Advanced Cryptography — Midterm Exam
Solution

Serge Vaudenay

3.5.2011

I A Crazy Cryptosystem

We define a new RSA-like public-key cryptosystem.

- For key generation, we generate two different prime numbers p and q of $\ell + 1$ bits and larger than 2^ℓ, and make $N = pq$. Then, we pick a random α between 0 and $p - 1$ and compute $a = 1 + \alpha p$. The public key is (a, N) and the secret key is p.
- To encrypt a message x of at most ℓ bits, the sender computes $y = xa^r \bmod N$ for a random r.
- To decrypt y, the receiver computes $x = y \bmod p$.

Q.1 Give the complexity of the three algorithms. What is the advantage with respect to RSA?

Key generation takes $\mathcal{O}(\ell^4)$, like in RSA. Encryption takes $\mathcal{O}(\ell^3)$, the cost of the exponentiation, like in RSA. Decryption takes $\mathcal{O}(\ell^2)$, the cost of a modular reduction. Decryption is much faster than with RSA.

Q.2 Show that the correctness property of the cryptosystem is satisfied.

We have

$$y \bmod p = (xa^r \bmod N) \bmod p = (xa^r) \bmod p = x$$

since $a \bmod p = 1$ and $x < p$.

Q.3 Show that the decryption problem is as hard as the key recovery problem.

Assume we have a decryption oracle $\mathcal{O}$. We can pick a random y and send it to $\mathcal{O}$. It will answer by $x = y \bmod p$. Then, we observe that p divides $y - x$. Since y is random, so is $\frac{y-x}{p}$, and it is likely to be coprime with q. So, $\mathsf{gcd}(y - x, N) = p$ with high probability. Therefore, we can do a key recovery by using $\mathcal{O}$: decryption and key recovery are equivalent.

Q.4 Show that key recovery is easy.

$\mathsf{gcd}(a - 1, N) = p$.

II The DDH Problem and Bilinear Maps

We consider a (multiplicatively denoted) finite group $G = \langle g \rangle$ generated by some g element. We assume that there is a map e from $G \times G$ to some group H such that

- $\#G = \#H$;
- $h = e(g, g)$ generates H;
- for all $a, b, c \in G$, $e(ab, c) = e(a, c)e(b, c)$.
- for all $a, b, c \in G$, $e(a, bc) = e(a, b)e(a, c)$.

We call e a *bilinear map.*

Q.1 Show that for all integers x, y, we have $e(g^x, g^y) = h^{xy}$.

We first show by induction on x that $e(g^x, b) = e(g, b)^x$. For $x = 0$, since $a = 1.a$, we have that $e(a, b) = e(1, b)e(a, b)$, so $e(1, b) = 1$. Then, assuming it holds for $x-1$, since $g^x = g^{x-1}.g$, we have

$$e(g^x, b) = e(g^{x-1}, b)e(g, b) = e(g, b)^{x-1}e(g, b) = e(g, b)^x$$

So, we have $e(g^x, b) = e(g, b)^x$ for all $x \geq 0$. Since x is taken modulo the order of g, it holds for any integer x.
Then, we show that $e(a, g^y) = e(a, g)^y$ in the same way. We deduce that

$$e(g^x, g^y) = e(g^x, g)^y = (e(g, g)^x)^y = h^{xy}$$

Q.2 Recall what is the Decisional Diffie-Hellman (DDH) problem in group G.

We consider any algorithm $\mathcal{A}$ fed with (U, X, Y, K) and which yields 0 or 1. We let

$$\mathsf{Adv}(\mathcal{A})(s) = \Pr_{\mathsf{exp}_1}[\mathcal{A}(U, X, Y, K) = 1] \Pr_{\mathsf{exp}_0}[\mathcal{A}(U, X, Y, K) = 1]$$

where experiment exp_b consists of

- *generate $U \leftarrow \mathsf{Gen}(1^s)$*
- *generate X, Y, K uniformly in $\langle g \rangle$*
- *if $b = 1$, replace K by the solution of $\mathsf{DHP}(U, X, Y)$*

where $\mathsf{DHP}(U, U^x, U^y) = U^{xy}$.
The DDH problem consists of building a probabilistic polynomial-time algorithm $\mathcal{A}$ such that $\mathsf{Adv}(\mathcal{A})(s)$ is not negligible.

Q.3 Show that the DDH problem in G is easy to solve when it is easy to compute e.

We define $\mathcal{A}(U, X, Y, K) = 1_{e(U,K)=e(X,Y)}$. Clearly, we have $\Pr_{\mathsf{exp}_1}[\mathcal{A}(U, X, Y, K) = 1] = 1$. To evaluate $\Pr_{\mathsf{exp}_0}[\mathcal{A}(U, X, Y, K) = 1]$, we notice that $\mathcal{A}(g^u, g^x, g^y, g^k) = 1$ if and only if $uk = xy$, which shall occur with a probability of $1/\#G$. So, $\mathsf{Adv}(\mathcal{A})(s) = 1 - \frac{1}{\#G(s)}$ which is certainly not negligible.

Q.4 Show that if the Discrete Logarithm problem is easy in H, then it is easy in G as well.

We observe that $e(g, g^x) = h^x$. So, if we can extract x from h and h^x, then we can extract x from g and g^x by computing $h = e(g, g^x)$.

III Almost Bent Functions

The exercise is inspired by Links between differential and linear cryptanalysis *by Chabaud and Vaudenay. Published in the proceedings of Eurocrypt'94 pp. 356–365, LNCS vol. 950, Springer 1995.*

In this exercise, we consider a function f mapping n bits to n bits. We define two functions DP^f and LP^f mapping two strings of n bits to a real number by

$$\mathsf{DP}^f(a,b) = \Pr[f(X \oplus a) \oplus f(X) = b]$$
$$\mathsf{LP}^f(\alpha,\beta) = (2\Pr[\alpha \cdot X = \beta \cdot f(X)] - 1)^2$$

where X is uniformly distributed in $\{0,1\}^n$, $\oplus$ represents the bitwise exclusive-OR of two bitstrings, and $u \cdot v$ represents the parity of the bitwise AND of two bitstrings, i.e.

$$(u_1,\ldots,u_n)\cdot(v_1,\ldots,v_n) = (u_1v_1 + \cdots + u_nv_n) \bmod 2$$

In this problem, we define

$$\mathsf{DP}^f_{\max} = \max_{(a,b)\neq(0,0)} \mathsf{DP}^f(a,b)$$
$$\mathsf{LP}^f_{\max} = \max_{(\alpha,\beta)\neq(0,0)} \mathsf{LP}^f(\alpha,\beta)$$

Our purpose is to minimize $\mathsf{DP}^f_{\max}$ and $\mathsf{LP}^f_{\max}$. We recall that $\mathsf{DP}^f(a,b)$ and $\mathsf{LP}^f(\alpha,\beta)$ are always in the $[0,1]$ interval, that $\mathsf{DP}^f(0,b) \neq 0$ if and only if $b = 0$, that $\mathsf{LP}^f(\alpha,0) \neq 0$ if and only if $\alpha = 0$, and that for all a, $\sum_b \mathsf{DP}^f(a,b) = 1$. We further recall the two link formulas between DP^f and LP^f coming from the Fourier transform:

$$\mathsf{DP}^f(a,b) = 2^{-n}\sum_{\alpha,\beta}(-1)^{(a\cdot\alpha)\oplus(b\cdot\beta)}\mathsf{LP}^f(\alpha,\beta)$$
$$\mathsf{LP}^f(\alpha,\beta) = 2^{-n}\sum_{a,b}(-1)^{(a\cdot\alpha)\oplus(b\cdot\beta)}\mathsf{DP}^f(a,b)$$

Part 1: Preliminaries

Q.1a Show that for all β, $\sum_\alpha \mathsf{LP}^f(\alpha,\beta) = 1$.

We have

$$\sum_\alpha \mathsf{LP}^f(\alpha,\beta) = \sum_\alpha 2^{-n}\sum_{a,b}(-1)^{(a\cdot\alpha)\oplus(b\cdot\beta)}\mathsf{DP}^f(a,b)$$
$$= 2^{-n}\sum_{a,b}\mathsf{DP}^f(a,b)(-1)^{b\cdot\beta}\sum_\alpha(-1)^{a\cdot\alpha}$$

but the inner sum is nonzero only for $a = 0$, in which case it is 2^n, so

$$\sum_\alpha \mathsf{LP}^f(\alpha,\beta) = \sum_b \mathsf{DP}^f(0,b)(-1)^{b\cdot\beta}$$

Now, $\mathsf{DP}^f(0,b)$ is nonzero only for $b = 0$, so

$$\sum_\alpha \mathsf{LP}^f(\alpha,\beta) = 1$$

Q.1b Show that $\sum_{a,b}(\mathsf{DP}^f(a,b))^2 = \sum_{\alpha,\beta}(\mathsf{LP}^f(\alpha,\beta))^2$.

Hint$_1$: $\sum_x \left(\sum_y g(x,y)\right)^2 = \sum_{x,y,z} g(x,y)g(x,z)$. Do not be afraid of big sums!

Hint$_2$: remember your other classes on the Fourier transform.

We apply again the link formula. We have

$$\begin{aligned}\sum_{a,b}\left(\mathsf{DP}^f(a,b)\right)^2 &= \sum_{a,b}\left(2^{-n}\sum_{\alpha,\beta}(-1)^{(a\cdot\alpha)\oplus(b\cdot\beta)}\mathsf{LP}^f(\alpha,\beta)\right)^2 \\ &= 2^{-2n}\sum_{a,b}\sum_{\alpha,\beta,\gamma,\delta}(-1)^{(a\cdot\alpha)\oplus(b\cdot\beta)}\mathsf{LP}^f(\alpha,\beta)(-1)^{(a\cdot\gamma)\oplus(b\cdot\delta)}\mathsf{LP}^f(\gamma,\delta) \\ &= 2^{-2n}\sum_{\alpha,\beta,\gamma,\delta}\mathsf{LP}^f(\alpha,\beta)\mathsf{LP}^f(\gamma,\delta)\sum_{a,b}(-1)^{(a\cdot(\alpha\oplus\gamma))\oplus(b\cdot(\beta\oplus\delta))}\end{aligned}$$

where the inner sum is nonzero only for $\alpha=\gamma$ and $\beta=\delta$, in which case it is 2^{2n}, so

$$\sum_{a,b}\left(\mathsf{DP}^f(a,b)\right)^2 = \sum_{\alpha,\beta}\left(\mathsf{LP}^f(\alpha,\beta)\right)^2$$

Part 2: APN functions

Q.2a Show that $\mathsf{DP}^f_{\max} \geq 2^{1-n}$. In the case of an equality, we say that f is *Almost Perfect Nonlinear (APN)*.

Hint: First show that $2^n\mathsf{DP}^f(a,b)$ is an even integer.

$\mathsf{DP}^f(a,b)$ *is 2^{-n} times the number of x's such that $f(x\oplus a)\oplus f(x)=b$. When x satisfies this property, so does $x\oplus a$. Hence, the number of x's is even. Therefore, $\mathsf{DP}^f(a,b)$ is an even number divided by 2^n. Since $\sum_b \mathsf{DP}^f(a,b)=1$, we can take any $a\neq 0$ and we deduce that there is at least one b such that $\mathsf{DP}^f(a,b)\neq 0$. So, $\mathsf{DP}^f(a,b)\geq 2^{1-n}$ with $a\neq 0$ from which we deduce $\mathsf{DP}^f_{\max}\geq 2^{1-n}$.*

Q.2b Show that f is an APN function if and only if for all a and b such that $(a,b)\neq(0,0)$, we have either $\mathsf{DP}^f(a,b)=2^{1-n}$ or $\mathsf{DP}^f(a,b)=0$.

Since $\mathsf{DP}^f(a,b)$ is an even integer divided by 2^n and bounded by 2^{1-n}, it can only be 2^{1-n} or 0. The converse is trivial.

Part 3: AB functions

Q.3a Show that $\sum_\alpha\sum_{\beta\neq 0}\left(\mathsf{LP}^f(\alpha,\beta)\right)^2 \geq 2^{1-n}(2^n-1)$.

Hint: use Q.1b and observe that $(\mathsf{DP}^f(a,b))^2 \geq 2^{1-n}\mathsf{DP}^f(a,b)$

In Q.1b, we have proven that

$$\sum_{\alpha,\beta}\left(\mathsf{LP}^f(\alpha,\beta)\right)^2=\sum_{a,b}\left(\mathsf{DP}^f(a,b)\right)^2$$

So,

$$\sum_{\alpha}\sum_{\beta\neq 0}\left(\mathsf{LP}^f(\alpha,\beta)\right)^2+\sum_{\alpha}\left(\mathsf{LP}^f(\alpha,0)\right)^2=\sum_{a\neq 0}\sum_{b}\left(\mathsf{DP}^f(a,b)\right)^2+\sum_{b}\left(\mathsf{DP}^f(0,b)\right)^2$$

which leads to

$$\sum_{\alpha}\sum_{\beta\neq 0}\left(\mathsf{LP}^f(\alpha,\beta)\right)^2=\sum_{a\neq 0}\sum_{b}\left(\mathsf{DP}^f(a,b)\right)^2$$

Then, since $(\mathsf{DP}^f(a,b))^2\geq 2^{1-n}\mathsf{DP}^f(a,b)$, *we obtain*

$$\sum_{\alpha}\sum_{\beta\neq 0}\left(\mathsf{LP}^f(\alpha,\beta)\right)^2\geq 2^{1-n}\sum_{a\neq 0}\sum_{b}\mathsf{DP}^f(a,b)=2^{1-n}(2^n-1)$$

Q.3b Show that $\mathsf{LP}^f_{\max}\geq\frac{\sum_\alpha\sum_{\beta\neq 0}\left(\mathsf{LP}^f(\alpha,\beta)\right)^2}{\sum_\alpha\sum_{\beta\neq 0}\mathsf{LP}^f(\alpha,\beta)}$ with equality if and only if for all α,β with $\beta\neq 0$, we have either $\mathsf{LP}^f(\alpha,\beta)=0$ or $\mathsf{LP}^f(\alpha,\beta)=\mathsf{LP}^f_{\max}$.

This is equivalent to show that $\sum_\alpha\sum_{\beta\neq 0}\mathsf{LP}^f(\alpha,\beta)\left(\mathsf{LP}^f_{\max}-\mathsf{LP}^f(\alpha,\beta)\right)\geq 0$ *with equality if and only if all terms in the sum are zero. Since all terms are positive, this is trivial.*

Q.3c Show that $\mathsf{LP}^f_{\max}\geq 2^{1-n}$. In the case of an equality, we say that f is *Almost Bent (AB).*

The previous inequality in Q.3b together with the results of Q.3a and Q.1a turns into $\mathsf{LP}^f_{\max}\geq 2^{1-n}$.

Q.3d Show that f is an AB function if and only if for all α and β such that $(\alpha,\beta)\neq(0,0)$, we have either $\mathsf{LP}^f(\alpha,\beta)=2^{1-n}$ or $\mathsf{LP}^f(\alpha,\beta)=0$.

If f is AB, then we have an equality case in the inequality of Q.3b. This leads to the result. The other direction is trivial.

Q.3e Show that if f is an AB function, then it is APN as well.

If f is AB, then $\sum_\alpha\sum_{\beta\neq 0}\left(\mathsf{LP}^f(\alpha,\beta)\right)^2=2^{1-n}(2^n-1)$. *So, thanks to Q.1b,* $\sum_{a\neq 0}\sum_b\left(\mathsf{DP}^f(a,b)\right)^2=2^{1-n}(2^n-1)$. *Just like in Q.3b, we have* $\mathsf{DP}^f_{\max}\geq\frac{\sum_{a\neq 0}\sum_b\left(\mathsf{DP}^f(a,b)\right)^2}{\sum_{a\neq 0}\sum_b\mathsf{DP}^f(a,b)}$ *which is equal to* 2^{1-n}. *So, f is APN.*

IV Analyzing Two-Time Pad

We consider the Vernam cipher defined by $\mathsf{Enc}_K(X) = x \oplus K$, where the plaintext X and the key K are two bitstrings of length n, independent random variables, and K is uniformly distributed. We assume that X comes from a biased source with a given distribution. The purpose of this exercise is to analyze the information loss when we encrypt two random plaintexts X and Y with the same key K. We assume that X, Y, and K are independent random variables, that X and Y are identically distributed, and that K is uniformly distributed.

Part 1: Preliminaries

Q.1a Show that for all x and y, $\Pr[\mathsf{Enc}_K(X) = x, \mathsf{Enc}_K(Y) = y] = 2^{-n} \Pr[X \oplus Y = x \oplus y]$.

We have $\Pr[\mathsf{Enc}_K(X) = x, \mathsf{Enc}_K(Y) = y] = \Pr[\mathsf{Enc}_K(X) \oplus \mathsf{Enc}_K(Y) = y, \mathsf{Enc}_K(X) = x] = \Pr[X \oplus Y = x \oplus y, K = x \oplus X]$. *Then,*

$$\begin{aligned}
\Pr[X \oplus Y = x \oplus y, K = x \oplus X] &= \sum_a \Pr[X \oplus Y = x \oplus y, K = x \oplus a, X = a] \\
&= \sum_a \Pr[X \oplus Y = x \oplus y, X = a] \Pr[K = x \oplus a] \\
&= 2^{-n} \sum_a \Pr[X \oplus Y = x \oplus y, X = a] \\
&= 2^{-n} \Pr[\Pr[X \oplus Y = x \oplus y]
\end{aligned}$$

since K is independent from (X, Y) and uniformly distributed.

Q.1b Deduce that the statistical distance between $(\mathsf{Enc}_K(X), \mathsf{Enc}_K(Y))$ and a uniformly distributed $2n$-bit string is the same as the statistical distance between $X \oplus Y$ and a uniformly distributed n-bit string.

We have

$$\begin{aligned}
d &= \frac{1}{2} \sum_{x,y} \left| \Pr[\mathsf{Enc}_K(X) = x, \mathsf{Enc}_K(Y) = y] - 2^{-2n} \right| \\
&= \frac{1}{2} \sum_{x,y} 2^{-n} \left| \Pr[X \oplus Y = x \oplus y] - 2^{-n} \right| \\
&= \frac{1}{2} \sum_{x,\delta} 2^{-n} \left| \Pr[X \oplus Y = \delta] - 2^{-n} \right| \\
&= \frac{1}{2} \sum_{\delta} \left| \Pr[X \oplus Y = \delta] - 2^{-n} \right|
\end{aligned}$$

which is the statistical distance between $X \oplus Y$ and a uniformly distributed random variable.

Q.1c Further show that this is similar for the Euclidean distance.

For the Euclidean distance, we have

$$\begin{aligned}
&\sum_{x,y}\left(\Pr[\mathsf{Enc}_K(X)=x,\mathsf{Enc}_K(Y)=y]-2^{-2n}\right)^2\\
&=\sum_{x,y}2^{-2n}\left(\Pr[X\oplus Y=x\oplus y]-2^{-n}\right)^2\\
&=\sum_{x,\delta}2^{-2n}\left(\Pr[X\oplus Y=\delta]-2^{-n}\right)^2\\
&=2^{-n}\sum_{\delta}\left(\Pr[X\oplus Y=\delta]-2^{-n}\right)^2
\end{aligned}$$

So the two Euclidean distances have a constant ratio (of 2^{-n}, for the squared Euclidean distance).

Part 2: Best distinguisher with a single sample

Q.2a What is the best advantage to distinguish $(\mathsf{Enc}_K(X),\mathsf{Enc}_K(Y))$ from a uniformly distributed $2n$-bit string using a single sample?

The best advantage is the statistical distance

$$\begin{aligned}
d&=\frac{1}{2}\sum_{x,y}\left|\Pr[\mathsf{Enc}_K(X)=x,\mathsf{Enc}_K(Y)=y]-2^{-2n}\right|\\
&=\frac{1}{2}\sum_{\delta}\left|\Pr[X\oplus Y=\delta]-2^{-n}\right|
\end{aligned}$$

Q.2b As an application, assume that X consists of a uniformly distributed random string of $n-1$ bits followed by a parity bit, i.e. a bit set to 1 if and only if there is an odd number of 1's amount the $n-1$ other bits. Describe an optimal distinguisher with a single query and compute its advantage.

Due to the parity bit, we have that $\Pr[X\oplus Y=\delta]=2^{1-n}$ when the δ is even, and $\Pr[X\oplus Y=\delta]=0$ otherwise. So,

$$d=\frac{1}{2}2^{n-1}(2^{1-n}-2^{-n})+\frac{1}{2}2^{n-1}\times 2^{-n}=\frac{1}{2}$$

Given $X\oplus Y$, the distinguisher simply outputs the parity of $X\oplus Y$.

Part 3: Best distinguisher with many samples

Q.3a How many samples do we need (roughly) to distinguish $(\mathsf{Enc}_K(X),\mathsf{Enc}_K(Y))$ from a uniformly distributed $2n$-bit string with a good advantage?

The rough number of samples is $N=1/C$ where C is the Chernoff information between $(\mathsf{Enc}_K(X),\mathsf{Enc}_K(Y))$ and the uniform string.

Q.3b Approximate this in terms of squared Euclidean distance.

The Chernoff information is approximated with the help of the Euclidean distance by

$$\begin{aligned} C &\approx \frac{2^{2n}}{8\ln 2} \sum_{x,y} \left(\Pr[\mathsf{Enc}_K(X) = x, \mathsf{Enc}_K(Y) = y] - 2^{-2n}\right)^2 \\ &= \frac{2^n}{8\ln 2} \sum_{\delta} \left(\Pr[X \oplus Y = \delta] - 2^{-n}\right)^2 \end{aligned}$$

This holds when $\Pr[\mathsf{Enc}_K(X) = x, \mathsf{Enc}_K(Y) = y]$ *is always close to* 2^{-2n}.