CSL7480 Cryptography **09 Jan 2023**

Lecture 3: Cryptography

Lecturer: Somitra Sanadhya *Scribe: Sumit Kumar Prajapati (B20CS074)*

3.1 History

There were two area of study -

- **Cryptography**: designing secure communication schemes
- **Cryptoanalysis**: attacking above schemes to break it

Nowadays, we study both under Cryptography

3.2 Attack Models

The goal of the attacker is to gain some information about plaintext m* corresponding to a given cipher text c*. We categorise the attack models based on the resources accessible to the attacker.

1. Ciphertext Only
   - The attacker has access to only ciphertext c*.
2. Known Plaintext Attack
   - The attacker has access to some (m_i, c_i) pairs s.t. $Enc_k(m_i) = c_i$ and $c_i \neq c*$.
3. Chosen Plaintext Attack (CPA)
   - The attacker has access to some (m_i, c_i) pairs s.t. $Enc_k(m_i) = c_i$ and $c_i \neq c*$. The attacker can choose specific m_i and ask for its corresponding encryption c_i.
4. Chosen Ciphertext Attack (CCA)
   - The attacker has access to some (m_i, c_i) pairs s.t. $Enc_k(m_i) = c_i$ and $c_i \neq c*$. Here, the attacker can choose specific either m_i (and ask for the corresponding encryption c_i) or c_i(and ask for the corresponding decryption m_i) for each pair.
   - CCA is further divided into two categories:
     - **CCA-1**: The c* will be given at the end of after all pair-exchanges and no further query will be allowed.
     - **CCA-2**: The c* can be asked by the attacker at any point of time. He/she can query more pairs even after c* is revealed.

Clearly, the power(resources) of the attacker increases from top to bottom in the list. Our ultimate goal is to design a cryptographic scheme which will be secure under **CCA-2** attack model.

3.3 Encryption Scheme Model

- An encryption scheme consists of three spaces:
  - **Key Space** ($\mathcal{K}$): The key k is generated at random using $KeyGen()$ algorithm. It may have some security parameter as input (for eg. length of the key).
  - **Message Space** ($\mathcal{M}$): 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.
  - **Ciphertext Space** ($\mathcal{C}$): The ciphertext $c = Enc(k, m)$ depends on:
    - m chosen according to D.
    - k chosen randomly (according to $KeyGen()$)
    - Enc may also use some randomness

    These induce a distribution C over the ciphertexts c.
- For correctness of the scheme, following must hold: $\forall k \in \mathcal{K}, \forall m \in \mathcal{M},\ Dec(k, Enc(k, m)) = m$.

3.4 'Unbreakable Cryptosystem'

- Intuitively, we might want to define perfect security of an encryption scheme as follows: **Given a ciphertext all messages are equally likely**.
- This can be formulated as:
$$\forall\ m_0, m_1 \in M, c \in C$$
$$Pr[M = m_0 | C = c] = Pr[M = m_1 | C = c]$$
- The probability here is over the randomness used in the $KeyGen()$ and Enc algorithms and the probability distribution over the message space.
- But this definition has a problem. It might be a priori known that the message m_0 is more likely than m_1. We do not want 'seeing the ciphertext' to change this information.
- We want the ciphertext to provide **no additional information** about the message.

3.4.1 Shannon's Secrecy

A cipher $(\mathcal{M}, \mathcal{K}, KeyGen, Enc, Dec)$ is Shannon secure w.r.t a distribution D over M if

$$\forall m' \in \mathcal{M}, \forall c \in \mathcal{C}$$

$$Pr[m \leftarrow \mathcal{D} : m = m'] = Pr[k \leftarrow KeyGen() : m = m' | Enc(m, k) = c]$$

where the first probability is taken over $\mathcal{M}$ chosen according to distribution D, over random keys K chosen in $\mathcal{K}$, and over the possible random choices of the (possibly) probabilistic encryption algorithm Enc, while the second probability is taken over $M \leftarrow D$.
It is Shannon secure if it is Shannon secure w.r.t **all distributions** D **over** $\mathcal{M}$.

3.4.2 Perfect Security

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

$$C_1 := \{Enc(m_1, k) \mid k \leftarrow KeyGen()\}$$

- What is the distribution of ciphertexts for m_2?

$$C_2 := \{Enc(m_2, k) \mid k \leftarrow KeyGen()\}$$

- For **perfect secrecy**:
  C_1 and C_2 must be identical for every pair m_1, m_2.
  $\Rightarrow$ Ciphertexts are *independent* of the plaintext(s)

  Definition: A Scheme $(\mathcal{M}, \mathcal{K}, KeyGen, Enc, Dec)$ is **perfectly secure** if

$$\forall m_1, m_2 \in \mathcal{M}, \forall c \in \mathcal{C}$$

$$Pr[k \leftarrow KeyGen() : Enc(m_1, k) = c] = Pr[k \leftarrow KeyGen() : Enc(m_2, k) = c]$$

  where both probabilities are taken over the choice of K in $\mathcal{K}$ and over the coin tosses of the (possibly) probabilistic algorithm $Enc()$

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

3.4.3 Which notion is better?

- We have two definitions: Shannon secrecy and Perfect 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?

3.4.4 Equivalence of Shannon Secrecy and Perfect Secrecy

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

$$Perfect\ Secrecy \Leftrightarrow Shannon\ Secrecy$$

<u>Simplifying Notation</u>

- We drop $KeyGen$ and D when clear from context.
- $Enc_k(m)$ will be shorthand for $Enc(k, m)$.

- For example:
  - $Pr_m[...] = Pr[m \leftarrow D : ...]$
  - $Pr_k[...] = Pr[k \leftarrow KeyGen() : ...]$
  - $Pr_{k,m}[...] = Pr[m \leftarrow D, k \leftarrow KeyGen() : ...]$

Proof: Perfect Secrecy $\Rightarrow$ Shannon Secrecy

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

$$Pr_k[Enc_k(m_1) = c] = Pr_k[Enc_k(m_2) = c]$$

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

$$Pr_{k,m}[m = m' \mid Enc_k(m) = c] = Pr_m[m = m']$$

$$L.H.S = Pr_{k,m}[m = m' \mid Enc_k(m) = c] \tag{3.1}$$

$$= \frac{Pr_{k,m}[m = m' \cap Enc_k(m) = c]}{Pr_{k,m}[Enc_k(m) = c]} \tag{3.2}$$

$$\tag{3.3}$$

$$= \frac{Pr_{k,m}[m = m' \cap Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]} (\because m = m') \tag{3.4}$$

$$\tag{3.5}$$

$\because Pr[m = m']$ is independent of k and $Pr[Enc_k(m') = c\ ]$ is independent of m

$$= \frac{Pr_{k,m}[m = m'].Pr[Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]} \tag{3.6}$$

$$= R.H.S \times \frac{Pr[Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]} \tag{3.7}$$

$$\tag{3.8}$$

Now, we will show that $\frac{Pr[Enc_k(m')=c]}{Pr_{k,m}[Enc_k(m)=c]} = 1$

The probability that we get a cipher-text c from any message m is the sum of the probabilities of each test in the message set $\mathcal{M}$ leading to c on encryption using Enc

$$\therefore Pr[Enc_k(m') = c] = \sum_{m=m"} Pr_m[m = m"] Pr_k[Enc_k(m") = c] \tag{3.9}$$

$$\tag{3.10}$$

$\because$ probability of getting ciphertext c is equal for every message in $\mathcal{M}$

$$= \sum_{m=m''} Pr_m[m = m''] Pr_k[Enc_k(m') = c] \tag{3.11}$$

$$= Pr_k[Enc_k(m') = c] \sum_{m=m''} Pr_m[m = m''] \tag{3.12}$$

$$= Pr_k[Enc_k(m') = c] \times 1 \text{ (QED)} \tag{3.13}$$

$$\tag{3.14}$$

Proof: Perfect Secrecy $\Leftarrow$ Shannon Secrecy

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

$$Pr_{k,m}[m = m' \mid Enc_k(m) = c] = Pr_m[m = m']$$

To Show: $\forall (m_1, m_2) \in \mathcal{M} \times \mathcal{M},\ \forall c \in \mathcal{C}$:

$$Pr_k[Enc_k(m_1) = c] = Pr_k[Enc_k(m_2) = c]$$

Now,

$$Pr_{k,m}[m = m_1 \mid Enc_k(m) = c] = \frac{Pr_{k,m}[m = m_1 \cap Enc_k(m) = c]}{Pr_{k,m}[Enc_k(m) = c]} \tag{3.15}$$

$$= \frac{Pr_{k,m}[m = m_1 \cap Enc_k(m_1) = c]}{Pr_{k,m}[Enc_k(m) = c]} (\because m = m_1) \tag{3.16}$$

$$= \frac{Pr_{k,m}[m = m_1].Pr[Enc_k(m_1) = c]}{Pr_{k,m}[Enc_k(m) = c]} \tag{3.17}$$

$$\tag{3.18}$$

$$\because Pr_{k,m}[m = m_1 \mid Enc_k(m) = c] = Pr_{k,m}[m = m'] \tag{3.19}$$

$$\frac{Pr_{k,m}[m = m_1].Pr[Enc_k(m_1) = c]}{Pr_{k,m}[Enc_k(m) = c]} = Pr_{k,m}[m = m'] \tag{3.20}$$

$$Pr[Enc_k(m_1) = c] = Pr_{k,m}[Enc_k(m) = c] \tag{3.21}$$

$$Pr[Enc_k(m_1) = c] = Pr_{k,m}[Enc_k(m) = c] \tag{3.22}$$

$$\text{Similarily, } Pr[Enc_k(m_2) = c] = Pr_{k,m}[Enc_k(m) = c] \tag{3.23}$$

$$\therefore Pr[Enc_k(m_1) = c] = Pr[Enc_k(m_2) = c]\ \forall m_1, m_2 \in \mathcal{M} \text{ (QED)}$$

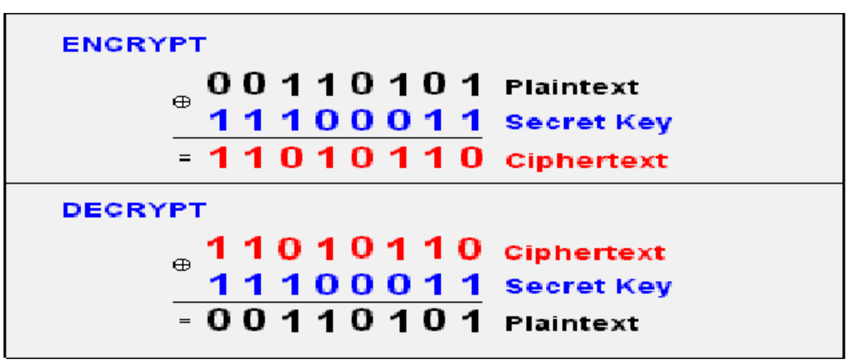
3.4.5 Perfect Security: Key Size Requirement

Theorem: For Perfect Security, $|\mathcal{K}| \geq |\mathcal{M}|$ must hold.
Proof by Contradiction

- Assume that there is a perfectly secure cipher with $|\mathcal{K}| < |\mathcal{M}|$.
- Choose any random $m_1 \in \mathcal{M}, k \in \mathcal{K}$ and let $c = Enc_k(m)$.
- Now let $M = \{Dec'_k(c)\}$ for all possible keys k'
- Clearly, $|M| \leq |\mathcal{K}|$
- Since $|\mathcal{K}| < |\mathcal{M}|$, this means $\exists\ m_2 \notin M$.
- Hence, $Pr[Enc_k(m_2) = c] = 0$
- But, $Pr[Enc_k(m_1) = c] > 0$
- Note: The above probability will be 1 for a deterministic encryption scheme.
- There exist m_1, m_2, c s.t. $Pr[Enc_k(m_1) = c] \neq Pr[Enc_k(m_2) = c]$.
- Contradiction.

3.4.6 One Time Pad: A perfect secure scheme

- let n be an integer = length of 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. A random key is used **only once**.
- The Encryption Scheme:
  - $KeyGen()$: samples a key uniformly at random $k \leftarrow \{0, 1\}^n \Rightarrow Pr_k[k = k'] = 2^{-n}$
  - $Enc(m, k) = m \oplus k$ (bit-by-bit xor)
    Let $m = m_1m_2...m_n$ and $k = k_1k_2...k_n$;
    Output $c = c_1c_2...c_n$ where $c_i = m_i \oplus k_i \forall i \in [n]$
  - $Dec(c, k) = c \oplus k$.
  - Return m where $m_i = c_i \oplus k_i \forall i$



Theorem: One Time Pad is a perfectly secure private-key encryption scheme.
Proof
Fix $m \in \{0,1\}^n$ and $c \in \{0,1\}^n$.

$$Pr_k[Enc(m) = c] = Pr[m \oplus k = c] \tag{3.24}$$
$$= Pr[k = m \oplus c] = 2^{-n} \tag{3.25}$$
$$\tag{3.26}$$

$\Rightarrow \forall (m_1, m_2) \in \{0,1\}^{n \times n}, \forall c :$
$Pr[Enc_k(m_1) = c] = Pr_k[Enc_k(m_2) = c] = 2^{-n}$ (QED)

The One Time Pad (OTP) scheme is also known as the **Vernam Cipher**.

References

- https://www.ics.uci.edu/~stasio/fall04/lect1.pdf
- https://www3.cs.stonybrook.edu/~omkant/L02-short.pdf
- https://www.cs.purdue.edu/homes/hmaji/teaching/Fall%202016/lectures/03.pdf