Perfectly-Secret Encryption

CSE 5351: Introduction to Cryptography

Reading assignment:

- Read Chapter 2
- You may skip proofs, but are encouraged to read some of them.

Outline

- Definition of encryption schemes
- Shannon's notion of perfect secrecy
- Shannon's Theorems
- Limitions of perfect secrecy
- Perfect indistinguishability

Symmetric-key encryption scheme

- An encryption scheme Π consists of three algorithms Gen, Enc, Dec and three spaces K, M, C.
- K, M, C: key space, message space, ciphertext space.
- Key generation algorithm Gen generates keys k according to some distribution (usually uniform distribution).
  We write $k \leftarrow Gen$.
- Encryption algorithm: $c \leftarrow Enc_k(m)$
- Decryption algorithm: $m := Dec_k(c)$

- Note: Gen and Enc are probabilistic algorithms, Dec is deterministic.

- Note: We don't need to explicitly specify K and C as they are implicitly defined by Gen and Enc, respectively.
- Correctness requirement: for any $k \in K$ and $m \in M$,

$$Dec_k\left(Enc_k(m)\right) = m.$$

- To use the scheme, Alice and Bob run Gen to generate a key $k \in K$, and keep it secret.

- Question: What is the security requirement?

Example encryption scheme

- Consider Caesar's shift cipher with $M = \{\mathrm{a,b,c,d}\}$ represented as $\{0,1,2,3\}$.
- Key generation: $k \leftarrow_u \{0,\ldots,25\}$.
- Encryption:
  - Randomly generate a bit $b \leftarrow \{0,1\}$.
  - Let $Enc_k(m) = (m + k + 5b) \bmod 26$.
  - I.e., $Enc_k(m) = \begin{cases} (m+k) \bmod 26 & \text{with probability } 1/2 \\ (m+k+5) \bmod 26 & \text{with probability } 1/2 \end{cases}$
- Decryption: ?

The notion of security

- Consider a ciphertext-only attack, where the adversary is an eavesdropper with a single ciphertext $c \leftarrow Enc_k(m)$.
- Adversary's possible objectives:
  1. To recover the secret key k.
  2. To recover the plaintext m.
  3. To recover any information about m.
- We will adopt and formalize the last one (#3).
- Informally, an encryption scheme is secure if from a ciphertext c no adversary can obtain any information about its plaintext m.

Shannon's notion of perfect secrecy

- Adversary: an eavesdropper with unlimited computing power and being able to see a single ciphertext.
- Encryption scheme: $(Gen, Enc, Dec, \mathcal{K}, \mathcal{M}, \mathcal{C})$
- Envision an experiment:
  - Alice generates a key $\mathsf{K} \leftarrow Gen$,
  - picks a message M from the message space $\mathcal{M}$ according to some probability distribution, and
  - obtains a ciphertext $\mathsf{C} = Enc_{\mathsf{K}}(\mathsf{M})$.
- M, K, C are random variables over $\mathcal{M}$, $\mathcal{K}$, $\mathcal{C}$, respectively.

- Notation:
  - $\Pr[\mathsf{M}=m]=$ probability that message m is picked.
  - $\Pr[\mathsf{K}=k]=$ probability that key k is generated by *Gen*.
  - $\Pr[\mathsf{C}=c]=$ probability that c is the ciphertext.
  - The distribution of M is a characteristic of M.
  - The distribution of K is determined by *Gen*.
  - The distribution of C is induced by *Enc* and depends on the distributions of M and C:

$$\Pr[\mathsf{C}=c]=\sum_{m\in M,\, k\in K}\Pr[\mathsf{M}=m]\cdot\Pr\left[\mathsf{K}=k\right]\cdot\Pr\left[Enc_k(m)=c\right]$$

- Conditional probabilities:
  - $\Pr[\mathsf{C}=c \mid \mathsf{M}=m] \;=\; \sum_{k\in K} \Pr[\mathsf{K}=k]\cdot\Pr[Enc_k(m)=c]$
  - $\Pr[\mathsf{M}=m \mid \mathsf{C}=c] \;=\; \dfrac{\Pr[(\mathsf{M}=m)\wedge(\mathsf{C}=c)]}{\Pr[\mathsf{C}=c]}$

$$=\sum_{k\in K}\Pr[(\mathsf{M}=m)\wedge(\mathsf{K}=k)\wedge(Enc_k(m)=c)] \Big/ \Pr[\mathsf{C}=c]$$

$$=\sum_{k\in K}\Pr[\mathsf{M}=m]\cdot\Pr[\mathsf{K}=k]\cdot\Pr[Enc_k(m)=c] \Big/ \Pr[\mathsf{C}=c]$$

(M and K and the randomness of *Enc* are assumed to be independent.)

Example encryption scheme

- Consider Caesar's shift cipher with $M = \{\text{a,b,c,d}\}$ represented as $\{0,1,2,3\}$.
- Key generation: $k \leftarrow_u \{0,\ldots,25\}$.
- Encryption: $Enc_k(m) = \begin{cases} (m+k) \bmod 26 & \text{with probability } 1/2 \\ (m+k+5) \bmod 26 & \text{with probability } 1/2 \end{cases}$
- Assume $\Pr[\mathsf{M} = m] = (m+1)/10$.
- Get familiar with these: $\Pr[Enc_k(m) = c]$, $\Pr[Enc_{\mathsf{K}}(m) = c]$,

  $\Pr[Enc_{\mathsf{K}}(\mathsf{M}) = c]$, $\Pr[\mathsf{C} = c]$, $\Pr[\mathsf{C} = c \mid \mathsf{M} = m]$, $\Pr[\mathsf{M} = m \mid \mathsf{C} = c]$

- Shannon's Definition of Perfect Scerecy:

  An encryption scheme is perfectly secret if for every probability distribution over M, every message $m \in M$, and every ciphertext $c \in C$ for which $\Pr[\mathsf{C} = c] > 0$, it holds:

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

- Lemma 1: An encryption scheme is perfectly secret if and only if for all $m, m' \in M$ and $c \in C$, it holds:

$$\Pr[Enc_{\mathsf{K}}(m) = c] = \Pr[Enc_{\mathsf{K}}(m') = c]$$

where $\Pr[Enc_{\mathsf{K}}(m) = c] = \sum_{k \in K} \Pr[\mathsf{K} = k] \cdot \Pr[Enc_k(m) = c].$

- To show a scheme not perfectly secret, it suffices to show a counterexample, i.e., to construct a distribution over M, a message $m \in M$, and a ciphertext $c \in C$ with $\Pr[\mathsf{C} = c] > 0$, such that:

$$\Pr[\mathsf{M} = m \mid \mathsf{C} = c] \neq \Pr[\mathsf{M} = m]$$

- Or construct two messages $m,\ m' \in M$ and a $c \in C$ such that

$$\Pr[Enc_{\mathsf{K}}(m) = c] \neq \Pr[Enc_{\mathsf{K}}(m') = c]$$

Example encryption scheme

- Consider Caesar's shift cipher with $M = \{\text{a,b,c,d}\}$ represented as $\{0,1,2,3\}$.
- Key generation: $k \leftarrow_u \{0,\ldots,25\}$.
- Encryption:
  - Randomly generate a bit $b \leftarrow \{0,1\}$.
  - Let $Enc_k(m) = (m + k + 5b) \bmod 26$.
- For every $m \in M$, $c \in C$, it holds:

$$\Pr\left[Enc_{\mathsf{K}}(m) = c\right] = \Pr\left[(m + \mathsf{K} + 5b) \bmod 26 = c\right]$$
$$= 1/2 \cdot \Pr\left[\mathsf{K} = (c - m) \bmod 26\right] + 1/2 \cdot \Pr\left[\mathsf{K} = (c - m - 5) \bmod 26\right]$$
$$= 1/26.$$

- This scheme is perfectly secret by Lemma 1.

Vernam's one-time pad encryption scheme

- $M = K = C = \{0,1\}^n$, n fixed.

  Key generation: $k \leftarrow_u \{0,1\}^n$.

  Encryption: $c := m \oplus k$.

- One-time pad: each key is used only once.

- The scheme is perfectly secret (against eavesdroppers having a single ciphertext). Reasons:
  - $\forall m, c \in \{0,1\}^n$, $Enc_k(m) = c$ iff $k = m \oplus c$.
  - Thus, $\Pr[Enc_{\mathsf{K}}(m) = c] = \Pr[\mathsf{K} = m \oplus c] = 1/2^n$.
  - Apply Lemma 1.

If a pad is used twice

- $M = K = C = \{0,1\}^n$, n fixed.

  Key generation: $k \leftarrow \{0,1\}^n$.

  Encryption: $c := m \oplus k$.

- If a key k is used to encrypt two messages:

  $$c := m \oplus k \quad \text{and} \quad c' := m' \oplus k$$

- From c and c', the adversary can tell something about the messages: $m \oplus m' = c \oplus c'$.
- The scheme is not secure against eavesdroppers with multiple ciphertexts.

If a pad is used twice

- We may regard the scheme as having $K=\{0,1\}^n$ and $M=C=\{0,1\}^{2n}$, with encryption algorithm:

$$Enc_k(m)=m\oplus(k\,\|\,k).$$

- It is not perfectly secret since

$$\Pr\left[\mathsf{M}=0^n0^n\mid\mathsf{C}=0^n1^n\right]=0\neq\Pr\left[\mathsf{M}=0^n0^n\right]>0$$

for the uniform distribution over M.

One-time pad for messages of varying length

- $M = C = \{0,1\} \cup \{0,1\}^n$, n fixed.

  $K = \{0,1\}^n$.

  Key generation: $k \leftarrow_u \{0,1\}^n$.

  Encryption: $c := E_k(m) := m \oplus k$

  where if $m \in \{0,1\}$ then only the first bit of k is used.

- Question: Is this scheme perfectly secret?

Shannon's Theorems

- Theorem 1: [a necessary condition for perfect secrecy]
  If an encryption scheme is perfectly secrect, then $|M| \leq |K|$.

- Thus, if $M = \{0,1\}^n$ and $K = \{0,1\}^l$, then $n \leq l$,
  i.e., keys must be at least as long as messages.

- Theorem 2: When $|M| = |K| = |C|$, the encryption scheme is perfectly secret if and only if both of the following hold:
  - Every key is generated by Gen with equal probability $1/|K|$;
  - For every $m \in M$ and $c \in C$, there is a unique $k \in K$ such that $Enc_k(m) = c$. (Encrypting a message m with different keys k will yield different ciphertexts c.)

Proof of $|M| \leq |K|$ (Theorem 1)

- Consider the uniform distribution over M.
  Let c be any ciphertext such that $\Pr[\mathsf{C} = c] > 0$.
  Let $M(c) = \{Dec_k(c) : k \in K\}$, the set of all messages that may be encrypted to c with non-zero probability.
  Clearly, $|M(c)| \leq |K|$.
- If $M(c) \neq M$, then there is a message $m \in M - M(c)$ for which $\Pr[\mathsf{M} = m \mid \mathsf{C} = c] = 0 \neq \Pr[\mathsf{M} = m]$, contradicting the assumption of perfect secrecy.
- Hence, $M(c) = M$, and thus $|M(c)| \leq |K|$.

Proof of Theorem 2

- Observation: $|M| = |C| \Rightarrow Enc$ is deterministic.
- Sufficiency:

The two conditions hold

$\Rightarrow$ For every $m \in M$ and $c \in C$, $\Pr\left[Enc_K(m) = c\right] = 1/|K|$

$\Rightarrow$ For all $m,\ m' \in M$ and $c \in C$,

$$\Pr\left[Enc_K(m) = c\right] = \Pr\left[Enc_K(m') = c\right]$$

$\Rightarrow$ Perfect secrecy (by Lemma 1).

- Necessity: Assume $|M| = |K| = |C|$ and perfect secrecy.
  - Consider any arbitrary (but fixed) $c \in C$.
  - Let $K(m) = \{k \in K : Enc_k(m) = c\}$, the set of all keys encrypting m to c. Note: $K(m) \cap K(m') = \varnothing$ if $m \neq m'$.
  - There is an $\bar{m} \in M$ with $\Pr[Enc_{\mathsf{K}}(\bar{m}) = c] \neq 0$, since $|M| = |C|$. By Lemma 1, $\Pr[Enc_{\mathsf{K}}(m) = c] \neq 0$ for every $m \in M$. Thus, $|K(m)| \geq 1$ for every $m \in M$.
  - This, together with $|M| = |K|$, implies $|K(m)| = 1$.
  - Let k_m be the unique key in $K(m)$ that encrypts m to c.
  - Then, $\Pr[Enc_{\mathsf{K}}(m) = c] = \Pr[\mathsf{K} = k_m]$.
  - Since $\Pr[Enc_{\mathsf{K}}(m) = c] = \Pr[Enc_{\mathsf{K}}(m') = c]$ for all $m,\ m' \in M$, $\Pr[\mathsf{K} = k_m] = \Pr[\mathsf{K} = k_{m'}] = 1/|K|$ for all $m,\ m' \in M$.

Applying Shannon's Theorem

- With Shannon's theorem, it is trivial to see that Vernam's one-time pad is perfectly secret.
- It is easy to design another perfectly secret encryption scheme.
- For example, take Caesar’s shift cipher:
  - $K = M = C = \{0,\ 1,\ \ldots,\ 25\} = \{a, b, \ldots, z\}.$
  - Key generation: $k \leftarrow_u K.$
  - Encryption: $E_k(m) = (m + k) \bmod 26$
- Caesar’s shift cipher is perfectly secret if it is used to encrypt only one letter.

Is it perfectly secret?

- Suppose we use Caesar's shift cipher to encrypt a message (any sequence of letters), but uniformly randomly generate a new key for each letter.
- Thus, $K = M = C = \{0,\ 1,\ ...,\ 25\}^* = \{a, b, ..., z\}^*$.
- To encrypt a message $m = m_1 m_2 \ldots m_t$:
  - Generate a key $k = k_1 k_2 \ldots k_t$, with $k_i \leftarrow_u K$ for each i.
  - Let $Enc_k(m) = c_1 c_2 \ldots c_t$, where $c_i = (m_i + k_i) \bmod 26$.
- Each plaintext letter m_i is perfectly protected, but not the entire message.

Limitations of Perfect Secrecy

- To achieve perfect secrecy:
  - keys must be as long as messages (if $K = \{0,1\}^l$ and $M = \{0,1\}^n$);
  - a new key must be generated for each message.

- It is desired to use a short key to encrypt multiple messages.
  - To this end, we need to relax the security requirement.
  - Unfortunately, it is hard to relax the conditions of perfect secrecy.
  - We will define a different notion of security that is equivalent to perfect secrecy and can be easily relaxed.

Perfect Indistinguishability Experiment $\mathrm{PrivK}^{\mathrm{eav}}_{A,\Pi}$

- Imagine an experiment on an encryption scheme Π:
  - The adversary A chooses two messages $m_0, m_1 \in M$, not necessarily of the same length.
  - Bob generates a key $k \leftarrow Gen$ and a bit $b \leftarrow_u \{0,1\}$. He computes and gives the ciphertext $c \leftarrow Enc_k(m_b)$ to A. (c is called the challenge ciphertext.)
  - A outputs a bit b', triying to tell whether c is the encryption of m_0 or m_1.
  - The output, $\mathrm{PrivK}^{\mathrm{eav}}_{A,\Pi}$, of the experiment is 1 iff $b = b'$ (i.e., A succeeds.)

Definition of Perfect Indistinguishability

- Encryption scheme: $\Pi = (Gen,\ Enc,\ Dec)$ with message space M.
- Adversary: an eavesdropper with unlimited computing power.
- We model the adversary as a probabilistic algorithm A that on input $m_0, m_1 \in M$ and $c \in C$ outputs a bit $b' \in \{0,1\}$.
- An encryption scheme is perfectly indistinguishable if for every adversary A and every two messages $m_0,\ m_1 \in M$,

$$\Pr\left[\mathrm{PrivK}^{\mathrm{eav}}_{A,\,\Pi}(m_0,\ m_1) = 1\right] \leq \frac{1}{2}$$

or, equivalently, for every A and every two $m_0,\ m_1 \in M$,

$$\Pr\left[\mathrm{PrivK}^{\mathrm{eav}}_{A,\,\Pi}(m_0,\ m_1) = 1\right] = \frac{1}{2}$$

- Some authors write $\Pr\left[\mathrm{PrivK}^{\mathrm{eav}}_{A,\,\Pi}(m_0,\ m_1)=1\right]$ as

  $\Pr\left[A\left(m_0,m_1,Enc_k(m_b)\right)=b:\ b\leftarrow_u\{0,1\},\ k\leftarrow_{Gen}K\right]$

  where $A\left(m_0,m_1,Enc_k(m_b)\right)$ indicate the output of A on input $m_0,m_1,Enc_k(m_b)$.

- Thus, an encryption scheme is perfectly indistinguishable if for every adversary A and every two messages $m_0,m_1\in M$,

  - $\Pr\left[A\left(m_0,m_1,Enc_k(m_b)\right)=b:\ b\leftarrow_u\{0,1\},\ k\leftarrow Gen\right]\ \le\ \frac{1}{2}$

  or, equivalently,

  - $\Pr\left[A\left(m_0,m_1,Enc_k(m_0)\right)=1:\ k\leftarrow Gen\right]$

    $=\ \Pr\left[A\left(m_0,m_1,Enc_k(m_1)\right)=1:\ k\leftarrow Gen\right]$

Remark

$$\Pr\left[\mathrm{PrivK}^{\mathrm{eav}}_{A,\,\Pi}(m_0,\ m_1)=1\right]$$

$$= \ \Pr\left[A\big(m_0, m_1, E_k(m_b)\big)=b:\ b\leftarrow_u \{0,1\},\ k\leftarrow Gen\right]$$

$$= \sum_{\substack{b\in\{0,1\}\\ k\in K}} \Pr[b]\cdot\Pr[k]\cdot\Pr\left[A\big(m_0, m_1, Enc_k(m_b)\big)=b\right]$$

$$= \sum_{\substack{b\in\{0,1\}\\ k\in K, c\in C}} \Pr[b]\cdot\Pr[k]\cdot\Pr\left[Enc_k(m_b)=c\right]\cdot\Pr\left[A\big(m_0, m_1, c\big)=b\right]$$

$$= \sum_{\substack{b\in\{0,1\}\\ c\in C}} \Pr[b]\cdot\Pr\left[Enc_K(m_b)=c\right]\cdot\Pr\left[A\big(m_0, m_1, c\big)=b\right]$$

- $A\big(m_0, m_1, Enc_k(m_b)\big)$ = output of A on input $m_0, m_1, Enc_k(m_b)$.

Equivalence of perfect secrecy and perfect indistinguishability

- Theorem: An encryption scheme is perfectly secret if and only if it is perfectly indistinguishable.

Perfect secrecy $\Rightarrow$ perfect indistinguishability

- If the encryption scheme is perfectly secret, then

$$\Pr\left[Enc_K(m_0)=c\right]=\Pr\left[Enc_K(m_1)=c\right] \text{ for all } m_0, m_1 \in M,\ c \in C.$$

- $\Pr\left[\mathrm{PrivK}^{\mathrm{eav}}_{A,\,\Pi}(m_0,\ m_1)=1\right] \qquad \left(=\Pr\left[A \text{ wins}\right]\right)$

$$=\sum_{i=0,1;\ c\in C}\Pr\left[b=i,\ Enc_K(m_i)=c,\ A(m_0,m_1,c)=i\right]$$

$$=\sum_{c\in C}\sum_{i=0,1}\Pr\left[b=i\right]\cdot\Pr\left[Enc_K(m_i)=c\right]\cdot\Pr\left[A(m_0,m_1,c)=i\right]$$

$$=\frac{1}{2}\sum_{c\in C}\left(\Pr\left[Enc(m_0)=c\right]\cdot\sum_{i=0,1}\Pr\left[A(m_0,m_1,c)=i\right]\right)=\frac{1}{2}$$

Perfect secrecy $\Leftarrow$ perfect indistinguishability

- If not perfectly secret, then there exist $m_0, m_1 \in M$ such that $\Pr\left[Enc_K(m_0) = c^*\right] \neq \Pr\left[Enc_K(m_1) = c^*\right]$ for some ciphertext $c^* \in C$.
- Define an adversary A as follows.
  - A chooses the above two messages m_0, m_1.
  - On a given challenge ciphertext c,

$$A(m_0, m_1, c) = \begin{cases} 0 & \text{if } \Pr\left[Enc_K(m_0) = c\right] > \Pr\left[Enc_K(m_1) = c\right] \\ 1 & \text{if } \Pr\left[Enc_K(m_0) = c\right] < \Pr\left[Enc_K(m_1) = c\right] \\ b' \leftarrow_u \{0,1\} & \text{otherwise} \end{cases}$$

- It can be verified that A succeeds with probability $>1/2$.
- The scheme is not perfectly indistinguishable.