Lecture 2: Shannon and Perfect Secrecy

Instructor: Omkant Pandey

Spring 2017 (CSE 594)

Last Class

- We discussed some historical ciphers
- ...and how to break them

- This class: a more formal treatment of ciphers.
- Specifically Shannon's treatment of secure ciphers

Symmetric Ciphers

- A symmetric cipher consists of:
  - A method for generating random keys k, denoted by KG
  - Encryption algorithm: Enc
  - Decryption algorithm: Dec
- Enc encrypts messages using a secret key:
  - $\mathsf{Enc}(k, m) \to c$
  - Enc may use randomness
  - c is called the ciphertext
- Dec should decrypt correctly:

$$\forall k, \forall m : \mathsf{Dec}(k, \mathsf{Enc}(k, m)) = m.$$

- The set of all messages m is called *message space* $\mathcal{M}$;
- c is called the *ciphertext* and set of all ciphertexts *ciphertext space* $\mathcal{C}$;
- The set of all keys k is called the *key space* $\mathcal{K}$.
- messages m are also known as *plaintexts*.

Security of a Cipher

What about security?

What should it mean **intuitively**?

First attempt: hide the key

- All ciphers in the frequency analysis recover the key...

  What if we just guarantee that key remains completely hidden?

- No reason why plaintext should be hidden!

- Example from Caesar Cipher:
  ATTACK = BUUBDL and DEFEND = EFGFOE

  Broken by checking patterns! don't need the key!

Second approach: hide the message

- What does it mean?
- Hide the full message only?
- Hide every letter of the message?
- What if the ciphertext reveals the frequency of the alphabets in the plaintext?
- Dangerous: May be enough to find out if the army will attack or defend?
- Hide *everything* about the message: all possible functions of the message.
  - Good starting point but impossible! Something about the message may already be known!
    (E.g., it is in English, starts with "Hello" and today's date, etc.)

Third approach: hide everything that is not already known!

- We cannot hide what may be *a priori* known about the message.
- The ciphertext must hide everything else!
- Adversary should not learn any **NEW** information about the message after seeing the ciphertext.
- How to capture it mathematically?

Shannon's Treatment

- Messages come from some *distribution*; let D be a random variable for sampling the messages from the message space $\mathcal{M}$.
- Distribution D is known to the adversary. This captures *a priori* information about the messages.
- The ciphertext $c = \mathsf{Enc}(m, k)$ depends on:
  - m chosen according to D
  - k is chosen randomly (according to KG)
  - Enc may also use some randomness
  - These induce a distribution C over the ciphertexts c.
- The adversary only observes c
  (for some $m \stackrel{D}{\leftarrow} \mathcal{M}$ and $k \stackrel{\mathsf{KG}}{\leftarrow} \mathcal{K}$, but m, k themselves)

Shannon's Treatment (continued)

- Knowledge about m **before** observing the output of C is captured by: D
- Knowledge about m **after** observing the output of C is captured by: $D|C$
- **Shannon secrecy**: distribution D and $D|C$ must be **identical**.
- Intuitively, this means that:

  C contains **no NEW information** about m

  ...in the standard sense of information theory.

Shannon Secrecy

Definition (Shannon Secrecy)

A cipher $(\mathcal{M}, \mathcal{K}, \mathsf{KG}, \mathsf{Enc}, \mathsf{Dec})$ is **Shannon secure w.r.t a distribution** D over $\mathcal{M}$ if for all $m' \in \mathcal{M}$ and for all c,

$$\Pr\left[m \leftarrow D : m = m'\right] =$$
$$\Pr\left[k \leftarrow \mathsf{KG}, m \leftarrow D : m = m' | \mathsf{Enc}(m, k) = c\right]$$

It is Shannon secure if it is Shannon secure w.r.t. all distributions D over $\mathcal{M}$.

Questions?

Perfect Secrecy

- Suppose you have two messages: $m_1 \in \mathcal{M}$ and $m_2 \in \mathcal{M}$.
- What is the distribution of ciphertexts for m_1?

$$C_1 := \{k \leftarrow \mathsf{KG},\ \text{output } \mathsf{Enc}(m_1, k)\}$$

- Likewise, for m_2, the ciphertext distribution is:

$$C_2 := \{k \leftarrow \mathsf{KG},\ \text{output } \mathsf{Enc}(m_2, k)\}$$

- **Perfect secrecy:**
  C_1 and C_2 must be **identical for every pair** of m_1, m_2.

  $\Rightarrow$ Ciphertexts are independent of the plaintext(s)!

Perfect Secrecy (conitinued)

Definition (Perfect Secrecy)

Scheme $(\mathcal{M}, \mathcal{K}, \mathsf{KG}, \mathsf{Enc}, \mathsf{Dec})$ is **perfectly secure** for every pair of messages m_1, m_2 in $\mathcal{M}$ and for all c,

$$\Pr\left[k \leftarrow \mathsf{KG} : \mathsf{Enc}(m_1, k) = c\right] = \Pr\left[k \leftarrow \mathsf{KG} : \mathsf{Enc}(m_2, k) = c\right]$$

- So much simpler than Shannon Secrecy!
- No mention of distributions, a priori or posteriori.
- Much easier to work with...

Which notion is better?

- OK, so we have two definitions: perfect secrecy and Shannon secrecy.
- Both of them intuitively seem to guarantee great security!

- Is one better than the other?
- If our intuition is right, shouldn't they offer "same level" of security?

Equivalence Theorem

Theorem (Equivalence Theorem)

A private-key encryption scheme is perfectly secure if and only if it is Shannon secure.

Proof: Simplifying Notation

- We drop KG and D when clear from context.
- $\mathsf{Enc}_k(m)$ will be shorthand for $\mathsf{Enc}(m, k)$
- For example:
  - $\Pr_m[\ldots]$ means $\Pr[m \leftarrow D : \ldots]$
  - $\Pr_k[\ldots]$ means $\Pr[k \leftarrow \mathsf{KG} : \ldots]$
  - $\Pr_{k,m}[\ldots]$ means $\Pr[k \leftarrow \mathsf{KG}, m \leftarrow D : \ldots]$

Proof: Perfect Secrecy ⇒ Shannon Secrecy

Given: $\forall (m_1, m_2) \in \mathcal{M} \times \mathcal{M}$ and every $c \in \mathcal{C}$:

$$\Pr_k[\mathsf{Enc}_k(m_1) = c] = \Pr_k[\mathsf{Enc}_k(m_2) = c]$$

Show: for every D over $\mathcal{M}$, $m' \in \mathcal{M}$, and $c \in \mathcal{C}$:

$$\Pr_{k,m}[m = m' | \mathsf{Enc}_k(m) = c] = \Pr_m[m = m']$$

Proof: Perfect Secrecy $\Rightarrow$ Shannon Secrecy (continued)

$$\begin{aligned}
\text{L.H.S.} &= \Pr_{k,m}[m = m' | \mathsf{Enc}_k(m) = c] \\
&= \frac{\Pr_{k,m}[m=m' \quad \cap \quad \mathsf{Enc}_k(m)=c]}{\Pr_{k,m}[\mathsf{Enc}_k(m)=c]} \\
&= \frac{\Pr_{k,m}[m=m' \quad \cap \quad \mathsf{Enc}_k(m')=c]}{\Pr_{k,m}[\mathsf{Enc}_k(m)=c]} \\
&= \frac{\Pr_m[m=m'] \cdot \Pr_k[\mathsf{Enc}_k(m')=c]}{\Pr_{k,m}[\mathsf{Enc}_k(m)=c]} \\
&= \text{R.H.S.} \times \frac{\Pr_k[\mathsf{Enc}_k(m')=c]}{\Pr_{k,m}[\mathsf{Enc}_k(m)=c]}
\end{aligned}$$

Proof: Perfect Secrecy ⇒ Shannon Secrecy (continued)

Show:

$$\frac{\Pr_k[\mathsf{Enc}_k(m') = c]}{\Pr_{k,m}[\mathsf{Enc}_k(m) = c]} = 1$$

Proof:

$$\begin{aligned}
\Pr_{k,m}[\mathsf{Enc}_k(m) = c] &= \sum_{m'' \in \mathcal{M}} \Pr_m[m = m''] \Pr_k[\mathsf{Enc}_k(m'') = c] \\
&= \sum_{m'' \in \mathcal{M}} \Pr_m[m = m''] \Pr_k[\mathsf{Enc}_k\underline{(m')} = c] \\
&= \Pr_k[\mathsf{Enc}_k(m') = c] \cdot \underbrace{\sum_{m'' \in \mathcal{M}} \Pr_m[m = m'']} \\
&= \Pr_k[\mathsf{Enc}_k(m') = c] \times 1. \quad \text{(QED)}
\end{aligned}$$

Proof: Perfect Secrecy $\Leftarrow$ Shannon Secrecy

We have to show: $\forall (m_1, m_2) \in \mathcal{M} \times \mathcal{M}$ and $\forall c$:

$$\Pr_k[\mathsf{Enc}_k(m_1) = c] = \Pr_k[\mathsf{Enc}_k(m_2) = c]$$

Fix any m_1, m_2, c as above.
Let D be the uniform distribution over $\{m_1, m_2\}$ so that:

$$\Pr_m[m = m_1] = \Pr_m[m = m_2] = 1/2.$$

By definition, the scheme is Shannon secure w.r.t. this D. Therefore,

$$\begin{aligned}\Pr_{k,m}[m = m_1 | \mathsf{Enc}_k(m) = c] &= \Pr_m[m = m_1], \text{ and} \\ \Pr_{k,m}[m = m_2 | \mathsf{Enc}_k(m) = c] &= \Pr_m[m = m_2]\end{aligned}$$

Proof: Perfect Secrecy $\Leftarrow$ Shannon Secrecy (continued)

Therefore: $\Pr_{k,m}[m = m_1 | \mathsf{Enc}_k(m) = c] = \Pr_{k,m}[m = m_2 | \mathsf{Enc}_k(m) = c]$

Consider the LHS:

$$\Pr_{k,m}[m = m_1 | \mathsf{Enc}_k(m) = c] = \frac{\Pr_{k,m}[m = m_1 \ \cap \ \mathsf{Enc}_k(m) = c]}{\Pr_{k,m}[\mathsf{Enc}_k(m) = c]}$$

$$= \frac{\Pr_m[m = m_1] \cdot \Pr_k \mathsf{Enc}_k(m_1) = c]}{\Pr_{k,m}[\mathsf{Enc}_k(m) = c]}$$

$$= \frac{\frac{1}{2} \cdot \Pr_k \mathsf{Enc}_k(m_1) = c]}{\Pr_{k,m}[\mathsf{Enc}_k(m) = c]}$$

Likewise, the RHS is:

$$\Pr_{k,m}[m = m_2 | \mathsf{Enc}_k(m) = c] = \frac{\frac{1}{2} \cdot \Pr_k \mathsf{Enc}_k(m_2) = c]}{\Pr_{k,m}[\mathsf{Enc}_k(m) = c]}$$

Cancel and rearrange. (QED)

Should we go over this proof again?

The *One Time Pad* : A perfect secure scheme

- Let n be an integer = length of the plaintext messages.
- Message space $\mathcal{M} := \{0,1\}^n$ (bit-strings of length n)
- Key space $\mathcal{K} := \{0,1\}^n$ (keys too are length n bit-strings)
- The key is as long as the message
- The algorithms are:
  - KG: samples a key uniformly at random $k \leftarrow \{0,1\}^n$
  - $\mathsf{Enc}(m,k)$: XOR bit-by-bit,
    Let $m = m_1 m_2 \ldots m_n$ and $k = k_1 k_2 \ldots k_n$;
    Output $c = c_1 c_2 \ldots c_n$ where $c_i = m_i \oplus k_i$ for every $i \in [n]$.
  - $\mathsf{Dec}(c,k)$: XOR bit-by-bit.
    Return m where $m_i = c_i \oplus k_i$ for every i.

Perfect Security of OTP

Theorem (Perfect security of OTP)

One Time Pad is a perfectly secure private-key encryption scheme.

- Let $a \oplus b$ for n-bit strings a, b mean bit-wise XOR.
- Then: $\mathsf{Enc}(m, k) = m \oplus k$ and $\mathsf{Dec}(c, k) = c \oplus k$.
- Ciphertext space is $\mathcal{C} := \{0,1\}^n$. Correctness: straightforward.
- Perfect secrecy: fix any $m \in \{0,1\}^n$ and $c \in \{0,1\}^n$.

$$\begin{aligned} \Pr_k[\mathsf{Enc}_k(m) = c] &= \Pr[m \oplus k = c] \\ &= \Pr[k = m \oplus c] = 2^{-n}. \\ \Pr_k[\mathsf{Enc}_k(m) = c] &= 0 \quad (\forall c \notin \{0,1\}^n) \end{aligned}$$

$\Rightarrow \forall (m_1, m_2) \in \{0,1\}^{n \times n}$ and $\forall c :$
$\Pr_k[\mathsf{Enc}_k(m_1) = c] = \Pr_k[\mathsf{Enc}_k(m_2) = c]$. (QED)

Some Remarks

- The One Time Pad (OTP) scheme is also known as the **Vernam Cipher**.
- The Caesar Cipher is just OTP for 1-alphabet messages!
- Mathematically:
  - XOR is the same as **addition modulo 2:**

    $a + b \mod 2$.
  - Caesar Cipher for 1-alphabet is **addition modulo 26**.
  - You can work modulo any number n
- As the name suggests, one key can be used only once.
- The key must be:
  - sampled uniformly every time, and
  - be as long as the message.

Key Length in Perfectly Secure Encryption

- If the key has to be as long as the message, it is a serious problem!
- Imagine encrypting your machine's hard drive with a OTP...
  - 80 GB long key to encrypt 80 GB data
  - 80 GB space to store this key in a safe place (other than your hard drive)
  - Key for OTP is uniform, so it cannot be compressed either!
  - This is never done in practice...
- OTP looks naïve, quite elementary: can't we design a more sophisticated scheme with shorter keys?

Shannon's Theorem

Theorem (Shannon's Theorem)

For every perfectly secure cipher $(\mathsf{Enc}, \mathsf{Dec})$ *with message space* $\mathcal{M}$ *and key space* $\mathcal{K}$*, it holds that* $|\mathcal{K}| \geqslant |\mathcal{M}|$.

Some Remarks:

- Message length is $n = \lg|\mathcal{M}|$ and key length is $\ell = \lg|\mathcal{K}|$.
- It follows that $\ell \geqslant n$, i.e., keys must be as long as the messages.

Shannon's Theorem

Theorem (Shannon's Theorem)

For every perfectly secure cipher $(\mathsf{Enc}, \mathsf{Dec})$ *with message space* $\mathcal{M}$ *and key space* $\mathcal{K}$*, it holds that* $|\mathcal{K}| \geqslant |\mathcal{M}|$.

Proof:

- Assume the contrary: i.e., $|\mathcal{K}| < |\mathcal{M}|$
- Fix any message m_0, and any key k_0. Let

$$c_0 = \mathsf{Enc}(m_0, k_0).$$

$$\implies Pr_k[\mathsf{Enc}(m_0, k) = c_0] > 0. \quad (1)$$

- What happens if we decrypt c_0 with each key one by one? We get a set of messages, which we denote by:

$$S = \{\mathsf{Dec}(c_0, k) : k \in \mathcal{K}\}.$$

- Note that $|S| \leqslant |\mathcal{K}|$ and $|\mathcal{K}| < |\mathcal{M}|$.

$$\implies |S| < |\mathcal{M}|.$$

Proof continued..

- This means, there exists a message $m_1 \in \mathcal{M}$ such that $m_1 \notin S$.
- What happens if we encrypt m_1 with a key $k \in \mathcal{K}$?
- Since $m_1 \notin S$, by definition:

$$\forall k \in \mathcal{K}: \qquad \mathsf{Enc}(m_1, k) \neq c_0.$$

$$\implies \Pr_k[\mathsf{Enc}(m_1, k) = c_0] = 0. \tag{2}$$

- Therefore, there exist m_0, m_1, c_0 such that:

$$\Pr_k[\mathsf{Enc}(m_0, k) = c_0] \neq \Pr_k[\mathsf{Enc}(m_1, k) = c_0].$$

- This contradicts perfect secrecy. (QED)

Exercise: Reusing OTP

- What could go wrong if you re-use a OTP anyway?
- If we could re-use then we could encrypt longer messages with shorter keys.
- Simply break the message in shorter parts.
- Therefore, by Shannon's Theorem, the resulting scheme will not be perfectly secure.
- Even worse — it will be open to the frequency attack!
(just like Vigènere Cipher)
- In fact, lots of neat examples where reusing OTP leaks clear patterns.
- Can you construct such examples?

Back to Key Length in Perfect Secrecy

- Shannon's Theorem on key length is pretty bad news for perfect ciphers.
- It means we really have to give up on perfect secrecy for practical applications, unless we absolutely need it.
- This is really the dawn of modern cryptography: we want to construct something that is "just as good for practical purposes."

- The modern approach focuses on what computers can do **efficiently**.
- For example, if we have a short, efficient computer program, which generates large "random looking" strings, we can use this program to generate strings that look like a OTP key.
- This is really what we will try to do — we will build some theory in the next few lectures and return to this issue again.
- In the next class, we will review notions of "efficient computation" and define what is called a "one way function."