"Sapienza" University of Rome

Faculty of Information Engineering, Informatics and Statistics

Department of Computer Science

Cryptography

Author

Alessio Bandiera

October 1, 2025

Contents

Information and Contacts

Personal notes and summaries collected as part of the *Cryptography* course offered by the degree in Computer Science of the University of Rome "La Sapienza".

Further information and notes can be found at the following link:
https://github.com/aflaag-notes. Anyone can feel free to report inaccuracies, improvements or requests through the Issue system provided by GitHub itself or by contacting the author privately:

- Email: **alessio.bandiera02@gmail.com**

- LinkedIn: **Alessio Bandiera**

The notes are constantly being updated, so please check if the changes have already been made in the most recent version.

Suggested prerequisites:

TODO

Licence:

These documents are distributed under the **GNU Free Documentation License**, a form of copyleft intended for use on a manual, textbook or other documents. Material licensed under the current version of the license can be used for any purpose, as long as the use meets certain conditions:

- All previous authors of the work must be **attributed**.

- All changes to the work must be **logged**.

- All derivative works must be **licensed under the same license**.

- The full text of the license, unmodified invariant sections as defined by the author if any, and any other added warranty disclaimers (such as a general disclaimer alerting readers that the document may not be accurate for example) and copyright notices from previous versions must be maintained.

- Technical measures such as DRM may not be used to control or obstruct distribution or editing of the document.

1

TODO

1.1 TODO

TODO ⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯ missing introduction

In this section, we will discuss **symmetric cryptography**, i.e. cryptosystems where a shared secret key is used for both encryption and decryption. Such encryption method is fast and efficient, but generally considered less secure. The core model of this approach is the **Secret Key Encryption (SKE)**, composed of:

- a shared *secret* key $K \in_R \mathcal{K}$ chosen uniformly at random

- an *encryption function* $\text{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$ that transforms plaintext into cyphertext — where $\mathcal{K}$ and $\mathcal{C}$ are the spaces of the keys and the cyphertexts, respectively

- a *decryption function* $\text{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$ that transforms a cyphertext into a plaintext

Clearly, to be useful at all SKEs must be **correct**, which means that if a message $m \in \mathcal{M}$ is encrypted through $K \in_R \mathcal{K}$ getting $c \in \mathcal{C}$, the decription process on c through K must produce the original message m

> **Definition 1.1: Correctness of SKEs**
>
> An SKE $\Pi = (\text{Enc}, \text{Dec})$ is said to be **correct** if it holds that
>
> $$\forall m \in \mathcal{M}, K \in_R \mathcal{K} \quad \text{Dec}(\text{Enc}(K, m)) = m$$

TODO ⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯ diagram

During the 19th century, the Dutch cryptographer Kerckhoff postulated his homonym principle, which is stated below.

> **Principle 1.1: Kerckhoff's principle**
>
> The security of a cryptographic system should depend solely on the secrecy of the key.

In other words, a system should be secure *even if everything is public but the key*. During his work, Shannon proposed a formal definition of **perfect secrecy** which fully follows the concept of Kerckhoff's principle.

> **Definition 1.2: Perfect secrecy**
>
> Given an SKE $\Pi = (\mathrm{Enc}, \mathrm{Dec})$, a random variable M over $\mathcal{M}$ and a random variable $C = \mathrm{Enc}(K, M)$ for some $K \in_R \mathcal{K}$, we say that Π has **perfect secrecy** if
>
> $$\forall m \in \mathcal{M}, c \in \mathcal{C} \quad \Pr[M = m] = \Pr[M = m \mid C = c]$$

In other words, as Shannon originally formulated it, this definition states that the probability of m being exactly the communicated message must not depend on the cyphertext c, which implies that c can be known by everyone. Hence, this definition requires the encrypted text c to *not reveal* anything about the plaintext m. The following lemma shows some properties about perfect secrecy.

Shannon provde that perfect secrecy is achievable by some cryptosystems, but with some limitations. For instance, consider the **One Time Pad (OTP)** SKE, in which we assume that $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0,1\}^n$ for some fixed length $n \in \mathbb{N}$, and the encryption and decryption functions are defined as follows:

$$\mathrm{Enc}(K, m) = K \oplus m \qquad \mathrm{Dec}(K, c) = K \oplus c$$

(which is always possible by invertibility of the XOR function). We will prove that OTP has perfect secrecy by first providing an alternative definition to the latter.

> **Lemma 1.1**
>
> Let $\Pi = (\mathrm{Enc}, \mathrm{Dec})$ be an SKE, M be a random variable over $\mathcal{M}$ and C be a random variable defined as $C = \mathrm{Enc}(K, M)$ for some $K \in_R \mathcal{K}$. The following three conditions are equivalent:
>
> 1. Π has perfect secrecy
>
> 2. M and C are independent
>
> 3. $\forall m, m' \in \mathcal{M}, c \in C \quad \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m) = c] = \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m') = c]$

Proof. We will prove the statements cyclically.

- $1 \implies 2$. By perfect secrecy of Π we have that

$$\Pr[M = m] = \Pr[M = m \mid C = c] = \frac{\Pr[M = m \wedge C = c]}{\Pr[C = c]}$$

therefore, by rearranging the terms we get that

$$\Pr[M = m \wedge C = c] = \Pr[M = m] \cdot \Pr[C = c]$$

- $2 \implies 3$. Fix $m \in \mathcal{M}$ and $c \in \mathcal{C}$; we have that

$$\begin{aligned}
\Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, m) = c] &= \Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, M) \mid M = m] \\
&= \Pr_{K \in_R \mathcal{K}}[C = c \mid M = m] && \text{(by definition)} \\
&= \Pr[C = c] && \text{(by independence of } M \text{ and } C\text{)}
\end{aligned}$$

Now fix another message $m' \in \mathcal{M}$; we can repeat the same steps and obtain that $\Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, m') = c] = \Pr[C = c]$ which concludes the proof.

- $3 \implies 1$. Fix $c \in \mathcal{C}$.

Claim: $\Pr[C = c] = \Pr[C = c \mid M = m]$

Proof of the Claim. By assuming property 3, we get that

$$\begin{aligned}
\Pr[C = c] &= \sum_{m' \in \mathcal{M}} \Pr[C = c \mid M = m'] \cdot \Pr[M = m'] && \text{(by the L.T.P.)} \\
&= \sum_{m' \in \mathcal{M}} \Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, M) = c \mid M = m'] \cdot \Pr[M = m'] \\
&= \sum_{m' \in \mathcal{M}} \Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, m') = c] \cdot \Pr[M = m'] \\
&= \sum_{m' \in \mathcal{M}} \Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, m) = c] \cdot \Pr[M = m'] && \text{(by property 3)} \\
&= \Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, m) = c] \cdot \sum_{m' \in \mathcal{M}} \Pr[m = m'] \\
&= \Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, m) = c] \\
&= \Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, M) = c \mid M = m] \\
&= \Pr_{K \in_R \mathcal{K}}[C = c \mid M = m]
\end{aligned}$$

$\square$

Finally, by Bayes' theorem we have that

$$\begin{aligned}
\Pr[M = m] &= \frac{\Pr[M = m \mid C = c] \cdot \Pr[C = c]}{\Pr[C = c \mid M = m]} \\
&= \Pr[M = m \mid C = c] && \text{(by the claim)}
\end{aligned}$$

which is precisely perfect secrecy.

$\square$

Thanks to this lemma, we can prove the following result through the third alternative definition of perfect secrecy.

Theorem 1.1

OTP has perfect secrecy.

Proof. Fix two messages $m, m' \in \mathcal{M}$, and a cyphertext $c \in \mathcal{C}$; by the definition of the OTP system and the properties of the XOR function, we have that

$$\Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, m) = c] = \Pr_{K \in_R \mathcal{K}}[K \oplus m = c] = \Pr_{K \in_R \mathcal{K}}[K = m \oplus c] = 2^{-n}$$

and through the same reasoning we can analogously show that $\Pr_{K \in_R \mathcal{K}}[\text{Enc}(K, m') = c] = 2^{-n}$ which proves that the third condition of the alternative definition of perfect secrecy provided in the previous lemma holds for the OTP SKE. $\square$

TODO _____ buco

We will now focus on the second goal of cryptography: *message integrity*, which is usually achieved through **Message Authentication Codes (MACs)**, which allows the receiver to determine if the message has been tempered with.

First, let's start with a simple model, that ignores secrecy and only cares about message integrity. This type of MACs use a deterministic **tagging function**, usually implemented though *hash functions*, but in general it is a function of the form

$$\text{Tag} : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$$

, whetre $\mathcal{T}$ is the tag space, i.e. the set of all tag strings.

TODO _____ paragrafo

1.2 Message Authentication Codes (MACs)

TODO _____ paragrafo

TODO _____ disegno

Definition 1.3: t-time ε-statistical security

A MAC $\Pi = (\text{Tag})$ is said to have t**-time ε-statistical security** if $\forall m, m_1, \ldots, m_t \in \mathcal{M}$ pairwise distinct, and $\forall \tau, \tau_1, \ldots, \tau_t \in \mathcal{T}$ it holds that

$$\Pr_{K \in_R \mathcal{K}}[\text{Tag}(K, m) = \tau \mid \text{Tag}(K, m_1) = \tau_1, \ldots, \text{Tag}(K, m_t) = \tau_t] \leq \varepsilon$$

In other words, this property states that even when t message-tag pairs $(m_1, \tau_1), \ldots, (m_t, \tau_t)$ obtained through the same key K are known, the probability of any message-tag pair (m, τ) being obtained through the same key K is at most ε. Clearly, we would like ε to

be as small as possible, and t to be as large as possible. However it is easy to see that $\forall t \in \mathbb{N}$ it is impossible to get $\varepsilon = 0$, since any random $\tau \in \mathcal{T}$ has always a $\frac{1}{|\mathcal{T}|}$ probability of being correct by random chance.

Furthermore, as we proved for perfect secrecy, we are going to show that the notion of "good" statistical security is achievable, even though it's highly inefficient in terms of key size.

> **Theorem 1.2**
>
> Any t-time $2^{-\lambda}$-statistically secure MAC must have a key of size $(t+1) \cdot \lambda$

A good enough 1-time statisticalllhy secure MAC can be achieved through **pairwise independent hash functions**, a family of hash functions where each pair of functions forms a pari of independent random variables.

> **Definition 1.4: Pairwise independency**
>
> Let $\mathcal{H} = \{h_K : \mathcal{K} \to \mathcal{T}\}_{K \in_R \mathcal{K}}$ be a family of hash functions; we say that $\mathcal{H}$ is **pairwise independent** if $\forall m, m' \in \mathcal{M}$ such that $m \neq m'$ it holds that the distribution $(h_K(m), h_K(m'))$ is uniform over $\mathcal{T} \times \mathcal{T}$ when $K \in_R \mathcal{K}$

Note that, in this definition, $h_K(m)$ and $h_K(m')$ are two random variables .

what the f does this def mean

> **Theorem 1.3**
>
> Any $\mathcal{H} = \{h_K : \mathcal{M} \to \mathcal{T}\}_{K \in_R \mathcal{K}}$ family of pairwise independent hash functions induces a 1-time $\frac{1}{|\mathcal{T}|}$-statistically secure MAC.

Proof. TODO □

todo

TODO

buco

1.2.1 Randomness extraction

TODO

buco

Let's first address the problem of *extracting* randomness from an unpredictable secure source (i.e. random variable) X. The first **extractor** has been introduced by Von Neumann, which yields a fair random coin from an unpredictable unfair one. Let $B \in \{0, 1\}$ be the random variable describing the unpredictable unfrair coin, such that for instance $\Pr[B = 0] = p < \frac{1}{2}$. Let $Y \in \{0, 1\}$ be the random variable describing the coin we want to extract; its vaule will be determined by the following procedure:

- sample two values b_1, b_2 from B independently

- if $b_1 = b_2$, Y is assigned no value — noted with $Y = ?$ and the process is repeated from the beginning

- if $b_1 \neq b_2$, $Y = 1$ if and only if $b_1 = 0$ and $b_2 = 1$

First, we observe that this process can go on indefinitely, never halting on any value for Y. But assuming the procedure does halt, we have that

$$\Pr[Y = 0] = \Pr[b_1 = 1 \wedge b_2 = 0] = \Pr[b_1 = 1] \cdot \Pr[b_2 = 0] = (1 - p)p$$

$$\Pr[Y = 1] = \Pr[b_1 = 0 \wedge b_2 = 1] = \Pr[b_1 = 0] \cdot \Pr[b_2 = 1] = p(1 - p)$$

Moreover, we see that the probability of $Y = ?$ happening for m consecutive tries is at most

$$\Pr[Y = ? \text{for } m \text{ tries}] = (\Pr[Y = ?])^m = (1 - \Pr[Y = 0 \vee Y = 1])^m \leq (1 - 2p(1 - p))^m$$

and as m goes to infinity, the latter probabijlity goes to 0, which implies that $\Pr[Y = 0]$ and $\Pr[Y = 1]$ will tend towards $\frac{1}{2}$ due to them having the same probability and them being the only two possible outcomes. Thus, we have achieved a fair coin Y.

TODO _____ `par`

> **Definition 1.5: Min-entropy**
>
> Given a random variable X, the **min-entropy** of X is defined as follows:
>
> $$H_{\min}(X) := -\log \max_x \Pr[X = x]$$

TODO _____ `todo`

Is there an extractor Ext that for any random variable X it outputs a uniform distribution $Y = \text{Ext}(X)$ such that $H_{\min}(X) \geq k$ for some $k > 0$? Even under this constraint, the answer is still negative, in fact it is impossible to define such extractor even if we restrict the question to extracting only *one bit* — i.e. $\text{Ext}(X) = b \in \{0, 1\}$ — and $k = n - 1$.

> **Proposition 1.1**
>
> Let X be a uniform random variable defined over $\{0, 1\}^n$; there is no extractor Ext such that $\text{Ext}(X) \in \{0, 1\}$ is uniform.

Proof. Let $\text{Ext} : \{0, 1\}^n \to \{0, 1\}$ be any extractor, and let $b \in \{0, 1\}$ be the output minimizing the cardinality of the preimage of the extractor, i.e. the set of inputs for which the extractor output b

$$b = \arg\min_{b' \in \{0, 1\}} \left| \text{Ext}^{-1}(b) \right|$$

By the pigeonhole principle, we have that

$$\left| \text{Ext}^{-1}(b) \right| \leq \frac{|\{0, 1\}^n|}{2} = \frac{2^n}{2} = 2^{n-1}$$

Let X be a random variable uniform over $\text{Ext}^{-1}(b)$; since X is uniform, we have that $H_{\min}(X) \geq n - 1$, however $\text{Ext}(X)$ cannot be uniform since the output is constant (in fact $\text{Ext}(X) \equiv b$). $\square$

TODO

la roba del par prima va capita

Hence, since we cannot define an extractor that yields a uniform distribution, the best we can achieve is a distribution that is *close enough* to a uniform one. But what does it mean to be "close enough" mathematically? First, let us provide the definition of **statistical distance**.

> **Definition 1.6: Statistical distance**
>
> Given two random variables X and X' defined over the same set, the **statistical distance** between X and X' is defined as follows:
> $$\text{SD}(X; X') = \frac{1}{2} \sum_x |\Pr[X = x] - \Pr[X' = x]|$$

We can now introduce the concept of *closeness* between distributions.

> **Definition 1.7: ε-closeness**
>
> Given two random variables X and X' defined over the same set, we say that X and X' are ε-**close**, written as $X \sim_\varepsilon X'$, if it holds that
> $$\text{SD}(X; X') \leq \varepsilon$$

We can describe the concept of ε-closeness between random variables as saying that no *unbounded adversary* A can distinguish whether a value x has been sampled from X or X', or in symbols

$$|\Pr[A(x) = 1 : x \leftarrow X] - \Pr[A(x) = 1 : x \leftarrow X']| \leq \varepsilon$$

> **Definition 1.8: Deterministic extractor**
>
> Given a random variable S, called *seed*, we say that $\text{Ext} : \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^\ell$ is a (k, ε)-extractor if for every variable X such that $H_{\min}(X) \geq k$ it holds that
> $$S \sim U_d \implies (S, \text{Ext}(S, X)) \sim_\varepsilon (S, U_\ell)$$

We observe that the condition $(S, \text{Ext}(S, X)) \sim_\varepsilon (S, U_\ell)$ implies that S must be *public*. This requirement is forced in order to avoid trivial extractors, such as $\text{Ext}(S, X) = S$.

why?