TD 4 Advanced cryptography 2024

Léo Colisson Palais

Exercice 1: Security of Regev's encryption

Solve the questions from the slides, i.e. show that for appropriately chosen parameters, the LWE-based encryption seen in the course is secure assuming the hardness of LWE.

Exercice 2: Worst case to average case reduction

One may be worried that the LWE problem may be hard in the worst case ("there exists a few hard instances"), but that it is still insecure on average ("if I pick a random LWE instance, it is easy to break on average"). In this exercise, we will show that if the LWE problem is hard in the worst case, it is also hard on average. Or said differently, if we can break LWE on average, we can break *all* LWE instances. In the following, $A_{\mathbf{s},\chi}$ will denote the LWE distribution sampling $\mathbf{a} \xleftarrow{\$} \mathbb{Z}_q^n$, $e \leftarrow \chi$ and returning $(\mathbf{a}, \mathbf{a}^T\mathbf{s} + e)$, where n and q are integers and χ is a distribution on $\mathbb{Z}_q$. For any set S, U_S will denote the uniform distribution on S, and we define $U \coloneqq U_{\mathbb{Z}_q^n \times \mathbb{Z}_q}$.

1. For any $\mathbf{t} \in \mathbb{Z}_q^n$, let $f_\mathbf{t} \colon \mathbb{Z}_q^n \times \mathbb{Z}_q \to \mathbb{Z}_q^n \times \mathbb{Z}_q$ be defined as $f_t(\mathbf{a}, b) \coloneqq (\mathbf{a}, b + \mathbf{a}^T\mathbf{t})$. For any $\mathbf{t}$ and $\mathbf{s}$, show that $f_\mathbf{t}(A_{\mathbf{s},\chi}) = A_{\mathbf{s}+\mathbf{t},\chi}$, i.e. that the distribution obtained by applying $f_\mathbf{t}$ to a sample of $A_{\mathbf{s},\chi}$ is statistically indistinguishable from a sample of $A_{\mathbf{s}+\mathbf{t},\chi}$.

2. Show that the image of the uniform distribution on $\mathbb{Z}_q^n \times \mathbb{Z}_q$ by $f_\mathbf{t}$ is the uniform distribution on $\mathbb{Z}_q^n \times \mathbb{Z}_q$, i.e. $f_\mathbf{t}(U_{\mathbb{Z}_q^n \times \mathbb{Z}_q}) = U_{\mathbb{Z}_q^n \times \mathbb{Z}_q}$.

3. We assume that there exists a polynomial-time algorithm W that distinguishes $U_{\mathbb{Z}_q^n \times \mathbb{Z}_q}$ from $A_{\mathbf{s},\chi}$ on average, i.e. such that there exists a set $S \subseteq \mathbb{Z}_q^n$ of elements where W guesses reasonably correctly, i.e. such that for any $s \in S$ and large enough n,

$$\left| \Pr_{x \leftarrow A_{\mathbf{s},\chi}} [\, W(x) = 1 \,] - \Pr_{x \xleftarrow{\$} \mathbb{Z}_q^n \times \mathbb{Z}_q} [\, W(x) = 1 \,] \right| \geq \frac{1}{n^c} \tag{1}$$

   for some integer $c > 1$, and such that S covers a non-negligible fraction of $\mathbb{Z}_q^n \times \mathbb{Z}_q$, i.e. there exists some integer $c' > 1$ such that, for large enough n:

$$\frac{|S|}{q^{n+1}} \geq \frac{1}{n^{c'}} \tag{2}$$

   Our goal is to use W to build an adversary able to break LWE for all instances.

   (a) Without loss of generality, we can assume that W outputs either 0 or 1. Show that

$$\Pr_{x \leftarrow D} [\, W(x) = 1 \,] = \mathop{\mathbb{E}}_{x \leftarrow D} [\, W(x) \,] \tag{3}$$

   (b) First, show that for any distribution D, one can efficiently estimate the quantity $\Pr_{x \leftarrow D} [\, W(x) = 1 \,]$, i.e. that there exists a procedure $\mathsf{Estimate}(D)$ running in time polynomial in n, such that the probability to have:

$$\left| \mathsf{Estimate}(D) - \Pr_{x \leftarrow D} [\, W(x) = 1 \,] \right| \geq \frac{1}{10n^c} \tag{4}$$

   is negligible in n.

   Hint: For this, you may want to use the inequality of Hoeffding, that (in particular) states that if N independent random variables $V_1, \ldots, V_N$ are bounded by 0 and 1, then for any $t \geq 0$,

$$\Pr\left[\left| \sum_i V_i - \mathbb{E}\left[\sum_i V_i \right] \right| \geq t \right] \leq 2 \exp\left(-\frac{2t^2}{N} \right) \tag{5}$$

(c) Show that with overwhelming[1] probability:

$$|\mathsf{Estimate}(U) - \mathsf{Estimate}(U)| \leq \frac{2}{10n^c} \tag{6}$$

and explain why this does not trivially simplify to 0.

(d) Show that for any $\mathbf{s} \in S$, with overwhelming probability:

$$|\mathsf{Estimate}(A_{\mathbf{s},\chi}) - \mathsf{Estimate}(U)| \geq \frac{8}{10n^c} \tag{7}$$

(e) We define now the algorithm $W'(D)$ where D is a distribution on $\mathbb{Z}_q^n \times \mathbb{Z}_q$ given as a black-box oracle to W', that will internally calls W and try to guess if D is the uniform distribution or a LWE distribution. More precisely, $W'(D)$ will repeat M times (M being a polynomial in n to be determined later) the following procedure: it will pick a random $\mathbf{t} \xleftarrow{\$} \mathbb{Z}_q$, compute $|\mathsf{Estimate}(f_{\mathbf{t}}(D)) - \mathsf{Estimate}(U)|$: if this value is greater than $\frac{1}{2n^c}$, it will output "LWE", otherwise it continues the loop. If at the end it has not returned before, it will output "Uniform".

i. Show that W' is correct when the input is a uniform distribution and when $M \geq 1$, i.e. with overwhelming probability, $W'(U) =$ "Uniform".

ii. Let $\mathbf{s} \in \mathbb{Z}_q^n$. Show that the probability that $W'(A_{\mathbf{s},\chi})$ returns "LWE" after 1 iteration of its inner loop is lower bounded by $\frac{8}{10n^c}$.

iii. Show that if we choose M large enough (but polynomial), W' is correct (with overwhelming probability) when the input is any LWE distribution, i.e. **for any** $\mathbf{s} \in \mathbb{Z}_q$, with overwhelming probability on the randomness of W', $W'(A_{\mathbf{s},\chi}) =$ "LWE". What value can we choose for M?
Hint: you may need to use the fact that $1 + x \leq e^x$.

iv. Conclude by showing that W' runs in polynomial time and that it therefore solves efficiently the LWE problem in the worst case by computing its advantage.

[1] I.e. the probability of not having this equation true is negligible over the randomness involved in Estimate.