Contents

Chapter 1

Introduction

Cryptography is the art and science of encrypting and decrypting a message.

1.1 Symmetric cipher

A symmetric cipher scheme Π can be viewed as a triplet $(\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ of algorithms. Suppose $\mathcal{M}$ be the set of all possible messages and $\mathcal{K}$ be the set of all keys. Gen chooses a key $k \in \mathcal{K}$ and then $\mathrm{Enc} : \mathcal{M} \times \mathcal{K} \to \mathcal{C}$ encrypts the message m with key k and returns the cipher c. Lastly, $\mathrm{Dec} : \mathcal{C} \times \mathcal{K} \to \mathcal{M} \cup \perp$ decrypts the cipher c with key k and returns either a message or an error, denoted as $\perp$. Without loss of generality we can assume that Gen picks k uniformly from $\mathcal{K}$. Futhermore, Enc can be randomized, however Dec is deterministic and for every message m and key k we must have

$$\mathrm{Dec}_k(\mathrm{Enc}_k(m)) = m$$

1.2 Kerckhoff's principle

Kerckhoff's principle assumes the following for every encryption scheme

1. The encryption and decryption is known to everyone.
2. The security of the scheme is only dependent on the key.

1.3 Attacks

Some possible attacks include (in increasing power)

Cihpertext only Attacker only knows the ciphertexts.

Known-plaintext Attacker knows one or more plaintext/ciphertext generated by the key.

Chosen-plaintext Attacker can obtain encryption of plaintexts of his choice.

Chosen-ciphertext Attacker can obtain decryption of ciphertexts of his choise.

Chapter 2

Perfectly Secret Encryption

2.1 perfectly secure encryption

Let K and M be two random variables, where K is the result of Gen and M is the message. We can assume that they are independent. Furthermore, $C = \mathrm{Enc}_K(M)$ is also a random varible. By the Kerckhoff's principle, we assume that the distribution on M and Enc is known and only K is unknown.

Definition (Perfectly secure encrption): An encryption scheme is perfectly secure if for all $c \in \mathcal{C}$ with $\mathbb{P}(C = c) > 0$:

$$\forall m \in \mathcal{M}, \quad \mathbb{P}(M = m \mid C = c) = \mathbb{P}(M = m) \tag{2.1}$$

Proposition 2.1. *An encryption scheme* Π *is perfectly secure if and only if*

$$\forall m, m' \in \mathcal{M}, \quad \mathbb{P}(\mathrm{Enc}_K(m) = c) = \mathbb{P}(\mathrm{Enc}_K(m') = c) \tag{2.2}$$

Proof. Suppose Π is perfectly secure then (assuming that $\mathbb{P}(M = m) > 0$)

$$\begin{aligned}\mathbb{P}(\mathrm{Enc}_K(m) = c) = \mathbb{P}(C = c \mid M = m) &= \frac{\mathbb{P}(M = m \mid C = c)\mathbb{P}(C = c)}{\mathbb{P}(M = m)} \\ &= \frac{\mathbb{P}(M = m)\mathbb{P}(C = c)}{\mathbb{P}(M = m)} = \mathbb{P}(C = c)\end{aligned}$$

Now if the equation holds for Π then (again assuming that $\mathbb{P}(M = m) > 0$)

$$\begin{aligned}\mathbb{P}(M = m \mid C = c) &= \frac{\mathbb{P}(C = c \mid M = m)\mathbb{P}(M = m)}{\mathbb{P}(C = c)} \\ &= \frac{\mathrm{Enc}_K(m)\mathbb{P}(M = m)}{\sum_{m^*} \mathbb{P}(C = c \mid M = m^*)\mathbb{P}(M = m^*)} \\ &= \frac{\mathbb{P}(M = m)}{\sum_{m^*} \mathbb{P}(M = m^*)} = \mathbb{P}(M = m)\end{aligned}$$

2.2 Prefect adversarial indistinguishability

An encryption scheme is **perfectly indistinguishable** if no adversary $\mathcal{A}$ can succeed with probability better than $\frac{1}{2}$. Formally, we run the following experiment $\mathrm{PrivK}^{eav}_{\mathcal{A},\Pi}$

1. $\mathcal{A}$ outputs a pair $m_0, m_1 \in \mathcal{M}$.
2. $k = \mathrm{Gen}$ and b - chosen from $\{0,1\}$ uniformly - then the **challenge ciphertext** $c = \mathrm{Enc}_k(m_b)$ is given to $\mathcal{A}$.
3. $\mathcal{A}$ tries to determine the which message was encrypted and then outputs b'.
4. $$\mathrm{PrivK}^{eav}_{\mathcal{A},\Pi} \begin{cases} 1 & b' = b \text{ then } \mathcal{A} \text{ succeeds} \\ 0 & b' \neq b \text{ then } \mathcal{A} \text{ fails} \end{cases}$$

Since $\mathcal{A}$ can guess randomly $\mathbb{P}\big(\mathrm{PrivK}^{eav}_{\mathcal{A},\Pi} = 1\big) \geq \frac{1}{2}$ and thus a scheme is perfectly indistinguishable if

$$\mathbb{P}\big(\mathrm{PrivK}^{eav}_{\mathcal{A},\Pi} = 1\big) = \frac{1}{2}, \quad \forall \mathcal{A}$$

Proposition 2.2. Π *is perfectly secret if and only if it is perfectly indistinguishable.*

Proof.

$$\mathbb{P}\big(\mathrm{PrivK}^{eav}_{\mathcal{A},\Pi} = 1\big) = \mathbb{P}(M = m \mid C = c)$$

2.3 One-time pad

Let $l \in \mathbb{N}^*$ and $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0,1\}^l$ then *one-time pad* scheme is describe as follows

- Gen is uniform.
- $\mathrm{Enc}_k(m) = k \oplus m$.
- $\mathrm{Dec}_k(c) = k \oplus c$.

Theorem 2.3. *One-time pad is perfectly secure.*

Proof.

$$\begin{aligned} \mathbb{P}(M = m \mid C = c) &= \frac{\mathbb{P}(C = c \mid M = m)\mathbb{P}(M = m)}{\sum_{m^*} \mathbb{P}(C = c \mid M = m^*)\mathbb{P}(M = m^*)} \\ &= \frac{\mathbb{P}(K = c \oplus m)}{\sum_{m^*} \mathbb{P}(K = c \oplus m^*)\mathbb{P}(M = m^*)}\mathbb{P}(M = m) \\ &= \mathbb{P}(M = m) \end{aligned}$$

Proposition 2.4. *If* Π *is perfectly secure then we must have* $|\mathcal{K}| \geq |\mathcal{M}|$.

Proof. Suppose $|\mathcal{K}| < |\mathcal{M}|$ and let $c \in \mathcal{C}$ be a ciphertext and define $\mathcal{M}(c)$ to the

$$\mathcal{M}(c) = \{m \,|\, m = \mathrm{Dec}_k(c) \text{ for some } k \in \mathcal{K}\}$$

Then $|\mathcal{M}(c)| \leq |\mathcal{K}| < |\mathcal{M}|$ and therefore there exists $m \in \mathcal{M}$ such that $m \notin \mathcal{M}(c)$ hence

$$\mathbb{P}(M = m \mid C = c) = 0 \neq \mathbb{P}(M = m)$$

Note that we assumed the distribution over $\mathcal{M}$ is uniform. ∎

Theorem 2.5 (Shannon's Theorem). Π *with* $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$ *is perfectly secure if and only if*

1. Gen *is uniform.*
2. $\forall m \in \mathcal{M}$ *and* $c \in \mathcal{C}$, $\exists! k \in \mathcal{K}$ *such that* $\mathrm{Enc}_k(m) = c$.

Proof.

Chapter 3

Private-Key

3.1 Asymptotic security

Definition: A scheme is (t, ϵ)-secure if any adversary running for time at most t succeeds in breaking the scheme with probability at ϵ at most.

Consider the following definitions

Definition: An **efficient algorithm** (it might be probabilistic) runs in polynomial time, that is, there is p such that for all $x \in \{0,1\}^*$, the alogrithm $A(x)$ terminates in at most $p(|X|)$ steps.

Definition: $f : \mathbb{N} \to \mathbb{R}^+$ is **negligible** if

$$\forall p \geq 0,\ \exists N \text{ s.t. } n \geq N \implies f(n) < \frac{1}{p(n)}$$

Proposition 3.1. *Suppose f, g are negligible then*

1. *$h = f + g$ is negligible.*
2. *for $p \geq 0$, $h = pf$ is negligible.*

We can now define **asymptotic security** as

Definition: A scheme is secure if for every probabilistic polynomial time adversary $\mathcal{A}$ carrying out at attack from some formally specified type, the probability of $\mathcal{A}$ succeeding is negligible.

Note that being negligible is asymptotic by definition.

3.2 Computational security

Let $\mathcal{M} = \{0,1\}^*$ and $\mathrm{Dec}_k(c)$ returns an error $\perp$ if c is invalid. If $\forall k \leftarrow \mathrm{Gen}(1^n)$, Enc is only defined for $m \in \{0,1\}^{l(n)}$, Π is a fixed-length private key encryption scheme for messages of length $l(n)$. Furthermore, unless specified, we assume that Enc and Dec are *stateless*. That is, each call is independent of previous calls. We revise the definition of $\mathrm{PrivK}^{eav}_{\mathcal{A},\Pi}(n)$ so that $|m_0| = |m_1|$. That is

1. $\mathcal{A}$ outputs a pair $m_0, m_1 \in \mathcal{M}$ such that $|m_0| = |m_1|$.
2. $k =$ Gen and b - chosen from $\{0,1\}$ uniformly - then the **challenge ciphertext** $c = \text{Enc}_k(m_b)$ is given to $\mathcal{A}$.
3. $\mathcal{A}$ tries to determine the which message was encrypted and then outputs b'.
4. $$\text{PrivK}^{eav}_{\mathcal{A},\Pi}(n) \begin{cases} 1 & b' = b \text{ then } \mathcal{A} \text{ succeeds} \\ 0 & b' \neq b \text{ then } \mathcal{A} \text{ fails} \end{cases}$$

Then the definition of indistinguishability becomes

Definition: Π has *indistinguishable encryptions in presence of eavesdropper* or it is *EAV-secure* if for any PPT $\mathcal{A}$ there is a negligible function such that

$$\mathbb{P}\big(\text{PrivK}^{eav}_{\mathcal{A},\Pi}(n) = 1\big) = \frac{1}{2} + negl(n)$$

Proposition 3.2. *If b is fixed in the aforementioned EAV-security is equivalent to*

$$\big|\mathbb{P}\big(\text{PrivK}^{eav}_{\mathcal{A},\Pi}(n, b=0) = 1\big) - \mathbb{P}\big(\text{PrivK}^{eav}_{\mathcal{A},\Pi}(n, b=1) = 1\big)\big| \leq negl(n)$$

Theorem 3.3. *Let Π be an EAV-secure fixed-length encryption scheme. Then for all PPT $\mathcal{A}$ and any bit m^i, $i \in \{1, \ldots, l\}$ there is a negligible function such that*

$$\mathbb{P}\big(\mathcal{A}(1^n, \text{Enc}_k(m)) = m^i\big) \leq \frac{1}{2} + negl(n)$$

That is, $\mathcal{A}$ can not determine any bit any better than guessing it.

Proof. proof by reduction ∎

Theorem 3.4. *Let Π be defined as above. Then for any PPT $\mathcal{A}$, there is a PPT algorithm $\mathcal{A}'$ such that for any $S \subset \{0,1\}^l$ and any function $f : \{0,1\}^l \to \{0,1\}$, there is a negligible function such that*

$$|\mathbb{P}(\mathcal{A}(1^n, \text{Enc}_k(m)) = f(m)) - \mathbb{P}(\mathcal{A}'(1^n) = f(m))| \leq negl(n)$$

That is, no $\mathcal{A}$ can do any better finding a function of the message if they had the ciphertext than when they do not.

Lastly, we must take into account any external information $h(m)$ about the plaintext that may be leaked.

Definition: Π is sementically secure if $\forall \mathcal{A}$, $\exists \mathcal{A}'$ (both are PPT) such that for any PPT algorithm *Samp* and polynomial time computable functions f and h,

$$|\mathbb{P}(\mathcal{A}(1^n, \text{Enc}_k(m), h(m)) = f(m)) - \mathbb{P}(\mathcal{A}'(1^n, |m|, h(m)) = f(m))|$$

is negligible. That is, given the additionally leaked information, no $\mathcal{A}$ can do better finding a function of the message if they have the ciphertext than when when they don not have the ciphertext but know about the length of the message.

Theorem 3.5. Π *is sementically secure if and only if it is EAV-secure.*

3.3 Pseudorandom generator

A **pseudorandom generator** G is an efficient deterministic algorithm for transforming a short uniform string called seed, into a longer uniform-looking, pseudorandom, output string.

Definition: Let a l be polynomial and G be a deterministic polynomial time algorithm such that $\forall n, s \in \{0,1\}^n$, $G(s)$ returns a string of length $l(n)$. Then G is a pseudorandom random generator if

1. $\forall n,\ l(n) \geq n$.

2. For any PPT algorithm D, there is a negligible function such that

$$|\mathbb{P}(D(G(s)) = 1) - \mathbb{P}(D(r) = 1)| \leq negl(n)$$

   where is r is taken uniformly from $\{0,1\}^{l(n)}$

A **stream cipher** is pair of deterministic alogrithm $(Init, GetBits)$ where

Definition:

Init takes as input a seed s and optional initialization vector IV, and outputs an initial state st_0.

GetBits takes as input state st_i and return a bit y and updated state st_{i+1}.

3.4 Proof by reduction

Assume X can not be solved by any polynomial time algorithm with negligible probability. Then to prove Π is secure we must show

1. Fix some efficient adversary $\mathcal{A}$ attacking Π with success probability $\epsilon(n)$.

2. Construct $\mathcal{A}'$ that attempts to solve X using $\mathcal{A}$ as a subroutine. Given an instance x of X, $\mathcal{A}'$ simulates Π for $\mathcal{A}$ such that

   (a) As far as $\mathcal{A}$ can tell, it is interacting with Π.

   (b) If $\mathcal{A}$ breaks Π, this should allow $\mathcal{A}'$ to solve x at least with probability $\frac{1}{p(n)}$ for some polynomial.

3. Taken together, they imply that $\mathcal{A}'$ can solve X with probability $\frac{\epsilon(n)}{p(n)}$. If ϵ is not negligible then $\frac{\epsilon}{p}$ is not negligible neither. Moreover, if $\mathcal{A}$ is efficient we obtain an efficient algorithm $\mathcal{A}'$ solving X with non-negligible probability, contradicting our assumptions.

4. Given our assumption about X there is no efficient adversary $\mathcal{A}$ that can succeed in breaking with non-negligible probability. Meaning that Π is computationally secure.

3.5 Security for multiple encryptions

Define $\text{PrivK}^{mult}_{\mathcal{A},\Pi}(n)$:

1. $\mathcal{A}$ is given input 1^n and output $M_0 = (m_{0,1}, \dots, m_{0,n})$ and $M_1 = (m_{1,1}, \dots, m_{1,n})$ with $|m_{0,i}| = |m_{1,i}|$, $\forall i$.
2. $k \leftarrow \text{Gen}(1^n)$ and $b \in \{0, 1\}$ are chosen uniformly. $C = (c_1, \dots, c_n)$ is constructed with $c_i \leftarrow \text{Enc}_k(m_{b,i})$
3. $\mathcal{A}$ outputs its guess b'.

Then Π is mult-secure if for all PPT $\mathcal{A}$ there is a negligible function such that

$$\mathbb{P}\big(\text{PrivK}^{mult}_{\mathcal{A},\Pi}(n)\big) \leq \frac{1}{2} + negl(n)$$

Proposition 3.6. $\text{PrivK}^{mult}_{\mathcal{A},\Pi}(n) \implies \text{PrivK}^{eav}_{\mathcal{A},\Pi}(n)$*, however, the converse is not true.*

Proposition 3.7. *If* Π *is stateless and* Enc *is deterministic then* Π *can not be mult-secure.*

3.6 CPA and CPA-security

Let $\text{PrivK}^{cpa}_{\mathcal{A},\Pi}(n)$

1. $k \leftarrow \text{Gen}(1^n)$ unknown to $\mathcal{A}$.
2. $\mathcal{A}$ is given 1^n and $\text{Enc}_k(\cdot)$ and output messages m_0, m_1 with $|m_0| = |m_1|$.
3. $b \in \{0, 1\}$ is chosen uniformly and $c \leftarrow \text{Enc}_k(m_b)$ is given to $\mathcal{A}$.
4. $\mathcal{A}$ ouptut b'.

Π is indistinguishable against CPA attacks, CPA-secure, if $\forall \mathcal{A}$ running in PPT there exists a negligible function such that

$$\mathbb{P}\big(\text{PrivK}^{cpa}_{\mathcal{A},\Pi}(n) = 1\big) \leq \frac{1}{2} + negl(n)$$

which then can be extended to multiple messages by $\text{PrivK}^{LR-cpa}_{\mathcal{A},\Pi}(n)$. Instead of outputting lists of messages in this scheme attacker can sequentially query Π.

1. $k \leftarrow \text{Gen}(1^n)$ and $b \in \{0, 1\}$ is uniformly chosen, both unknown to $\mathcal{A}$.
2. $\mathcal{A}$ is given 1^n and $LR_{k,b}(\cdot, \cdot)$.
3. $\mathcal{A}$ ouptut b'.

where $LR_{k,b}(m_0, m_1) = \text{Enc}_k(m_b)$ with $|m_0| = |m_1|$.

Π has indistinguishable multiple encryptions under CPA, CPA-secure for multiplie messages, if for any PPT $\mathcal{A}$ there exists a negligible function such that

$$\mathbb{P}\Big(\text{PrivK}^{LR-cpa}_{\mathcal{A},\Pi}(n) = 1\Big) \leq \frac{1}{2} + negl(n)$$

Theorem 3.8. *CPA-secure is equivalent to CPA-secure for multiple messages.*

Corollary 3.9. *CPA-security for fixed-length messages can be extended to arbitrary length.*

3.7 Pseudorandom function

$F : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$ is a key function $F(k,x)$. We assume that $k \in \{0,1\}^{l_{key}(n)}$, $x \in \{0,1\}^{l_{in}(n)}$, and $F_k(x) \in \{0,1\}^{l_{out}(n)}$ with $l_{key}(n) = l_{in}(n) = l_{out}(n) = n$. F is a psuedorandom if the function for any uniformly chose k, F_k, is indistinguishable from a function chosen uniformly from the set of all functions having the same domain and range. That is, F is a pseudorandom function if for all polynomial-time distinguisher D, there is a negligible function such that

$$\left|\mathbb{P}\big(D^{F_k(\cdot)}(1^n) = 1\big) - \mathbb{P}\big(D^{f(\cdot)}(1^n) = 1\big)\right| \leq negl(n)$$

where the first probability is taken over uniform $k \in \{0,1\}^n$ and D and the second probability is taken over uniform f and D. Note that D only evaluate F_k or f polynomially many times. F is a keyed permutation if $l_{in} = l_{out}$ and $\forall k \in l_{key}(n)$, F_k is a permutation. F is efficient if there is polynomial-time algorithm to compute $F_k(x)$, $\forall k, x$ and $F_k^{-1}(y)$, $\forall k, y$.

Proposition 3.10. *If F is a permutation pseudorandom and $l_{in}(n) \geq n$ then F is also a pseudorandom function.*

Definition: F is strong pseudorandom permutation if for all PPT distinguisher D, there is a negligible function such that

$$\left|\mathbb{P}\Big(D^{F_k(\cdot),F_k^{-1}(\cdot)}(1^n) = 1\Big) - \mathbb{P}\Big(D^{f(\cdot),f^{-1}(\cdot)}(1^n) = 1\Big)\right| \leq negl(n)$$

Definition: Block ciphers are secure instances of strong pseudorandom permutations with some fixed key length and block length.

–Construction of pseudorandom generators and stream ciphers from pseudorandom functions.

Definition: CPA encryption with pseudorandom function can be achieved with

- Gen takes 1^n and outputs $k \in \{0,1\}^n$ uniformly.
- Enc chooses uniform $r \in \{0,1\}^n$ and outputs $c = \langle r, F_k(r) \oplus m\rangle$, where F is a pseudorandom function.
- Dec takes k and $c = \langle r, s\rangle$ and outputs

$$m = F_k(r) \oplus s$$

Theorem 3.11. *If F is a pseudorandom function, then the above Construction is a CPA-secure for messages of length n.*

3.8 Modes of operation

3.8.1 Stream cipher

Synchronized mode

It is good for single communication and messages that are send and recieved in order without being loss. – insert shape Let $G_\infty(s,l)$ where s is the seed and l is the desired length, then

$$c = G_\infty\big(k, 1^{|m|}\big) \oplus m, \quad m = c \oplus G_\infty\big(k, 1^{|m|}\big)$$

Note that the message m can be composed of $m_1||m_2||\ldots$ and then we can calculate c by XORing the corresponding portion of G_∞ with m_i.

Proposition 3.12. *If stream cipher is indistinguishable this is EAV-secure.*

3.8.2 Unsynchronized mode

We let G_∞ to take an initialization vector $IV \in \{0,1\}^n$ as well. Note that IV and k are uniformly chosen.

$$c = \left\langle IV, G_\infty\left(k, IV, 1^{|m|}\right) \oplus m\right\rangle$$

Proposition 3.13. *If there exists a weak pseudorandom function such that*

$$F_K(IV) = G_\infty\left(k, IV, 1^l\right)$$

then the unsynchronized mode is CPA-secure.

3.8.3 Block-cipher

Let F be a block cipher with length n. We assume that $|m|$ is a multiple of n. That is $m = m_1||\ldots||m_l$ with $m_i \in \{0,1\}^n$ - if not we can append 1 followed by sufficiently many zeros.

Electronic Code Block mode (EBC)

$$\begin{aligned} c &= \langle F_K(m_1), \ldots, F_K(m_l)\rangle \\ m &= \left\langle F_K^{-1}(c_1), \ldots, F_K^{-1}(c_l)\right\rangle \end{aligned}$$

it is not even EAV-secure, because of repeated portions.

Cipher Block Chaining (CBC)

$$\begin{aligned} c &= \langle c_0 = IV \in \{0,1\}^n, F_K(m_1 \oplus c_0), \ldots, F_K(m_l \oplus c_{l-1})\rangle \\ m &= \left\langle F_K^{-1}(c_1) \oplus c_0, \ldots, F_K^{-1}(c_l) \oplus c_{l-1}\right\rangle \end{aligned}$$

Proposition 3.14. *If F is a pseudorandom permutation then this is CPA secure.*

Then chained (stateful) CBC variant uses that last c_l for encryption of new message m' instead of a fresh new IV. However, chained CBC is vulenrable to CPA. Consider the following scheme

1. Attacker knows $m_1 \in \{m_1^0, m_1^1\}$.
2. Attacker observes $\langle IV, c_1, c_2, c_3\rangle$.
3. Attacker chooses $m_4 = IV \oplus m_1^0 \oplus c_3$ and observes c_4.
4. $m_1 = m_0^1 \iff c_1 = c_4$.

Output feedback mode (OFB)

$$c = \left\langle IV, m_1 \oplus F_K(IV), m_2 \oplus F_K(F_K(IV)), \ldots, m_l \oplus F_K^{(l)}(IV) \right\rangle$$

Proposition 3.15. *OFB is CPA-secure if F is a pseudorandom function. To improve the efficiency we can calculate $F_K^{(n)}(IV)$ ahead of time. Stateful variant is secure as well.*

Counter mode (CTR)

Let $ctr \in \{0,1\}^n$ be a uniformly chosen number then

$$c = \langle ctr, m_1 \oplus F_K(ctr+1), \ldots, m_l \oplus F_K(ctr+l) \rangle$$

Proposition 3.16. *CTR is CPA-secure if F is a pseudorandom function.*

Stateful version is secure as well. It can be parallelized.