Advanced Cryptography — Midterm Exam

Serge Vaudenay

3.5.2011

- duration: 3h30
- any document allowed
- a pocket calculator is allowed
- communication devices are not allowed
- readability and style of writing will be part of the grade
- it is unlikely we will answer any technical question during the exam
- do not forget to put your full name on your copy!

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?
Q.2 Show that the correctness property of the cryptosystem is satisfied.
Q.3 Show that the decryption problem is as hard as the key recovery problem.
Q.4 Show that key recovery is easy.

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 $e(g^x, g^y) = h^{xy}$.
Q.2 Recall what is the Decisional Diffie-Hellman (DDH) problem in group G.
Q.3 Show that the DDH problem in G is easy to solve when it is easy to compute e.
Q.4 Show that if the Discrete Logarithm problem is easy in H, then it is easy in G as well.

III Almost Bent Functions

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

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.

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.

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 $\mathsf{DP}^f(a,b) = 0$.

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 $(\mathsf{DP}^f(a,b))^2 \geq 2^{1-n}\mathsf{DP}^f(a,b)$

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 $\mathsf{LP}^f(\alpha,\beta) = \mathsf{LP}^f_{\max}$.

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)*.

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 $\mathsf{LP}^f(\alpha,\beta) = 0$.

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

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, $\Pr[\mathsf{Enc}_K(X) = x, \mathsf{Enc}_K(Y) = y] = 2^{-n} \Pr[X \oplus Y = x \oplus y]$.
- **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.
- **Q.1c** Further show that this is similar for the 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?
- **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.

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