lambda) \to (\mathsf{pp}, \mathsf{msk})$**:** generate the public parameters and the master key.
$\mathsf{CPABE.Keygen}(\mathsf{msk}, z) \to \mathsf{sk}$**:** given the attributes of Bob, generate his secret key.
$\mathsf{CPABE.Enc}(\mathsf{pp}, \varphi, m) \to c$**:** encrypt a message m for a formula φ.
$\mathsf{CPABE.Dec}(\mathsf{sk}, c)$**:** obtain m is $\varphi(z)$ holds and $\perp$ otherwise.
We want that a user who holds attributes not satisfying φ learns nothing about m.
By well choosing a predicate, construct one CP-ABE scheme with the above syntax based on predicate encryption.
Give an argument for the security.

An index ind *is a formula:* $\mathsf{ind} = \varphi$*. A key* k *is a set of attributed:* $k = z$*. We set* $P(z, \varphi)$ *to* true *if and only if* z *satisfies* φ*. The algorithms for CP-ABE match the ones of predicate encryption:* $\mathsf{CPABE.Setup} = \mathsf{Setup}$*,* $\mathsf{CPABE.Keygen} = \mathsf{Keygen}$*,* $\mathsf{CPABE.Enc} = \mathsf{Enc}$*, and* $\mathsf{CPABE.Dec} = \mathsf{Dec}$*. Clearly, the correctness property is satisfied.*
If $\varphi(z)$ *is not satisfied, since* Dec *returns* $\perp$*, the holder does not learn more than* $\perp$*. So, the confidentiality of* m *is preserved.*

Q.3 Given a modulus N and a length n, we assume that, when $\mathsf{ind} = (a_1, \ldots, a_n)$ and $k = (b_1, \ldots, b_n)$ are in $\mathbf{Z}_N^n$, we have a predicate encryption scheme for inner product (IP), i.e., for the predicate

$$P(k, \mathsf{ind}) \iff a_1 b_1 + \cdots + a_n b_n \equiv 0 \pmod{N}$$

We denote this scheme by $(\mathsf{IP.Setup}_N^n, \mathsf{IP.Keygen}_N^n, \mathsf{IP.Enc}_N^n, \mathsf{IP.Dec}_N^n)$.
Based on an IP scheme with the above syntax, construct one predicate encryption scheme where users are associated to a variable x and messages are encrypted with a polynomial f of bounded degree d, and the decryption works if x is a root of f. More precisely, construct a scheme $(\mathsf{POL.Setup}_N^d, \mathsf{POL.Keygen}_N^d, \mathsf{POL.Enc}_N^d, \mathsf{POL.Dec}_N^d)$ in which $\mathsf{POL.Keygen}_N^d(\mathsf{msk}, x)$ gives sk, $\mathsf{POL.Enc}_N^d(\mathsf{pp}, f, m)$ gives c, and $\mathsf{POL.Dec}_N^d(\mathsf{sk}, c) = m$ if and only if $f(x) = 0$.

We set $n = d + 1$*. We set* ind *to the coefficients of* $f(x)$ *and* $k = (1, x, x^2, \ldots, x^d)$*. We let* $\mathsf{POL.Setup}_N^d = \mathsf{IP.Setup}_N^n$*,* $\mathsf{POL.Keygen}_N^d(\mathsf{msk}, x) = \mathsf{IP.Keygen}_N^n(\mathsf{msk}, k)$*,* $\mathsf{POL.Enc}_N^d(\mathsf{pp}, f, m) = \mathsf{IP.Enc}_N^n(\mathsf{pp}, \mathsf{ind}, m)$*, and* $\mathsf{POL.Dec}_N^d = \mathsf{IP.Dec}_N^n$*.*

Q.4 We now give several variables to the participants. Extend the previous construction to multivariate polynomials.

We now set k *to the evaluation of all monomials of degree at most* d*. So,* $n = \binom{d+s-1}{s-1}$*. Then, we expand the polynomial and use the coefficients in same order in* ind*.*

Q.5 We consider a predicate $P(a, b)$ which is a CNF (conjunctive normal form) of terms of form $a_i = b_j$. I.e., $P(a, b) = \bigwedge_u \bigvee_v (a_{i_{u,v}} = b_{j_{u,v}})$. ($\vee$ is a notation for the OR and $\wedge$ is a notation for the AND.) Given some (secret) random r_u, we consider the polynomial $f(a, b) = \sum_u r_u \prod_v (a_{i_{u,v}} - b_{j_{u,v}})$.

Q.5a If $P(a, b)$ is true, show that $f(a, b) = 0$.

If $P(a,b)$ is true, for all u, there exists at least one v such that $a_{i_{u,v}} = b_{j_{u,v}}$. So, $\prod_v (a_{i_{u,v}} - b_{j_{u,v}}) = 0$ and we obtain $f(a,b) = 0$.

Q.5b Given a and b fixed such that $P(a,b)$ is false, show that $\Pr[f(a,b) = 0]$ is small, over the distribution of the r_u's.
HINT: assume that N is prime.

If $P(a,b)$ is false, there exists some u such that for all v, $a_{i_{u,v}} \neq b_{j_{u,v}}$. Let $s = \prod_v (a_{i_{u,v}} - b_{j_{u,v}})$. By isolating the term in the sum corresponding to the index u, we write $f(a,b) = r_u s + Z$ with Z independent from r_u. For N prime, $s \neq 0$ implies $s \in \mathbf{Z}_N^*$. So, $f(a,b) = 0$ is equivalent to $r_u = -\frac{Z}{s}$. We deduce $\Pr[f(a,b) = 0] = \frac{1}{N}$.

Q.5c From the IP scheme, show that we can construct a predicate encryption scheme for the predicate P. How large is n in the above construction?

We use the previous construction for multivariate polynomials with the polynomial f. The length n corresponds to the number of monomials. It is $n = \binom{d+s-1}{s-1}$ where d is the size of the largest clause (i.e., the maximal number of v's for the same u).

3 Distribution Fitting

Let p be a prime number and $\ell \leq \log_2 p$. Let $r = p \bmod 2^\ell$. Let $V \in_U \mathbf{Z}_p$ be uniformly distributed. Let $X = V \bmod 2^\ell$. We want to distinguish the distribution of X from the uniform distribution over $\{0, \ldots, 2^\ell - 1\}$.

Q.1 Compute the distribution of X: depending on $x \in \{0, \ldots, 2^\ell - 1\}$, provide a formula to compute $\Pr[X = x]$ in terms of x, r, p, ℓ.

Let $p = q2^\ell + r$ be the Euclidean division of p by 2^ℓ, i.e., $r = p \bmod 2^\ell$. For $x < r$, the set of preimages of x by $v \mapsto v \bmod 2^\ell$ is $\{x, 2^\ell + x, \ldots, q2^\ell + x\}$. So, $\Pr[X = x] = \frac{q+1}{p}$.
For $x \geq r$, the set of preimages of x by $v \mapsto v \bmod 2^\ell$ is $\{x, 2^\ell + x, \ldots, (q-1)2^\ell + x\}$. So, $\Pr[X = x] = \frac{q}{p}$.
By using $\frac{q}{p} = 2^{-\ell} - \frac{r}{p2^\ell}$, we obtain $\Pr[X = x] = 2^{-\ell} - \frac{r}{p2^\ell} + \frac{1}{p}1_{x<r}$.

Q.2 Given a single sample, compute the best advantage of a distinguisher to distinguish the distribution of X from a uniform one.

The best advantage is the statistical distance, i.e.

$$\mathsf{Adv} = \frac{1}{2} \sum_x \left| \Pr[X = x] - 2^{-\ell} \right|$$

So, we have

$$\begin{aligned} \mathsf{Adv} &= \frac{r}{2} \left| \frac{1}{p} - \frac{r}{p2^\ell} \right| + \frac{2^\ell - r}{2} \left| -\frac{r}{p2^\ell} \right| \\ &= \frac{r}{2} \left(\frac{1}{p} - \frac{r}{p2^\ell} \right) - \frac{2^\ell - r}{2} \times \frac{r}{p2^\ell} \\ &= \frac{r}{p}(1 - r2^{-\ell}) \end{aligned}$$

Q.3 We assume that p is selected arbitrarily among prime numbers of k bits. With ℓ fixed and in the worst case for p, how large should k be so that the best advantage given a single sample is lower than $2^{-\ell}$?

r is arbitrary between 0 and $2^\ell - 1$. In the worst case, $r(1 - r2^{-\ell})$ is reached for $r = \frac{1}{2}2^\ell$. So, we can use the bound $\mathsf{Adv} \leq \frac{2^\ell}{4p}$. Thus, we need $k \geq 2\ell - 2$ to ensure $\mathsf{Adv} \leq 2^{-\ell}$.

Q.4 Assume the identified condition is satisfied, approximate the Chernoff information by the squared Euclidean imbalance and estimate the number of samples needed to distinguish the distribution of X from a uniform one.

Since we made sure that the distribution of X is close to uniform, we can use

$$C(X,U) \approx \frac{2^\ell}{8\ln 2}\sum_x \left(\Pr[X=x]-2^{-\ell}\right)^2$$

We have

$$\begin{aligned} C(X,U) &\approx \frac{2^\ell}{8\ln 2}\sum_x \left(\Pr[X=x]-2^{-\ell}\right)^2 \\ &= \frac{2^\ell}{8\ln 2}\times r\left(\frac{1}{p}-\frac{r}{p2^\ell}\right)^2 + \frac{2^\ell}{8\ln 2}(2^\ell-r)\left(-\frac{r}{p2^\ell}\right)^2 \\ &= \frac{2^\ell}{8\ln 2}\times \frac{r}{p^2}\left(1-\frac{r}{2^\ell}\right) \end{aligned}$$

The worst case is for $r \approx \frac{1}{2}2^\ell$. The number of samples to distinguish X from a uniform distribution can be approximated to

$$\frac{1}{C(X,U)} \approx \frac{p^2}{2^{2\ell}} \times \mathsf{cte} \approx 2^{2(k-\ell)} \times \mathsf{cte}$$

for $p \approx 2^k$.