1 Symmetric Encryption

Symmetric Key Encryption consists of the following triplet of algorithms:

- Key Generation
- Encryption
- Decryption

1.1 Correctness Property

It ensures that if a message is encrypted using a key, then the same message is returned as the output when the cipher text is decrypted.

1.2 Security Property

We studied various properties that a SE should possess in order to be securely used for achieving confidentiality.

The first property that we studied is called perfect or unconditional or information-theoretic security. In achieving perfect security, we consider an adversary who has an unbounded computational power (i.e., it possesses infinite amount of computing power to break an SE).

Definitions of Perfect Security:

**Definition 1:** A symmetric encryption is said to be perfectly secure if \( \forall m \in M, c \in C, \) and \( \forall \) probability distributions over \( M \)

\[
Pr(M = m) = Pr(M = m|C = c)
\]

**Definition 2:** A symmetric encryption is said to be perfectly secure if \( \forall m \in M, c \in C, \) and \( \forall \) probability distributions over \( M, \)

\[
Pr(C = c) = Pr(C = c|M = m)
\]
Definition 3: A symmetric encryption is said to be perfectly secure if \( \forall m_0, m_1 \in M, c \in C, \text{ and } \forall \text{ probability distributions over } M, \)

\[
Pr(C = c|M = m_0) = Pr(C = c|M = m_1)
\]

This means that all messages are equally likely to be the plain-text corresponding to a given ciphertext. In other words, it is impossible to distinguish between the cipher-text corresponding to any two messages. If a SE scheme is secure in one definition then it means it is secure in all. From Bayes theorem it can be proved Definition 1 \(\Rightarrow\) Definition 2.

Why is OTP perfectly secure?

In OTP \(K=M=C\{0, 1 \}^l\) i.e, the message space is the same as the key space which is the same as the cipher text space. Now suppose \(l=2\). \(2^l = 2^2 = 4\) possibilities. Suppose you are given a cipher-text, and try to find out the corresponding plain-text by XORING with all the 4 possible keys. You get 4 different messages, i.e. all the messages are likely. Probability is \(\frac{1}{4}\). You dont learn anything extra even though you know the cipher text. Therefore OTP is PS.

We cannot use OTP in practice because the key is large as long as the message.

Encryption schemes which are perfectly secure (i.e.,secure against an adversary who is computationally unbounded) are of little practical use. So we consider an adversary who is limited in power or computationally bounded.

2 Computationally bounded adversary

Definition: Realistic Adversary Intuitively, a realistic adversary is an algorithm that runs in polynomial amount of time. In our "computational security" model, any adversary that runs in exponential amount of time can always break the scheme (with a probability 1), but since its running time is exponential, an exponential time adversary can safely be ignored for all practical purposes.

Definition: Polynomial time Algorithm An algorithm running on an input of length \(n\), is said be polynomial-time if its running time is \(f(n) = O(n^c)\). In other words, for some positive constant \(c\), for some \(a, \forall \ n \in \mathbb{N} \ f(n) < an^c\).

Definition: Probabilistic Polynomial time (PPT) Algorithm An algorithm that is randomized and runs in polynomial amount of time is referred to as a probabilistic polynomial time (PPT) algorithm.

Definition: Secure Cryptographic Scheme A cryptographic scheme is said to be secure if for all polynomial time (PPT) adversaries, the success probabilities are negligible. i.e. if probability of all adversaries are negligible then the scheme is secure.
3 SECURITY NOTIONS FOR SYMMETRIC KEY ENCRYPTION

i) Key Recovery (KR): Key recovery takes place when an entity other than the two communicating parties is able to learn the shared key. Using this key, the attacker can decrypt any messages exchanged between the communicating parties. Any encryption should be secure against key recovery. The notion of security against KR is defined by the following experiment between a PPT adversary A and a challenger possessing the shared key.

\[ \text{EXP}_{SE}^{KR}(A) \]

where SE is Symmetric Encryption, KR is Key Recovery and A is PPT adversary.

![Experiment of KR on SE](image)

Figure 1: \( \text{EXP}_{SE}^{KR}(A) \)

1. The challenger selects a message \( m \) at random from the message space \( M \), and encrypts it using a random key \( k \), then sends the resulting cipher text \( c \) to the adversary; 2. The adversary does some operations (what the adversary does should be polynomially bounded) and sends back a guess of the key \( k' \). If the guessed key is the same as the actual key, the experiment returns a ‘1’, otherwise a ‘0.’ The advantage of A is to be able to make the right guess for the key, which is:

\[ \text{Adv}_{SE}^{KR}(A) = P_r(\text{Exp}_{SE}^{KR}(A) = 1) \]

A SE is said to be secure against key recovery if: \( \forall A, \text{Adv}_{SE}^{KR}(A) = \text{negl}(|K|) = \epsilon \) KR is only a necessary property but not sufficient. In other words, any SE must be secure against KR, however, a SE secure against KR need not really be secure. For example,
it might be possible to learn the plaintext given the ciphertext, without even recovering the key.

ii) One-Way Security (OW): One-Way security means that it should be infeasible to learn the message given the ciphertext. We define it using the following experiment between a PPT adversary $A$ and the challenger possessing the key:

$$\text{Exp}^{\text{OW}}_{\text{SE}}(A)$$

Figure 2: $\text{Exp}^{\text{OW}}_{\text{SE}}(A)$

1. The challenger selects a message $m$ uniformly at random from the message space $M$, and encrypts it using the random key $k$, then sends the resulting cipher text $c$ to the adversary; 2. The adversary sends back a guess of the message $m'$. 3. If $m = m'$, a ‘1’ is returned, otherwise a ‘0’ The advantage of $A$ is to be able to make the right guess, which is:

$$\text{Adv}^{\text{OW}}_{\text{SE}}(A) = \Pr(\text{Exp}^{\text{OW}}_{\text{SE}}(A) = 1)$$

A SE is said to be OW-secured if:

$$\forall A, \text{Adv}^{\text{OW}}_{\text{SE}}(A) \leq \epsilon$$

**Theorem**: (OW $\Rightarrow$ KR) If SE is secure against the OW notion, then it is also secure against the KR notion. Proof in Lecture 3

KR does not imply OW. OW-security is not a sufficient property

iii) Semantic Security (SEM): Given the cipher text, the adversary cannot learn any partial information about the plain text. Formally, we define the advantage of $A$, as

$$\text{Adv}^{\text{SEM}}_{\text{SE}}(A) = \Pr(\text{Exp}^{\text{SEM}}_{\text{SE}}^{-1}(A) = 1) - \Pr(\text{Exp}^{\text{SEM}}_{\text{SE}}^{-0}(A) = 1)$$
This means whether you are given some cipher text or not, you can guess the message. A SE is said to be SEM-secure, if:

\[ \forall A, \text{Adv}^{\text{SEM}}_{\text{SE}}(A) < \epsilon \]

iv) **INDISTINGUISHABILITY (IND):** An adversary cannot distinguish between the cipher texts of two different messages. When two messages \( m_0 \) and \( m_1 \) are given to the adversary, it can guess if \( C \) (cipher text) belongs to \( m_0 \) or \( m_1 \). The challenger resides in either World-0 or World-1.

A SE would be called IND(indistinguishable), if:

\[ \forall A, \text{Adv}^{\text{IND}}_{\text{SE}}(A) < \epsilon \]

Advantages of IND

\[ \text{Adv}^{\text{IND}}_{\text{SE}}(A) = \Pr(\text{Exp}^{\text{IND}-1}_{\text{SE}}(A) = 1) - \Pr(\text{Exp}^{\text{IND}-0}_{\text{SE}}(A) = 1) \]

v) **INDISTINGUISHABILITY(IND-G)** \( \text{Exp}^{\text{IND}-G}_{\text{SE}}(A) \):

![Figure 3: \( \text{Exp}^{\text{IND}-G}_{\text{SE}}(A) \)](image)

A SE is said to be secure against IND-G if:

\[ \text{Adv}^{\text{IND}-G}_{\text{SE}}(A) = \Pr(d = b) < \frac{1}{2} + \epsilon \]

**Theorem:** IND ⇔ IND-G

The relations among KR, OW, SEM, IND and IND-G are as follows:

SEM ⇔ IND ⇔ IND-G ⇒ OW ⇒ KR

SEM satisfies Perfect Security (\( \Pr(M = m) = \Pr(M = m \mid C = c) \)).
4 Active Adversarial Attacks

Now we look at stronger adversaries - **Active Adversary**. **Definition: Active Adversary**: An active adversary is modeled as having access to special oracles called the encryption oracle or decryption oracle which retains the secret encryption key and encrypts or decrypts arbitrary plaintexts or ciphertexts at the adversary’s request.

**Definition: CPA - Adaptive chosen plain-text attack**: In this, the adversary has more power. It has access to an Encryption Oracle. The attacker can choose an arbitrary number of plaintexts to be encrypted and obtain the corresponding ciphertexts.

**Definition: CCA - Chosen Cipher-text attack**: Contrary to the previous model, the adversary, in these attacks, has access to a decryption oracle and chooses an arbitrary number of ciphertexts to obtain the corresponding plaintexts.

**Definition: CCA2 Adaptive Chosen Cipher-text Attack**: This attack is a more interactive version of the previous one. The adversary tries to gain partial information by making queries to the decryption oracle based on the results of previous decryption.

CCA2 $\Rightarrow$ CCA $\Rightarrow$ CPA CCA2 is the strongest security notion with respect to an active adversary.
4.1 Indistinguishability under Adaptive Chosen Plaintext Attack (IND-CPA)

A cryptosystem is said to have achieved indistinguishability under adaptive chosen plaintext attack, if a PPT adversary, which can query the encryption oracle a reasonable number of times, chooses two plaintexts for the challenger to encrypt and has only a negligible advantage over random guessing in distinguishing which plaintext belongs to which ciphertext.

A cryptosystem is indistinguishable under chosen plaintext attack if every probabilistic polynomial time adversary has only a negligible advantage over random guessing.

Figure 6 shows IND-CPA as follows:

\[
\text{Adv}^{\text{IND-CPA}}(A) < \frac{1}{2} + \epsilon
\]

Proposition 1: Any deterministic encryption scheme is not IND-CPA.
4.2 Indistinguishability under Chosen Ciphertext Attack (IND-CCA)

A cryptosystem is said to have achieved indistinguishability under chosen ciphertext attack, if a PPT adversary first queries the decryption oracle a reasonable number of times, chooses two plaintexts for the challenger to encrypt and has only a negligible advantage over random guessing in distinguishing which ciphertext belongs to which plaintext.

4.3 Indistinguishability under Adaptive Chosen Ciphertext Attack (IND-CCA2)

A cryptosystem is said to have achieved indistinguishability under adaptive chosen ciphertext attack, if a PPT adversary, which can queries the decryption oracle a reasonable number of times, chooses two plaintexts for the challenger to encrypt and has only a negligible advantage over random guessing in distinguishing which ciphertext belongs to which plaintext.

A scheme is IND-CCA/IND-CCA2 secure if no adversary has a non-negligible advantage in winning the above game.
5 Pseudo Random Function (PRF) and Pseudo Random Permutation (PRP)

Definition- (Random Function (RF)): A random function is a function picked at random from a function family \( f : D\{0,1\}^l \rightarrow R\{0,1\}^L \).

Note that there are a total of \( 2^L \times 2^L \times 2^L \times \ldots \times 2^L = 2^{L \times 2^l} \) functions in the above function family.

Definition- (Pseudo Random Function (PRF)): A random instance of a function family \( (F : K \times D \Rightarrow R, k \leftarrow K, F_k : D\{0,1\}^l \Rightarrow R\{0,1\}^L) \) will be called a PRF if it is a close approximation to the RF.

To capture how close an approximation a PRF is to a random function, we define two adversarial experiments, as shown in the following figure

\[ Exp_f^{PRF^{-1}}(A), \text{ where the adversary } A \text{ interacts with an oracle initialized with a PRF} \]
Figure 8: Random Function (RF)

$F_K$ and $Exp_f^{PRF-0}(A)$, where $A$ interacts with an oracle initialized as an ideal random function. The adversarial goal is to be able to tell which world the oracle resides in. The advantage of $A$ interacting with such an oracle is defined as

$$Adv_f^{PRF-0}(A) = \Pr(Exp_f^{PRF-1}(A) = 1) - \Pr(Exp_f^{PRF-0}(A) = 1)$$

The function $F_K$ would be said to be a PRF if, for all adversaries $A$,

$$Adv_f^{PRF}(A) \leq \epsilon$$

Notice: $F_k^{-1}$ might not exist.

**Construction 1**: Construct a SE from PRF. KeyGen: $k \leftarrow K$

1. Choose key $k$ from key space $K$ as the private key.
2. Choose a random $r$ from $D$ of PRF $F_k()$ as $C_1$.
3. Make the plaintext $m \leftarrow R$ and encrypt it with $F_k(r)$ as $C_2$ by exclusive or.
4. In decryption, get $m$ by exclusive or between $C_2$ and $F_k(C_1)$.

**Theorem 1**: Construction 1 is IND-CPA if $F_k()$ is a PRF

**Definition 3 (Random Permutation (one-one and onto))**: A permutation picked at random from a permutation family $\pi : (D : \{0, 1\}^L) \rightarrow (D : \{0, 1\}^L)$

Note that there are a total of $2^L!$ permutations in the above family.

**Definition 4 (Pseudo Random Permutation (PRP) Family)**: A permutation family, for some key $k$ picked at random from the key space $K$, $\pi_k : \{0, 1\}^L \rightarrow \{0, 1\}^L$ is said to be PRP if it is a close enough approximation to the random permutation.
The advantage of $A$ interacting with such an oracle is defined as

$$Adv_{PRP}^{RP}(A) = Pr(Exp_{RP}^{PRP-1}(A) = 1) - Pr(Exp_{RP}^{PRP-0}(A) = 1)$$

The permutation $\pi_k$ would be said to be a PRP if, for all adversaries $A$, $Adv_{RP}^{PRP}(A) \leq \epsilon$

**Theorem 2**: If $\pi-k$ is a PRP, then CBC-PRP is IND-CPA.

**Theorem 3**: CBC-PRP is not secure against IND-CCA2.
6 HASH FUNCTIONS

**Definition 1- Hash Function Family:** Any mapping of the form $H : K \times D \rightarrow R$ is said to be a cryptographic hash function family, where the domain, $D$, may be arbitrarily long while the range, $R$, is small. A Hash Function is an instance of the hash function family. Generally a hash function does not have a key. Thus when we model the security of hash functions, we will simply provide the key to the adversary.

There are three properties/security notions for hash functions: 1. One-Way Security (Pre-image resistance) 2. Weak Collision Resistance (second pre-image resistance) 3. Strong collision Resistance

**Definition 2- One – Wayness (OW) (pre – image resistance):** Formally, the notion of one-wayness is defined using the following experiment/game between an adversary $A$ and a challenger possessing the shared key.

$$Exp_{H}^{OW}(A)$$

![Figure 11: $Exp_{H}^{OW}(A)$](image)

The adversarial game is :

1. The challenger chooses key $k$ and value $x$ in the domain $D$ and then computes $y = H_{k}(x)$.
2. The challenger sends a key $k$ and value $y$ to the adversary
3. The adversary then sends $x'$ to the challenger.
4. If $y = H_{k}(x')$ return 1, else return 0.

The advantage of adversary $A$ is:

$$Adv_{H}^{OW}(A) = Pr(Exp_{H}^{OW}(A) = 1)$$
A hash function, $H$, is said to be secure against one-wayness if:

$$\forall A, Adv_{H}^{OW}(A) \leq \epsilon$$

**Definition 3- Weak Collision Resistance (CR1) (secondpre – imageresistance):**

Formally, we define it using the following experiment between an adversary $A$ and the challenger possessing the key:

$$Exp_{H}^{CR1}(A)$$

![Figure 12: $Exp_{H}^{CR1}(A)$](image)

The adversarial game is:

1. The adversary $A$ sends value $x_1 \in D$ to the challenger.
2. The challenger then selects a $k$ at random from $K$. The challenger also computes $y = H_k(x)$. Then the challenger sends $y$ and $k$ back to the adversary.
3. The adversary sends $x_2$ ($x_2 \neq x_1$) to the challenger.
4. If $H_k(x_2) = y$ return 1, else return 0.

A hash function, $H$, is said to be CR1-secured if

$$Adv_{H}^{CR1}(A) = Pr(Exp_{H}^{CR1}(A) = 1)$$

$$\forall A, Adv_{H}^{CR1}(A) \leq \epsilon$$

**Definition 4- ((Strong) Collision Resistance (CR2)):** Formally, we define it using the following experiment between an adversary $A$ and the challenger possessing the key:

$$Exp_{H}^{CR2}(A)$$
A hash function, $H$, is said to be CR2-secured if
\[ \forall A, Adv^{CR2}_H(A) \leq \epsilon \]

It is proved that: $\text{CR2} \Rightarrow \text{CR1} \Rightarrow \text{OW}$

**Theorem 1 (CR2 $\Rightarrow$ CR1):** If $H$ is secure against the CR2 notion, then it is also secure against the CR1 notion.

**Definition 5 (Pre-image Set):** A pre-image set is the set of all $x \in D$ that map to the same value $y \in R$. The pre-image set of $y \in R$ is denoted by $H^{-1}_k(y)$ and it consists of all $x \in D$ that map to $y$.

\[ H^{-1}_k(y) = \{ x \in D : H_k(x) = y \} \]

Let’s define a set $S_k$ (for a given key $k$ in the keyspace) as,

\[ S_k = \{ x \in D : |H_k^{-1}(H_k(x))| = 1 \} \]

**Theorem 2 (CR1 $\Rightarrow$ OW):** If $H$ is weakly collision resistant, then it is also secure against OW.

13-14
Definition 1 \((\text{Merkle – Damgard Cascaded Hash Function Construction})\): Merkle-Damgard construction is a particular method of constructing a hash function, where if the underlying compression function \(h\) is collision resistant, then the hash function \(H\), is also collision-resistant.

The construction is as follows. A message \(m\) is split into \(n\) blocks, \(m_1, ..., m_n\) (each block being of a fixed length), with appropriate padding whenever needed. In case the message \(m\) is not \('b'\) bits it is padded to derive the same. So we can say : \(\text{pad}(m) \Rightarrow H(\text{pad}(m))\). The first block, and the IV, is passed as an input to a compression function \(h\) to yield the first chaining vector \(CV_1 = h(\text{IV}, m_1)\). \(CV_1\) and \(m_2\), are passed as input to \(h\) again to yield the second chaining vector \(CV_2 = h(CV_1, m_2)\). This chaining goes on until \(m_n\), and \(CV_{n-1}\), are passed to \(h\), which outputs \(CV_n = h(CV_{n-1}, m_n)\). This entire construction is denoted by \(H\), therefore the output of hash function becomes \(H(m) = CV_n\). A key property of this construction is that it has an avalanche effect.
Properties of padding

1. If $|m_1| = |m_2|$, then $|\text{pad}(m_1)| = |\text{pad}(m_2)|$. In case we pick the two messages $m_1$ and $m_2$ and pad the same to make them equal, then the length of the pad is going to be same.

2. If $|m_1| \neq |m_2|$, then at least the last block of $\text{pad}(m_1)$, and $\text{pad}(m_2)$ should be different.

**Theorem 1 (If h is CR2, then H is CR2):** If an Attacker 'A' can break the CR2 property of H then another Attacker 'B' can be created which can break the CR2 property of h.

**Message Authentication:** Authentication is used to verify that a party is indeed who he/she claims to be.

**Definition 2 (Message Authentication Code (MAC)):** Data appended to the message, which allows us to verify the message. Key Generation: $k \leftarrow K$

Probabilistic algorithm

$\text{MAC}():$ takes input message $m$, and key $k$. Therefore, $\sigma = \text{MAC}(m, k)$. $(m, \sigma)$ is sent as the output. This is a deterministic algorithm Verify: The receiver calculates the MAC for $m$, and checks it against the received $\sigma$. If the two are the same, the value returned is 1, otherwise a 0. This algorithm is deterministic.

7 Security Notions for Message Authentication

(I) Key Recovery (KR):

(II) No-Message Attack (NMA):

(III) Existential forgery under adaptive chosen message attacks (CMA):

One can easily show that: $\text{CMA} \Rightarrow \text{KR}$.

A PRF is a CMA-secure MAC

A PRF, using a key $k$, can be used as a (deterministic) MAC using following steps:

KeyGen: $k \leftarrow K$

MAC: $\sigma = MAC_k(m) = f_k(m)$. $(m, \sigma)$ is sent to the recipient.

Verify: Return $f_k(m) == \sigma$

**Theorem 2** ($f$-PRF $\Rightarrow MAC_f$-CMA) If $f$ is a PRF, then MAC scheme using $f$ is CMA-secure.
**CBC-MAC** One can encode longer messages by dividing a message into l-bit long blocks and chaining them, similar to CBC-DES. To verify a message and MAC pair, the receiver computes the MAC on the message and compare it with the received MAC value.

Differences between CBC MAC scheme and CBC encryption are as follows:

- CBC MAC produces one output block for any length message, whereas CBC encryption produces N l-bit output blocks for N l-bit plaintext blocks.
- Initialisation is not required for CBC MAC (i.e, IVs are not used), since the MAC is deterministic and no randomization is needed as in CBC encryption.

**Theorem 1**: If $F_K$ is a PRF (or CMA-secure), then CBC-MAC is also CMA-secure.

**Proposition 1**: CBC-MAC can be used to authenticate only a fixed length messages (i.e it can not be used to authenticate arbitrarily long messages of variable lengths)

Proofs in Lecture 9.