# Notes for EECS 475

Cryptography

Albon Wu

January 15, 2025

## Contents

# 0 Introduction

Essential concepts in cryptography, precise attack models and security definitions, and constructions of real-world cryptosystems.

Professor: Mahdi Cheraghchi

# 1 The Cryptographic Methodology

Suppose Alice wants to communicate a message to Bob that she wants to be safe from an eavesdropper. She should do three things:

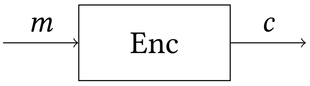
1. Form a realistic model of the scenario, adjusting as necessary to allow for possible solutions.
2. Precisely define the desired function and security properties of a potential solution.
3. Construct and analyze a solution, ideally proving it satisfies the desired properties.

Step 3 is the most mathematical and requires careful attention. But obviously your system is still vulnerable if you formally prove everything but your model does not fully capture reality. In cryptography, we often prove things (e.g., RSA) given well-established assumptions.

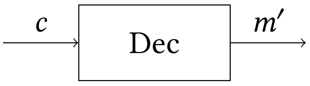
## 1.1 Modeling encryption

Here's a specific case of encryption:

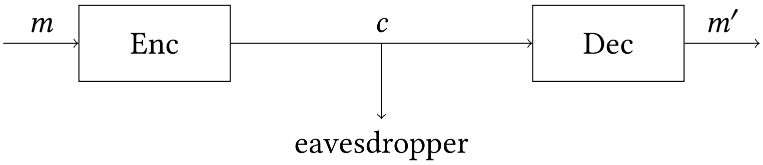
Sender A is represented by an algorithm Enc(·) that takes a "plaintext" message $m$ from a (finite) set of possible messages $\mathcal{M}$ ("message space") and outputs a "ciphertext" $c$ from a finite set $\mathcal{C}$ ("ciphertext space").



Receiver B is represented by an algorithm Dec(·) that takes some $c \in \mathcal{C}$ and outputs some $m' \in \mathcal{M}$.

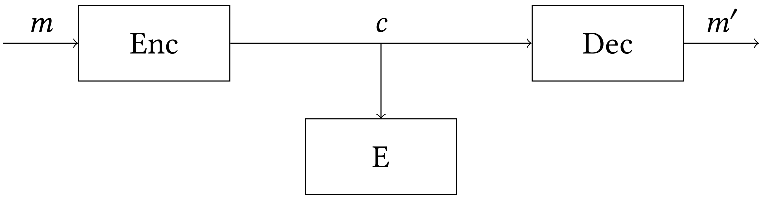


So here is our model so far:



The eavesdropper is represented by an algorithm E(·) that takes $c \in \mathcal{C}$ and outputs...

...what does it output?



At the very least, the system should be "correct" in that if there is no adversary, legitimate users should be able to communicate. That is,

$$\forall m \in \mathcal{M} \qquad \mathrm{Dec}(\mathrm{Enc}(m)) = m.$$

The tricky part is formulating in a mathematical sense what “security” means. At a minimum, E shouldn’t be able to always recover the message $m$ from $c$.

But what prevents E := Dec? It brings up an interesting problem in our model - there is no way to distinguish a legitimate party from an attacker.

The correct way is to use a secret that only the legitimate parties know. What is not correct is to keep the decryption algorithm itself secret - secrets don’t last if you can’t change them quickly.

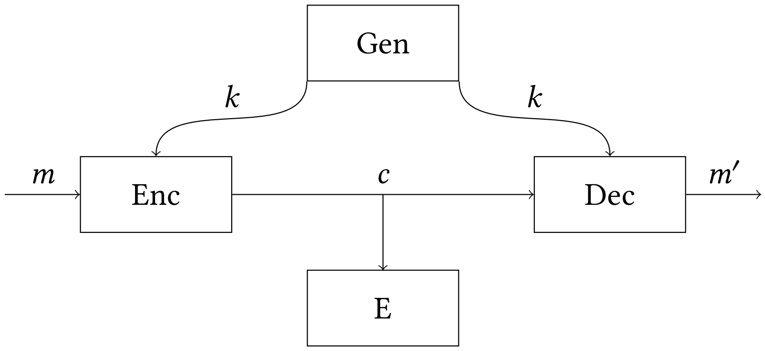
> “Not only is it bad, it is terrible.” (Madhi)

In cryptography, everything is open source. There are no secrets in knowledge; security by obscurity is to be avoided. This is called Kerckhoffs’s principle: the system should remain secure even if all its algorithms are known to the public.

> “The enemy knows the system.” (Shannon)

Instead, we give the Dec box an extra (secret) input key. Making the key itself secret is a problem for later.

So some algorithm Gen(·) (the “key generator”) is responsible for producing and sharing that key. Note that Gen is not secret, even if its output is.



This is called “private (or symmetric) key encryption”. Our updated definition of correctness is then:

$$\forall m \in \mathcal{M}, \forall k \in \mathcal{K} \qquad \mathrm{Dec}_k(\mathrm{Enc}_k(m)) = m$$

where $\mathcal{K}$ is the “key space.”

Security should mean:

1. E won’t be able to recover $m$ from $c$
2. E won’t be able to recover $k$
3. E won’t recover *any part of* or anything about $m$

## 1.2 Shannon and perfect secrecy

**Definition 1.2.1.** An encryption scheme (Gen, Enc, Dec) is **Shannon secret** if for any message distribution $D$ over message space $\mathcal{M}$, any fixed $\overline{m} \in \mathcal{M}$ and fixed ciphertext $\overline{c} \in \mathcal{C}$,

$$\Pr_{\substack{m \leftarrow D \\ k \leftarrow \mathrm{Gen}()}} (m = \overline{m} \mid \mathrm{Enc}_k(m) = \overline{c}) = \Pr_{m \leftarrow D}(m = \overline{m}).$$

If you are an eavesdropper on a Shannon secret communication, you are no more certain about the message contents after eavesdropping than before.

The RHS is also called *a priori*, while the LHS is *a posteriori.*

This condition provides a very strong security guarantee but is very cumbersome and requires that we deal with every possible D.

$$\Pr_{m,k}(m = \overline{m} \mid \mathrm{Enc}_k(m) = \overline{c}) = \frac{\Pr_{m,k}(m = \overline{m} \wedge \mathrm{Enc}_k(m) = \overline{c})}{\Pr_{m,k}(\mathrm{Enc}_k(m) = \overline{c})}$$

$$= \frac{\Pr_{m,k}(m = \overline{m} \wedge \mathrm{Enc}_k(\overline{m}) = \overline{c})}{\Pr_{m,k}(\mathrm{Enc}_k(m) = \overline{c})}$$

where in the last equality we replace $m$ by $\overline{m}$. By independence of $m, k$:

$$= \Pr_m(m = \overline{m}) \cdot \frac{\Pr_k(\mathrm{Enc}_k(\overline{m}) = \overline{c})}{\Pr_{m,k}(\mathrm{Enc}_k(m) = \overline{c})}.$$

So a scheme is Shannon secret if and only if the right term is 1 for all $D$ and all $\overline{m} \in \mathrm{supp}(D) \subseteq \mathcal{M}$. Note that the message distribution is not relevant here.

The fraction equals 1 if:

$$\Pr_k(\mathrm{Enc}_k(\overline{m}) = \overline{c}) = \Pr_{m,k}(\mathrm{Enc}_k(m) = \overline{c}) \qquad \forall \overline{m} \in \mathrm{supp}(D), \forall \overline{c} \in C. \tag{$*$}$$

Let $D$ be the uniform distribution over $\mathcal{M}$. This means the LHS is constant for all $\overline{m}$. This means that $\forall m_0, m_1 \in \mathcal{M}, \forall \overline{c} \in C$:

$$\Pr_k(\mathrm{Enc}_k(m_0) = \overline{c}) = \Pr_k(\mathrm{Enc}_k(m_1) = \overline{c}).$$

This means that if we encrypt distinct messages, they are equally likely to be a given ciphertext; that is, the distribution of the ciphertext is the same. This is called **perfect secrecy**; it is implied by Shannon secrecy.

Conversely, if perfect secrecy holds, then ($*$) holds.

$$\Pr_{\substack{m \leftarrow D \\ k \leftarrow \mathrm{Gen}}}(\mathrm{Enc}_k(m) = \overline{c}) = \sum_{m' \in \mathcal{M}} \Pr_k(\mathrm{Enc}_k(m') = \overline{c}) \cdot \Pr_m(m = m')$$

conditioning over all possible messages and applying total probability. By perfect secrecy, this equals

$$\sum_{m \in \mathcal{M}} \Pr_k(\mathrm{Enc}_k(\overline{m}) = \overline{c}) \cdot \Pr_m(m = m')$$

for arbitrary $m' \in \mathcal{M}$, so we can pull out the first sum term factor:

$$\Pr_k(\mathrm{Enc}_k(\overline{m}) = \overline{c}) \cdot \sum_{m' \in \mathcal{M}} \Pr(m = m')$$

$$= \Pr_k(\mathrm{Enc}_k(\overline{m}) = \overline{c})$$

as desired. So we have just shown:

**Theorem 1.2.1.** *Shannon secrecy is equivalent to perfect secrecy.*