

One-time pad [Vigenere cipher where key is as long as the message!]

$$K = \{0,1\}^n \quad \text{Encrypt}(k,m): \text{output } c = k \oplus m$$

$$M = \{0,1\}^n \quad \text{Decrypt}(k,c): \text{output } m = k \oplus c$$

$$C = \{0,1\}^n \quad \leftarrow \text{bitwise exclusive OR operation (addition mod 2)}$$

Correctness: Take any $k \in \{0,1\}^n$, $m \in \{0,1\}^n$:

$$\text{Decrypt}(k, \text{Encrypt}(k, m)) = k \oplus (k \oplus m) = (k \oplus k) \oplus m = m \quad (\text{since } k \oplus k = 0^n)$$

Is this secure? How do we define security?

- Given a ciphertext, cannot recover the key?

NOT GOOD! Says nothing about hiding message. $\text{Encrypt}(k, m) = m$ would be secure under this definition, but this scheme is totally insecure intuitively!

- Given a ciphertext, cannot recover the message.

NOT GOOD! Can leak part of the message. $\text{Encrypt}(k, (m_0, m_1)) = (m_0, m_1 \oplus k)$. This encryption might be considered secure but leaks half the message. [Imagine if message was "username: alice || password: 123456"]

- Given a ciphertext, cannot recover any bit of the message.

NOT GOOD! Can still learn parity of the bits (or every pair of bits), etc. Information still leaked...

↳ this might be the string that is leaked!

- Given a ciphertext, learn nothing about the message.

GOOD! But how to define this?

Coming up with good definitions is difficult! Definitions have to rule out all adversarial behavior (i.e., capture broad enough class of attacks)

↳ Big part of crypto is getting the definitions right. Pre-1970s: cryptography has relied on intuition, but intuition is often wrong! Just because I cannot break it does not mean someone else cannot...

How do we capture "learning nothing about the message"?

If the key is random, then ciphertext should not give information about the message.

Definition. A cipher (Encrypt, Decrypt) satisfies perfect secrecy if for all messages $m_0, m_1 \in \mathcal{M}$, and all ciphertexts $c \in \mathcal{C}$:

$$\Pr[k \xleftarrow{\$} K : \text{Encrypt}(k, m_0) = c] = \Pr[k \xleftarrow{\$} K : \text{Encrypt}(k, m_1) = c]$$

probability that encryption of m_0 is c , where the probability is taken over the random choice of the key k

Perfect secrecy says that given a ciphertext, any two messages are equally likely.

⇒ Cannot infer anything about underlying message given only the ciphertext (i.e., "ciphertext-only" attack)

Theorem. The one-time pad satisfies perfect secrecy.

Proof. Take any message $m \in \{0,1\}^n$ and ciphertext $c \in \{0,1\}^n$. Then,

$$\begin{aligned} \Pr[k \xleftarrow{\$} \{0,1\}^n : \text{Encrypt}(k, m) = c] &= \Pr[k \xleftarrow{\$} \{0,1\}^n : k \oplus m = c] \\ &= \Pr[k \xleftarrow{\$} \{0,1\}^n : k = m \oplus c] \\ &= \frac{1}{2^n} \end{aligned}$$

This holds for all messages m and ciphertexts c , so one-time pad satisfies perfect secrecy.

Are we done? We now have a perfectly-secure cipher!

No! Keys are very long! In fact, as long as the message... [if we can share keys of this length, can use same mechanism to share the message itself]

"One-time" restriction

Malleable

Issues with the one-time pad:

- One-time: Very important. Never reuse the one-time pad to encrypt two messages. Completely broken!

Suppose $c_1 = k \oplus m_1$ and $c_2 = k \oplus m_2$

$$\begin{aligned} \text{Then, } c_1 \oplus c_2 &= (k \oplus m_1) \oplus (k \oplus m_2) \\ &= m_1 \oplus m_2 \end{aligned}$$

← can leverage this to recover messages
← learn the xor of two messages!

One-time pad reuse:

- Project Verona (U.S. counter-intelligence operation against U.S.S.R during Cold War)

↳ Soviets reused some pages in codebook ~ led to decryption of ~ 3000 messages sent by Soviet intelligence over 37-year period [notably exposed espionage by Julius and Ethel Rosenberg]

- Microsoft Point-to-Point Tunneling (MS-PTP) in Windows 98/NT (used for VPN)

↳ Same key (in stream cipher) used for both server → client communication AND for client → server communication ↳ (RC4)

- 802.11 WEP: both client and server use same key to encrypt traffic

many problems just beyond one-time pad reuse (can even recover key after observing small number of frames!)

- Malleable: one-time pad provides no integrity; anyone can modify the ciphertext:

$$m \leftarrow k \oplus c$$

← replace c with $c \oplus m'$

$$\Rightarrow k \oplus (c \oplus m') = m \oplus m' \leftarrow \text{adversary's change now xored into original message}$$

Theorem (Shannon). If a cipher satisfies perfect secrecy, then $|K| \geq |M|$.

Intuition: Every ciphertext can decrypt to at most $|K| < |M|$ messages. This means that ciphertext leaks information about the message (not all messages equally likely). Cannot be perfectly secret.

Proof. We will use a "counting" argument. Suppose $|K| < |M|$. Take any ciphertext $c \leftarrow \text{Encrypt}(k, m)$ for some $k \in K, m \in M$. This ciphertext can only decrypt to at most $|K|$ possible messages (one for each choice of key). Since $|K| < |M|$, there is some message $m' \in M$ such that

$$\forall k \in K : \text{Decrypt}(k, c) \neq m'$$

By correctness of the cipher,

$$\forall k \in K : \text{Encrypt}(k, m') \neq c$$

This means that

$$\left. \begin{array}{l} \Pr[k \xleftarrow{\$} K : \text{Encrypt}(k, m') = c] = 0 \\ \Pr[k \xleftarrow{\$} K : \text{Encrypt}(k, m) = c] > 0 \end{array} \right\} \text{Cannot be perfectly secret!}$$

Take-away: Perfect secrecy requires long keys. Very impractical (except in the most critical scenarios - exchanging daily codebooks)

If we want something efficient/usable, we need to compromise somewhere.

- Observe: Perfect secrecy is an information-theoretic (i.e., a mathematical) property

Even an infinitely-powerful (computationally-unbounded) adversary cannot break security

We will relax this property and only require

security against computationally-bounded (efficient) adversaries