Perfect secrecy

perfectly secret

- Perfectly secret :Encryption schemes that are provably secure even against an adversary with unbounded computational power.

Probability review

- *Random variable (r.v.):* variable that takes on (discrete) values with certain probabilities

- Probability distribution for a r.v. specifies the probabilities with which the variable takes on each possible value
  - Each probability must be between 0 and 1
  - The probabilities must sum to 1

Probability review

- *Event*: a particular occurrence in some experiment
  - Pr[E]: probability of event E

- Conditional probability: probability that one event occurs, *given that* some other event occurred
  - Pr[A | B] = Pr[A and B]/Pr[B]

- Two random variables X, Y are *independent* if for all x, y: Pr[X=x | Y=y] = Pr[X=x]

Probability review

- Law of total probability: say $E_1, \ldots, E_n$ are a *partition* of all possibilities. Then for any A:

  $$\Pr[A] = \Sigma_i \Pr[A \text{ and } E_i] = \Sigma_i \Pr[A \mid E_i] \cdot \Pr[E_i]$$

Notation

- $\mathcal{K}$ (key space) – set of all possible keys

- $\mathcal{C}$ (ciphertext space) – set of all possible ciphertexts

Probability distributions

- Let M be the random variable denoting the value of the message
  - M ranges over $\mathcal{M}$
  - Context dependent!
  - Reflects the likelihood of different messages being sent, given the attacker's prior knowledge
  - E.g.,

$$\Pr[M = \text{“attack today”}] = 0.7$$
$$\Pr[M = \text{“don't attack”}] = 0.3$$

Probability distributions

- Let K be a random variable denoting the key
  - K ranges over $\mathcal{K}$
- Fix some encryption scheme (Gen, Enc, Dec)
  - Gen defines a probability distribution for K:

$$\Pr[K = k] = \Pr[\text{Gen outputs key } k]$$

Probability distributions

- Random variables M and K are *independent*
  - Require that parties don’t pick the key based on the message, or the message based on the key

Probability distributions

- Fix some encryption scheme (Gen, Enc, Dec), and some distribution for M
- Consider the following (randomized) experiment:
  1. Generate a key k using Gen
  2. Choose a message m, according to the given distribution
  3. Compute $c \leftarrow \mathrm{Enc}_k(m)$
- This defines a distribution on the ciphertext!
- Let C be a random variable denoting the value of the ciphertext in this experiment

Example 1

- Consider the shift cipher
  - So for all $k \in \{0, \ldots, 25\}$, $\Pr[K = k] = 1/26$

- Say $\Pr[M = \text{'a'}] = 0.7$, $\Pr[M = \text{'z'}] = 0.3$

- What is $\Pr[C = \text{'b'}]$?
  - Either M = 'a' and $K = 1$, or M = 'z' and $K = 2$
  - $\Pr[C=\text{'b'}] = \Pr[M=\text{'a'}] \cdot \Pr[K=1] + \Pr[M=\text{'z'}] \cdot \Pr[K=2]$
    $= 0.7 \cdot (1/26) + 0.3 \cdot (1/26)$
    $= 1/26$

Example 1

We can calculate conditional probabilities as well. For example, what is the probability that the message $\mathtt{a}$ was encrypted, given that we observe ciphertext $\mathtt{B}$? Using Bayes' Theorem (Theorem A.8) we have

$$\begin{aligned}\Pr[M = \mathtt{a} \mid C = \mathtt{B}] &= \frac{\Pr[C = \mathtt{B} \mid M = \mathtt{a}] \cdot \Pr[M = \mathtt{a}]}{\Pr[C = \mathtt{B}]} \\ &= \frac{\Pr[C = \mathtt{B} \mid M = \mathtt{a}] \cdot 0.7}{1/26}.\end{aligned}$$

Note that $\Pr[C = \mathtt{B} \mid M = \mathtt{a}] = 1/26$, since if $M = \mathtt{a}$ then the only way $C = \mathtt{B}$ can occur is if $K = 1$ (which occurs with probability $1/26$). We conclude that $\Pr[M = \mathtt{a} \mid C = \mathtt{B}] = 0.7$. $\diamondsuit$

Example 2

- Consider the shift cipher, and the distribution on M given by

Consider the shift cipher again, but with the following distribution over $\mathcal{M}$:

$$\Pr[M = \texttt{kim}] = 0.5,\ \Pr[M = \texttt{ann}] = 0.2,\ \Pr[M = \texttt{boo}] = 0.3.$$

What is the probability that $C = \texttt{DQQ}$? The only way this ciphertext can occur is if $M = \texttt{ann}$ and $K = 3$, or $M = \texttt{boo}$ and $K = 2$, which happens with probability $0.2 \cdot 1/26 + 0.3 \cdot 1/26 = 1/52$.

We can also compute the probability that ann was encrypted, conditioned on observing the ciphertext DQQ? A calculation as above using Bayes' Theorem gives $\Pr[M = \texttt{ann} \mid C = \texttt{DQQ}] = 0.4$. $\diamondsuit$

Perfect secrecy (informal)

- "Regardless of any *prior* information the attacker has about the plaintext, the ciphertext should leak no *additional* information about the plaintext"

a scheme to be perfectly secret, observing this ciphertext should have *no effect* on the adversary's knowledge regarding the actual message that was sent; in other words, the *a posteriori* probability that some message $m \in \mathcal{M}$ was sent, conditioned on the ciphertext that was observed, should be no different from the *a priori* probability that m would be sent. This means that the ciphertext reveals nothing about the underlying plaintext, and the adversary learns absolutely nothing about the plaintext that was encrypted. Formally:

Perfect secrecy (informal)

- Attacker's information about the plaintext = attacker-known *distribution* of M

- Perfect secrecy means that observing the ciphertext should not change the attacker's knowledge about the distribution of M

Perfect secrecy (formal)

- Encryption scheme (Gen, Enc, Dec) with message space $\mathcal{M}$ and ciphertext space $\mathcal{C}$ is *perfectly secret* if for every distribution over $\mathcal{M}$, every m $\in \mathcal{M}$, and every c $\in \mathcal{C}$ with Pr[C=c] > 0, it holds that

Pr[M = m | C = c] = Pr[M = m].

- I.e., the distribution of M does not change conditioned on observing the ciphertext

Equivalent formulation

the distribution of the ciphertext does not depend on the plaintext, i.e., for any two messages $m,m^0 \in M$ the distribution of the ciphertext when m is encrypted should be identical to the distribution of the ciphertext when m^0 is encrypted. That is, for every $m,m^0 \in M$, and every $c \in C$, we have

$$\mathbf{Pr[Enc}_{K}(m) = c] = \mathbf{Pr[Enc}_{K}(^{m0}) = c]$$

it is impossible to distinguish an encryption of m from an encryption of m0, since the distributions of the ciphertext are the same in each case.

Perfect secrecy

- *An encryption scheme* (Gen,Enc,Dec) *with message space* $\mathcal{M}$ *is perfectly secret if and only if Equation below holds for every* $m, m^0 \in \mathcal{M}$ *and every* $c \in \mathcal{C}$.

$$\mathbf{Pr[Enc}_K\mathbf{(}m\mathbf{) = }c\mathbf{] = Pr[Enc}_K\mathbf{(}^{m0}\mathbf{) = }c\mathbf{]}$$

- *for any scheme, any distribution on M, any m ▯ M for which Pr[M = m] > 0, and any c ▯ C, we have*

$$\begin{aligned}\Pr[C = c \mid M = m] &= \Pr[\mathsf{Enc}_K(M) = c \mid M = m] \\ &= \Pr[\mathsf{Enc}_K(m) = c \mid M = m] \\ &= \Pr[\mathsf{Enc}_K(m) = c],\end{aligned}$$

$$\Pr[M = m \mid C = c] \cdot \Pr[C = c] = \Pr[C = c \mid M = m] \cdot \Pr[M = m]. \quad (2.3)$$

Perfect secrecy

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

$$\begin{aligned}
\Pr[\mathsf{Enc}_K(m) = c] &= \Pr[C = c \mid M = m] \\
&= \Pr[C = c] \\
&= \Pr[C = c \mid M = m'] = \Pr[\mathsf{Enc}_K(m') = c]
\end{aligned}$$

Perfect (adversarial) indistinguishability

- Adversary passively observing a ciphertext and then trying to guess which of two possible messages was encrypted.
- An adversary A first specifies two arbitrary messages m0,m1 ∈ M.
- A key k is generated using Gen.
- One of the two messages specified by A is chosen (each with probability 1/2) and encrypted using k; the resulting ciphertext is given to A.
- A outputs a “guess” as to which of the two messages was encrypted; A succeeds if it guesses correctly

Perfectly indistinguishable

The adversarial indistinguishability experiment $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi}$**:**

1. *The adversary* $\mathcal{A}$ *outputs a pair of messages* $m_0, m_1 \in \mathcal{M}$.
2. *A key* k *is generated using* Gen, *and a uniform bit* $b \in \{0,1\}$ *is chosen. Ciphertext* $c \leftarrow \mathsf{Enc}_k(m_b)$ *is computed and given to* $\mathcal{A}$. *We refer to* c *as the* challenge ciphertext.
3. $\mathcal{A}$ *outputs a bit* b'.
4. *The output of the experiment is defined to be* 1 *if* $b' = b$, *and* 0 *otherwise. We write* $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1$ *if the output of the experiment is* 1 *and in this case we say that* $\mathcal{A}$ succeeds.

it is trivial for A to succeed with probability 1/2 by outputting a random guess

Perfectly indistinguishable

- *Encryption scheme* Π = (Gen,Enc,Dec) *with message space* M *is* perfectly indistinguishable *if for every* A *it holds that*

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

- *Encryption scheme Π is perfectly secret if and only if it is perfectly indistinguishable.*
- *Vigen`ere cipher is not perfectly indistinguishable*

One-time pad

- Patented in 1917 by Vernam
  - Recent historical research indicates it was invented (at least) 35 years earlier

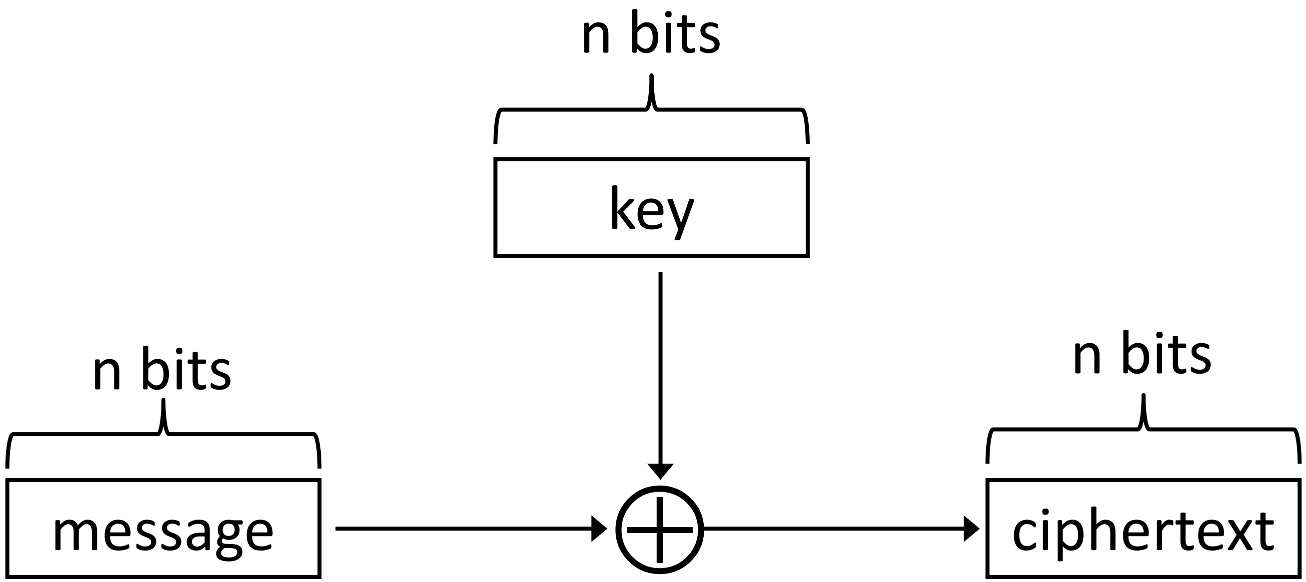
- Proven perfectly secret by Shannon (1949)

One-time pad

- Let $\mathcal{M} = \{0,1\}^n$
- Gen: choose a uniform key $k \in \{0,1\}^n$
- $Enc_k(m) = k \oplus m$
- $Dec_k(c) = k \oplus c$

- Correctness:
$$Dec_k(\ Enc_k(m)\ ) = k \oplus (k \oplus m)$$
$$= (k \oplus k) \oplus m = m$$

One-time pad



Perfect secrecy of one-time pad

- *The one-time pad encryption scheme is perfectly secret.*

Fix an integer $\ell > 0$. The message space $\mathcal{M}$, key space $\mathcal{K}$, and ciphertext space $\mathcal{C}$ are all equal to $\{0,1\}^\ell$ (the set of all binary strings of length ℓ).

- Gen: the key-generation algorithm chooses a key from $\mathcal{K} = \{0,1\}^\ell$ according to the uniform distribution (i.e., each of the 2^ℓ strings in the space is chosen as the key with probability exactly $2^{-\ell}$).
- Enc: given a key $k \in \{0,1\}^\ell$ and a message $m \in \{0,1\}^\ell$, the encryption algorithm outputs the ciphertext $c := k \oplus m$.
- Dec: given a key $k \in \{0,1\}^\ell$ and a ciphertext $c \in \{0,1\}^\ell$, the decryption algorithm outputs the message $m := k \oplus c$.

Secrecy of the one-time pad

PROOF We first compute $\Pr[C = c \mid M = m]$ for arbitrary $c \in \mathcal{C}$ and $m \in \mathcal{M}$ with $\Pr[M = m] > 0$. For the one-time pad, we have

$$\begin{aligned}\Pr[C = c \mid M = m] &= \Pr[K \oplus m = c \mid M = m] \\ &= \Pr[K = m \oplus c \mid M = m] \\ &= 2^{-\ell},\end{aligned}$$

Using the above result, we see that for any $c \in C$ we have

$$\begin{aligned}\Pr[C = c] &= \sum_{m \in \mathcal{M}} \Pr[C = c \mid M = m] \cdot \Pr[M = m] \\ &= 2^{-\ell} \cdot \sum_{m \in \mathcal{M}} \Pr[M = m]] \\ &= 2^{-\ell},\end{aligned}$$

Secrecy of the one-time pad

where the sum is over $m \in \mathcal{M}$ with $\Pr[M = m] \neq 0$. Bayes' Theorem gives:

$$\begin{aligned}
\Pr[M = m \mid C = c] &= \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\Pr[C = c]} \\
&= \frac{2^{-\ell} \cdot \Pr[M = m]}{2^{-\ell}} \\
&= \Pr[M = m].
\end{aligned}$$

We conclude that the one-time pad is perfectly secret.

Drawbacks of one time pad

- The key is as long as the message
- Is only secure if used once
- It is easy to see that encrypting more than one message with the same key leaks a lot of information.

$$c \oplus c^0 = (m \oplus k) \oplus (m^0 \oplus k) = m \oplus m^0$$

Limitations of Perfect Secrecy

- that any perfectly secret encryption scheme must have a key space that is at least as large as the message space

THEOREM 2.11 *If* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is a perfectly secret encryption scheme with message space* $\mathcal{M}$ *and key space* $\mathcal{K}$*, then* $|\mathcal{K}| \geq |\mathcal{M}|$.

Shannon's Theorem

THEOREM 2.12 (Shannon's theorem) *Let* (Gen, Enc, Dec) *be an encryption scheme with message space* $\mathcal{M}$, *for which* $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$. *The scheme is perfectly secret if and only if:*

1. *Every key* $k \in \mathcal{K}$ *is chosen with (equal) probability* $1/|\mathcal{K}|$ *by* Gen.
2. *For every* $m \in \mathcal{M}$ *and every* $c \in \mathcal{C}$, *there is a unique key* $k \in \mathcal{K}$ *such that* $\mathsf{Enc}_k(m)$ *outputs* c.

Summary

Discussed

- Probabilities
- Perfect Secrecy
- One-Time pad
- Shannon's Theorem on perfect secrecy