CSA E0 235: Cryptography August 14, 2019

Scribe For Lecture 3

Instructor: Arpita Patra *Submitted by: Virti Savla*

1 Recap

In the previous lecture, we learned about Symmetric Key Encryption(SKE) in modern way. We also saw definition of perfect security.We assumed that the attacker has unbounded computational power and the attacker can launch Ciphertext Only Attack(COA). We also discussed construction of one time pad, proof of why one time pad is perfectly secure and the limitations of one time pad.

1.1 Formal definition for Perfect Security

Definition 1.1 *An encryption scheme (Gen, Enc, Dec) over a message space M is perfectly secret if for every probability distribution over M, every message m ∈ M, and every ciphertext c ∈ C for which Pr[C = c] >0, the following holds.*

$$\Pr[\mathrm{M} = \mathrm{m}|\mathrm{C} = \mathrm{c}] = \Pr[\mathrm{M} = \mathrm{m}].$$

This means that, the probability of adversary knowing the Message (m) remains same before and after the adversary sees the Ciphertext (c).Ciphertext does not reveal any extra information about the plaintext.

2 Equivalent definitions of Perfect Security

2.1 Definition : Perfect Security

Definition 2.1.1 *An encryption scheme (Gen, Enc, Dec) over a message space M is perfectly secret if and only if for every probability distribution over M, every m_0 , m_1 ∈ M, and every c ∈ C:*

$$\Pr[\mathrm{C} = \mathrm{c}|\mathrm{M} = \mathrm{m}_0] = \Pr[\mathrm{C} = \mathrm{c}|\mathrm{M} = \mathrm{m}_1].$$

2.2 Definition : Shannon's Theorem

Definition 2.2.1 *Let (Gen, Enc, Dec) be an encryption scheme over a message space M for which $|M| = |K| = |C|$. This scheme is perfectly secret if and only if:*

1. Every key k ∈ K is chosen with equal probability 1/|K| by algorithm Gen.
2. For every m ∈ M and every c∈C, there exists a single key k∈K such that Enc $_{\mathrm{k}}$(m) outputs c.

Proof : Let us prove that if condition (1) , (2) hold and also if |M| = |K| = |C| then the scheme is perfectly secure.

Why the condition |M|=|K|=|C| ?
It is inherent to any perfectly secure scheme that |K| ≥ |M|.(Read section 4.1)
Now, consider if |M| > |C| Using a particular key k∈K, two messages will be mapped to same ciphertext. This violates the correctness of Decryption algorithm. Thus, |C| ≥ |M|.
Thus, |M| = |K| = |C| is the ideal optimal case.

Since there is exactly one key k∈K such that Enc $_k$(m) outputs c where m∈M and c∈C.

$$\mathrm{Pr[C = c|M = m] = Pr[K{=}k] =}\ \mathbf{1/|K|}$$

$$\blacksquare ...1a$$

By Law of Total Probability,

$$Pr[C = c] = \sum_{minM} Pr[C = c|M = m]Pr[M = m] = 1/|K| * \sum_{m \in M} Pr[M = m] = 1/|k|$$

$$\blacksquare ...2a$$

So,we have shown that Pr[C =c] = Pr[C = c|M = m](from 1a and 2a). Multiplying both sides by Pr[M=m]/Pr[C=c],

$$\frac{Pr[C = c|M = m] * Pr[M = m]}{Pr[C = c]} = \frac{Pr[M = m] * Pr[C = c]}{Pr[C = c]}$$

Therefore, Pr[M = m|C = c] = Pr[M = m]. (By using Bayes theorem,
L.H.S = Pr[M=m|C=c]) Thus, the scheme is perfectly secure.

Lets prove the other direction.

If the scheme is perfectly secure then statement (1) and statement (2) hold.

Let $\mathrm{K_i}$ be set of all keys that map message $\mathrm{m_i}$ to some ciphertext c.Thus we define the subsets $\mathrm{K_i}$ over the Key space. Then we claim the following:

Claim 1 : $\mathrm{K_i} \neq \varphi$ (The set $\mathrm{K_i}$ containing all keys that map $\mathrm{m_i}$ to some ciphertext c should not be null).

Let $\mathrm{K_i}$ be set that contains keys that encrypt message $\mathrm{m_i}$ to ciphertext c such that, $\mathrm{Enc_k(m_i)}$ = c where $\mathrm{k \in K_i}$.For any perfectly secure scheme,

$$\mathrm{Pr[C = c|M = m_0] = Pr[C = c|M = m_1]} \neq 0.$$

Its true for any $\mathrm{m_i,m_j \in M}$.This means that there exists some key that encrypts the message m_i to give ciphertext c.Thus, there must be atleast one key k∈K that encrypts messages m_i to give some ciphertext c. Thus we claim that $\mathrm{K_i} \neq \varphi$.

Claim 2 : $\mathrm{K_i \cap K_j} = \varphi$ (Same key cannot be used to encrypt more than one message and give same ciphertext c).

Assume that, key k lies in the intersection of two subsets k_i and k_j. Thus, same key k maps both the messages $\mathrm{m_i}$ and $\mathrm{m_j}$ to same ciphertext c. For correctness (in decryption algorithm) to hold good, our assumption would be contradicted since it would be impossible to decrypt the ciphertext c correctly.

Since, $\mathrm{K_i} \neq \varphi$ and $\mathrm{K_i} \cap \mathrm{K_j} = \varphi$ and $|\mathrm{K}| = |\mathrm{M}|$ each set $\mathrm{K_i}$ contains only 1 key i.e $|\mathrm{K_i}| = 1$.

(a)From definition of perfect secrecy,

$$\mathrm{Pr}[\mathrm{C} = \mathrm{c}|\mathrm{M} = \mathrm{m_i}] = \mathrm{P\ r}[\mathrm{C} = \mathrm{c}|\mathrm{M} = \mathrm{m_j}]$$

which implies that,

$$\mathrm{Pr}[\mathrm{K} = \mathrm{k_i}] = \mathrm{Pr}[\mathrm{K} = \mathrm{k_j}].$$

Thus each Key in key space is equiprobable and is chosen with probabilty $1/|\mathrm{K}|$ by Gen algorithm.
This proves condition (1)

(b) Also, there exists a unique key k for every m $\in$M and every c $\in$ C such that $\mathrm{Enc_k(m)} = \mathrm{c}$.
This proves condition (2)

2.3 Advantages of Shannon's Theorem

1. Its easy to check conditions (1) and (2) of shannon's theorem.
2. There is no need of any probability calculation unlike original perfect security definition.

3 Adversarial indistinguishability experiment

Consider a hypothetical game between adversary and a challenger. The adversary is able to carry out Ciphertext Only Attack(COA) i.e the adversary sees the ciphertext and tries to guess the plaintext. The experiment is as follows:
(i)The adversary chooses a pair of message $\mathrm{m_0}$,$\mathrm{m_1}$ $\in$M
(ii)The challenger then chooses a random key k generated by Gen.Challenger also tosses a coin to select a random bit b$\in$\{0,1\} Then a cipher text $\mathrm{c{=}Enc_k(m_b)}$ is computed and given to Adversary A.
(iii)The adversary then guesses the message which was sent and ouputs a message b'$\in$\{0,1\}.
(iv)The adversary succeeds only when b'=b, also defined as

$$\mathrm{PrivK^{eav}_{A,\pi}}{=}1$$

The adversary should not be able to distinguish between the two message after observing the ciphertext c.Thus, the probability of guessing the bit b' correctly before and after seeing the ciphertext must be same.Before seeing the ciphertext the probability of guessing b' correct is 1/2 as there are only 2 messages in the message space. After seeing the ciphertext c the probability of guessing the message correct must be no greater than 1/2. Thus here we state that perfect secrecy is achieved if no adversary A can succeed with probability any better than half.Hence, this definition is equivalent to definition 1 of perfect secrecy.

Definition 3.1 *An encryption scheme (Gen, Enc, Dec) over a message space M is perfectly secret if for every adversary A it holds that*

$$\Pr[\mathrm{PrivK}^{\mathrm{eav}}_{\mathrm{A},\pi}{=}1]{=}\frac{1}{2}$$

This definition is equivalent to definition 1.1 of perfect secrecy.

4 Drawbacks of One Time Pad/Limitations of Perfect security

The main reason behind why we want to get rid of perfect security is the limitations of perfect security. The following two drawbacks of one time pad are inherent to any Perfectly Secure Scheme.
(1) The key space K must be atleast as large as the message space M i.e. $|\mathrm{K}| \geq |\mathrm{M}|$.
(2) Keys cannot be reused.
(3) The length of key must be as large as length of message.
(4)Variable length messages in message space are not allowed.For scheme to be perfectly secure all messages in message space must be of same length. If message space contains variable length messages padding must be used.
In following subsections let us look at each limitation in detail.

4.1 Key Space must be as large as Message Space

Theorem 1 *In any perfectly secure cipher scheme,* $|K| \geq |M|$.
Proof : Assume that, if $|\mathrm{K}| < |\mathrm{M}|$. Take uniform distribution over message space M. Consider arbitary message $\mathrm{m}{\in}\mathrm{M}$. Let c be the ciphertext corresponding to the encryption of message m using key k such that $\mathrm{k}{\in}\mathrm{K}$.

$$\text{Therefore, } \mathrm{c}{=}\mathrm{Enc}_{\mathrm{k}}(\mathrm{m}).$$

(Enc is randomized and $\Pr[\mathrm{C}{=}\mathrm{c} > 0]$.)

$$\text{By correctness, } \mathrm{Dec}_{\mathrm{k}}(\mathrm{c}){=}\mathrm{m}.$$

Consider a set M(c) of all messages that correspond to ciphertext c

$$\mathrm{M(c)} = \{\ \mathrm{m'} \mid \mathrm{m'} = \mathrm{Dec}_{\mathrm{k'}}(\mathrm{c}) \text{ for some } \mathrm{k'} \in \mathrm{K}\ \}$$

Now, $|\mathrm{M(c)}| \leq |\mathrm{K}|$, Since for a message m such that $\mathrm{m}{\in}\mathrm{M(c)}$ we can identify one key $\mathrm{k}{\in}\mathrm{K}$ for which $\mathrm{m}{=}\mathrm{Dec}_{\mathrm{k}}(\mathrm{c})$ assuming that decryption is deterministic.

$$|\mathrm{M(c)}| \leq |\mathrm{K}| < |\mathrm{M}| \text{ (According to our assumption } |\mathrm{K}| < |\mathrm{M}|)$$

This means there is a message m' such that $\mathrm{m'}{\neq}$ m and $\mathrm{m'}{\in}\mathrm{M}$ and $\mathrm{m'} \notin \mathrm{M(c)}$.
Thus, we infer that $\Pr[\mathrm{M}{=}\mathrm{m'} \mid \mathrm{C}{=}\mathrm{c}] = 0 \neq \Pr[\mathrm{M}{=}\mathrm{m'}]$ as $\Pr[\mathrm{M}{=}\mathrm{m'}] > 0$.
Hence, the scheme is **Not perfectly secure**. Thus, in any perfectly secure scheme,
$|\mathrm{K}| \geq |\mathrm{M}|$

4.2 Keys cannot be reused

Consider two messages m_1, m_2 such that $m_1 \neq m_2$. In One Time Pad, c = k $\oplus$ m. Consider same key k is used to encrypt message m_1 and m_2 such that c_1 = k $\oplus$ m_1 and c_2 = k $\oplus$ m_2. Thus, $c_1 \oplus c_2 = m_1 \oplus m_2$. The attacker can simply mount frequency analysis and break the scheme.
If the scheme is perfectly secure we have,
Pr[M_1= $m_1 \wedge M_2 = m_2 \mid C_1 = c_1 \wedge C_2 = c_2$]= Pr[$M_1 = m_1 \wedge M_2 = m_2$].
Let the key k be used twice then,

$$\Pr[C_1 = c_1 \wedge C_2 = c_2 \mid M_1 = m_1 \wedge M_2 = m_2] = \Pr[C_1 = c_1 \wedge C_2 = c_2 \mid M_1 = m_1' \wedge M_2 = m_2']$$

Now take two message pairs (m_1,m_1) and (m_1,m_2)

$$\Pr[C_1 = c_1 \wedge C_2 = c_2 \mid M_1 = m_1 \wedge M_2 = m_1] \neq \Pr[C_1 = c_1 \wedge C_2 = c_2 \mid M_1 = m_1' \wedge M_2 = m_2']$$

Thus, the scheme is not perfectly secure.
If the keys are reused then the scheme is not perfectly secure can be proved by the Definition 3.1 Consider, the message space consist of vector of messages M=\{(m_1,m_1),(m_1,m_2)\} The attacker chooses one message vector from M and gives to challenger.The challenger returns the ciphertext for the two messages to the attacker. If the attacker chooses message set (m_1,m_1) the ciphertext (c_1,c_2) returned by the challenger will be same i.e $c_1 = c_2$ If the attacker chooses message set (m_1,m_2) then the ciphertext (c_1,c_2) returned by challenger will be different i.e $c_1 \neq c_2$. Thus, on observing the ciphertext returned by challenger the attacker can guess the messages (m_1,m_2) with probability 1. Thus, Pr[$\mathrm{PrivK}^{\mathrm{eav}}_{\mathrm{A},\pi}$=1]=1 $>\frac{1}{2}$
Thus, the scheme is not perfectly secure.

4.3 Length of key must be atleast as large as length of message

In One Time Pad, suppose our key k $\in$ K is a binary sequence of k-bits and our message m$\in$M is a binary sequence of m-bits , then we have |K| = 2^k and also |M| = 2^m . Theorem 1 states that

$$|\mathrm{K}| \geq |\mathrm{M}|$$

which implies that

$$2^k \geq 2^m$$

.

$$\text{Hence, k} \geq \text{m}$$

i.e. the length of the key must be as large as the message length.

4.4 Messages in the message space cannot be of variable length

Perfect security is violated if the length of the message is revealed to the attacker.If the attacker knows the message space in advance, then by just observing the length of ciphertext attacker can find the plaintext. Consider the following example. If M={yes,no}, then simply by learning the length of the ciphertext attacker can find out plaintext. Thus, if message space is variable padding must be implemented.

5 Conclusion

We conclude that perfect secrecy is attainable and there are schemes where ciphertext reveals nothing about the plaintext even to the Adversary having unbounded computational power.However, there are some limitations inherent to all perfectly secure schemes. Inorder to restrict the simplest form of attack Ciphertext Only attack, we had to deal with limitation like keys cannot be reused, variable length message not allowed(padding required) etc.Inorder to restrict stronger attacks like Chosen Plaintext Attack and Chosen Ciphertext Attack we will have to impose many other limitations.These limitations of perfect security gave rise to the world of "Computational Security".In Computational Security model some of our assumptions are relaxed.
(1)Our assumption of the adversary having unbounded computational power is now reduced to being polynomially bounded.
(2) We also allow break with negligible probability.
Computational security compromises efficiency, however it helps us bypass the inherent limitations of perfect security by allowing usage of short length keys and also allowing the reusability of keys.

References.

[1] Jonathan Katz and Yehuda Lindell Introduction to Modern Cryptography, second edition.CRC Press, 2014.