Advanced Cryptography — Midterm Exam

Serge Vaudenay

22.4.2009

1 RSA Public-Key Recovery

Given an integer e and a few (x_i, y_i) pairs such that $y_i = x_i^e \bmod N$ for some unknown common N of known bit-length ℓ, we consider the problem of recovering N. We assume that $0 \leq x_i, y_i < N$ and that i ranges from 1 to n.

1. Using Buffon's needle problem we can show that the probability that two independent uniformly distributed integers in $\{0, 1, \ldots, 2^\ell - 1\}$ are coprime tends towards $\frac{6}{\pi^2}$ as ℓ goes to infinity. We take independent uniformly distributed integers $X_1, \ldots, X_n$ in $\{0, 1, \ldots, 2^\ell - 1\}$. Show that the probability that $\mathsf{gcd}(X_1, \ldots, X_n) > 1$ is less than $\left(1 - \frac{6}{\pi^2}\right)^{\frac{n}{2}}$ as ℓ goes to infinity.
   Hint: consider $\frac{n}{2}$ disjoint pairs of form (X_{2i-1}, X_{2i}).
2. We now take iid random integers $X_1, \ldots, X_n$ in $\{0, 1, \ldots, 2^\ell - 1\}$ which are uniformly distributed among the multiples of N. Show that $\mathsf{gcd}(X_1, \ldots, X_n) = N$ except with negligible probability as n increases.
3. Deduce that we can recover N by computing $\mathsf{gcd}(x_1^e - y_1, \ldots, x_n^e - y_n)$. What is its complexity in terms of ℓ, e, and n?

2 DP and LP Tricks

Consider a function f from $A = \{0, 1\}^p$ to $B = \{0, 1\}^q$. We define DP^f and LP^f as functions from $A \times B$ to $\mathbf{R}$ as usual by

$$\mathsf{DP}^f(a, b) = \Pr_X[f(a \oplus X) \oplus f(X) = b]$$

$$\mathsf{LP}^f(a, b) = \left(2 \Pr_X[a \cdot X = b \cdot f(X)] - 1\right)^2$$

1. Show that for any $b \neq 0$ we have $\mathsf{DP}^f(0, b) = 0$. Give a necessary and sufficient condition about f so that
   $$\forall a \neq 0 \quad \mathsf{DP}^f(a, 0) = 0$$
2. Show that for any $a \neq 0$ we have $\mathsf{LP}^f(a, 0) = 0$.
3. We define a function g from B to $\mathbf{R}$ by $g(y) = \Pr[f(X) = y]$ for all $y \in B$ where X is uniformly distributed in A. Show that for any function h we have
   $$E(h(f(X))) = E(g(Y)h(Y))$$
   where Y is uniformly distributed in B.

4. Deduce that
$$\mathsf{LP}^f(0,b) = \left(E\left(g(Y)(-1)^{b\cdot Y}\right)\right)^2$$
where Y is uniformly distributed in B.
5. Show that
$$g(y) = 2^{-q}\sum_{b\in B}(-1)^{b\cdot y}E\left((-1)^{b\cdot f(X)}\right)$$
where X is uniformly distributed in A.
6. Deduce that
$$\forall b \neq 0 \quad \mathsf{LP}^f(0,b) = 0$$
if and only if $g(y) = 2^{-q}$ for all $y \in B$.
7. Deduce that
$$\forall b \neq 0 \quad \mathsf{LP}^f(0,b) = 0$$
if and only if f is balanced, i.e. all elements in B are equally taken as images by f.

3 Applied Crypto-polymorphism

The CONFIKER worm is permanently updating itself by looking for updates over the Internet. Once it has found the update, it checks if the update code has a correct RSA signature with modulus N and public exponent e. One problem is that the value of N in the code of the worm is large enough to be used by anti-virus software to detect the presence of the worm. The worm conceptor attended to a lecture on cryptography and would like to obfuscate N using cryptographic tricks.

1. Recall how the RSA signature verification works for a message m with signature σ. (Assume for example PKCS#1v1.5 with deterministic formatting rules for m.)
2. Once the worm installs, it picks a random prime number p, computes $N' = pN$ and discards p and N. The value of N' remains in the worm code. Show that a signature σ of an update code m can still be verified using e and N' instead of e and N.
   Can an anti-virus software detect the presence of the RSA key?
3. Assume that the anti-virus software conceptor has analyzed the code of the worm on two independent infected machines and extracted N'_1 and N'_2. Show that he can deduce the value of N.
   With the value of N, show that we can still detect the presence of the worm based on the value of N' in the code. (Assume that N' can easily be extracted from the code.)

4 Distinguishing Sources

We consider a source producing iid random variables $X_i \in \{0,1,\ldots,2^\ell - 1\}$ for $i = 1,\ldots,q$. For this, we consider two distributions:

- the uniform distribution P_0
- the distribution P_1 induced by $X_i = Y_i \bmod 2^\ell$ where Y_i is uniformly distributed in $\{0,1,\ldots,p-1\}$ and $p > 2^\ell$. (Note that P_0 can be considered as a particular case of P_1 with $p = 2^\ell$.)

We assume that ℓ is large, e.g. $\ell \geq 80$ and we let $r = p \bmod 2^\ell$.

1. Given $x \in \{0, 1, \ldots, 2^\ell - 1\}$, show that

$$P_1(x) = \begin{cases} \left(1 - \frac{r}{p}\right) 2^{-\ell} + \frac{1}{p} & \text{if } x < r \\ \left(1 - \frac{r}{p}\right) 2^{-\ell} & \text{if } x \geq r. \end{cases}$$

2. Describe a distinguisher using $q = 1$ which achieves the optimal advantage.
3. For $q = 1$, what is the best advantage for distinguishing P_0 from P_1? Express it as a formula in terms of ℓ, p, and r.
4. Deduce that for $p \leq c2^\ell$ with c small and $r2^{-\ell}$ neither too small nor too close to 1, then P_0 and P_1 can be distinguished using a single sample.
5. Describe a distinguisher using an arbitrarily fixed q which achieves the optimal advantage.
6. Compute the squared Euclidean distance between P_0 and P_1.
7. Assuming that P_1 is close to P_0, approximate the Chernoff information between P_0 and P_1. Deduce that $C(P_0, P_1) \leq \frac{2^\ell}{2p \ln 2}$ whatever r.
8. Deduce that for p larger than $2^{2\ell}$ the two distributions are indistinguishable in practice.