Cryptography

Jack

Last Updated: March 15, 2025

[WIP] I will be compiling my notes on cryptography here.

Contents

§1 Introduction

Various notes from chapter 1 of [KL08].

Definition 1.1 (Kerchoffs' Principle)

Kerchoffs' Principle states that the security of a cryptographic system must not depend on the secrecy of the cipher.

In other words, a cipher should be secure even if the algorithm is public.

A few simple attack scenarios:

- **Ciphertext-only attack**: The adversary observes ciphertexts and must try to determine the plaintext.
- **Known-plaintext attack**: The adversary can learn one or more plaintexts encrypted under the same key. They must try to determine the plaintext of a different ciphertext they have not seen before.
- **Chosen-plaintext attac**k: The adversary can learn encryptions of plaintexts of its choice. They must try to determine the plaintext of a different ciphertext like before.
- **Chosen-ciphertext attac**k: The adversary can *additionally* learn the plaintexts of ciphertexts of its choice. Their goal is the same as above.

Note that the first two are **passive** while the last two are **active**.

Exercise 1.2. The first two scenarios are quite realistic. Can you think of some examples?

Exercise 1.3. Think of, or research, real world scenarios of the latter two attack scenarios.

Historical cryptographic ciphers are weak by modern standards, but they give us a few important lessons:

- **Sufficient key space principle**: it is a necessary condition to have a large key space (the domain from which keys are chosen). Otherwise, we can brute force all the keys.
- **Designing secure ciphers is hard!**: there were many ciphers such as the Vigenere cipher that were insecure (for example, to cryptanalysis). It is the goal of modern cryptography to rigorously define and prove security.

Modern cryptography comes with a few principles:

- **Clear, rigorous definitions**.

- **Clearly stated assumptions; the more minimal assumptions the better**.
- **Rigorous proofs of security with respect to principles 1 and 2**.

Example 1.4

To take an example, we give an idea of just how hard it is to rigorously define a *secure encryption scheme.*

- **Try 1: Secure if no adversary can find the key**

  But what if the adversary simply learns the plaintext?

- **Try 2: Secure if no adversary can learn the plaintext**

  But what if the adversary learns 50% of the plaintext? Or the length of the plaintext? Is this definition clear enough? What "percentage" is okay to learn?

- **Try 3: Secure if the adversary cannot determine *any* character of the plaintext**

  What if our plaintext is an integer like our salary and the adversary learns the *range* of our salary? Surely this isn't what we wanted with an "encryption" scheme.

- **Try 4: Secure if the adversary cannot derive any *meaningful* information from the ciphertext**

  Close, but no cigar. What exactly does meaningful *mean*? Our encryption scheme could be used in multiple different contexts and "meaningful" could have different meanings in each. For a *definition*, this attempt is not enough.

- **Try 5: Secure if the adversary cannot compute *any* function of the plaintext from the ciphertext**

  This is a rigorous definition: we have replaced "meaningful" with a more meaningful term.

Remark 1.5. Albeit, we often allow a carve out for the length of a message to be learned. Can you think of a way to avoid this?

Remark 1.6. For most of these notes, we will consider adversaries that are **efficient**: that is, those that run in polynomial time.

§2 Perfectly-Secret Encryption

Various notes from chapter 2 of [KL08].

Definition 2.1 (Perfectly Secret Encryption Scheme)

Even an adversary with *unbounded* computational power cannot break a **perfectly secret encryption scheme**.

First, we must formally define an encryption scheme.

Definition 2.2 (Encryption Scheme)

Consists of three algorithms:

- Gen, which outputs a key k according to a distribution. The key space is denoted by $\mathcal{K}$ and is finite.
- $\mathsf{Enc}(k, m)$, which encrypts m under k. The space of possible ciphertexts is denoted by $\mathcal{C}$.
- $\mathsf{Dec}(k, m)$, which decrypts m under k.

And also a message space $\mathcal{M}$ where $|\mathcal{M}| > 1$.

Definition 2.3 (Perfectly Correct)

For all $k \in \mathcal{K}$, $m \in \mathcal{M}$, if $c \leftarrow \mathsf{Enc}_k(m)$ then $\mathsf{Dec}_k(c) = m$ with probability 1.

Unless stated otherwise, we will be working with perfectly correct encryption schemes.

Note that we will reference distributions over $\mathcal{K}, \mathcal{M}$, and $\mathcal{C}$. The distribution over the key space is given by Gen. The distribution over the message space models how not all messages have equal probability of being sent.

$\mathcal{K}$ and $\mathcal{M}$ are independent distributions, as messages are chosen independent of keys. However, $\mathcal{C}$ is fully determined by the distributions over $\mathcal{K}$ and $\mathcal{M}$. The distribution of $\mathcal{K}$ is fixed for a given encryption scheme since it is defined by Gen, but $\mathcal{M}$ may vary depending on the parties.

Definition 2.4 (Perfect Secrecy)

An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is **perfectly secret** if for all $m \in \mathcal{M}, c \in \mathcal{C}$ such that $\Pr[C = c] > 0$, we have

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

Remark 2.5. Intuitively, this says that observing a ciphertext does not change the a priori probability of a message being sent.

Remark 2.6. In the future we may omit conditions such as $\Pr[C = c] > 0$ simply for convention and ease of use; however, it is implicitly there. See exercise 2.6.

Lemma 2.7

An equivalent definition of perfect encryption states

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

Another equivalent and useful formulation of perfect secrecy is *perfect indistinguishability*.

Definition 2.8 (Perfect Indistinguishability)

Perfect Indistinguishability states that given $m_0, m_1 \in \mathcal{M}$, then

$$\mathcal{C}(m_0) = \mathcal{C}(m_1)$$

where $\mathcal{C}(m_i)$ denotes the distribution of ciphertexts for the encryption of m_i.

Phrased differently,

Lemma 2.9

An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ over $\mathcal{M}$ is perfectly secret iff for all $m_0, m_1 \in \mathcal{M}, c \in \mathcal{C}$,

$$\Pr[C = c \mid M = m_0] = \Pr[C = c \mid M = m_1]$$

The final equivalent formulation of perfect secrecy is *adversarial indistinguishability*.

Definition 2.10 (Adversarial Indistinguishability)

The intuition is that an adversary cannot distinguish between the encryption of two messages better than simply guessing. It is a taste of future "game based" definitions of security.

We define an experiment $\text{PrivK}^{\text{eav}}$ in the setting of private key encryption and an eavesdropping adversary. It is defined for a scheme $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ over $\mathcal{M}$ for arbitrary $\mathcal{A}$. One execution of the experiment is denoted $\text{PrivK}^{\text{eav}}_{\mathcal{A},\Pi}$ and defined as

1. $\mathcal{A}$ chooses a pair of messages $m_0, m_1 \in \mathcal{M}$.
2. $k \leftarrow \mathcal{K}$ and $b \leftarrow 0, 1$. $c := \mathsf{Enc}(m_b)$ and given to $\mathcal{A}$.
3. $\mathcal{A}$ outputs b'.
4. If $b = b'$ then the output of the experiment is 1. Otherwise it is 0.

Lemma 2.11 (Adversarial Indistinguishability)

An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ over $\mathcal{M}$ is perfectly secure if for all $\mathcal{A}$,

$$\Pr[\text{PrivK}^{\text{eav}}_{\mathcal{A},\Pi}] = \frac{1}{2}.$$

§2.1 One Time Pad is Perfectly Secret

The one time pad encryption scheme is given by the following algorithm:

1. Fix an integer $l > 0$. Then $\mathcal{M}, \mathcal{K}, \mathcal{C}$ are all equal to $\{0, 1\}^l$.
2. Gen chooses a random string in $\mathcal{K}$ uniformly.
3. $\mathsf{Enc}_k(m) = c := m \oplus k$.
4. $\mathsf{Dec}_k(c) = m := c \oplus k$

Intuitively, the one time pad (proposed in 1917 by Verman) is perfectly secret because for any ciphertext, every plaintext could have been the original. Now since the keys are uniform randomly selected, each plaintext is equally likely. This intuition wasn't formally proved until 25 years later when Shannon introduced the idea of perfect secrecy.

Theorem 2.12

The one-time pad encryption scheme is perfectly secure.

Proof 2.13

We will aim to show that an adversary cannot do better than guessing which plaintext a ciphertext came from.

Fix an arbitrary distribution over $\mathcal{M}$ and choose $m \in \mathcal{M}, c \in \mathcal{C}$. Observe that

$$\begin{aligned}
\Pr[C = c \mid M = m] &= \Pr[M \oplus K = c \mid M = m] \\
&= \Pr[m \oplus K = c] \\
&= \Pr[K = c \oplus m] \\
&= \frac{1}{2^l}
\end{aligned}$$

Where the last step holds because the key is chosen uniformly out of $\frac{1}{2^l}$ strings. Now since the distribution over messages was arbitrary, and the message itself was arbitrary, we see that

$$\Pr[C = c | M = m_0] = \frac{1}{2^l} = \Pr[C = c | M = m_1]$$

which proves perfect secrecy by Lemma 2.9. □

It is crucial to note that the length of the key must be equal to the length of the message, which in reality is impractical. Moreover, the key can only be used *once*.

References

[KL08] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography, Second Edition*. Chapman & Hall/CRC, 1st edition, 2008.