

Contents

4.1	Negligible function	3
4.2	One-Way Functions	4
4.2.1	Impagliazzo's Worlds	5
4.3	Computational Indistinguishability	6
4.4	Pseudorandom Generators	8
4.5	Stretching a PRG	9
4.6	Hardcore predicate	12
4.7	One Way Permutation	13
4.8	Computationally secure encryption	14
4.9	Pseudorandom functions	17
4.9.1	GGM Tree	18
4.10	CPA SECURITY	21
4.11	Domain extension	25
4.11.1	Electronic CodeBook (ECB)	25
4.11.2	Cipher block chaining (CBC)	26
4.12	Message Authentication and UFCMA-security	30
4.13	Domain extension	32
4.13.1	\mathcal{H} family using Galois Fields	36
4.14	Domain extension for PRFs/MACs	38
4.14.1	XOR mode	38
4.14.2	CBC MAC	39
4.14.3	XOR MAC	39
4.15	Chosen Ciphertext Attack security	41
4.16	Authenticated encryption	42
4.16.1	Three approaches to authenticated encryption	43
4.17	Authenticated encryption (Age of Ultron)	45
4.18	Pseudorandom permutations	46
4.18.1	Feistel Network	47
4.19	Public key encryption recap	48
4.19.1	Trapdoor permutation	49
4.19.2	TDP examples	49
4.20	Textbook RSA	50
4.20.1	Trapdoor Permutation from Factoring	51
4.20.2	Rabin's Trapdoor permutation	51
4.21	PKE schemes over DDH assumption	54
4.21.1	El Gamal scheme	54
4.21.2	Cramer-Shoup PKE scheme	57
4.22	Construction of a CCA-secure PKE	60
4.22.1	Instantiation of U-HPS (Universal Hash Proof System)	64

4.23	Digital signatures	67
4.23.1	Public Key Infrastructure	68
4.23.2	Waters Signature	70

Lesson 4

4.1 Negligible function

What is exactly a negligible function? Below here there is a possible interpretation of this notion:

“In real life, we can just consider adversaries with limited computational power; even if every non-perfectly secure authentication scheme can resist to unbounded computational power, the true unbounded computational power doesn’t exist at all. So, it’s reasonable to consider just bounded adversaries.

So consider a scheme Π where the only attack against it is brute-force attack. We consider Π to be secure if it cannot be broken by a brute-force attack in polynomial time.

The idea of **negligible probability** encompasses this exact notion. In Π , let’s say that we have a polynomial-bounded adversary. Brute force attack is not an option.

But instead of brute force, the adversary can try (a polynomial number of) random values and hope to guess the right one. In this case, we define security using negligible functions: The probability of success has to be smaller than the reciprocal of any polynomial function.

And this makes a lot of sense: if the success probability for an individual guess is a reciprocal of a polynomial function, then the adversary can try a polynomial amount of guesses and succeed with high probability. If the overall success rate is $\frac{1}{poly(\lambda)}$ then we consider this attempt a feasible attack to the scheme, which makes the latter insecure.

So, we require that the success probability must be less than the reciprocal of every polynomial function. This way, even if the adversary tries $poly(\lambda)$ guesses, it will not be significant since it will only have tried: $\frac{poly(\lambda)}{superpoly(\lambda)} \cdot 1$

As λ grows, the denominator grows far faster than the numerator and the success probability will not be significant.²”

¹If we design a function hard for $superpoly(\lambda)$ possible attempts and the attacker completed $poly(\lambda)$ attempts, he has just $\mathcal{P}[\frac{poly(\lambda)}{superpoly(\lambda)}]$ of finding the key to break the scheme

²<https://crypto.stackexchange.com/questions/5832/what-exactly-is-a-negligible-and-non-negligible-function>

Definition

Let $\nu : \mathbb{N} \rightarrow [0, 1]$ be a function. Then it is deemed **negligible** iff:

$$\forall p(\lambda) \in \text{poly}(\lambda) \implies \nu(\lambda) \in o\left(\frac{1}{p(\lambda)}\right)$$

Exercise 1. Let $p(\lambda), p'(\lambda) \in \text{poly}(\lambda)$ and $\nu(\lambda), \nu'(\lambda) \in \text{negl}(\lambda)$. Then prove the following:

1. $p(\lambda) \cdot p'(\lambda) \in \text{poly}(\lambda)$
2. $\nu(\lambda) + \nu'(\lambda) \in \text{negl}(\lambda)$

Solution 1 (1.1). T0D0 1: Questa soluzione usa disuguaglianze deboli; per essere negligibile una funzione dev'essere strettamente minore di un polinomiale inverso. Da approfondire

We need to show that for any $c \in \mathbb{N}$, then there is n_0 such that $\forall n > n_0 \implies h(n) \leq n^{-c}$.

So, consider an arbitrary $c \in \mathbb{N}$. Then, since $c + 1 \in \mathbb{N}$, and both f and g are negligible, there exists n_f and n_g such that:

$$\begin{aligned} \forall n \geq n_f &\implies f(n) \leq n^{-(c+1)} \\ \forall n \geq n_g &\implies g(n) \leq n^{-(c+1)} \end{aligned}$$

Fix $n_0 = \max(n_f, n_g)$. Then, since $n \geq n_0 \geq 2$, $\forall n \geq n_0$ we have:

$$\begin{aligned} h(n) &= f(n) + g(n) \\ &\leq n^{-(c+1)} + n^{-(c+1)} \\ &= 2n^{-(c+1)} \\ &\leq n^{-c} \end{aligned}$$

Thus we conclude $h(n)$ is negligible.

4.2 One-Way Functions

A One-Way Function (OWF) is a function that is “easy to compute”, but “hard to invert”.

Definition 1. Let $f : \{0, 1\}^{n(\lambda)} \rightarrow \{0, 1\}^{n(\lambda)}$ be a function. Then it is a OWF iff:

$$\forall \text{PPT } \mathcal{A} \exists \nu(\lambda) \in \text{negl}(\lambda) : \Pr [\text{GAME}_{f, \mathcal{A}}^{\text{OWF}}(\lambda) = 1] \leq \nu(\lambda) \quad (4.1)$$

◇

Do note that the game does not look for the equality $x' = x$, but rather for the equality of their images computed by f .

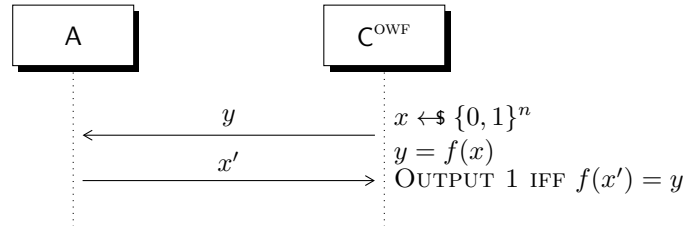


Figure 4.1: One-Way Function hardness

Exercise 2.

1. Show that there exists an inefficient adversary that wins GAME^{OWF} with probability 1
2. Show that there exists an efficient adversary that wins GAME^{OWF} with probability 2^{-n}

One-Way puzzles

A one-way function can be thought as a function which is very efficient in generating puzzles that are very hard to solve. Furthermore, the person generating the puzzle can efficiently verify the validity of a solution when one is given.

For a given couple $(\mathcal{P}_{\text{GEN}}, \mathcal{P}_{\text{VER}})$ of a puzzle generator and a puzzle verifier, we have:

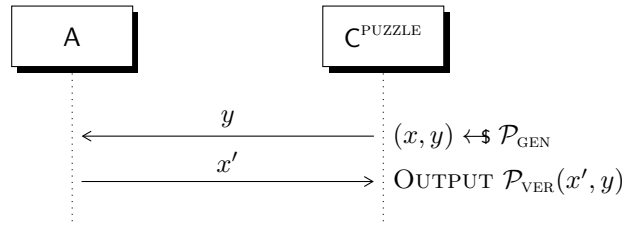


Figure 4.2: The puzzle game

So, we can say that one-way puzzle is a problem in NP, because the verifier enables efficient solution verification, but not in P because finding a solution from scratch is impractical by definition.

4.2.1 Impagliazzo's Worlds

Suppose to have Gauss, a genius child, and his professor. The professor gives to Gauss some mathematical problems, and Gauss wants to solve them all.

Imagine now that, if using one-way functions, the problem is $f(x)$, and its solution is x . According to Impagliazzo, we live in one of these possible worlds:

- *Algorithmica*: $P = NP$, meaning all efficiently verifiable problems are also efficiently solvable.

The professor can try as hard as possible to create a hard scheme, but he won't succeed because Gauss will always be able to efficiently break it using the verification procedure to compute the solution

- *Heuristica*: NP problems are hard to solve in the worst case but easy on average.

The professor, with some effort, can create a game difficult enough, but Gauss will solve it anyway; here there are some problems that the professor cannot find a solution to

- *Pessiland*: NP problems are hard on average but no one-way functions exist
- *Minicrypt*: One-way functions exist but public-key cryptography is impractical
- *Cryptomania*: Public-key cryptography is possible: two parties can exchange secret messages over open channels

4.3 Computational Indistinguishability

Distribution ensembles $X = \{X_{\lambda \in \mathbb{N}}\}$ and $Y = \{Y_{\lambda \in \mathbb{N}}\}$ are distribution sequences.

Definition 2. X and Y are **computationally indistinguishable** ($X \approx_c Y$) if \forall PPT $D, \exists \nu(\lambda) \in \text{negl}(\lambda)$ such that:

$$|\Pr[D(X_{\lambda}) = 1] - \Pr[D(Y_{\lambda}) = 1]| \leq \nu(\lambda)$$

◇

T0D0 2: AP181129-2344: There may be room for improvement, but I like how it's worded: it puts some unusual perspective into the cryptographic game, and it could be a good thing since it closely precedes our first reduction, and the whole hybrid argument mish-mash.

Suppose we have this mental game: a Distinguisher D receives the value z . This value has been chosen by me, the Challenger, among X_{λ} and Y_{λ} , and the Distinguisher has to *distinguish* which was the source of z . What does this formula mean?

This formula means that, fixed 1 as one of the sources, the *probability* that D says “1!” when I pick z from X_{λ} is **not so far** from the *probability* that D says “1!” when I pick z from Y_{λ} .

So, this means that, when this property is verified by two random variables, there isn't too much *difference* between the two variables in terms of exposed information (reachable by D), otherwise the distance between the two probabilities should be much more than a *negligible* quantity.

What's the deep meaning of this formula? This is something to do.

Lemma 1. \forall PPT $f, X \approx_c Y \implies f(x) \approx_c f(y)$ ◇

Proof. We want to show that $f(x) \approx_c f(y)$. So, let's suppose this property is not true.

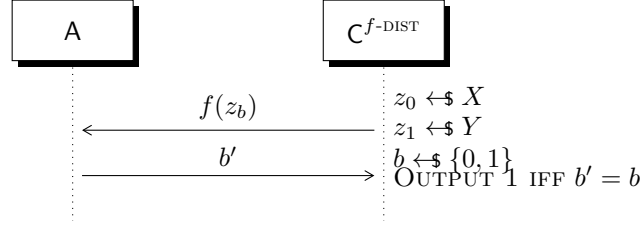


Figure 4.3: A distinguisher for f

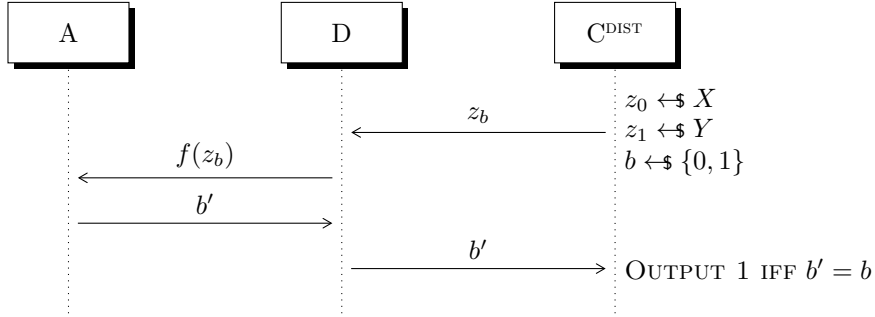


Figure 4.4: Distinguisher reduction

Assume \exists PPT f, A and some $p'(\lambda) \in \text{poly}(\lambda)$ such that

$$|\Pr[D'(f(x)) = 1] - \Pr[D'(f(y)) = 1]| \geq \frac{1}{p'(\lambda)} \quad (4.2)$$

In the game depicted by figure 4.3, the adversary can make use of f , because it is PPT. Thus, if such a distinguisher exists, then it can be used to distinguish X from Y , as shown in figure 4.4. So, if A could efficiently distinguish between $f(x)$ and $f(y)$, then D can efficiently distinguish between x and y . This contradicts the hypothesis $X \approx_c Y$, which in turn means $f(X) \approx_c f(Y)$. \square

4.4 Pseudorandom Generators

A deterministic function $G : \{0,1\}^\lambda \rightarrow \{0,1\}^{\lambda+l(\lambda)}$ is called a PseudoRandom Generator, or PRG iff:

- G runs in polynomial time
- $|G(s)| = \lambda + l(\lambda)$
- $G(U_\lambda) \approx_c U_{\lambda+l(\lambda)}$

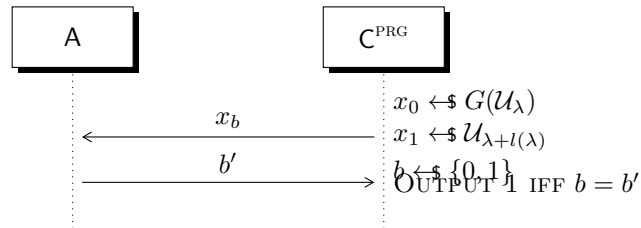


Figure 4.5: The pseudorandom game

So, if we take $s \leftarrow \mathcal{U}_\lambda$, the output of G will be indistinguishable from a random draw from \mathcal{U}_λ .

Lesson 5

4.5 Stretching a PRG

Consider this algorithm that uses G to construct G^l , as depicted in figure 4.6:

1. Let $s_0 \leftarrow \{0, 1\}^\lambda$
2. $\forall i \in [l(\lambda)]$, let $G(s_{i-1}) = (s_i, b_i)$, where b_i is the extra bit generated by G
3. Compose $(b_1, b_2, \dots, b_{l(\lambda)}, s_{l(\lambda)})$ and output the result; the outputted string will be $\lambda + l(\lambda)$ bits long

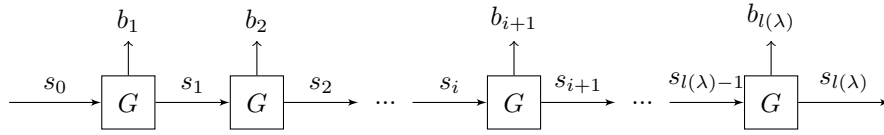


Figure 4.6: Constructing $G^{l(\lambda)}$ from $G(\lambda)$

To prove that this construct is a valid PRG, we will make use of a known technique for proving many other results, which relies heavily on reductions, and is called **the hybrid argument**.

Lemma 2. (Hybrid Argument) Let $X = \{X_n\}, Y = \{Y_n\}, Z = \{Z_n\} : X \approx_c Y \wedge Y \approx_c Z$. Then $X \approx_c Z$. \diamond

Proof. \forall PPT D , by using the triangle inequality:

$$\begin{aligned}
 & |\Pr[D(X_n) = 1] - \Pr[D(Z_n) = 1]| \\
 &= |(\Pr[D(X_n) = 1] - \Pr[D(Y_n) = 1]) + (\Pr[D(Y_n) = 1] - \Pr[D(Z_n) = 1])| \\
 &\leq |\Pr[D(X_n) = 1] - \Pr[D(Y_n) = 1]| + |\Pr[D(Y_n) = 1] - \Pr[D(Z_n) = 1]| \\
 &\leq \text{negl}(n) + \text{negl}(n) \\
 &= \text{negl}(n) \quad \square
 \end{aligned}$$

In essence, the hybrid argument proves that computational indistinguishability is transitive across a “hybrid” game, or by extension more than one. This property will be very useful in all future proofs.

Theorem 3. If there exists a PRG $G(\lambda)$ with one bit stretch, then there exists a PRG $G^{l(\lambda)}$ with polynomial stretch relative to its input length:

$$G : \{0, 1\}^\lambda \rightarrow \{0, 1\}^{\lambda+1} \implies \forall l(\lambda) \in \text{poly} \lambda \exists G^l : \{0, 1\}^\lambda \rightarrow \{0, 1\}^{\lambda+l(\lambda)} \quad \diamond$$

Proof. First off, do observe that, since both G and l are polynomial in λ , then $G^{l(\lambda)}$, which combines G l -many times is polynomial in λ too. To prove that $G^{l(\lambda)}$ is indeed a PRG, we will apply the hybrid argument. The hybrids are defined as:

- $H_\lambda^0 := G^{l(\lambda)}(U_\lambda)$, which is the original construct
- $H_\lambda^i := \begin{cases} b_1, \dots, b_i \leftarrow \{0, 1\} \\ s_i \leftarrow \{0, 1\}^{\lambda+i} \\ (b_{i+1}, \dots, b_{l(\lambda)}, s_{l(\lambda)}) := G^{l(\lambda)-i}(s_i) \end{cases}$
- $H_\lambda^{l(\lambda)} := U_{\lambda+l}$

Focusing on two subsequent generic hybrids, as shown in figures 4.7 and 4.8, it can be observed that the only difference between the two resides in how b_{i+1} is generated: in H^i it comes from an instance of G , whereas in H^{i+1} is chosen at random. H_λ^0 is the starting point where all bits are pseudorandom, which coincides with the $G^{l(\lambda)}$, and $H_\lambda^{l(\lambda)}$ will generate a totally random string.

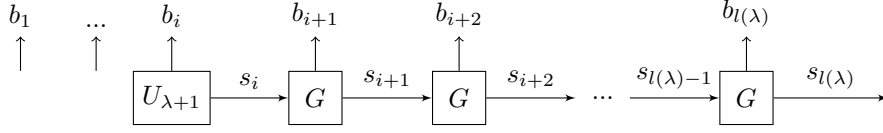


Figure 4.7: H_λ^i

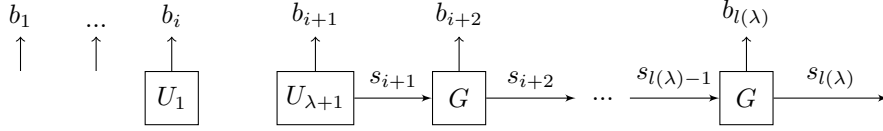


Figure 4.8: H_λ^{i+1}

So let's fix a step i in the gradual substitution, and consider the following function f_i :

$$f_i(s_{i+1}, b_{i+1}) = (b_1, \dots, b_i, b_{i+1}, b_{i+2}, \dots, b_{l(\lambda)}, s_{l(\lambda)})$$

where the first i bits are chosen uniformly at random, and the remaining ones are obtained by subsequent applications of G ($\forall j \in \{i+2, \dots, l(\lambda)\} \implies G(s_{j-1}) = (s_j, b_j)$). It can be observed that:

- $f_i(G(U_\lambda))$ has the same distribution of H_λ^i
- $f_i(U_{\lambda+1})$ has the same distribution of H_λ^{i+1}

Since by PRG definition $G(U_\lambda) \approx_c U_{\lambda+1}$, by lemma 1, we can deduce that $f_i(G(U_\lambda)) \approx_c f_i(U_{\lambda+1})$, which in turn, by how f is defined, implies $H^i \approx_c H^{i+1}$. This holds for an arbitrary choice of i , so by extension:

$$G^{l(\lambda)}(U_\lambda) = H_\lambda^0 \approx_c H_\lambda^1 \approx_c \dots \approx_c H_\lambda^{l(\lambda)} = U_{\lambda+l(\lambda)}$$

which proves that G^l is indeed a PRG. \square

Proof. (Contradiction): This is an alternate proof that, instead of looking for a function f to model hybrid transitioning, aims for a contradiction.

Suppose G^l is not a PRG; then there must be a point in the hybrid chain $H_\lambda^0 \approx_c \dots \approx_c H_\lambda^l$ where $H_\lambda^i \not\approx_c H_\lambda^{i+1}$. Thus there exists a distinguisher D' able to tell apart H_λ^i from H_λ^{i+1} , as shown in figure 4.9:

$$\exists i \in [0, l], \exists \text{PPT } D' : |\Pr[D'(H_\lambda^i) = 1] - \Pr[D'(H_\lambda^{i+1}) = 1]| \geq \frac{1}{\text{poly}(\lambda)}$$

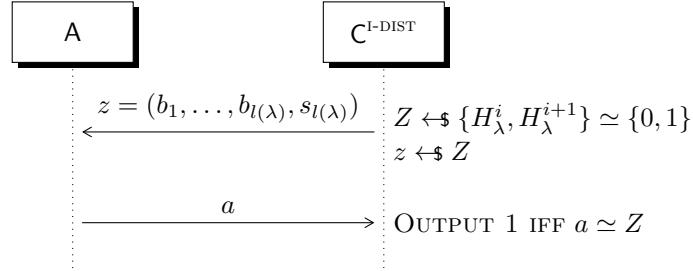


Figure 4.9: Distinguisher for H_λ^i and H_λ^{i+1}

If such a distinguisher exists, it can be also used to distinguish an output of G from a $\lambda + 1$ uniform string by “crafting” a suitable bit sequence, which will distribute exactly as the hybrids in question, as shown in the reduction in figure 4.10. This contradicts the hypothesis of f being a PRG, which by definition is to be indistinguishable from a truly random distribution. Therefore, G^l is indeed a PRG.

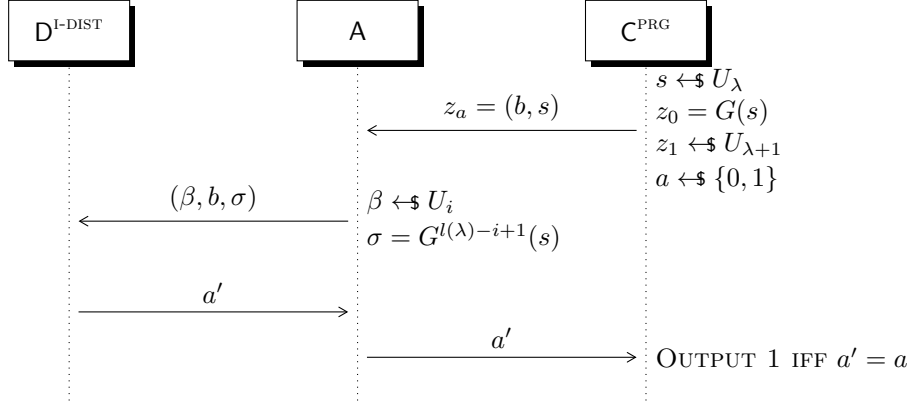


Figure 4.10: Reducing to a distinguisher for G , where $\beta = (b_1, \dots, b_{i-1})$ and $\sigma = (b_{i+1}, \dots, b_{l(\lambda)}, s_{l(\lambda)})$

□

4.6 Hardcore predicate

We've seen how to reuse a PRG in order to obtain an arbitrary length of pseudorandom bits starting from a PRG with one bit stretch. How to construct a 1-bit PRG?

Consider a typical one-way function f , s.t. $f(x) = y$.

Question 1. Which bits of the input x are hard to compute given $y = f(x)$?

Is it always true that, given f , the first bit of $f(x)$ is hard to compute $\forall x$?

Example 1. Given an OWF f , then $f'(x) = x_0 || f(x)$ is a OWF.

Definition 1. A polynomial time function $\mathsf{hc} : \{0, 1\}^n \rightarrow \{0, 1\}$ is **hard core** for a given function $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$ if, \forall PPT A :

$$\Pr(A(f(x)) = \mathsf{hc}(x) | x \leftarrow \{0, 1\}^n) \in \text{negl}(\lambda)$$

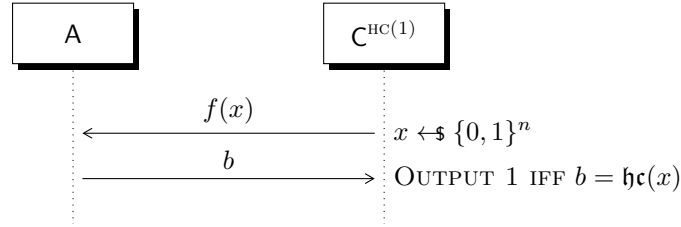


Figure 4.11: The hardcore game, f and hc are known

There is also an alternative definition:

Definition 2. A polynomial time function $\mathsf{hc} : \{0, 1\}^n \rightarrow \{0, 1\}$ is hard-core for a function f iff:

$$(f(x), h(x)) \approx_c (f(x), b)$$

where $x \leftarrow \{0, 1\}^n$ and $b \leftarrow \{0, 1\}$.

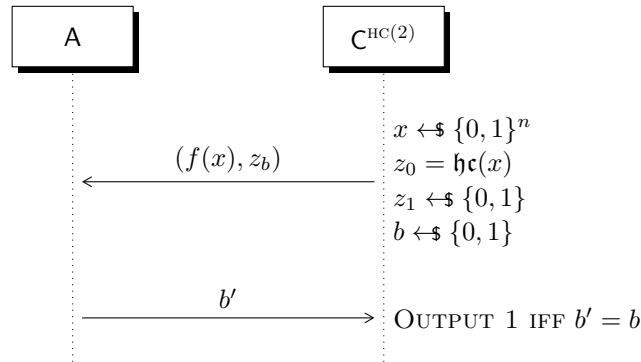


Figure 4.12: Another hardcore game, f and hc are known

Having made this definition, some observations are in order:

Claim 1. There is no *universal* hardcore predicate hc .

Proof. Suppose there exists such a predicate $\mathfrak{H}\mathfrak{C}$. Let $f'(x) = \mathfrak{H}\mathfrak{C}(x) || f(x)$ for a given function f . Then $\mathfrak{H}\mathfrak{C}$ cannot be a hardcore predicate of f' , because any image obtained by f reveals the predicate image itself. This contradicts the universality assumption. \square

Nevertheless, it has been proven always possible to construct a hardcore predicate for a OWF, given another OWF:

Theorem 4 (Goldreich-Levin, '99). *Let f be a OWF and consider $g(x, r) = (f(x), r)$ for $r \in \{0, 1\}^n$. Then g is a OWF and:*

$$h(x, r) = \langle x, r \rangle = \sum_{i=1}^n x_i r_i \mod 2 = \bigoplus_{i=1}^n x_i r_i$$

is hard core for g . \diamond

Proof. T0D0 3: Eh.... did we do this? Hope not...

\square

Exercise 5. Prove that g is a OWF if f is a OWF (Hint: do a reduction).

4.7 One Way Permutation

$f : \{0, 1\}^n \rightarrow \{0, 1\}^n$ is an OWF and

$$\forall x, |x| = |f(x)| \wedge x \neq x' \Rightarrow f(x) \neq f(x')$$

Corollary 1. If $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$ is a OWP, then for $g(), h()$ as in the GL theorem,

$$G(s) = (g(s), h(s))$$

is a PRG.

Proof. By GL, if f is an OWP, so is g . This means that if we want to invert g , since g depends on f we have to invert a OWP.

Moreover h is hardcore for g . Hence

$$G(U_{2n}) \equiv (g(U_{2n}), h(U_{2n})) \equiv \underbrace{(f(U_n), U_n, h(U_{2n}))}_{\text{definition 1 of hard core pred.}} \approx_c (f(U_n), U_n, U_1) \equiv U_{2n+1}$$

\square

We are stretching just 1 bit, but we know we can stretch more than one.

Lesson 6

4.8 Computationally secure encryption

Question 2. How to define the concept of **computationally secure encryption** ?

Find a task/scheme that is computationally hard for an attacker to break (supposing the attacker is $\text{poly}\lambda$, we want a scheme which requires an amount of time near ,as much as possible, to *superpoly*(λ) to be broken).

This scheme should have these properties:

- **one wayness** w.r.t. key (given $c = \text{Enc}(k, m)$, it should be hard to recover k)
- **one wayness** w.r.t. message (given $c = \text{Enc}(k, m)$, hard to obtain the message)
- **no information leakage** about the message

Consider the following experiment for $\Pi = (\text{Enc}, \text{Dec})$, named

$$GAME_{\Pi, \mathcal{A}}^{\text{ind}}(\lambda, b)$$

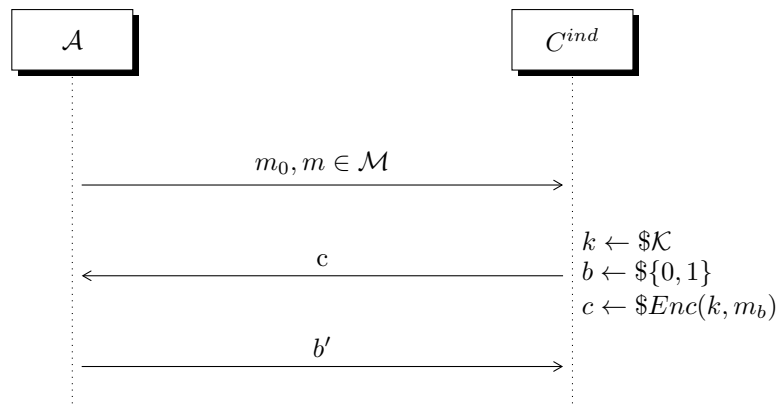


Figure 4.13

In the image 4.13 b' means that if $m_{b'} = m_b$ the adversary wins.

Definition 3. We say that Π is computationally **one time secure** if

$$Game_{\Pi, \mathcal{A}}^{ind}(\lambda, 0) \approx_c Game_{\Pi, \mathcal{A}}^{ind}(\lambda, 1)$$

or, alternatively $\forall .PPT.\mathcal{A} \exists \nu(\lambda) \in \text{negl}(\lambda)$

$$|\mathcal{P}[Game_{\Pi, \mathcal{A}}^{ind}(\lambda, 0) = 1] - \mathcal{P}[Game_{\Pi, \mathcal{A}}^{ind}(\lambda, 1) = 1]| \leq \nu(\lambda)$$

3

◇

This last definition is compliant with the three properties before exposed, in particular:

if a scheme is **one time secure** \Rightarrow the scheme has each one of these 3 properties

- **compliance with point 1** : suppose point 1 is not valid, and k is not hard to discover for \mathcal{A} . But then \mathcal{A} is able to perfectly distinguish m and m_0 with $\mathcal{P}[1]$ every time, and so the scheme couldn't be one time secure;
- **compliance with point 2** : suppose point 2 is not valid, and then the encrypted message can be easily discovered by \mathcal{A} . But then , as before, \mathcal{A} can win every game with $\mathcal{P}[1]$, so the scheme couldn't be one time secure;
- **compliance with point 3** : suppose point 3 is not valid, and some information about m is leaked in c , for example the first bit of c is the same bit of m . \mathcal{A} could forge $m_0 == m$ such that they have the same bits but just the first is different. When \mathcal{A} obtains c , he can look at the first bit and distinguish which was the message encrypted. Thus, the scheme wouldn't be one time secure. TO BE REVIEWED.

What is not **two time secure** ?

Construction 1. Let $G : \{0, 1\}^\lambda \rightarrow \{0, 1\}^l$. Consider the following schema Π_\oplus :

- $\mathcal{K} = \{0, 1\}^\lambda \Rightarrow k \leftarrow \mathcal{K}$
- $Enc(k, m) = G(k) \oplus m, m \in \{0, 1\}^l$
- $Dec(k, c) = c \oplus G(k) = m$

◇

This construction isn't 2-time secure. Assume the pair

$$(\bar{m}, \bar{c} = G(k) \oplus \bar{m})$$

is known. Now , given $c = G(k) \oplus m$, where c and m are unknown, we can force the schema and do the following

$$\bar{c} = G(k) \oplus m = c \oplus m \oplus \bar{m} \Rightarrow c \oplus \bar{c} = m \oplus \bar{m}$$

and obtain m .

Theorem 6. If G is a PRG, then Π_\oplus is computationally one-time secure ◇

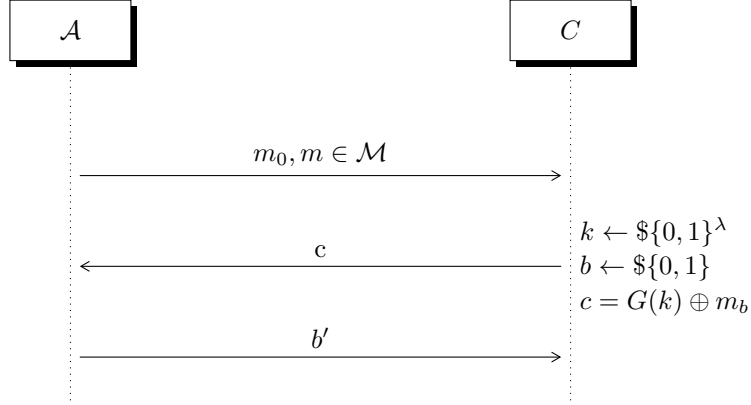


Figure 4.14: Game for Π_{\oplus} schema ($Game_{\Pi_{\oplus}, \mathcal{A}}$)

Proof. We need to show that

$$Game_{\Pi_{\oplus}, \mathcal{A}}^{ind}(\lambda, 0) \approx_c Game_{\Pi_{\oplus}, \mathcal{A}}^{ind}(\lambda, 1)$$

So first consider the following **hybrid game** :

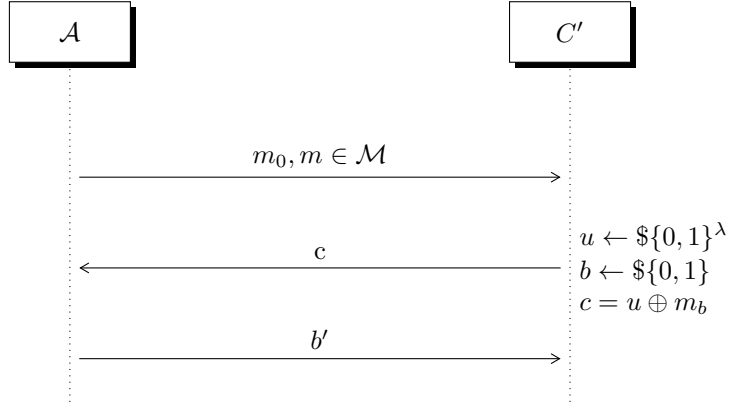


Figure 4.15: Hybrid game ($\mathcal{HYB}_{\Pi_{\oplus}, \mathcal{A}}(\lambda, b)$)

Lemma 3. $\mathcal{HYB}_{\Pi_{\oplus}, \mathcal{A}}(\lambda, 0) \equiv \mathcal{HYB}_{\Pi_{\oplus}, \mathcal{A}}(\lambda, 1)$ \diamond

This is true because distribution of c doesn't depend on $b \in \{0, 1\}$.

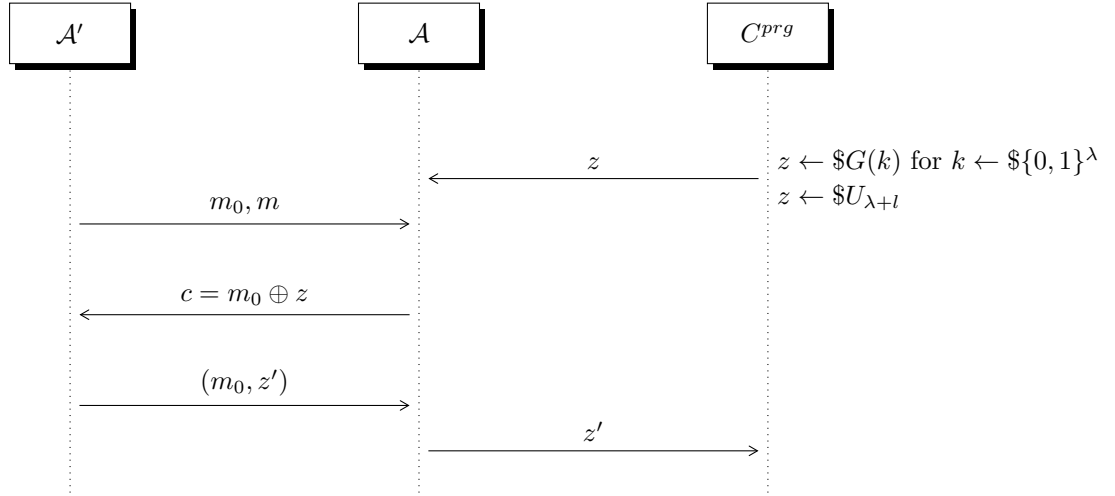
Lemma 4. $\forall b \in \{0, 1\}, \mathcal{HYB}_{\Pi_{\oplus}, \mathcal{A}}(\lambda, b) \approx_c Game_{\Pi, \mathcal{A}}^{ind}(\lambda, b)$ \diamond

Proof. Simple reduction to PRG, supposing that the statement isn't true.

This means that there exists \mathcal{A}' capable of distinguish $c = m_b \oplus G(k)$ and $c = m_b \oplus u$.

We will prove this first with $b = 0$ (the other case is the same).

³ $Game^{ind}$ refers to the indistinguishability of the messages sent by the attacker during the game



If \mathcal{A}' could distinguish these two sources, then C^{prg} could be distinguished, but this is impossible. \square

Now, for the two lemmas just seen, we have

$$Game_{\Pi, \mathcal{A}}^{ind}(\lambda, 0) \approx_c \mathcal{HYB}_{\Pi \oplus, \mathcal{A}}(\lambda, 0) \equiv \mathcal{HYB}_{\Pi \oplus, \mathcal{A}}(\lambda, 1) \approx_c Game_{\Pi, \mathcal{A}}^{ind}(\lambda, 1)$$

\square

4.9 Pseudorandom functions

A random function

$$R : \{0, 1\}^n \rightarrow \{0, 1\}^l$$

. is a function that takes in input x and :

- returns a new $R(x) = y \leftarrow \{0, 1\}^l$ if x has never been saw before and records that value (so **two distinct inputs can collide**)
- returns the recorded $R(x)$ otherwise

. We could generate these functions, but they occupy too much space: supposing all the possible outputs of R have been generated and stored in an array in memory, the occupied bits in memory are $2^n l$.



In particular, the family $\mathcal{R} = \{R : \{0, 1\}^n \rightarrow \{0, 1\}^l\}$, also indicated as $\mathcal{R}(\lambda, n, l)$, containing all the possible random functions has cardinality $2^{2^n l}$.

Intuition: a pseudo-random function is indistinguishable (computationally speaking) from a truly random one.

Call $\mathcal{F} = \{F_k : \{0, 1\}^{n(\lambda)} \rightarrow \{0, 1\}^{l(\lambda)}\}_{k \in \{0, 1\}^\lambda}$ the family of pseudorandom functions with key k . To give a definition of pseudo-random function, consider the following games:

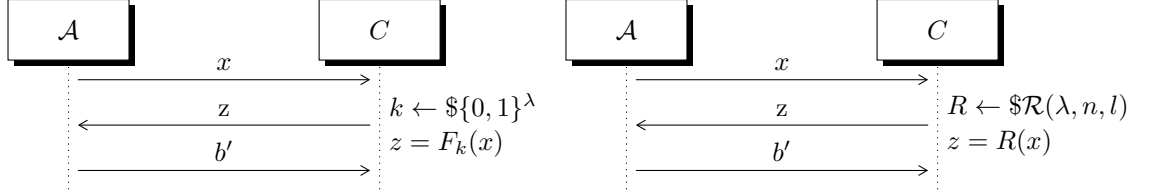


Figure 4.16: $Real_{\mathcal{F}, \mathcal{A}}(\lambda)$ vs $Rand_{\mathcal{R}, \mathcal{A}}(\lambda)$

where $b' \in \{0, 1\}$ is a convention and 1 is assigned to *Real* or *Rand*; so in this game the adversary **recognizes** which machine he is talking with.

Definition 4. \mathcal{F} is a PRF family if

$$Real_{\mathcal{F}, \mathcal{A}}(\lambda) \approx_c Rand_{\mathcal{R}, \mathcal{A}}(\lambda)$$

◇

Exercise 7. Show that no PRG is secure against **unbounded attackers**.

Exercise 8. Show the same (as above) for PRF.

4.9.1 GGM Tree

Construction 2. Let $G : \{0, 1\}^\lambda \rightarrow \{0, 1\}^{2\lambda}$ be a PRG and let us write

$$G(k) = (G_0(k), G_1(k))$$

Now consider this tree, called **GGM tree**, which describes the use of $G(k)$:

GGM TREE IMAGE

Build $\mathcal{F} = \{F_k : \{0, 1\}^n \rightarrow \{0, 1\}^\lambda\}$ such that

$$F_k(x) = G_{x_n}(G_{x_{n-1}} \dots G_{x_2}(G_{x_1}(k)))$$

◇

For example, in the tree with height $n = 3$, for $x = 001$ we have $F_k(001)$, which is $G_0(G_0(G_1(k)))$.

Lesson 7

WARNING : I was absent

Theorem 9. *If G is a PRG, then F_{GGM} is a PRF.* \diamond

(this definition should mean implicitly that the k key has been chosen by the challenger before the proof starts)

Proof. TO BE REVIEWED. Now use the induction on the height n of the GGM tree for $\nu(\lambda) \in \text{poly}\lambda$.

Base Case $\Rightarrow n = 1$ follows by security of PRG G function, because

$$(F_k(0), F_k(1)) = (G_0(k), G_1(k)) \approx_c U_{2\lambda}$$

This means that, chosen k , the two values returned by F are indistinguishable from 2 values taken at random and inserted in the truth table of a possible random function. Since these values are indistinguishables, the source functions are indistinguishables and F_k is pseudorandom.

Now, for the **Inductive step** :

Lemma 5. *Let $F' : \{0, 1\}^{n-1} \rightarrow \{0, 1\}^\lambda$ be a PRF. Now define $F_k(x, y) = G_x(F'_k(y))$ where $F_k : \{0, 1\}^n \rightarrow \{0, 1\}^\lambda$. If $\{F'_k\}$ is a PRF so is $\{F_k\}$.* \diamond

Consider the following images:

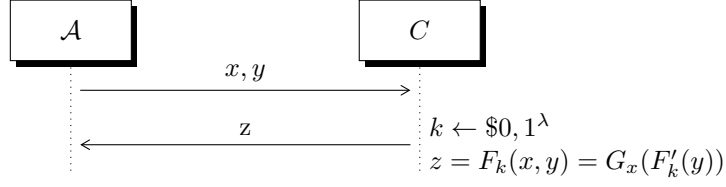


Figure 4.17: $\mathcal{HYB}_{\mathcal{F}, \mathcal{A}}^0(\lambda)$

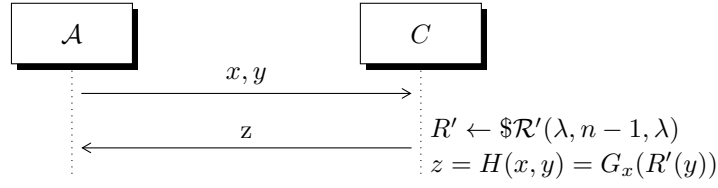


Figure 4.18: $\mathcal{HYB}_{\mathcal{R}', G, \mathcal{A}}^1(\lambda)$

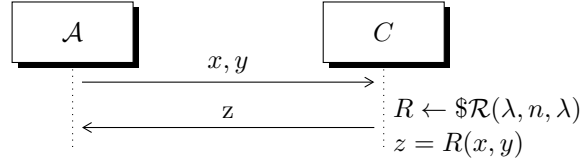
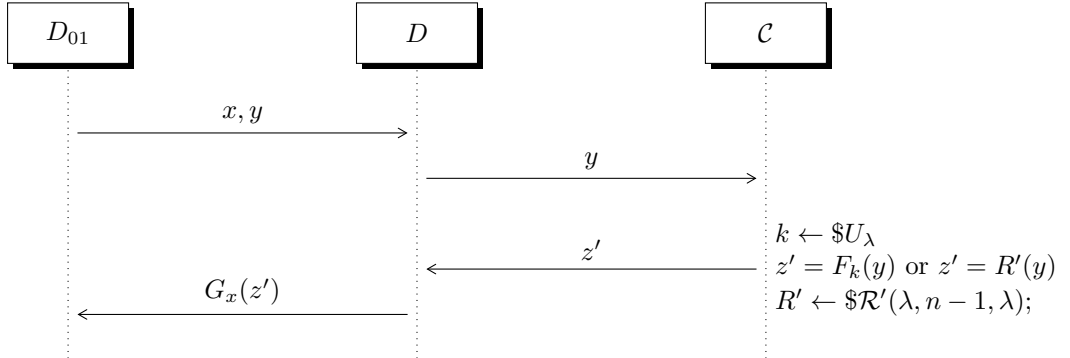


Figure 4.19: $\mathcal{HYB}_{\mathcal{R}, \mathcal{A}}^2(\lambda)$

Lemma 6. $\mathcal{HYB}^0 \approx_c \mathcal{HYB}^1 \approx_c \mathcal{HYB}^2$. ◇

Claim 2. $\mathcal{HYB}^0 \approx_c \mathcal{HYB}^1$

Assume $\exists.PPT.D_{01}$ that can distinguish F_k and H ; then there may exist a distinguisher D as in the image which breaks the assumption made by inductive step.



Claim 3. $\mathcal{HYB}^1 \approx_c \mathcal{HYB}^2$

For this proof we use the following simple lemma.

Lemma 1. If $G : \{0, 1\}^\lambda \rightarrow \{0, 1\}^{2\lambda}$ is a PRG, then for any $t(\lambda) \in \text{poly}\lambda$

$$(G(k_1), \dots, G(k_t)) \approx_c (U_{2\lambda}, \dots, U_{2\lambda})$$

for $k_1, \dots, k_t \leftarrow \U_λ

Assume that it exists a distinguisher $D_{1,2}$ which is capable of distinguish $H(x, y)$ and $R(x, y)$.

...TO REVIEW, NOT UNDERSTOOD AT ALL ...

□

4.10 CPA SECURITY

Suppose G is a PRG.

Given $\text{Enc}(k, m) = G(k) \oplus m$, and a known (\bar{m}, \bar{c}) where $c = G(k) \oplus m$, then this function is not 2 time secure, since $c \oplus \bar{c} = m \oplus \bar{m}$, easy to invert. (??? some glue arguments missing ???)

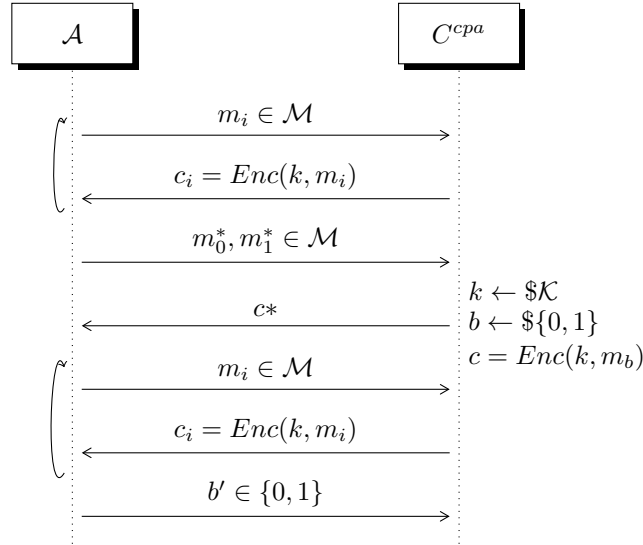


Figure 4.20: $\text{Game}_{\Pi, A}^{cpa}(\lambda, b)$

The adversary wins if finds which m_b^* message was previously encrypted. For sure m_i can be equal to m_0 or m_1 .

The two cyclic requests, before and after the starred messages, can be arbitrarily long (also 0, there is no constraint).

Definition 5. A scheme is CPA-secure if $\text{Game}_{\Pi, A}^{cpa}(\lambda, 0) \approx_c \text{Game}_{\Pi, A}^{cpa}(\lambda, 1)$ ◇

Observation 1. No deterministic scheme can achieve CPA security. ◇

This is because the Adversary is not limited in what he can ask to the Challenger: if he asks m_0 and m_1 before sending the starred messages, he will know

in advance the encrypted form of the messages, and he will be able to distinguish with $\mathcal{P}[\text{Game}^{cpa}] = 1$ the two games.

So, a way to obtain a CPA-secure encryption scheme consists of returning different cyphertexts for the same message, and we can build a scheme which tries to generate this kind of output using PRFs.

Consider the following SKE scheme Π .
Let $\mathcal{F} = \{F_k : \{0, 1\}^n \rightarrow \{0, 1\}^l\}$ be a PRF:

- $k \leftarrow \$U_\lambda$
- $\text{Enc}(k, m)$, picking a random $r \leftarrow \$\{0, 1\}^n$ with an output of $c = (c_1, c_2) = (r, F_k(r) \oplus m) \leftarrow \$\{0, 1\}^{n+l}$
- $\text{Dec}(k, (c_1, c_2)) = F_k(c_1) \oplus c_2$

. In this scheme, the only secret thing is k , which gives a *flavour* to the random function; Adversary can see $c = (r, F_k(r) \oplus m)$, so he can see r .

Theorem 10. *If \mathcal{F} is a family of PRF functions, then Π is CPA-secure* \diamond

To prove this, we have to prove that $\text{Game}_{\Pi, \mathcal{A}}^{cpa}(\lambda, 0) \approx_c \text{Game}_{\Pi, \mathcal{A}}^{cpa}(\lambda, 1)$.

Proof. Consider hybrid arguments

$$\mathcal{HYB}_0 \equiv \text{Game}_{\Pi, \mathcal{A}}^{cpa}(\lambda, 0)$$

and

$$\mathcal{HYB}_1$$

, like \mathcal{HYB}_0 but with another distribution of $\text{Enc}(k, m_b)$:

- picking $r \leftarrow \$\{0, 1\}^n$
- $R \leftarrow \$\mathcal{R}(\lambda, n, l)$
- output obtained is $(r, R(r) \oplus m_b)$

and

$$\mathcal{HYB}_2(\lambda, b)$$

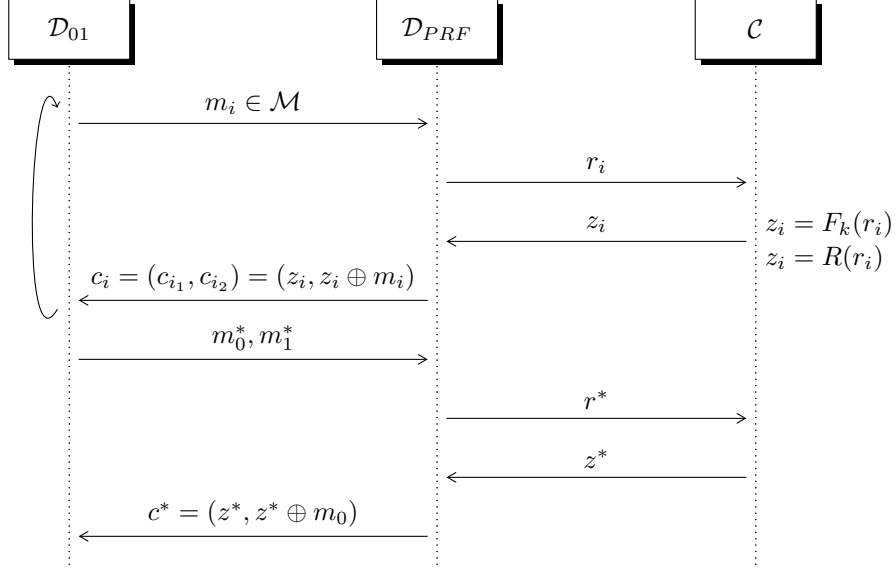
which simply outputs $(r_1, r_2) \leftarrow \$U_{n+l}$.

Lemma 7. $\mathcal{HYB}_0 \approx_c \mathcal{HYB}_1$ for each $b \in \{0, 1\}$ \diamond

Proof. Suppose these two hybrids are distinguishable; but then we can use the related distinguisher to break the starting assumption saying \mathcal{F} is a PRF.

Since they are two similar games (but just for the encryption function), we can use the CPA game for this reduction.

Fix $b = 0$ and build the following



After the last message, D_{01} will reply with a $b' \in \{0, 1\}$ which expresses which one of the encryption functions has been used; since this result can be used for solving the PRF game, the starting assumption fails and the lemma is proven. \square

Lemma 8. $\mathcal{HYB}_1 \approx_c \mathcal{HYB}_2$ \diamond

Proof. Even if the output of the first hybrid is $(r_i, R(r_i) \oplus m_b)$, the distribution of the second member of this couple doesn't depend on m_b , hence we can assume that $R(r_i) \oplus m_b \equiv R(r_i)$.

The difference between \mathcal{HYB}_1 and \mathcal{HYB}_2 comes if we play the game once more and we pick up the same c_1 of the first game: while in \mathcal{HYB}_2 r_2 is independent from r_1 and will be different from the previous r'_2 with probability near to 1, in \mathcal{HYB}_1 (if r_i is the same of the previous game) $R(r_i)$ will be the same of the previous game (because $R()$ is a random function).

Anyway, we can show that this "collision" happens with very low probability.

Call **REPEAT** this event of collision of r_i between 2 consecutive games. To show the statement of the lemma, it suffices to show that $P[\text{REPEAT}] \in \text{negl}(\lambda)$, therefore the two distributions are indistinguishable (or distinguishable with a negligible probability).

In fact, if I make q queries to \mathcal{HYB}_1 ,

$$\begin{aligned}
P[REPEAT] &= P[\exists i, j \in q \text{ such that } r_i = r_j] \leq \\
&\leq \sum_{i, j \wedge i \neq j} \mathcal{P}[r_i = r_j] = Col(U_n) = \\
&= \sum_{i \wedge j, i \neq j} \sum_{e \in \{0,1\}^n} \mathcal{P}[r_1 = r_2 = e] = \\
&= \sum_{i \wedge j, i \neq j} \sum_{e \in \{0,1\}^n} \mathcal{P}[r = e]^2 = \\
&= \binom{q}{2} 2^n \frac{1}{2^{2n}} = \\
&= \binom{q}{2} 2^{-n} \leq \\
&\leq q^2 2^{-n} \in \text{negl}(\lambda)
\end{aligned}$$

□

Since $\mathcal{HYB}_0 \equiv Game(\lambda, 0)$ but the same proofs can be made for $Game(\lambda, 1)$,
 $Game(\lambda, 0) \equiv \mathcal{HYB}_0(Game(\lambda, 0)) \approx_c \mathcal{HYB}_2 \approx_c \mathcal{HYB}'_0(Game(\lambda, 1)) \equiv Game(\lambda, 1)$

□

Lesson 8

4.11 Domain extension

How can we encrypt long messages , say $m = (m_1, m_2, \dots, m_t)$ where for $i \in [t], m_i \in \{0, 1\}^n$?

Mode of operation: let's build P_k , a blockcypher.

4.11.1 Electronic CodeBook (ECB)

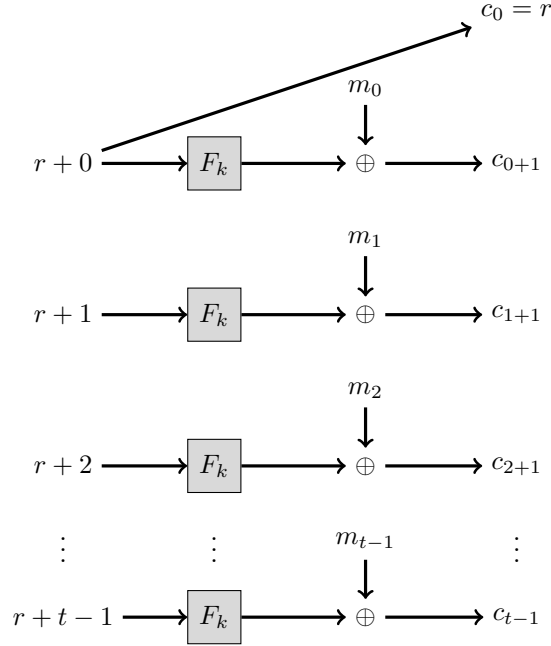
(In notes, this name is associated with CTR-mode, or **counter mode**, but why?)

The principle is that I have many blocks of information and I encrypt all of them individually (maybe in parallel):

$$c_i = P_k(m_i) \forall i \in [t]$$

and outputs c_1, c_2, \dots, c_t .

If I encrypt the same message, I obtain the same cyphertext. So it's not CPA-secure as the encryption function is **deterministic**.



In the image, each message is long exactly **n bits**, and

$$r \leftarrow \{0, 1\}^n$$

so the sums made over r are done in **mod** 2^n , since r can be a number in the range $[0, 2^n - 1]$.

This is also called streamcipher, cause as input flows ECB produces the ciphertext. How to decrypt this schema? Since I know r_i , if I compute $F_k(r_i)$ I can use the \oplus -operation to calculate the message.

4.11.2 Cipher block chaining (CBC)

CBC IMAGE

The schema to cipher here is

$$\forall i \in [1, t] c_i = P_k(C_{i-1} \oplus m_i)$$

How to decipher?

$$P_k^{-1}(i) = c_{i-1} \oplus m_i$$

so $m_i = P_k^{-1}(c_i) \oplus c_{i-1}$.

Theorem 11. Assume \mathcal{F} (F_k in the image above is part of this family) is a PRF, then CTR-mode yields a CPA-secure SKE for variable length messages.

◇

Variable length messages means that every message

$$m = (m_1, \dots, m_t)$$

has t subsets, and t can change from message m to another message $m' = (m'_1, \dots, m'_{t'})$. Consider the following hybrids:

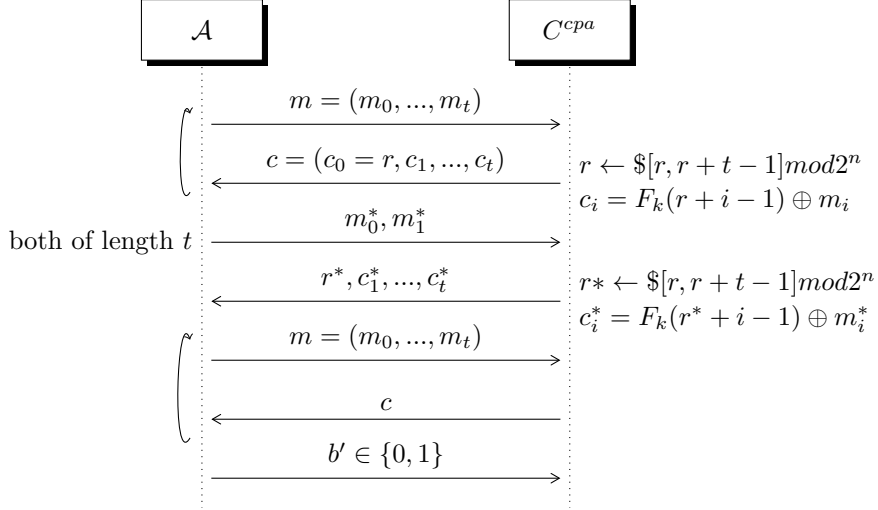


Figure 4.21: $\text{Game}_{CTR, \mathcal{A}}^{cpa}(\lambda, b)$

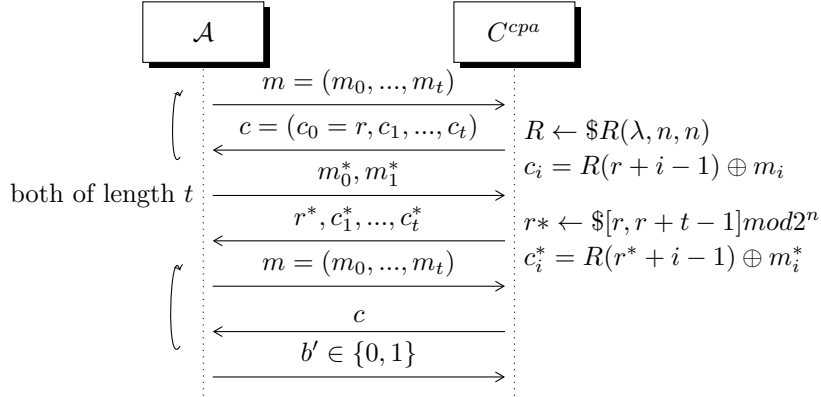


Figure 4.22: $\mathcal{HYB}_1(\lambda, b)$

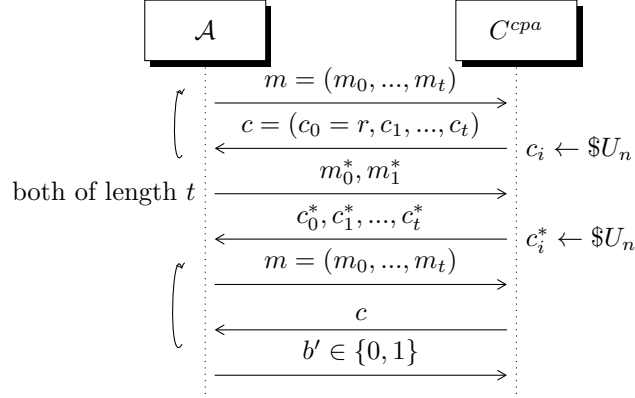


Figure 4.23: $\mathcal{HVB}_2(\lambda, b)$

Now we want to show that $\text{Game}_{\Pi, \lambda}^{\text{cpa}}(\lambda, 0) \approx_c \text{Game}_{\Pi, \lambda}^{\text{cpa}}(\lambda, 1)$

Proof.

Exercise 12.

Lemma 9. *Show that $\text{Game}_{\Pi, \lambda}^{\text{cpa}}(\lambda, b) \approx_c \mathcal{HVB}_1(\lambda, b), \forall b \in \{0, 1\}$* \diamond

(Since this $\text{Game}(\lambda, b)$ is a CPA scheme and the second one is very similar, we can use a distinguisher which plays the CPA game; since this is a lemma, our precondition to "break" during the reduction is the precondition contained in the theorem statement)

Lemma 10. $\mathcal{HVB}_1(\lambda, b) \approx_c \mathcal{HVB}_2(\lambda), \forall b \in \{0, 1\}$ \diamond

The two hybrids are identical but for the encryption function, which is completely random in the second one while the first one uses a random function.

Since $R(r + i) \oplus m_i \approx R(r + i)$ (because m_i doesn't affect the distribution of the result at all), if $R(r^*)$ behaves like a true random extractor, the two hybrids are indistinguishable.

Now, if I examine a simple query and the consecutive challenge query (q, q^*) in a game in \mathcal{HVB}_2 , the responses will be chosen at random with very low probability of being the same; but in a game in \mathcal{HVB}_1 , the problem comes if the challenger creates an $r^* + j'$ for q^* which overlaps some chosen $r_i + j$ previously chosen for q .

This is bad, because $R(r_i + j) = R(r^* + j')$ since R is a random function, and the probability of outputting the same response ciphertext will be much higher in \mathcal{HVB}_1 .

It's possible to show that these collisions happen very few times (with **negligible** probability) in \mathcal{HVB}_1 and that the two hybrids are distinct and distinguishable for a negligible factor.

Proof. Let :

- $q = \#$ of encryption queries
- $t_i = \#$ of blocks for the i -th query
- $t^* = \#$ of blocks for challenge

and let the function R run: we will have $R(r^*), \dots, R(r^* + t^* - 1)$ and $R(r_i), \dots, R(r_i + t_i - 1)$ sequences of random function outputs.

Definition 6. OVERLAP event: $\exists i, j, j', r_i + j = r^* + j'$ (here $j \leq t_i$ and $j' \leq t^*$), for some query q_i . \diamond

If **OVERLAP** does not happen, the sequence $(R(r^*), \dots, R(r^* + t^* - 1))$ is made of **uniform** and **independent** values. Thus $\mathcal{HYB}_1(b)$ is identical to $\mathcal{HYB}_2(b)$ for all $b \in \{0, 1\}$.

Now it suffices to show that $\mathcal{P}[\text{OVERLAP}] \in \text{negl}(\lambda)$.

For simplicity, assume $q = (\text{length of each query})$ and also $t_i = t^* = q$. Let OVERLAP_i be the event that the i -th query $r_i, \dots, r_i + q - 1$ partially or totally overlaps the challenge sequence $r^*, \dots, r^* + q - 1$.

Fix some r^* .

One can see that OVERLAP_i happens if

$$r^* - q + 1 \leq r_i \leq r^* + q - 1$$

, which means that r_i should be chosen *at least* in a way that :

- the sequence $r^*, \dots, r^* + q - 1$ comes before the sequence $r_i, \dots, r_i + q - 1$, and they overlap just for the last element $r^* + q - 1 = r_i$ or
- the sequence $r_i, \dots, r_i + q - 1$ comes before the sequence the sequence $r^*, \dots, r^* + q - 1$, and they overlap just for the last element $r_i + q - 1 = r^*$

So now

$$\mathcal{P}[\text{OVERLAP}_i] = \frac{(r^* + q - 1) - (r^* - q + 1) + 1}{2^n} = \frac{2q - 1}{2^n}$$

It is obvious that, for definition of **OVERLAP**,

$$\mathcal{P}[\text{OVERLAP}] \leq \sum_{i=1}^q \mathcal{P}[\text{OVERLAP}_i] \leq 2 \frac{q^2}{2^n} \in \text{negl}(\lambda)$$

□

Since $\mathcal{HYB}_1 \approx_c \mathcal{HYB}_2$ and $\mathcal{HYB}_1 \equiv \text{Game}(b)$ with b equal to 0 and 1 respectively, we can state that

$$\text{Game}(0) \approx_c \mathcal{HYB}_2 \approx_c \text{Game}(1)$$

□

Lesson 9

4.12 Message Authentication and UFCMA-security

First, remember the $Tag()$ function and how a MAC works. Now $Tag()$ is defined using a key k , and we call it $Tag_k()$.

In particular we are looking for a cool property of a message authentication protocol, called **universal unforgeability against chosen-message attacks**, which prevents the attacker from generating a valid couple (m^*, ϕ^*) after some queries containing messages and receiving the related tags.

This property is defined through a game called

$$Game_{\Pi, \mathcal{A}}^{ufcma}(\lambda)$$

and played in the following manner:

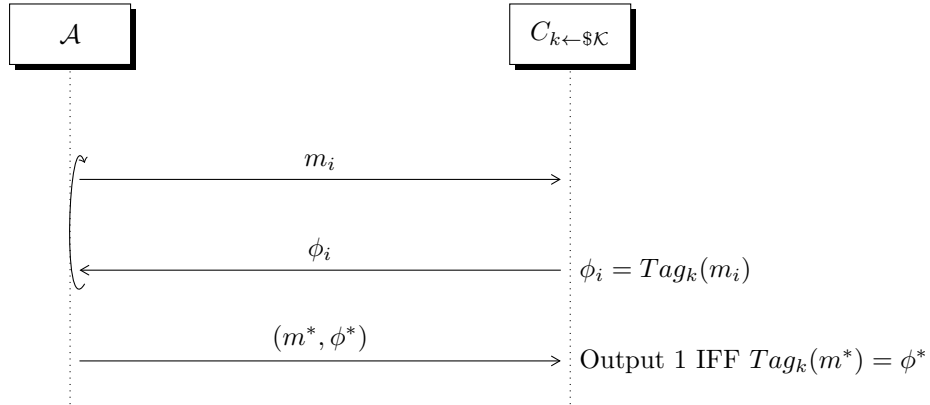


Figure 4.24: $Game_{\Pi, \mathcal{A}}^{ufcma}(\lambda)$

with m^* which must be a fresh new message never used before and with the Adversary \mathcal{A} which doesn't know the key k used in the tag.

Definition 7. Π is **ufcma-secure** if $\forall .PPT.\mathcal{A}$

$$\mathcal{P}[Game_{\Pi, \mathcal{A}}^{ufcma}(\lambda) = 1] \in \text{negl}(\lambda)$$

◇

Now consider the following theorem:

Theorem 13. Let $\mathcal{F} = \{F_k : \{0, 1\}^n \rightarrow \{0, 1\}^l\}_{k \in \{0, 1\}^\lambda}$ be a PRF family. Then Π which uses $\text{Tag}_k() = F_k()$ is a UFCMA-secure MAC with n -bit domain. \diamond

We show that the scheme which uses a PRF is indistinguishable from a scheme which uses a random function, and that a MAC scheme which uses a random function is breakable by an efficient attacker in negligible time.

Proof. Consider the two following hybrids:

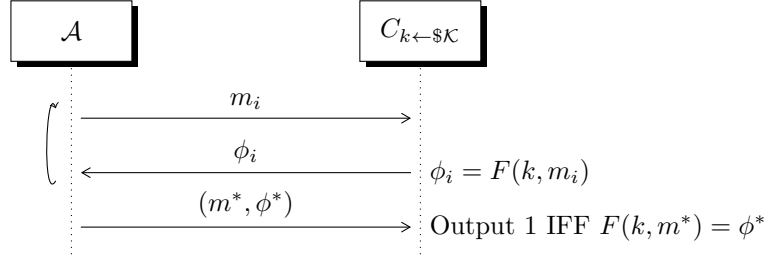


Figure 4.25: $\text{Game}_{\mathcal{F}, \mathcal{A}}^{ufcma}(\lambda)$

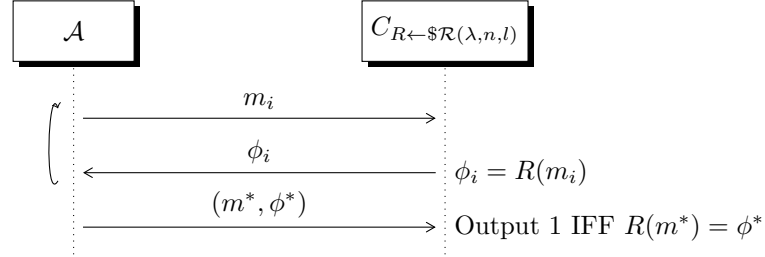
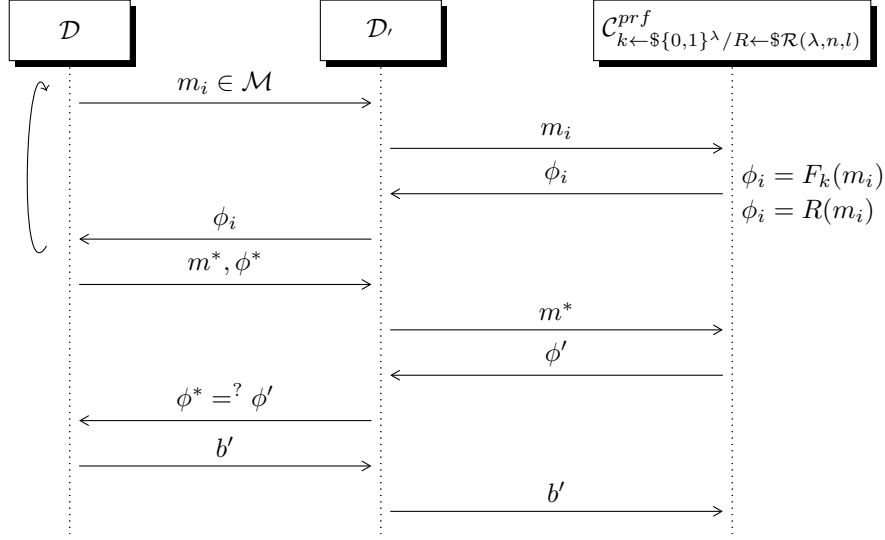


Figure 4.26: $\mathcal{HYB}_1^{ufcma}(\lambda)$

Lemma 11. $\text{Game}_{\mathcal{F}, \mathcal{A}}^{ufcma}(\lambda) \approx_c \mathcal{HYB}_1^{ufcma}(\lambda)$ \diamond

Assume that these two are distinguishable by D . So we could have D' which is capable, with the following game, of distinguishing a pseudo-random function from a true random function:



After \mathcal{D}' receive ϕ' from \mathcal{C} , he has to use the distinguisher \mathcal{D} to distinguish the PRF function from the random one. So he says to \mathcal{D} if $\phi^* = \phi'$ or not, and \mathcal{D} now is capable of understanding which game he is playing.

Lemma 12. *For all efficient adversaries $\mathcal{P}[\mathcal{HYB}_1(\lambda) = 1] \leq 2^{-l}$* \diamond

This is true because attacker has to predict the output $R(m^*)$ on a fresh input m^* and to send to the Challenger this couple to win the game. This can happen, at most, with probability $\frac{1}{2^l}$. \square

This was for fixed length messages, since we can encrypt messages long n . In the next subsection, we will see how to create a UFCMA-secure MAC for variable length messages.

4.13 Domain extension

Assume $m = (m_1, \dots, m_t) \in (\{0, 1\}^n)^t$ for some $t \geq 1$. How can we tag m given just $\text{Tag}_k : \{0, 1\}^n \rightarrow \{0, 1\}^l$?

We can try various attempts:

- XOR all blocks: $\phi = \text{Tag}_k(\bigoplus_{i=1}^t m_i)$.
This is not secure cause given (m, ϕ) it is possible to find $m^* \neq m$ s.t.

$$\bigoplus m_i = \bigoplus m_i^*$$

and output this the couple (m^*, ϕ) to win.

- Let $\phi_i = \text{Tag}_k(m_i)$ and the final message of the challenge has this form : $(m, \phi = (\phi_1, \dots, \phi_t))$.
This is not secure. Given $m = (m_1, \dots, m_t)$, this message has an unique $\phi = (\phi_1, \dots, \phi_t)$, and if I swap m_1 and m_t I obtain a fresh new message,

with a fresh new tag $\phi' = (\phi_t, \phi_2, \dots, \phi_{t-1}, \phi_1)$. Using this new couple, the game is won.

- Try with $\phi_i = \text{Tag}_k(i||m_i)$, authenticating the position of the block .
But this is not secure, and can be showed in **just 2 queries**. (I solved this in class during the break, with $t + 1$ queries: t for retrieving the partial tag of the submessage+position , and the last query to merge all the obtained results in a fresh new message. But this solution can be improved.)
(UPDATE: I send $m = m_{1_1}, \dots, m_{1_t}$ and I save the corresponding ϕ . Then I send $m' = m_{2_1}, \dots, m_{2_t}$ and I save the ϕ' . Now I forge the new fresh prince of Bel Air $m^* = m_1, m_2, m_2, \dots, m_2$ and I can forge also a valid ϕ^* because I have all the signed parts of this new tag.)

Now I feel cool in Los Angeles and I want to explore a new **IDEA**:
the design of a shrinking functions family

$$\mathcal{H} = \{h_s : \{0, 1\}^{nt} \rightarrow \{0, 1\}^n\}_{s \in \{0, 1\}^\lambda}$$

which can be used to shrink variable length messages and then apply a PRF on them.

This idea is cool, and I consider the induced family

$$\mathcal{F}(\mathcal{H}) = \{F_k(h_s(.))\}$$

Question 3. Which are the properties of this family?

The main problem are *collisions* , since for each $m \in \{0, 1\}^{nt}$ it should be hard to find $m' \neq m$ such that $h_s(m) = h_s(m')$.

But we know that collisions exist, because we are trying to create a function

$$\text{Tag}_{k,s}(m) = F_k(h_s(m))$$

which maps elements in $\{0, 1\}^{nt}$ to the elements in $\{0, 1\}^t$, and since the second set is smaller, for the pidgeonhole principle, there will be elements of $\{0, 1\}^t$ which will be reached by more than one element of $\{0, 1\}^{nt}$.

To overcome this problem, we can consider 2 ways:

- assume collisions are hard to find given $s \in \{0, 1\}^\lambda$ publicly, and we have a *collision resistant hashing*;
- let s be secret, and assume collisions are hard to find because it is hard to know how h_s works.

Definition 8. \mathcal{H} is called **universal** family if

$$\forall x, x' \in \{0, 1\}^{nt} \text{ such that } x \neq x'$$

$$\mathcal{P}_{s \leftarrow \{0, 1\}^\lambda} [h_s(x) = h_s(x')] \leq \varepsilon$$

◇

For $\varepsilon = 2^{-n}$ we call it **perfectly universal**.
For $\varepsilon \in \text{negl}(\lambda)$ we call it **almost universal** (or **AU**).

Exercise 14. Show that any pairwise independent hash function is perfectly universal. (should I use *Col* for solving this? What is the difference and when I should use *Col* instead of one-shot-probability?) **ASK FOR SOLVING PROPERLY** (Thoughts: when I ask *what's the probability that , chosen 2 distinct x -es, their hashes are the same on a certain value?* , maybe I have to use one-shot, because one-shot refers to the prob. that the two inputs collide on a specific value, even if not specified.

Instead, if I consider *what's the prob. that , chosen 2 distinct x -es, their hashes are the same?*, maybe I have to calculate all the possible collisions, because I want to know if the 2 inputs can collide in general.)

Theorem 15. Assuming \mathcal{F} is a PRF with n -bit domain and \mathcal{H} is AU, then $\mathcal{F}' = \mathcal{F}(\mathcal{H})$ is a PRF (and , if used in a MAC as tag function, makes it UFCMA) on nt -bit domain (for $t \geq 1$). \diamond

We want to show that $\mathcal{F}' = \mathcal{F}(\mathcal{H})$ is a PRF , so we want to show that

$$\text{Real}_{\mathcal{F}, \mathcal{A}}(\lambda) \approx_c \text{Rand}_{\mathcal{R}', \mathcal{A}}(\lambda)$$

Proof. Consider these 3 experiments/games:

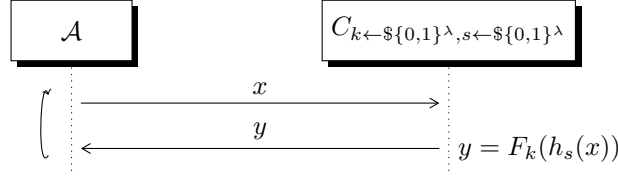


Figure 4.27: $\text{Real}_{\mathcal{F}, \mathcal{A}}(\lambda)$

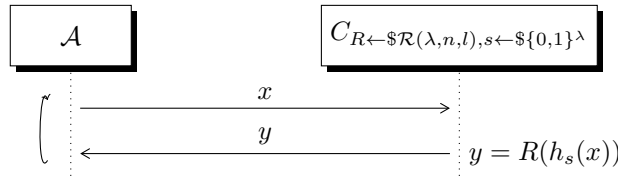


Figure 4.28: $\mathcal{HYB}_{\mathcal{R}, \mathcal{A}}(\lambda)$

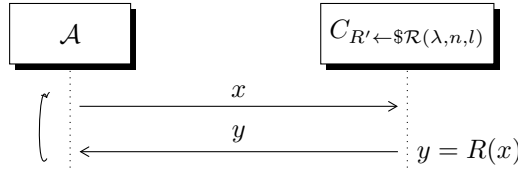


Figure 4.29: $\text{Rand}_{\mathcal{R}', \mathcal{A}}(\lambda)$

Lemma 13. $Real \approx_c \mathcal{HYB}$

◇

Exercise 16. Prove it!

Lemma 14.

$$\mathcal{HYB} \approx_c Rand$$

◇

These 2 experiments are very similar but for the encryption step. If I send 2 consecutive queries to both the experiments, while $Rand$ will give me

- the same y with the same x or
- the same y with different $x_1 \neq x_2$ but with very low probability,

\mathcal{HYB} could return the same y :

- with the same x input or
- with different $x_i \neq x_j \Rightarrow h_s(x_i) \neq h_s(x_j)$ but with very low probability (because R is a random function), or
- with different $x_i \neq x_j$ but with the same $h_s(x_i) = h_s(x_j)$,

We want to show that the last item doesn't happen too often, and that these 2 experiments are distinct for a negligible factor.

Proof. Let **BAD** be the event that

$$\exists i, j \in [q] \text{ with } i \neq j \text{ s.t. } h_s(x_i) = h_s(x_j)$$

As long as **BAD** doesn't happen, the function R is run as a sequence of distinct points $R(h_s(x_1)), \dots, R(h_s(x_q))$.

So in this case the distribution if the 2 games is identical, and it suffices to show that $\mathcal{P}[BAD] \in \text{negl}(\lambda)$.

The event **BAD**, which happens during a game made of q queries and q replies, is the same as collecting all the q queries, choosing the seed and then looking for collisions. If interpreted in this way,

$$\begin{aligned} \mathcal{P}[BAD] &= \mathcal{P}_s[\exists x_i \neq x_j, h_s(x_i) = h_s(x_j)] \leq \\ &\leq \sum_{i,j} \underbrace{\mathcal{P}[h_s(x_i) = h_s(x_j)]}_{h_s \text{ is AU by definition}} \leq \\ &\leq \binom{q}{2} \text{negl}(\lambda) \in \text{negl}(\lambda) \end{aligned}$$

□

So now we have $Real \approx_c \mathcal{HYB} \approx_c Rand$

□

Now we want to show a possible shrinking/hash function which can be used in a UFCMA-secure MAC. In order to be used in a UFCMA-secure MAC as the input of a PRF, this function must be part of an **almost universal** family.

4.13.1 \mathcal{H} family using Galois Fields

Construction 3. Take $\mathbb{F} = GF(2^n)$, a *Galois field* of 2^n elements. Let $m = (m_1, \dots, m_t) \in \mathbb{F}^t$ and $s = (s_1, \dots, s_t) \in \mathbb{F}^t$. We state that

$$h_s(m) = \sum_{i=1}^t s_i m_i = \langle s, m \rangle = q_m(s)$$

◇

A generic *galois field* has very interesting properties, like the following :

- addition of two elements is like applying the XOR operation on their binary forms;
- multiplication of two elements is like the product *mod* 2^n

Now, since this family must be almost universal to be used as part of UFCMA-secure MAC protocol, collisions must happen a negligible amount of time.

Suppose we have a collision with two different messages:

$$\sum_{i=1}^t m_i s_i = \sum_{i=1}^t m'_i s_i$$

Let $\delta_i = m_i - m'_i$, assuming without loss of generality that $\delta \neq 0$. Now we have that, when a collision happens,

$$\begin{aligned} \sum_{i=1}^t m_i s_i = \sum_{i=1}^t m'_i s_i &\Leftrightarrow \sum_{i=1}^t m_i s_i - \sum_{i=1}^t m'_i s_i = 0 \Leftrightarrow \\ &\sum_{i=1}^t \delta_i s_i = 0 \end{aligned}$$

Taking m, m' such that $m \neq m'$ means that m is distinct from m' at least for one subsequence $m_i \neq m'_i$.

So we can assume, without loss of generality, that $i = 1$ is an index (or the only index) which m and m' differ on.

So we can split the last summation in 2 parts, choosing $\delta_1 s_1$ as the first element and $\sum_{i=2}^t \delta_i s_i$ as second element:

$$\begin{aligned} \delta_1 s_1 + \sum_{i=2}^t \delta_i s_i &= 0 \Leftrightarrow \\ \delta_1 s_1 &= - \sum_{i=2}^t \delta_i s_i \Leftrightarrow \\ s_1 &= \frac{- \sum_{i=2}^t \delta_i s_i}{\delta_1} \end{aligned}$$

and this means that when a collision happens s_1 must be exactly equal to the second member of the equation, which is an element of \mathbb{F} . But since every seed is chosen at random among \mathbb{F} , what's the probability of picking the element s_1 which zeroes the above equation ?
This probability is just $\frac{1}{|\mathbb{F}|} = \frac{1}{2^n} \in \text{negl}(\lambda)$.

\mathcal{H} with Galois fields elements and polynomials

Construction 4. Take $\mathbb{F} = GF(2^n)$, a *Galois field* of 2^n elements.
Let $m = (m_1, \dots, m_t) \in \mathbb{F}^t$ and $s \leftarrow \mathbb{F}^t$. We state that

$$h_s(m) = \sum_{i=1}^t s^{i-1} m_i$$

◇

Exercise 17. Prove that this construction is **almost universal**.

(possible proof: to be almost universal, looking at the definition, collisions with $m \neq m'$ must be negligible.

So consider a collision as above: it must be true that

$$\sum_{i=1}^t m_i s^{i-1} = \sum_{i=1}^t m'_i s^{i-1} \Leftrightarrow \sum_{i=1}^t m_i s^{i-1} - \sum_{i=1}^t m'_i s^{i-1} = 0 \Leftrightarrow q_{m-m'}(s) = 0$$

How can we make a polynomial equal to 0? We have to find the **roots** of the polynomial, which we know are at most the **grade** of the polynomial. So, the grade of this polynomial is $t - 1$, and the probability of picking a root from \mathbb{F} as seed of $h_s(\cdot)$ is

$$\mathcal{P}[s = \text{root}] = \frac{t-1}{2^n} \in \text{negl}(\lambda)$$

)

Lesson 10

4.14 Domain extension for PRFs/MACs

Almost universal approach : I have a family $\mathcal{F}(\mathcal{H})$ with \mathcal{H} AU and with PRF $f \in \mathcal{F}$.

Computational AU : we want to build a family \mathcal{H} using some other PRFs. We expect to have:

- $\mathcal{P}[h_s(m) = h_s(m'), s \leftarrow \{0, 1\}^\lambda, (m, m') \leftarrow A(1^\lambda)] \in \text{negl}(\lambda)$;
- We need two PRFs. One is F_k , and the other is F_{sj}

=====

something related to $f_s(1, \cdot)$ and $f_s(0, \cdot)$, but I didn't get it.

=====

4.14.1 XOR mode

Assume that we have this function

$$h_s(m) = F_s(m_1||1) \oplus \dots \oplus F_s(m_t||t)$$

so that the input to the PRF $F_s(\cdot)$ is $n + \log_2 t$ bytes long.

Lemma 15. *Above \mathcal{H} is computational AU if \mathcal{F} is a PRF.* ◇

Exercise 18. Prove this !

Possible proof:

we have to show that

$$\mathcal{P}[h_s(m) = h_s(m')] \in \text{negl}(\lambda)$$

with $m \neq m'$.

This means that

$$\begin{aligned} \mathcal{P}[F_s(m_1||1) \oplus \dots \oplus F_s(m_t||t) = F_s(m'_1||1) \oplus \dots \oplus F_s(m'_t||t)] = \\ = \mathcal{P}[F_s(m_i||i) \oplus F_s(m'_i||i) = \alpha] = \bigoplus_{j=1, j \neq i}^t F_s(m_j||j) \oplus F_s(m'_j||j) \end{aligned}$$

for each $i \in [1, t]$.

But α is one unique random number chosen over 2^n possible candidates, so the collision probability is negligible.

4.14.2 CBC MAC

This is part of the standard, used in TLS. It's used with a PRF F_s , setting the starting vector $IV = 0^n = c_0$ and running this PRF as part of CBC. The output of the CBC process is just the last block:

$$h_s(m_1, \dots, m_t) = F_s(m_t \oplus F_s(m_{t-1} \oplus \dots \oplus F_s(m_2 \oplus F_s(m_1 \oplus IV))))$$

Lemma 16. *CBC MAC defines completely an AU family.*

(not proven) ◇

We can use this function to create an **encrypted CBC**, or **E-CBC** :

$$E - CBC_{K,S}(m) = F_k(h_s^{CBC}(m))$$

Theorem 19. *Actually if \mathcal{F} is a PRF, CBC-MAC is already a MAC with domain nt for arbitrary but fixed $t \in \mathbb{N}$.*

(not proven) ◇

4.14.3 XOR MAC

Instead of $\mathcal{F}(\mathcal{H})$ now the $Tag()$ function outputs $\phi = (\eta, F_k(\eta) \oplus h_s(m))$ where $\eta \leftarrow \mathcal{H}$ is random and it's called *nonce* .

When I want to authenticate, I should send the

$$(m, (\eta, F_k(\eta) \oplus h_s(m)))$$

couple.

When I want to verify a message and I get the couple $(m, (\eta, v))$, I just check that $v = F_k(\eta) \oplus h_s(m)$. It should be hard to find a value called a such that, given $m \neq m'$,

$$h_s(m) \oplus a = h_s(m')$$

In fact, since an adversary who wants to break this scheme has to send a valid couple (m^*, ϕ^*) after some queries, he could:

- ask for message m and store the tag $(\eta, F_k(\eta) \oplus h_s(m))$
- try to find $a = h_s(m) \oplus h_s(m')$ and modify the previous stored tag adding $v \oplus a$,

so now he could send the authenticated message

$$(m', (\eta, F_k(\eta) \oplus h_s(m')))$$

which is a valid message.

=====

WHAT IS IT ABOUT?

Lemma 17. *XOR mode gives computational AXU (Almost Xor Universal).*
(not proven) ◇

WAHT DOES IT MEAN AXU?? Has something to do with the first 2 definitions given in the start of this subsection ? =====

Theorem 20. *If \mathcal{F} is a PRF and \mathcal{H} is computational AXU, then XOR-MAC is a MAC.*
(not proven) ◇

=====

NOT CLEAR WHAT TO DO

Exercise 21. Now with variable input lenght:

- AXU based XOR mode
 - $\mathcal{F}(\mathcal{H})$ is insecure with polynomial construction $h_s(m) = q_m(s)$, but can be fixed.
 - CBC-MAC not secure. (exercise)
 - E-CBC is secure.
- =====

4.15 Chosen Ciphertext Attack security

In this kind of attack, the adversary has a decryption capability added to the previous encryption capability. A scheme which is CCA-secure is also **non-malleable**, since the attacker cannot modify the obtained ciphertexts to obtain new valid ciphertexts.

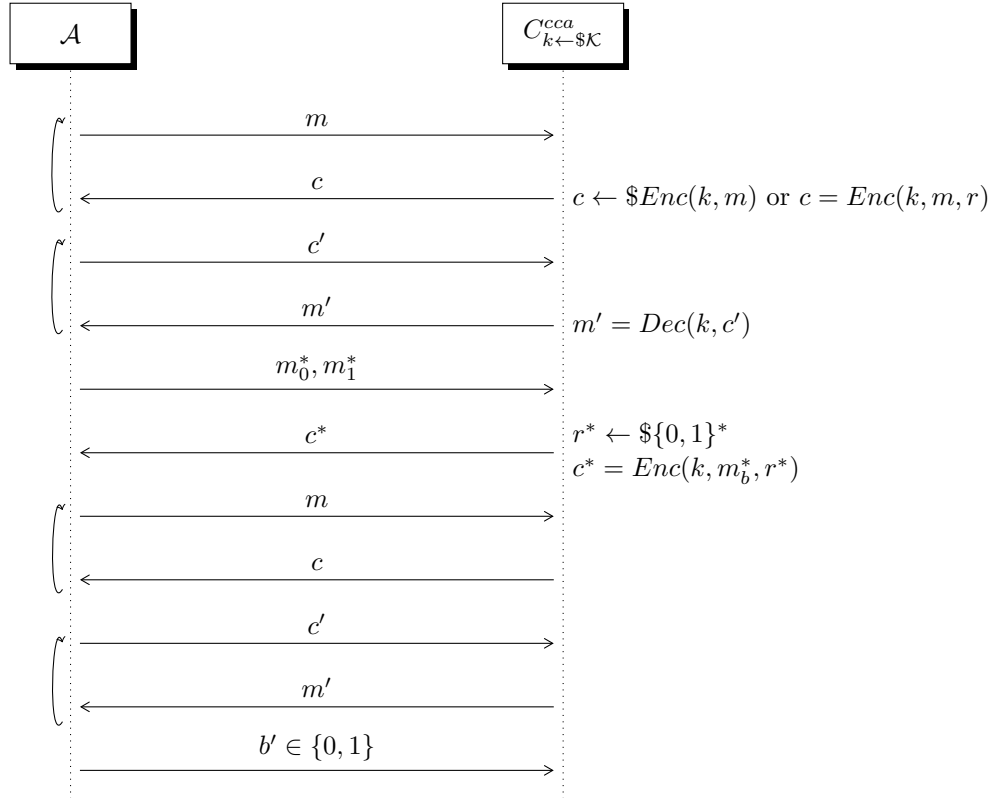
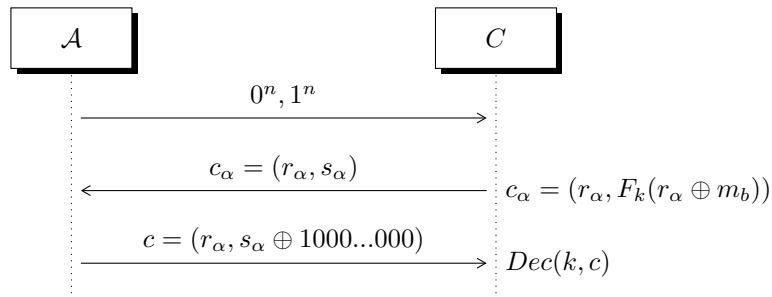


Figure 4.30: $\text{Game}_{\Pi, \mathcal{A}}^{cca}(\lambda, b)$

Exercise 22. Show that $(r, F_k(r) \oplus m)$ which is CPA-secure, is not CCA-secure.



Proof. with

$$\begin{aligned}
Dec(k, c) &= F_k(r_\alpha) \oplus s_\alpha \oplus 1000...000 = \\
&F_k(r_\alpha) \oplus F_k(r_\alpha) \oplus m_b \oplus 1000...000 = \\
&m_b \oplus 1000...000
\end{aligned}$$

At this point we have that the output is

- 1000...000 if m_b was 000...000
- 0111...111 if m_b was 0111...111

□

4.16 Authenticated encryption

Idea: what if we combine the target of authenticity with the target of encryption?

The first property is satisfied when the receiver is able to understand if the received message was sent exactly by the trusted sender; the last property is satisfied when no information of the sent message is contained in the ciphertext, thus only the chosen receiver can fully read the original sent message.

If we build up a schema with these 2 properties, we can obtain a new schema which should be *cpa-secure* (to enforce the encryption/privacy property) and essentially secure against forgeries of chosen message attacks (to enforce the authentication property).

In particular, a successful forgery can be obtained winning the following game:

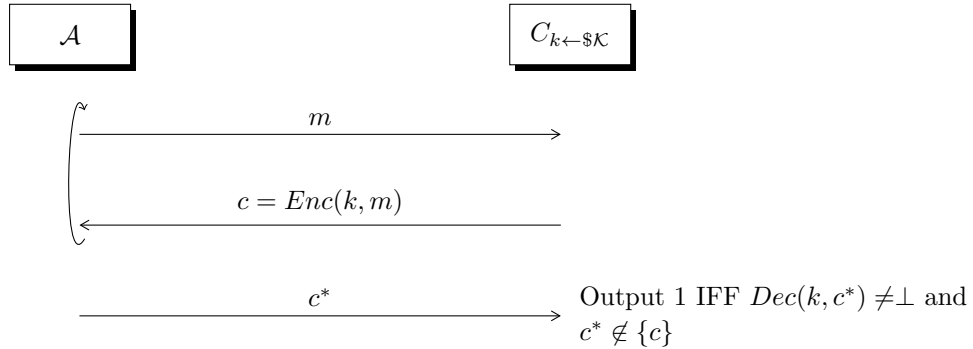


Figure 4.31: $Game_{\Pi, \mathcal{A}}^{auth}(\lambda)$

meaning that the c^* challenge ciphertext should be a **fresh** and a valid one ciphertext, considering that the $Dec(\cdot)$ function is defined as

$$Dec : \mathcal{K} * \mathcal{C} \rightarrow M \cup \{\perp\}$$

where \perp represents invalid/meaningless messages.

In the end, we want that our *authenticated encryption schema* is cpa-secure and has this last property, called **strong unforgeability** or **auth**.

Theorem 23. Let Π be an SKE (shared key encryption). If Π has **CPA** + **auth** security, then it also has **CCA** security. \diamond

Exercise 24. Prove it!

Hint: consider the experiment where $Dec(k, c)$:

- if c not fresh (i.e. output of previous encryption query m , output m)
- else output \perp

=====

TO DO AND TO PROVE

Approach: reduce cca to cpa; given D^{cca} , we can build D^{cpa} . D^{cca} will ask decryption queries, but D^{cpa} can answer just with these two properties shown above, so it can reply just if he asked these (c, m) before to its challenger \mathcal{C} .

=====

4.16.1 Three approaches to authenticated encryption

Let Π_1 be a **cpa-secure** SKE and Π_2 be a **auth** MAC schema.

We have 3 ways to combine these 2 schema to obtain a new one:

1. **Encrypt-and-MAC** :

$$\begin{aligned} c &\leftarrow \text{\$Enc}(k_1, m) \\ \phi &= \text{\$Tag}(k_2, m) \\ c^* &= (c, \phi) \end{aligned}$$

2. **MAC-then-encrypt** :

$$\begin{aligned} \phi &= \text{\$Tag}(k_2, m) \\ c &\leftarrow \text{\$Enc}(k_1, \phi \| m) \\ c^* &= c \end{aligned}$$

=====

DA RIVEDERE: come i ϕ sono randomici o no? A guardare la prima parte della prossima lezione , sembrerebbe di si.

3. **Encrypt-then-MAC** :

$$\begin{aligned} c &\leftarrow \text{\$Enc}(k_1, m) \\ \phi &\leftarrow \text{\$Tag}(k_2, c) \\ c^* &= (c, \phi) \end{aligned}$$

=====

In a paper is clearly stated that the first solution doesn't work taking random combination of CPA-secure and MAC schemes, because there are couple which, mixed, are not CCA-secure. Furthermore the second solution doesn't work for

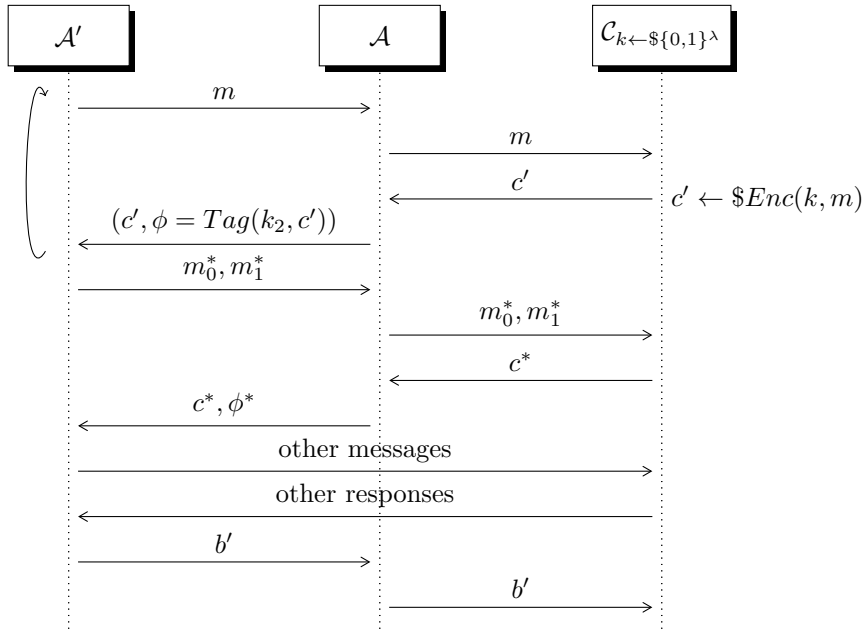
the same above reason, even if it's now part of the standard.

Instead, the third solution works always for a CPA-secure scheme Π_1 and strong unforgeability scheme Π_2 which, combined, they generate the final scheme Π .

Theorem 25. *If Π is made combining Π_1 cpa-secure and Π_2 auth-secure as stated in the third point, then Π is cpa-secure and auth-secure .* \diamond

Proof. By reduction, we negate that Π has both the properties and we find the contradiction.

Suppose now that Π hasn't the cpa-security property.



Here we are supposing that \mathcal{A} can break the cpa-security of a generic Π_1 scheme used by \mathcal{C} , while \mathcal{A}' can break the cpa-security of a generic scheme Π . \mathcal{C} can generate the $Tag(k_2, .)$ function, randomly choosing k_2 and simulating Π .

=====

REVIEW

Professor says that we have to show that $Game^{cpa}(\lambda, 0) \approx_c Game^{cpa}(\lambda, 1)$, but why??? Isn't this proof enough?

=====

Proved for the cpa-security property, now we have to prove, in a similar way, that the auth property must be holded by Π if Π_2 is an auth-secure scheme.

Exercise 26. Prove it!

Similar to the cpa-security proof.

□

Lesson 11

4.17 Authenticated encryption (Age of Ultron)

Last time we proved CPA-security of Π . Today we will explore the *auth* property. Consider Π as

$$\begin{aligned} Enc &: \{0,1\}^\lambda * \mathcal{M} \rightarrow \mathcal{C} \\ Tag &: \{0,1\}^\lambda * \mathcal{C} \rightarrow \Phi \end{aligned}$$

Lemma 18. *If $Tag(.,.)$ is **EUF-CMA**, then Π has *auth*-property.* \diamond

What is **EUFCMA** ?

It's a property similar to **uf-cma**, but now I want that the challenge message (m^*, ϕ^*) is made by a fresh m^* and a valid **fresh** ϕ^* .

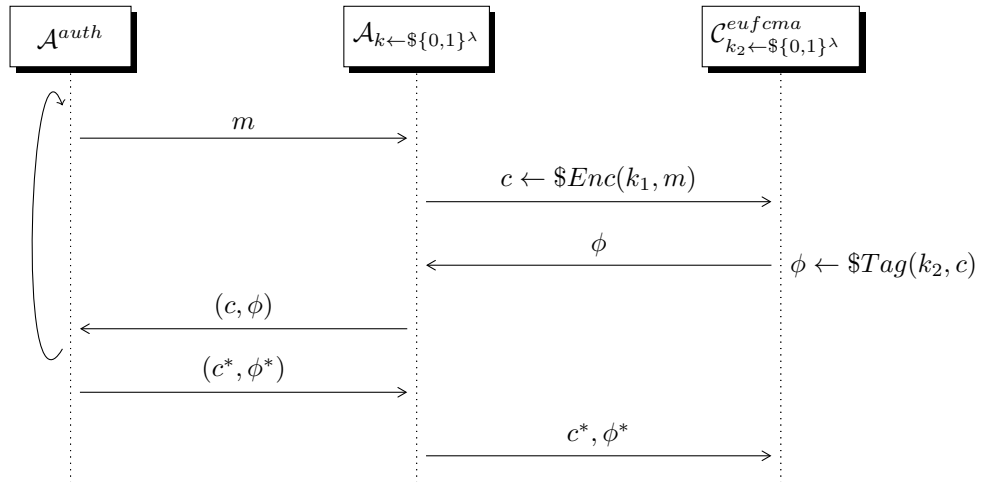
The difference is that in **ufcma** we didn't care about the freshness of ϕ^* .

Proof. Suppose Π has not the *auth* property.

So we have an \mathcal{A}' which can win the **auth** challenge of Π .

On the other hand, we have a Π_2 schema which uses an **eu-f-cma** $Tag(.,.)$ function.

So, by reduction, we show that



The game returns 1 IFF

$$c^* \leftarrow Enc(k_1, .) \wedge \phi^* \leftarrow Tag(k_2, c^*) \wedge (c^*, \phi) \notin \{(c, \phi)\}$$

From \mathcal{A}^{auth} perspective, all the couples (c_i, ϕ_i) received are made with the following schema:

$$c_i \in Enc(k_1, m \in \mathcal{M}) \wedge \phi_i \leftarrow \$Enc(k_2, c_i)$$

Since \mathcal{A}^{auth} wins $Game^{auth}$, the challenge couple (c^*, ϕ^*) which breaks $Game^{auth}$ will be produced to be decrypted as

$$Dec(k, (c^*, \phi^*)) \rightarrow Dec(k_1, c^*) \in \mathcal{M} \wedge Dec(k_2, \phi^*) = c^*$$

But if this happens, then \mathcal{A} can use the same challenge couple of \mathcal{A}^{auth} to win $Game^{ufcma}$, which is impossible.

It could happen that, for $c^* = c$ previously seen, ϕ^* is a new fresh tag, never seen before. Just in this case the *auth* game would be valid because (c^*, ϕ^*) would have never been seen before, but **not** the *ufcma* game, because c^* was previously sent to the challenger.

□

Now we want an *ufcma* secure scheme able to resist against message-tag challenge couples where the tag is fresh but the message has been already requested to the challenger.

4.18 Pseudorandom permutations

Pseudorandom permutations are like PRFs, but efficiently invertible. Consider the following family of functions:

$$\mathcal{F} = \{F_k : \{0, 1\}^n \rightarrow \{0, 1\}^n\}_{k \in \{0, 1\}^\lambda}$$

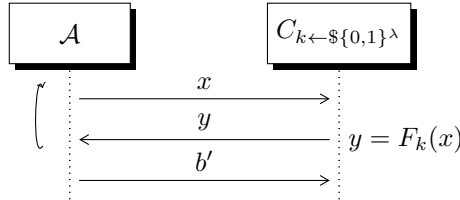


Figure 4.32: $Real_{\mathcal{F}, \mathcal{A}}(\lambda)$

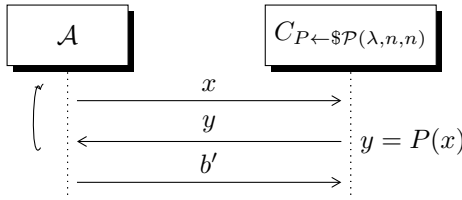


Figure 4.33: $Ideal_{\mathcal{P}, \mathcal{A}}(\lambda)$

and the two games are indistinguishable

$$Real_{\mathcal{F}, \mathcal{A}}(\lambda) \approx_c Ideal_{\mathcal{F}, \mathcal{A}}(\lambda)$$

4.18.1 Feistel Network

Let $F : \{0, 1\}^n \rightarrow \{0, 1\}^n$ and ψ_F the invertible function

$$\begin{aligned}\psi_F(\overbrace{x}^{\text{n bits}}, \overbrace{y}^{\text{n bits}}) &= (y, x \oplus F(y)) = (x', y') \\ \psi_F^{-1}(\overbrace{x'}^{\text{n bits}}, \overbrace{y'}^{\text{n bits}}) &= (F(x') \oplus y', x') = (x, y)\end{aligned}$$

Is this function pseudorandom?

This is not pseudorandom, because the first n bits of the output of ψ_F are always equal to y , while in a PRF the probability that, given two different (x, y) and (x', y) in input, the first bits are equal is very low.

=====

FEISTEL IMAGE

=====

The $l - th$ level outputs something like

$$\psi_F[l](x, y) = \psi_{F_{k_l}}(\psi_{F_{k_{l-1}}}(\dots(\psi_{F_{k_1}}(x, y))\dots))$$

Two XORed rounds of this function don't create a PRP. In particular, imagine 2 queries (x, y) and (x', y) such that

$$\psi_{F, F'}(x, y) \oplus \psi_{F, F'}(x', y) = (x \oplus F(y) \oplus x' \oplus F(y), \dots)$$

Since for 2 random queries with the same y , the first member of the output is always equal to $x \oplus x'$ with probability 1, this XORed rounds cannot constitute a PRP (which, instead, for 2 queries with the same second member outputs a first member equal to $x \oplus x'$ with negligible probability).

Lemma 19. *For every **unbounded** distinguisher making $q \in \text{poly}(\lambda)$ queries, the following are statistically close as long as y_1, \dots, y_q are **y-nique**, i.e. $\forall i \neq j, y_i \neq y_j$* \diamond

figures

Lesson 15

4.19 Public key encryption recap

$\text{Game}_{\Pi, \mathcal{A}}^{\text{pke-cca}}$:

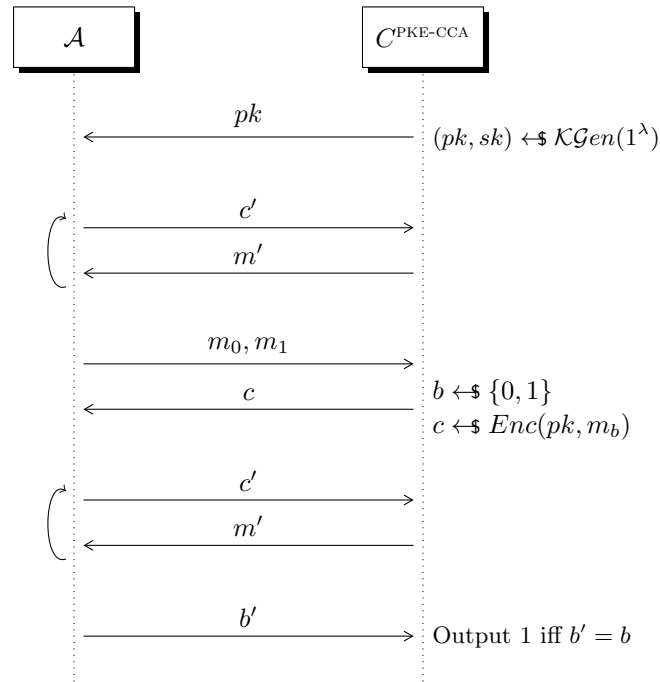


Figure 4.34: CCA on a PKE scheme

(Reminder): Encryptions are made like this: $\text{Enc}(k, m) = (r, F_k(r) \oplus m)$, $r \simeq \text{Unif}(\{0, 1\}^\lambda)$.

Every time an encryption is made, a fresh value r is picked UAR.

4.19.1 Trapdoor permutation

A TDP (TrapDoor Permutation) is a OWP family structured has these features:

- A key pair is chosen UAR by a key generator algorithm: $(pk, sk) \leftarrow \mathcal{KGen}(1^\lambda)$
- There is a function family $f_{pk} \subseteq (V_{pk} \rightarrow V_{pk})$ such that:
 - Computing f_{pk} is efficient
 - Sampling from the domain ($x \leftarrow V_{pk}$) is efficient
- There is an efficient function g_{sk} that “inverts” f_{pk} (sk is the “trapdoor”):

$$g(sk, f(pk, x)) = x$$

- No efficient adversary is able to invert f_{pk} without knowing sk

Note: Since pk is public, any adversary gets the capability of encrypting messages for free, without requiring an external challenger/oracle!

Therefore, if left deterministic, a TDP is not CPA-secure.

Here, in this scheme, we combine randomness and the notion of hardcore predicate \mathfrak{hc} :

- $(pk, sk) \leftarrow \mathcal{KGen}(1^\lambda)$
- $r \leftarrow \Xi_{pk}$
- $c := \text{Enc}(pk, m) = (f_{pk}(r), \mathfrak{hc}(r) \oplus m)$
- Correctness: $\text{Dec}(sk, c) = \mathfrak{hc}(g_{sk}(c_1)) \oplus c_2$

Theorem: if (\mathcal{KGen}, f, g) is a TDP, and \mathfrak{hc} is hardcore for f , then the above scheme is CPA-secure.

Proof: (Exercise)

4.19.2 TDP examples

One example stems from the factoring problem: let's look again at \mathbb{Z}_n^\times , where $n = pq$, $p, q \in \mathbb{P}$;

Theorem (Chinese remainder, or CRT): The following isomorphisms to \mathbb{Z}_n^\times are true:

- $\mathbb{Z}_n \simeq \mathbb{Z}_p \times \mathbb{Z}_q$
- $\mathbb{Z}_n^\times \simeq \mathbb{Z}_p^\times \times \mathbb{Z}_q^\times$

Note that the theorem is more general, and holds for any p, q that are co-prime.

How to use this theorem for constructing a PKE scheme:

Reminder (Euler's theorem): $\forall x \in \mathbb{Z}_n \implies x^{\varphi(n)} = x \pmod n$

Reminder: $\forall p, q \in \mathbb{P} \implies \varphi(pq) = (p-1)(q-1)$

So let a be the public key such that $\gcd(a, \varphi(n)) = 1$, then $\exists! b \in \mathbb{Z}_n : ab = 1 \pmod{\varphi(n)}$, d will be our private key.

Define encryption as $f(a, m) = m^a \bmod n$, and then decryption as $g(b, c) = c^b \bmod n$

Observe that $g(b, f(a, m)) = (m^a)^b = m^{ab} = m \bmod n$, because $ab = 1 \bmod \varphi(n)$

So we conjecture that the above is a valid TDP-based PKE scheme. This is actually called the “RSA assumption”:

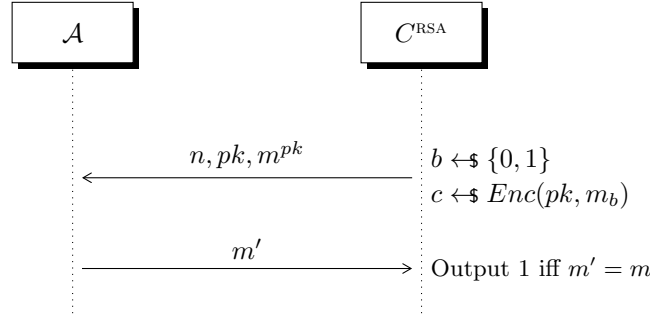


Figure 4.35: Depiction of the RSA assumption

Relation to the factoring problem: $\text{RSA} \implies \text{FACT}$

Proof: Given p, q , an adversary can compute $\varphi(n) = (p-1)(q-1)$, and then find the inverse of the public key in \mathbb{Z}_{pq}^\times .

It hasn't been proven that $\text{FACT} \implies \text{RSA}$

4.20 Textbook RSA

This is an insecure toy example of the more complex *RSA* (Rivest Shamir Adleman) algorithm. The key generation algorithm: $\text{KGen} = \text{GenRSA}(1^\lambda)$ outputs $P_k = (n, e)$ and $S_k = d$, then we have

$$\text{Enc}(P_k, m) = m^e \bmod n$$

$$\text{Dec}(S_k, c) = c^d \bmod n$$

Since the output of Enc is deterministic this is **not CPA secure**! However it can be used with HARD-CORE Predicate.

Preprocess the message to add randomness:

$$\hat{m} = r || m \text{ where } r \leftarrow \$_{\{0,1\}^l}$$

now Enc is not deterministic.

Facts:

1. $l \in \text{super}(\log(\lambda))$ otherwise it is possible to bruteforce in PPT.
2. If $m \in \{0, 1\}$ then I can prove it CPA secure under RSA (just use standard TDP)
3. If m is "in the middle" ($\{0, 1\} \leq m \leq \{0, 1\}^l$) RSA is believed to be secure and is standardized (PKCS#1,5)
4. Still not CCA secure!

4.20.1 Trapdoor Permutation from Factoring

Let's look at $f(x) = x^2 \pmod{n}$ where $f : \mathbb{Z}_n^* \rightarrow \mathbb{QR}_n(\subset \mathbb{Z}_n^*)$, this is not a permutation in general.

Now let's consider the Chinese Remainder Theorem (CRT) representation:

$$\begin{aligned} x &= (x_p, x_q) \rightarrow x_p \equiv x \pmod{p}, x_q \equiv x \pmod{q} \\ f(x) &= x^2 \pmod{p}; x \leftarrow \$_{Z_p^*} \end{aligned}$$

Since Z_p^* is cyclic I can always write:

$$\begin{aligned} Z_p^* &= \{g^0, g^1, g^2, \dots, g^{\frac{p-1}{2}-1}, g^{\frac{p-1}{2}}, \dots, g^{p-2}\} \\ \mathbb{QR}_p &= \{g^0, g^2, g^4, \dots, \underbrace{g^{p-3}}_{g^{\frac{p+1}{2}-1} \text{ in } Z_p^*}, \underbrace{g^0}_{g^{\frac{p-1}{2}} \text{ in } Z_p^*}, \dots\} \\ |\mathbb{QR}_p| &= \frac{p-1}{2} \end{aligned}$$

Moreover since $(g^{\frac{p-1}{2}})^2 \equiv 1 \pmod{p}$ then $g^{\frac{p-1}{2}} \equiv -1 \pmod{p}$.

Now assume $p \equiv 3 \pmod{4}$ ($[*]p = 4t + 3$ CRT), then squaring $\text{Mod } p$ is a permutation because, given $y = x^2 \pmod{p}$ if I compute:

$$\begin{aligned} (y^{t+1})^2 &= \underbrace{y^{2t+2}}_{[*] \ 2t+2 = \frac{p-3}{2} + 2 = \frac{p+1}{2} = \frac{p-1}{2} + 1} = (x^2)^{\frac{p-1}{2}+1} = 1x^2 = x^2 \\ &\implies x = \pm y^{t+1} \end{aligned}$$

But only 1 among the above $\pm y^{t+1}$ is a square, this is because $\frac{p-1}{2}$ is odd.

Lemma 20. $\forall z, z \in \mathbb{QR}_p \text{ IFF } -z \notin \mathbb{QR}_p$ \diamond

=====

4.20.2 Rabin's Trapdoor permutation

Now we study a one way function built on previous deductions about number theory and modular arithmetic.

The *Rabin trapdoor permutation* is defined as

$$f : \mathbb{Z}_n^* \rightarrow \mathbb{Z}_n^* f(x) = x^2 \pmod{n}$$

where $n = p * q$ for primes $p, q \equiv 3 \pmod{4}$.

T0D0 4: label related to a previous construction

We can observe that the image of this function is a subset of \mathbb{Z}_n^* .

For the **Chinese remainder theorem**, we can consider values modulo $N = pq$ as values modulo p and values modulo q .
So now consider

$$f : \mathbb{Z}_p^* \rightarrow \mathbb{Z}_p^* f(x) = x^2 \bmod p$$

The image of the function f exactly matches the definition of **quadratic residues**, since

$$f\text{'s image set} = \mathbb{QR}_p = \{y : y \equiv x^2 \bmod p, x \in \mathbb{Z}_p^*\}$$

We can observe the same for \mathbb{Z}_q^* .

So for **Chinese remainder theorem** it is possible to state that f maps as follows

$$x = (x_p, x_q) \mapsto (x_p^2, x_q^2)$$

As before, the image of f is exactly

$$\mathbb{QR}_n = \{y : \exists x : y = x^2 \bmod n\}$$

For the previous observations, it's possible to state the following **Fact**:

$$y \leftarrow \mathbb{QR}_n \Leftrightarrow y \in \mathbb{QR}_p \wedge y \in \mathbb{QR}_q$$

This is important because if we try to invert the function f , among the hypothetical 4 possible values

$$f^{-1}(y) = \{(x_p, x_q), (-x_p, x_q), (x_p, -x_q), (-x_p, -x_q)\}$$

only 1 out of these above 4 values is a quadratic residue because only one of $-x_k, x_k$ is a quadratic residue for $k = q, p$.

T0D0 5: label to previous fact

Therefore, we have that the Rabin's TDP is a permutation, and that the cardinality of \mathbb{QR}_n is $\frac{|\mathbb{Z}_n^*|}{4}$.

Furthermore, with the following claim we can state that the Rabin cryptosystem is as secure as the hardness of factoring.

Claim 1. *Given x, z such that $x^2 \bmod n \equiv z^2 \bmod n \equiv y \bmod n$,*

$$x \neq \pm z \Rightarrow \text{we can factor } n$$

◇

Proof. For the previous statement 4.20.2, since $f^{-1}(y)$ has only one value out of four, summing two distinct elements x, z in the domain of function f generates only 2 possible values:

$$x + z \in \{(0, 2x_q), (2x_p, 0)\}$$

Now assume $x + z = (2x_p, 0)$ without loss of generality, since the proof for the other case is the same.

We have that $x + z \equiv 0 \pmod{q}$ and $x + z \not\equiv 0 \pmod{p}$.

But then $\gcd(x + z, n) = q$, and we obtain q . \square

Theorem 27. *Squaring mod n (where n is a bloom integer⁴) is a **trapdoor permutation** under factoring.* \diamond

In other words

Factoring is hard \Rightarrow inverting $f(x)$ is hard

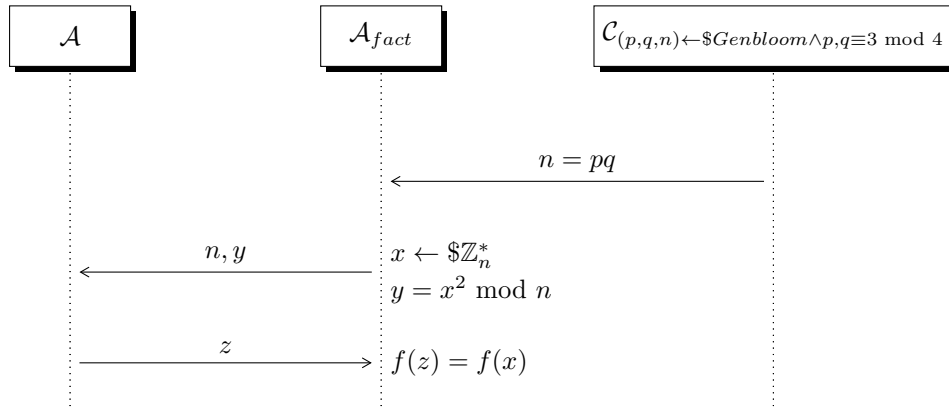
The following proof is by contraddiction.

Proof. Assume that exists an adversary PPT such that

$$\mathcal{P}[x^2 = y : (p, q, n) \leftarrow \text{\$Genbloom}(1^\lambda), y \leftarrow \text{\$QR}_n, x \leftarrow \mathcal{A}(y)] \geq \frac{1}{\text{poly}(\lambda)}$$

But then we can build the following reduction:

T0D0 6: check the correctness of the image



\square

⁴ a bloom integer n is $n = p, q$ for $p, q = 3 \pmod{4}$, as the definition of Rabin's TDP

Lesson 16

4.21 PKE schemes over DDH assumption

4.21.1 El Gamal scheme

Let's define a new $\Pi = (\text{KGen}, \text{Enc}, \text{Dec})$. Generate the needed (**public**) parameters $(G, g, q) \leftarrow \text{GroupGen}(1^\lambda)^5$.

The KeyGen algorithm is defined as follows:

- Pick $x \leftarrow \mathbb{Z}_q$
- Output the key pair (pk, sk) as (g^x, x)

The encryption routine $\text{Enc}(pk, m)$ will:

- Pick $r \leftarrow \mathbb{Z}_q$
- Output $c = (c_1, c_2) = (g^r, pk^r \cdot m)^6$

The decryption routine $\text{Dec}(sk, c)$ will:

- Compute $\hat{m} = c_1^{-sk} \cdot c_2$

Correctness of this scheme follows from some algebraic steps:

$$\begin{aligned}\hat{m} &= \text{Dec}(sk, \text{Enc}(pk, m)) \\ &= \text{Dec}(x, \text{Enc}(g^x, m)) \\ &= \text{Dec}(x, (g^r, (g^x)^r \cdot m)) \\ &= (g^r)^{-x} \cdot (g^x)^r \cdot m \\ &= m\end{aligned}$$

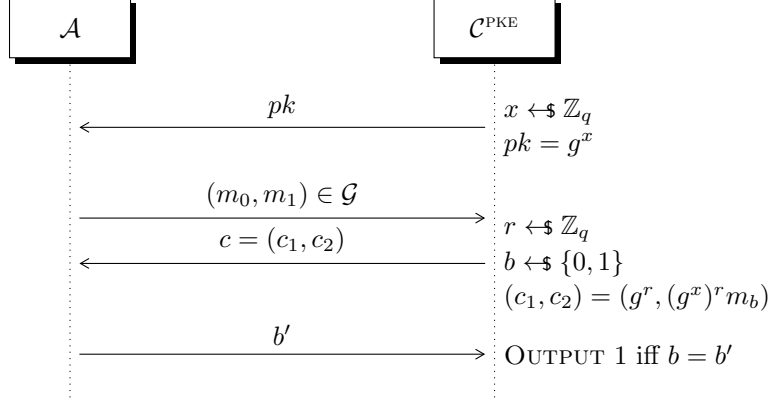
Theorem 28. *Assuming DDH, the El Gamal scheme is CPA-secure.* \diamond

Proof. Consider the two following games $H_0(\lambda, b)$ and $H_1(\lambda, b)$ defined as follows. Observe that b can be fixed without loss of generality.

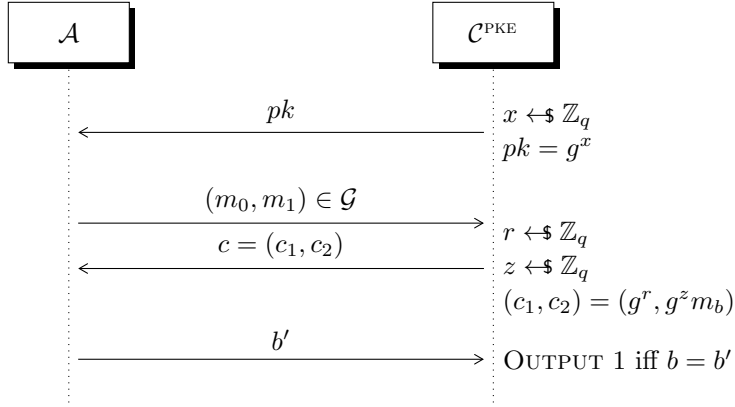
⁵G could be any "valid" group such as \mathbb{QR}_p or an Elliptic Curve

⁶We need r because we want to re-randomize c

$H_0(\lambda, b) :$



$H_1(\lambda, b) :$



Note: it is important to note that we can measure the advantage of \mathcal{A} , so fixed its output $Adv_{\mathcal{A}}(\lambda) = \left| \underbrace{Pr[\mathcal{A} \rightarrow 1 | b = 0]}_{\text{"A loses"}} - \underbrace{Pr[\mathcal{A} \rightarrow 1 | b = 1]}_{\text{"A wins"}} \right|$. Since b is fixed the above formula will give a value λ *innegl*, generally the advantage of an adversary is: $\frac{1}{2} + \lambda$ (random guessing + a negligible factor).

Proof technique: $H_0(\lambda, 0) \approx_c H_0(\lambda, 1) \equiv H_1(\lambda, 0) \approx_c H_1(\lambda, 1)$

$$\implies H_0(\lambda, 0) \approx_c H_1(\lambda, 1)$$

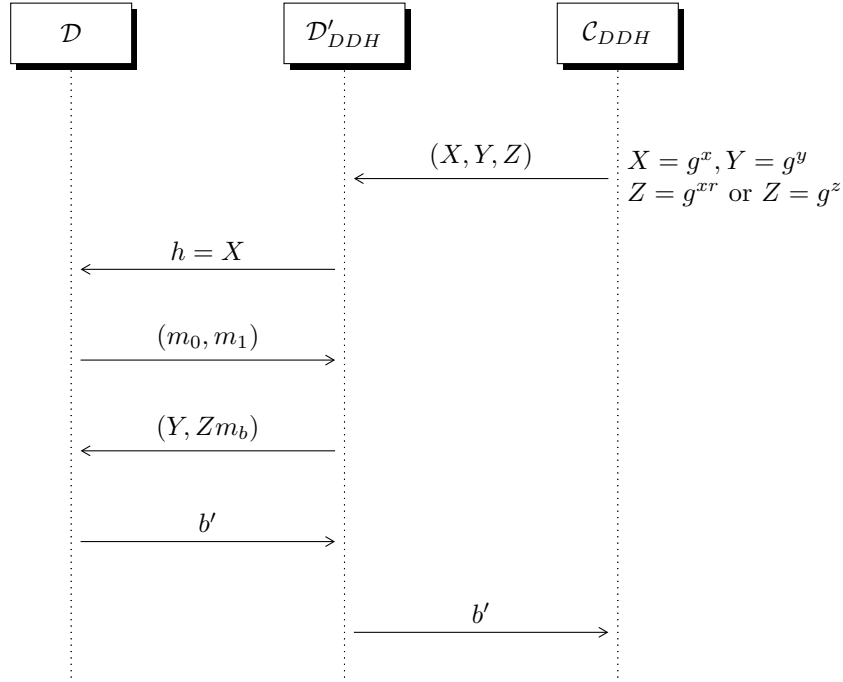
Lemma 21. $\forall b \in \{0, 1\}, H_0(\lambda, 0) \approx_c H_0(\lambda, 1)$

Fix b . (Reduction to DDH)

Assume \exists PPT D which is able to distinguish $H_0(\lambda, b)$ and $H_1(\lambda, b)$ with non negl. probability. \diamond

Proof. Consider the following Game:

Contradiction: \mathcal{D} should be able to compute \log_g to distinguish the message. \square



Lemma 22. $H_1(\lambda, 0) \equiv H_1(\lambda, 1)$ \diamond

Proof. This follows from the fact that: $(g^x, (g^r, g^z m_0)) \equiv (g^x, (g^r, U_\lambda) \equiv (g^x, (g^r, g^z m_1)))$ \square

Lemma 23. $H_1(\lambda, 1) \equiv H_0(\lambda, 1)$ \diamond

This is proved in the exact same way as **Lemma 20**. As a matter of fact it is the second part of the proof (where b is fixed to 1). \square

Properties of of El Gamal PKE scheme

Some useful observations can be made about this scheme:

- It is **homomorphic**: Given two ciphertexts (c_1, c_2) and (c'_1, c'_2) , then doing the product between them yields another valid ciphertext:

$$\begin{aligned} & (c_1 \cdot c'_1, c_2 \cdot c'_2) \\ &= (g^{r+r'}, h^{r+r'}(m \cdot m')) \end{aligned}$$

thus, decrypting $c \cdot c'$, gives $m \cdot m'$.

- It is **re-randomizable**: Given a ciphertext (c_1, c_2) , and $r' \leftarrow \mathbb{Z}_q$, then computing $(g^{r'} \cdot c_1, h^{r'} \cdot c_2)$ results in a “fresh” encryption for the same message: the random value used at the encryption step will change from the original r to $r + r'$

These properties of the El Gamal scheme can be desirable in some use cases, where a message must be kept secret to the second party. In fact, there are some PKE schemes which are designed to be **fully homomorphic**, i.e. they are homomorphic for any kind of function.

Consider the following use case: a client C has an object x and wants to apply a function f over it, but it lacks the computational power to execute it. There is another subject S , which is able to efficiently compute f , so the goal is to let it compute $f(x)$ but the client wishes to keep x secret from him. This can be achieved using a FH-PKE scheme as follows:

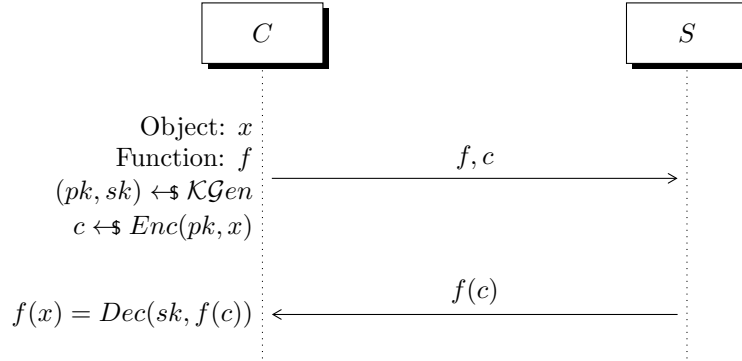


Figure 4.36: Delegated secret computation

However one important consideration must be made: All these useful characteristics expose an inherent malleability of any fully homomorphic scheme: any attacker can manipulate ciphertexts efficiently, and with some predictable results. This compromises even CPA security of such schemes.

4.21.2 Cramer-Shoup PKE scheme

This scheme is based on the standard DDH assumption, and has the advantage of being CCA secure. A powerful tool, called **Designated Verifier Non-Interactive Zero-Knowledge**, or alternatively **Hash-Proof System**, is used here.

Proof systems

Let L be a Turing-recognizable language in NP , and a predicate $\mathcal{R} \in X \times Y \rightarrow \{0, 1\}$ such that:

$$L := \{y \in Y : \exists x \in X \mathcal{R}(x, y) = 1\}$$

where x is called a “witness” of y .

In our instance, let $y = pq, x = (p, q)$

$\Pi = (\text{Setup}, \text{Prove}, \text{Verify})$

$(\omega, \tau) \leftarrow \text{Setup}(1^\lambda)$, where ω is the **Common Reference String**, and τ is the **trapdoor**.

Additional notes:

- ω is public ($= pk$)
- τ is part of the secret key

- $\tau = (x, y) : \mathcal{R}(x, y) = 1$
- There is presumably a common third-party, which samples from the setup and publishes ω , while giving τ to only B.

T0D0 7: B operations in the image

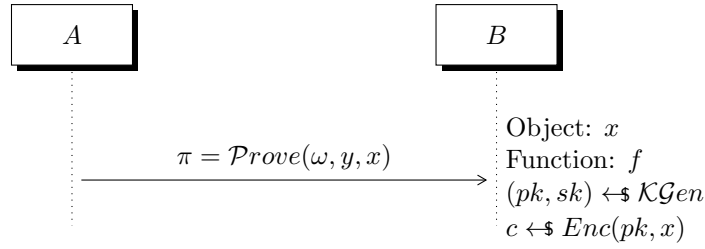


Figure 4.37: Overview of Cramer-Shoup operation

Proof system - purpose: a way to convince B that A knows something
 Can compute the proof in two different ways, this is the core notion of ZK
 No $\tau \implies ZK$

Properties

- (*implicit, against malicious Bob*) **Zero-knowledge**: Proof for x can be simulated without knowing x itself
- (*stronger, against malicious ALice*) **Soundness**: It is hard to produce a valid proof for any $y \notin L$
- *honest people* **Completeness**: $\forall y \in L, \forall (\omega, \tau) \leftarrow \mathcal{Setup}(1^\lambda) :$
 $\text{Prove}(\omega, x, y) = \text{Verify}(\tau, y)$

T0D0 8: to review and understand/better

t-universality

Definition 9. Let Π be DV-NIKZ⁷.

We say it is *t-universal* if for any distinct

$$y_1, \dots, y_t \text{ s.t. } y_i \notin L (\forall i \in [t])$$

we have

$$(\omega, \text{Ver}(\tau, y_1), \dots, \text{Ver}(\tau, y_t)) = (\omega, v_1, \dots, v_t)$$

where $(\omega, \tau) \leftarrow \mathcal{Setup}(1^\lambda)$ and $v_1, \dots, v_t \leftarrow \mathcal{P}$ where \mathcal{P} should be the proofs' space. \diamond

⁷Designated verifier non-interactive zero-knowledge

Enriching DV-NIKZ

Can we enrich DV-NIKZ with labels $l \in \{0, 1\}^*$?

Suppose to have the following:

$$L' = L \parallel \{0, 1\}^* = \{(y, l) : y \in L \wedge l \in \{0, 1\}^*\}$$

Then our scheme changes : $Prove(\omega, (y, l), x) = \Pi; Ver(\tau, (y, l))$ and , for t -*universality* , now we can consider 2 distinct (y_i, l_i) .

Membership Hard Language (MH)

Definition 10. Language L is **MH** if $\exists \bar{L}$ such that:

1. $L \cap \bar{L} = \emptyset$
2. \exists PPT $Samp$ outputting $y \leftarrow \$\mathcal{Y}$ together with $x \in \bar{\mathcal{X}}$ such that

$$R(y, x) = 1$$

(it's possible to say that $Samp(1^\lambda) \leftarrow \\mathcal{Y})

3. \exists PPT \bar{Samp} outputting $y \leftarrow \$\bar{L}$
4. $\{y : (y, x) \leftarrow \$Samp(1^\lambda)\} \approx_c \{y : y \leftarrow \$\bar{Samp}(1^\lambda)\}$

◇

Lesson 17

4.22 Construction of a CCA-secure PKE

This section exposes a construction of a CCA-secure PKE scheme, using hash-proof systems, membership-hardness, and the n -universality property.

Let Π_1, Π_2 be two distinct hash-proof systems for some NP language L and the range of $Prove_2$ supports labels ($L' = L || \{0, 1\}^\ell$).

Construct the CCA scheme as follows: $\Pi := (\mathcal{KGen}, Enc, Dec)$

- $(\overbrace{(\omega_1, \omega_2)}^{pk}, \overbrace{(\tau_1, \tau_2)}^{sk}) \leftarrow \mathcal{KGen}(1^\lambda)$, $(\omega_1, \tau_1) \leftarrow \mathcal{Setup}_1(1^\lambda), (\omega_2, \tau_2) \leftarrow \mathcal{Setup}_2(1^\lambda)$
- Encryption routine: $Enc((\omega_1, \omega_2), m)$
 - $(y, x) \leftarrow \mathcal{Sample}_1(1^\lambda)$
 - $\pi_1 \leftarrow \mathcal{Prove}_1(\omega_1, y, x)$
 - $l := \pi_1 \cdot m$
 - $\pi_2 \leftarrow \mathcal{Prove}_2(\omega_2, (y, l), x)$
 - $c := (c_1, c_2) = ((y, l), \pi_2)$
- Decryption routine: $Dec((\tau_1, \tau_2), (c_1, c_2))$
 - $\hat{\pi}_2 = \mathcal{Verify}_2(\tau_2, c_1)$
 - IF $\hat{\pi}_2 \neq c_2$ THEN OUTPUT FALSE
 - Recall: $c_1 = (y, l)$
 - $\hat{\pi}_1 = \mathcal{Verify}_1(\tau_1, y)$
 - OUTPUT $l \cdot \hat{\pi}_1^{-1}$

Correctness (assume $\hat{\pi}_i = \pi_i \forall i$):

$$\begin{aligned}
 \hat{m} &= Dec(sk, Enc(pk, m)) \\
 &= Dec((\tau_1, \tau_2), Enc((\omega_1, \omega_2), m)) \\
 &= Dec((\tau_1, \tau_2), ((y, l), \pi_2)) \\
 &= l \cdot \hat{\pi}_1^{-1} \\
 &= \pi_1 \cdot m \cdot \hat{\pi}_1^{-1} \\
 &= m
 \end{aligned}$$

Some additional notes (may be incorrect):

- The message space of the second prover is the range of the first prover.
- The message space of the first prover is a multiplicative group
- The message space of the second prover is a polylogarithmic language in (λ)

Theorem 29. *Assuming π_1 is 1-universal, π_2 is 2-universal and L is Membership-hard, then the above scheme is CCA-secure.* \diamond

Proof. Five different games will be defined, from $\text{GAME}_{\Pi,A}^0$ up to $\text{GAME}_{\Pi,A}^4$; the first game will be an analogous formalization of how the above PKE scheme works. It shall be proven that, for either fixed b in $\{0, 1\}$:

$$\text{GAME}_{\Pi,A}^0(\lambda, b) = \text{GAME}_{\Pi,A}^1(\lambda, b) \approx_C \text{GAME}_{\Pi,A}^2(\lambda, b) \approx_S \text{GAME}_{\Pi,A}^3(\lambda, b) = \text{GAME}_{\Pi,A}^4(\lambda, b)$$

and finally that $\text{GAME}_{\Pi,A}^4(\lambda, 0) = \text{GAME}_{\Pi,A}^4(\lambda, 1)$, therefore concluding that $\text{GAME}_{\Pi,A}^0(\lambda, 0) \approx_C \text{GAME}_{\Pi,A}^0(\lambda, 1)$, and proving this scheme is CCA-secure.

T0D0 9: Non sono per niente sicuro riguardo all'origine di x ed y , né tantomeno dove sia definito il sampler per essi

The games are defined as follows:

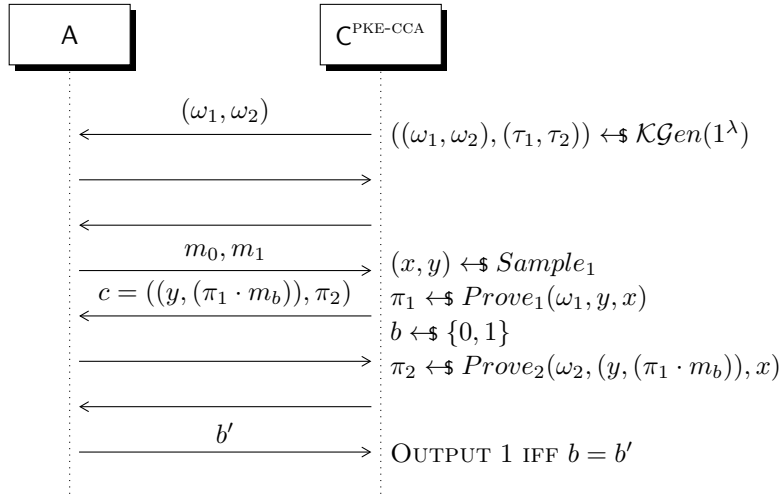


Figure 4.38: Original CCA game

T0D0 10: Non è stato chiaro sull'origine di x ed y , inoltre mi manca da scrivere le query di decifratura

□

Lemma 24.

$$\forall b, G_0(\lambda, b) \equiv G_1(\lambda, b)$$

◇

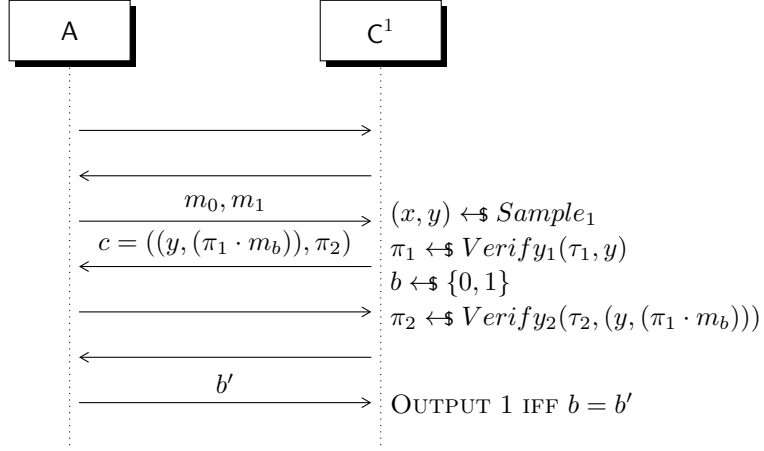


Figure 4.39: Use verifiers

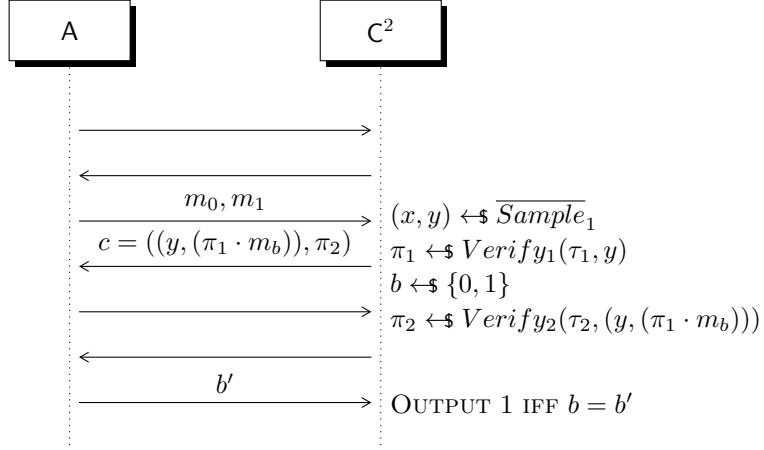


Figure 4.40: Sample statements outside the language

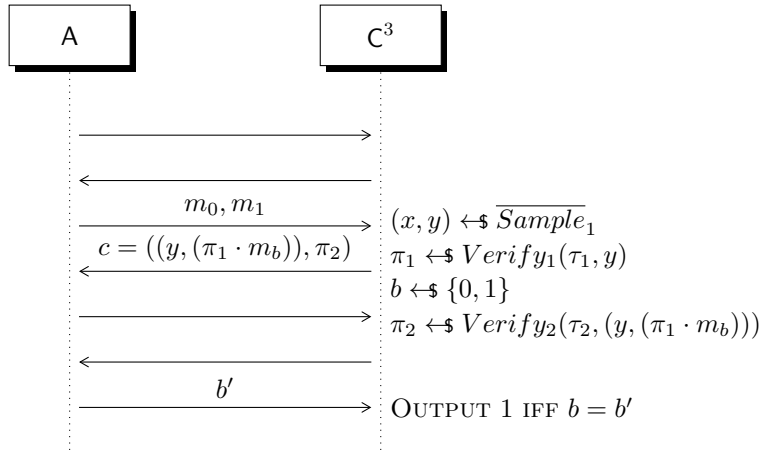


Figure 4.41: Modify decryption queries

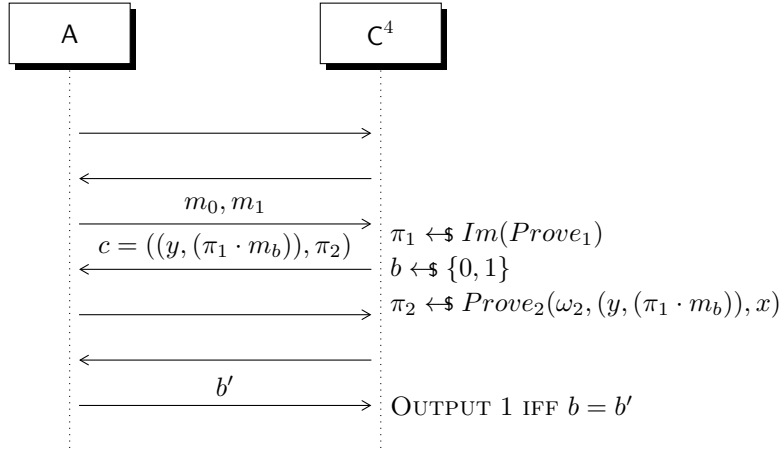


Figure 4.42: π_1 is chosen UAR

Proof. This follows by the correctness of Π_1 and Π_2

$$\Pi_1 = \tau_1 = \text{Ver}_1(\tau, y)$$

$$\Pi_2 = \tilde{\Pi}_2 = \text{Ver}_2(\tau, y)$$

with probability 1 over the choice of $(\omega_1, \tau_1) \leftarrow \text{\$Setup}_1(1^\lambda)$

$$\Pi_1 \leftarrow \text{Prove}_1(\omega, y, x), (\omega_2, \tau_1) \leftarrow \text{\$Setup}_2(1^\lambda)$$

$$\Pi_2 \leftarrow \text{Prove}(\omega_2, (y, l), x) \forall y \in L, l \in \mathcal{P}_1$$

□

Lemma 25.

$$\forall b, G_1(\lambda, b) \approx_c G_2(\lambda, b)$$

◇

Proof. Straight forward reduction from membership hardness.

□

Lemma 26.

$$\forall b, G_2(\lambda, b) \approx_c G_3(\lambda, b)$$

◇

T0D0 11: I'm not completely sure the next proof is complete

Proof. Recall that the difference between G_2 and G_3 is that

$$(g^{(i)}, l^{(i)}, \Pi_2^{(i)}) \text{ such that } y^{(i)} \notin L$$

are answered \perp in G_3 , instead in G_2

$$\perp \text{ comes out as output} \Leftrightarrow \tilde{\Pi}_2^{(i)} = \text{Ver}(\tau, y^{(i)} \neq \Pi_2^{(i)})$$

It's possible to distinguish **two cases**, looking at $c = (y, l, \Pi_2)$:

1. if $(y^{(i)}, l^{(i)}) = (y, l)$ and $\tilde{\Pi}_2^{(i)} = \Pi_2^{(i)}$, it outputs \perp if in the decryption scheme $(\Pi_2^{(i)} \neq \tilde{\Pi}_2^{(i)})$ it outputs \perp .
2. ⁸ otherwise $(y^{(i)}, l^{(i)}) \neq (y, l)$ if $y^{(i)} \notin L$ we want that $\Pi_2^{(i)}$ doesn't output exactly $Ver_2(\tau, (y^{(i)}, l^{(i)}))$, but it should output \perp .

EVENT BAD: If we look at the view of \mathcal{A} , the only information he knows is

$$(\omega_2, \tilde{\Pi}_2 = Ver_2(\tau_2, y))$$

for $y \notin L$.

The value

$$Ver_2(\tau_2, (y^{(i)}, x^{(i)}))$$

for $y^{(i)} \in L$ and $(y^{(i)}, l^{(i)}) \neq (y, l)$ is random.

So,

$$\mathcal{P}[BAD] = 2^{-|\mathcal{P}_2|}$$

□

Lemma 27.

$$\forall b, G_3(\lambda, b) \equiv G_4(\lambda, b)$$

◇

Proof. If we look at the view of \mathcal{A} , the only information known about τ_1 is ω_1 , since the decryption oracle only computes for $y^{(i)} \in L$

$$Ver_1(\tau, y^{(i)}) = Prove_1(\omega_1, y^{(i)}, x^{(i)})$$

By 1-universality, $\Pi = Ver_1(\tau, y)$ for any $y \in L$ is random.

□

Lemma 28.

$$G_4(\lambda, 0) \equiv G_4(\lambda, 1)$$

◇

Proof. The challenge ciphertext is independent of b .

□

REVIEW: referencing something from another part of lesson 17

4.22.1 Instantiation of U-HPS (Universal Hash Proof System)

MHL(Membership Hard Language) from DDH

$\exists r$ using a DDH language such that $L_{DDH} = \{(c_1, c_2), \exists r | c_1 = g_1^r, c_2 = g_2^r\}$

Given a group \mathcal{G} of order q with (g_1, g_2) as generators, we will have (g_1, g_2, c_1, c_2) but if we impose $g_1 = g$ and $g_2 = g^a$ then the previous construction becomes (g, g^a, c_1, c_2) . But for definition $c_1 = g_1^{r_1}$ and $c_2 = g_2^{r_2}$ then I can write $(g, g^a, g^{r_1}, g^{ar_2})$.

Now we can define our U-HPS $\Pi := (\text{Setup}, \text{Prove}, \text{Verify})$

⁸when decryption oracle doesn't output the challenge

- *Setup*(1^λ): Pick $x_1, x_2 \leftarrow \mathcal{Z}_q$ and define:
 $\omega = h_1 = (g_1^{x_1}, g_2^{x_2})$, $\tau = (x_1, x_2)$
- *Prove*($\omega, \underbrace{(c_1, c_2)}_y, r$) output $\Pi = \omega^2$
- *Verify*($\tau, \underbrace{(c_1, c_2)}_y$) output $\tilde{\Pi} = c_1^{x_1} c_2^{x_2}$

Correctness: $\Pi = \omega^2 = (g_1^{x_1} g_2^{x_2})^r = g_1^{rx_1} g_2^{rx_2} = c_1^{x_1} c_2^{x_2} = \tilde{\Pi}$

Theorem 30. *Above construction defines a 1-universal DVNIZK for L_{DDH}* \diamond

Proof. We want to prove that if we take any $(c_1, c_2) \notin L_{DDH}$ the distribution $(\omega = h_1, \tilde{\Pi} = \text{Verify}(\tau, (c_1, c_2)))$ *random* uniformly distributed.

Define a **MAP** $\mu(\underbrace{x_1, x_2}_{\text{random}}) = (\omega, \Pi) = (g_1^{x_1}, g_2^{x_2}, c_1^{x_1}, c_2^{x_2})$ it suffices to prove that μ is injective. This can easily be done with some constrains: $\mu'(x_1, x_2) = \log_{g_1}(\mu(x_1, x_2)) = (\log_{g_1}(\omega), \log_{g_1}(\Pi))$
For $r_1 \neq r_2$ then $c_1 = g_1^{r_1}, c_2 = g_2^{r_2} = g^{\alpha r_2}$. For $\alpha = \log_{g_2} g_1$ then $\Pi = c_1^{x_1} c_2^{x_2} = g_1^{r_1 x_1} g_2^{r_1 x_1 + \alpha r_2 x_2}$.

$$\mu'(x_1, x_2) = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ r_1 & r_2 \alpha \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Since $\text{Det} \begin{pmatrix} 1 & \alpha \\ r_1 & r_2 \alpha \end{pmatrix} = \alpha(r_2 - r_1) \neq 0$ the map is injective.

□

- *Setup*(1^λ):
 - Pick $x_3, x_4, x_5, x_6 \leftarrow \mathcal{Z}_q$ and define:
 - * $\omega = (h_2, h_3, s) = (g_1^{x_3}, g_2^{x_4}, g_1^{x_5}, g_2^{x_6}, s)$ where s is a **seed** for a $\text{CRH} \rightarrow \mathcal{H} = \{H_s\}$
- *Prove*($\omega, (c_1, c_2, l), r$)
 - Compute $\beta = H_s(c_1, c_2, l) \in \mathcal{Z}_p$
 - Output $\Pi = h_2^r h_3^{r\beta}$
- *Verify*($\tau, (c_1, c_2, l)$)
 - Compute $\beta = H_s(c_1, c_2, l) \in \mathcal{Z}_q$
 - Output $\tilde{\Pi} = c_1^{x_3 + \beta x_5} c_2^{x_4 + \beta x_6}$

Correctness:

$$\Pi = h_2^r h_3^{r\beta} = (g_1^{x_3} g_2^{x_4})^r (g_1^{x_5} g_2^{x_6})^{r\beta} = c_1^{x_3} c_2^{x_4} c_1^{\beta x_5} c_2^{\beta x_6} = c_1^{x_3 + \beta x_5} c_2^{x_4 + \beta x_6} = \tilde{\Pi}$$

Theorem 31. *The above construction define a 2-universal DVNIZK for L_{DDH}* \diamond

Proof. Same goal and procedure as before

- Take any $(c_1, c_2) \notin L_{DDH}$
- Fix $(c_1, c_2, l) \neq (c'_1, c'_2, l')$ s.t. $(c_1, c_2), (c'_1, c'_2) \notin L_{DDH}$ which means:

- $(c_1, c_2) = (g_1^{r_1} g_2^{r_2})$
- $(c'_1, c'_2) = (g_1^{r'_1} g_2^{r'_2})$
- $\beta = H_s(c_1, c_2, l)$
- $\beta' = H_s(c'_1, c'_2, l')$

- Let's define a MAP

$$\begin{aligned}
\mu'(x_3, x_4, x_5, x_6) &= (\omega, \widetilde{\Pi} = Ver(\tau, (c_1, c_2, l)), \widetilde{\Pi}' = Ver(\tau, (c'_1, c'_2, l'))) = \\
&= (\underbrace{(h_2, h_3)}_{\omega}, c_1^{x_3+\beta x_5} c_2^{x_4+\beta x_6}, c'_1^{x_3+\beta' x_5} c'_2^{x_4+\beta' x_6}) = \\
&= ((g_1^{x_3} g_2^{x_4}, g_1^{x_5} g_2^{x_6}), g_1^{r_1 x_3 + \beta r_1 x_5} g_2^{r_2 x_4 + \beta r_2 x_6}, g_1^{r'_1 x_3 + \beta' r'_1 x_5} g_2^{r'_2 x_4 + \beta' r'_2 x_6})
\end{aligned}$$

But I can rewrite g_2 as $g_2 = g_1^\alpha$ since they are generators. So:

$$\begin{aligned}
&((g_1^{x_3+\alpha x_4}, g_1^{x_5+\alpha x_6}), g_1^{r_1 x_3 + \beta r_1 x_5} g_1^{\alpha(r_2 x_4 + \beta r_2 x_6)}, g_1^{r'_1 x_3 + \beta' r'_1 x_5} g_1^{\alpha(r'_2 x_4 + \beta' r'_2 x_6)}) = \\
&= ((g_1^{x_3+\alpha x_4}, g_1^{x_5+\alpha x_6}), g_1^{r_1 x_3 + \alpha r_2 x_4 + \beta r_1 x_5 + \alpha \beta r_2 x_6}, g_1^{r'_1 x_3 + \alpha r'_2 x_4 + \beta' r'_1 x_5 + \alpha \beta' r'_2 x_6}) =
\end{aligned}$$

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} = \begin{pmatrix} 1 & \alpha & 0 & 0 \\ 0 & 0 & 1 & \alpha \\ r_1 & \alpha r_2 & \beta r_1 & \alpha \beta r_2 \\ r'_1 & \alpha r'_2 & \beta' r'_1 & \alpha \beta' r'_2 \end{pmatrix} \cdot \begin{pmatrix} x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix}$$

$$\text{Since Det} \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} = \alpha^2(r_2 - r_1)(r'_2 - r'_1)(\beta - \beta') \neq 0$$

IFF:

$\overbrace{r_2 \neq r_1, r'_2 \neq r'_1, \beta \neq \beta'}^{\text{for construction}} \rightarrow$ this last condition is true because we picked H_s collision resistant.

Otherwise H_s computed on different elements $((c_1, c_2, l)$ and (c'_1, c'_2, l')) will have to output the same $\beta(= \beta')$. \square

Lesson 18

Cramer-Shoup scheme construction

From the above two proof systems we can construct a PKE scheme, which is attributed to Cramer and Shoup:

T0D0 12: split definition from correctness

- $(pk, sk) \leftarrow \mathcal{KGen}$, where:
 - $pk := (h_1, h_2, h_3) = (g_1^{x_1} g_2^{x_2}, g_1^{x_3} g_2^{x_4}, g_1^{x_5} g_2^{x_6})$
 - $sk := (x_1, x_2, x_3, x_4, x_5, x_6)$
- Encryption procedure:
 - $r \leftarrow \mathbb{Z}_q$
 - $\beta = H_s(c_1, c_2, c_3) = (g_1^r, g_2^r, h_1^r m)$
 - $Enc(pk, m) = (c_1, c_2, c_3, (h_2 h_3^\beta)^r)$
- Decryption procedure:
 - Check that $c_1^{x_3 + \beta x_5} c_2^{x_4 + \beta x_6} = c_4$. If not, output FALSE.
 - Else, output $\hat{m} = c_3 c_1^{-x_1} c_2^{-x_2}$

4.23 Digital signatures

In this section we explore the solutions to the problem of authentication with the use of an asymmetric key scheme. Some observations are in order:

- In a symmetric setting, a verifier routine could be banally implemented as recomputing the signature using the shared secret key and the message. Here, Bob cannot recompute σ as he's missing Alice's secret key (and for good reasons too...). Thus, the verifying routine must be defined otherwise

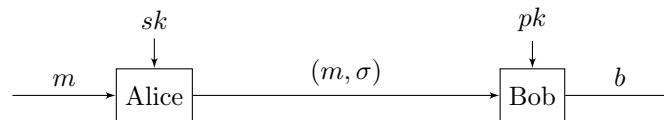


Figure 4.43: Asymmetric authentication

- In a vaguely similar manner to how an attacker could encrypt messages by itself in the asymmetric scenario, because the public key is known to everyone, any attacker can verify any signed messages, for free

Nevertheless, proving that a DS scheme is secure is largely defined in the same way as in the symmetric scenario, with the UF-CMA property:

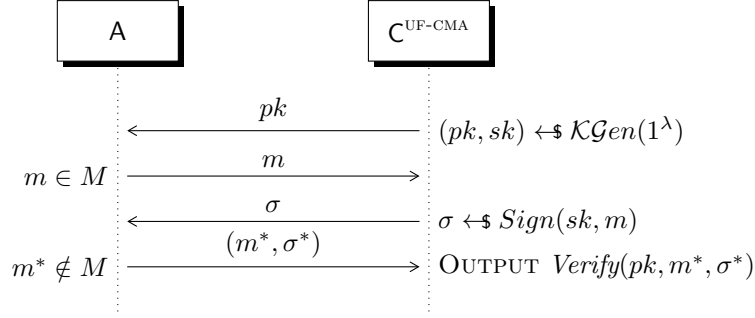


Figure 4.44: Unforgeable digital signatures

4.23.1 Public Key Infrastructure

The problem now is that Alice has a public key, but she wants "the blue check" over it, so Bob is sure that public key comes only from the true Alice.

To obtain the blue check, Alice needs an universally-trusted third party, called *Certification Authority*, which will provide a special *signature* to Alice for proving her identity to Bob.

The scheme works as follows:

The CA message pk' is also called as $cert_A$, the signature of the CA for Alice.

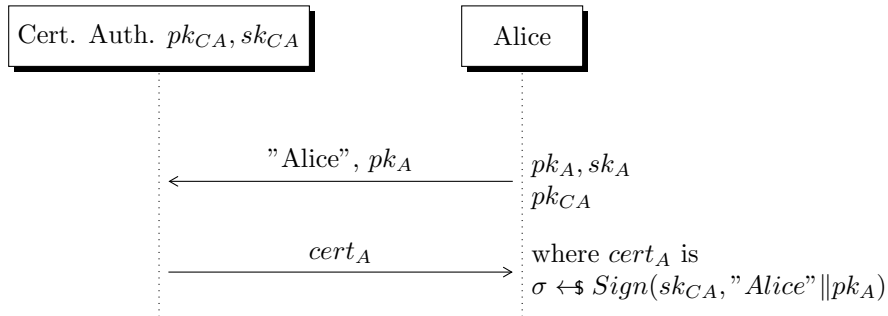
Now, when Bob wants to check the validity of the Alice's public key:

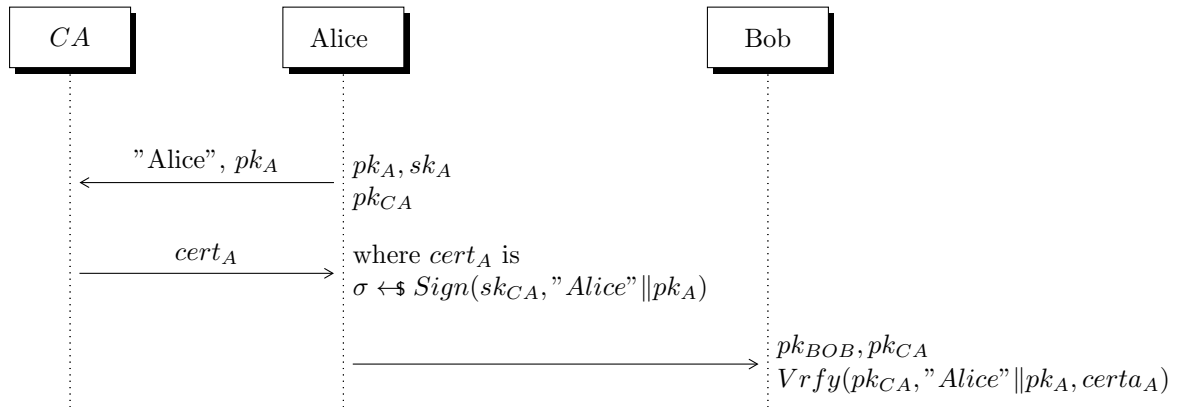
How can Bob recognize a valid certificate from an expired/invalid one?

The infrastructure provides some servers which contain the lists of the currently valid certificates.

Theorem 32. *Signatures are in Minicrypt.* \diamond

This is a counterintuitive result, not proven during the lesson, but very interesting because it implies that we can create valid signatures only with hash functions, without considering at all public key encryption.





In the next episodes:

- Digital Signatures from TDP*
- Digital Signatures from ID Scheme*
- Digital Signatures from CDH

Where * appears, something called *Random Oracle Model* is used in the proof. Briefly, this model assumes the existence of an ideal hash function which behaves like a truly random function (outputs a random y as long as x was never queried, otherwise gives back the already taken y).

Lesson 19

4.23.2 Waters Signature

This is a standard model for PKI which **does not need a Random Oracle**.
(or at least I think so)

Let's define a **Bilinear Group** as $(\mathcal{G}, \mathcal{G}_t, q, g, \hat{e}) \leftarrow \text{BilGen}(1^\lambda)$ where:

- $\mathcal{G}, \mathcal{G}_t$ are prime order groups (order q).
- g is a random generator of \mathcal{G} .
- (\mathcal{G}, \cdot) is a multiplicative group and \mathcal{G}_t is called target group.
- \hat{e} is a (bilinear) MAP: $\mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}_t$ efficiently computable. Defined as follows:

→ Take a generator g for \mathcal{G} and an element h in \mathcal{G} .
 $\forall g, h \in \mathcal{G}, a, b \in \mathbb{Z}_q \hat{e}(g^a, h^b) = \hat{e}(g, h)^{ab} = \hat{e}(g^{ab}, h)$
 and $\hat{e}(g, g) \neq 1$ (Non degenerative)

Venturi said something here related to Weil pairing over an elliptic curve. I found this. Interesting but not useful.

Assumption: CDH is HARD in \mathcal{G} .

Observation: DDH is EASY in \mathcal{G} !

Proof. $(g, g^a, g^b, g^c) \approx_c (g, g^a, g^b, g^{ab}) \implies \hat{e}(g^a, g^b) =? \hat{e}(g, g^c)$ now just move the exponents and I can transform the first element in $\hat{e}(g, g^{ab})$ and check if it is equal to $\hat{e}(g, g^c)$. □

Now $\text{KGen}(1^\lambda)$ will:

- Generate some params = $(\mathcal{G}, \mathcal{G}_t, g, q, \hat{e}) \leftarrow \text{BilGen}(1^\lambda)$
- Pick $a \leftarrow \mathbb{Z}_q$ then $g_1 = g^a$
- Pick $g_2 = g^b$ and $g_2, u_0, u_1, \dots, u_k \leftarrow \mathcal{G}$.
- Then output:

- $P_k = (\text{params}, g_1, g_2, u_0, \dots, u_k)$
- $S_k = g_2^a = g^{ab}$

$\text{Sign}(S_k, m)$:

- Divide the message m of length k in bits and not it as follows: $m = (m[1], \dots, m[k])$

- Now define $\alpha(m) = u_0 \prod_{i=1}^k u_i^{m[i]}$
- Pick $r \leftarrow \mathbb{Z}_q$ and output the signature $\sigma = (\underbrace{g_2^a}_{S_k} \cdot \alpha(m)^r, g^r) = (\sigma_1, \sigma_2)$

$\text{Ver}(P_k, m, (\sigma_1, \sigma_2))$

- Check $\hat{e}(g, \sigma_1) = \hat{e}(\sigma_2, \alpha(m)) = \hat{e}(g_1, g_2)$

Correctness: "Just move the exponents"

I can separate the second part for bilinearity

$$\hat{e}(g, \sigma_1) = \hat{e}(g, g_2^a \cdot \alpha(m)^r) = \hat{e}(g, g_2^a) \cdot \hat{e}(g, \alpha(m)^r) = \hat{e}(g, g_2^a) \cdot \hat{e}(g, \alpha(m)^r) = \hat{e}(g_1, g_2) \cdot \hat{e}(\sigma_2, \alpha(m))$$

We can say that we are moving the exponents from the "private domain" to the "public domain".

Theorem 33. The Waters' signature scheme is UF-CMA ◇

Proof. The trick is to "program the 'u's" (Venturi)

T0D0 13: Sequence is incomplete/incorrect, have to study it more...

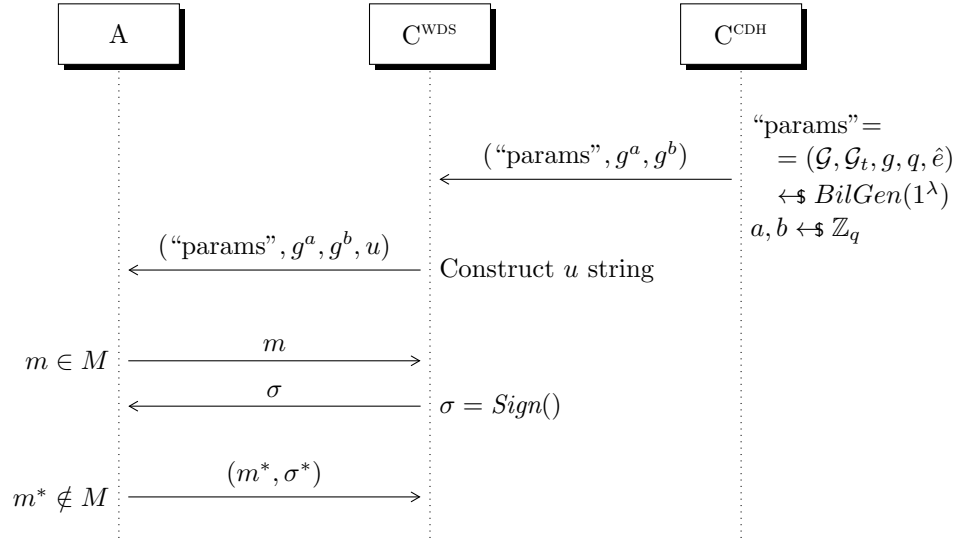


Figure 4.45: Reducing Waters' scheme to CDH

T0D0 14: The following explanation is roundabout, will rectify later

The following describes how the Waters' challenger constructs the u string. The main idea is, given k as the message length, to choose each single bit of u from 1 up to k such that:

$$\alpha(m) = g_2^{\beta(m)} g^{\gamma(m)}, \quad \beta(m) = x_0 + \sum_{i=1}^k m_i x_i, \quad \gamma(m) = y_0 + \sum_{i=1}^k m_i y_i$$

where $x_0 \leftarrow \mathbb{Z} \{-kl, \dots, 0\}$, $x_{1-k} \leftarrow \mathbb{Z}\{0, \dots, l\}$, $y_{0-k} \leftarrow \mathbb{Z}_q$

In particular: $l = 2q_s$, where q_s is the number of sign queries made by the adversary.

Therefore, let $u_i = g_2^{x_i} g^{y_i} \quad \forall i \in [0, k]$. Then:

$$\alpha(m) = g_2^{x_0} g^{y_0} \prod_{i=1}^k (g_2^{x_i} g^{y_i})^{m_i} = g_2^{x_0 + \sum_{i=1}^k m_i x_i} g^{y_0 + \sum_{i=1}^k m_i y_i} = g_2^{\beta(m)} g^{\gamma(m)}$$

T0D0 15: Partizioni, doppi if.... qui non ci ho capito 'na mazza

Step 1: $\sigma = (\sigma_1, \sigma_2) = (g_2^a \alpha(m)^{\bar{r}}, g^{\bar{r}})$, for $\bar{r} \leftarrow \mathbb{Z}_q \quad \bar{r} = r - a\beta^{-1}$

$$\begin{aligned} \sigma_1 &= g_2^a \alpha(m)^{\bar{r}} \\ &= g_2^a \alpha(m)^{r - a\beta^{-1}} \\ &= g_2^a (g_2^{\beta(m)} g^{\gamma(m)})^{r - a\beta^{-1}} \\ &= g_2^a g_2^{\beta(m)r - a} g^{\gamma(m)r - \gamma(m)a\beta^{-1}} \\ &= g_2^{\beta(m)r} g^{\gamma(m)r} g^{-\gamma(m)\beta^{-1}} \end{aligned}$$

□