Foundations of Cryptography: basic notions,security definitions and proofs **25 Fall**

Lecture 1 — September 10

Lecturer: Tianren Liu *Scribe: Jiaqing Zhu*

1.1 history of cryptography

- (Caesar cipher) shift each letter by 3 positions,e.g. $A \to D, B \to E$,corresponding mod 26.
- (Shify cipher) shift each letter by k positions,where k is the key.
- (monoalphabetic substitution cipher) permute the alphabet,where the permutation is the key.
- (Vigenere cipher) choose t permutation $\Pi_1, \dots, \Pi_t$,then encrypt the i-th letter by $\Pi_{i \bmod t}$.
- (Enigma machine) polyalphabetic substitution cipher,transmit each word,the gear rotates by a unit and cause the permutation to change,after the back gear turn by a round,causing the forehead gear to rotate by a unit,with a very long period 26^3.

Attack monoalphabetic substitution cipher by frequency analysis. Guess the t and then analysis each sub-cipher with frequency to attack Vigenere cipher.

1.2 perfect secrecy and perfect indistinguishability

Definition 1.1. *(encryption scheme) An encryption scheme is a tuple of algorithms* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *and message space $\mathscr{M}$,key space $\mathscr{K}$,ciphertext space $\mathscr{C}$.*

Gen is a probabilistic key generation algorithm output a key according to some distribution over $\mathscr{K}$. Enc is An encryption algorithm,which takes a key $k \in \mathscr{K}$ and a message $m \in \mathscr{M}$ as input,and outputs a ciphertext $c \in \mathscr{C}$. Dec is a deterministic decryption algorithm,which takes a key $k \in \mathscr{K}$ and a ciphertext $c \in \mathscr{C}$ as input,and outputs a message $m \in \mathscr{M}$.

(Perfect Correctness) $\forall m \in \mathscr{M}, k \in \mathscr{K}$, $\mathsf{Dec}(k, \mathsf{Enc}(k, m)) = m$. Afterwards,$m, c, k$ will be corresponding elements of $\mathscr{M}, \mathscr{C}, \mathscr{K}$ respectively.

(perfect secrecy) An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ over message space $\mathscr{M}$ has perfect secrecy if for every probability distribution U over $\mathscr{M}$,every $\alpha \in \mathscr{M}$ and every $\beta \in \mathscr{C}$,we have $\Pr[m = \alpha | c = \beta] = \Pr[m = \alpha]$,where M is the random variable denoting the message and C is the random variable denoting the ciphertext when the key is chosen according to Gen and the message is chosen according to the given distribution. Namely if there is an Eve "glimpse" the process and get the cipher of communication between Alice and Bob,Eve knows the cipher but cannot get "anything" about what the message is as he doesnot see the cipher. $\forall f, \forall$ distribution over $\mathscr{M}$,$\exists Eav$ that $\Pr[Eav(c) = f(m)] = \Pr[Eav' = f(m)]$,while the latter get no information about c.

Definition 1.2. *(Indistinguishability Game)Adversary $\mathscr{A}$ choose two $m_0, m_1 \in \mathscr{M}$,a k is generated by* Gen *and uniformly choose $b \in \{0,1\}$,then compute $c = \mathsf{Enc}(k, m_b)$ and give c to the adversary $\mathscr{A}$,who then outputs a bit b'.The adversary wins if $b' = b$.*

Definition 1.3. *(Perfect Indistinguishability) An encryption scheme* (Gen, Enc, Dec) *over message space* $\mathscr{M}$ *has perfect indistinguishability if for every adversary* $\mathscr{A}$*,we have* $\Pr[\mathscr{A}\ wins] = \frac{1}{2}$.

The equivalence of perfect secrecy and perfect indistinguishability:

$$\Pr[c=\beta|m=\alpha]=\frac{\Pr[c=\beta]\Pr[m=\alpha|c=\beta]}{\Pr[m=\alpha]}$$

$\Pr[c=\beta|m=\alpha]=\Pr[c=\beta] \Longleftrightarrow \Pr[m=\alpha]=\Pr[m=\alpha|c=\beta]$ It means that $\mathscr{A}$ look at c,since $\Pr[c|m=m_0]=\Pr[c|m=m_1]$,he cannot tell which b is.

From the perspective of indistinguishability game,if change $1/2$ to $1/2+\varepsilon$,it will be the notion of statistic security, if both change $1/2$ to $1/2+\varepsilon$ and change $\mathscr{A}$ to computability bounded power machine,it will be the notion of computability security.

1.3 One time pad

$|\mathscr{M}|=|\mathscr{K}|=|\mathscr{C}|$, Gen uniformly choose k,$\mathsf{Enc}(m,k)=m\oplus k$, $\mathsf{Dec}(c,k)=c\oplus k$.

$$\begin{aligned}\Pr[c=\beta|m=\alpha]&=\Pr[k=m\oplus c|m=\alpha]\\&=\frac{1}{|\mathscr{K}|}=\frac{1}{|\mathscr{M}|}\\\Pr[c=\beta]&=\sum_{\alpha}\Pr[c=\beta|m=\alpha]\Pr[m=\alpha]\\&=\sum_{\alpha}\frac{1}{|\mathscr{M}|}\Pr[m=\alpha]\\&=\frac{1}{|\mathscr{M}|}\end{aligned}$$

And then with Bayes Theorem $\Pr[m=\alpha|c=\beta]=\frac{\Pr[c=\beta|m=\alpha]\Pr[m=\alpha]}{\Pr[c=\beta]}$,one time pad has perfect secrecy.

But this scheme has problems,since it asks the size $|\mathscr{K}|=|\mathscr{M}|$ and it can be used only one time,which is impractical.In fact, for each perfect secrecy scheme,we have $|\mathscr{K}|\geq|\mathscr{M}|$.Thus,we may want to consider a weaker security notion that enables $|\mathscr{K}|\leq|\mathscr{M}|$.

1.4 Semantic Security

We consider a scheme with a parameter λ,it will influence the construction of scheme with

- length of key
- adversary and encryptor's computational power
- adversary's advantage(namely the ε term in probability to win the indistinguishability Game $1/2+\varepsilon$)

We consider the compute power of adversary and encryptor is polynomial in λ.

Under the assumption that $\mathbf{P}\neq\mathbf{NP}$,otherwise $\exists k, \mathsf{Enc}(k,m_0)=c$ is a **NP** problem,can solve indistinguishability game in polynomial time.

Definition 1.4. *(negligible function) A function* $\mathrm{negl}:\mathbb{N}\to\mathbb{R}^+$ *is negligible if for every polynomial* $\mathrm{poly}(\cdot)$*,there exists a* n_0 *such that for all* $n>n_0$*,we have* $\mathrm{negl}(n)<\frac{1}{\mathrm{poly}(n)}$.

Make some modification to previous security definitions:

1.(Eav Security)$\forall$PPT(probabilistic polynomial Turing machine)adversary $\mathscr{A}$,$\forall f$, $\forall$distribution U over $\mathscr{M}$,there exists a ppt simulator $\mathscr{A}'$ such that $|\Pr[\mathscr{A}(1^\lambda, c) = f(m)] - \Pr[\mathscr{A}'(1^\lambda) = f(m)]| = \text{negl}(\lambda)$,where m is chosen according to the given distribution and c is the cipher of m.

2.(Semantic Security)We say g is the previous knowledge of $\mathscr{A}$ about m,and $f(m)$ is some property of m that $\mathscr{A}$ want to compute, $\forall$PPTadversary $\mathscr{A}$,$\forall f$, $\forall$distribution U over $\mathscr{M}$,there exists a ppt simulator $\mathscr{A}'$ such that $|\Pr[\mathscr{A}(1^\lambda, c, g(m)) = f(m)] - \Pr[\mathscr{A}'(1^\lambda, g(m)) = f(m)]| = \text{negl}(\lambda)$,where m is chosen according to the given distribution and c is the cipher of m.

3.(modification to Indistinguishability Game) $\forall$PPTadversary $\mathscr{A}$,we have $\Pr[\mathscr{A}$ wins$] \leq \frac{1}{2}$+negl(λ).

We prove that the three are equivalent.

$2 \Rightarrow 1$:take g nothing.

$1 \Rightarrow 3$:take $f(m_0) = 0, f(m_1) = 1$ and the distribution is restrict to the two points with equal probability.

$3 \Rightarrow 2$:Proof by contradiction. Assume $\exists \mathscr{A}, g, f$ and distribution U over $\mathscr{M}$ such that $\forall \mathscr{A}'$, $|\Pr[\mathscr{A}(1^\lambda, c, g(m)) = f(m)] - \Pr[\mathscr{A}'(1^\lambda, g(m)) = f(m)]| > \text{negl}(\lambda)$. Then we can construct an adversary(here call it distinguisher) D to break indistinguishability game:

Let $\mathscr{A}$ to be the adversary that knowledge about m is $g(m)$,get a $c = \mathsf{Enc}(k, m)$ and output $f(m)$,take use of $\mathscr{A}$,we construct $\mathscr{A}'$: knowledge about m is $g(m)$,simulate $\mathscr{A}$ with $c' = \mathsf{Enc}(k, m_0)$. The distinguisher D works as follows:

(1) Choose m_0, m_1 and send to the challenger.

(2) The challenger choose $b \in \{0, 1\}$ uniformly at random and send $c = \mathsf{Enc}(k, m_b)$ to D.

(3) D simulate $\mathscr{A}$ with $c, g(m_1)$ and get $f(m_b)$,if $f(m_b) = f(m_0)$,output 0,else output 1.

analysis:
When challenger choose $b = 1$,D simulate $\mathscr{A}$ with knowledge $g(m_1)$ and $c = \mathsf{Enc}(k, m_1)$,so it works the same as $\mathscr{A}$, if $b = 0$,D simulate $\mathscr{A}$ with knowledge $g(m_1)$ and $c = \mathsf{Enc}(k, m_0)$ and work the same as $\mathscr{A}'$, since $|\Pr[\mathscr{A}(1^\lambda, c, g(m)) = f(m)] - \Pr[\mathscr{A}'(1^\lambda, g(m)) = f(m)]| > \text{negl}(\lambda)$,the distinguisher D can distinguish the two cases with probability $\frac{1}{2} + \text{negl}(\lambda)$,break the indistinguishability game.

1.5 begin of PRG

Definition 1.5. *PRG(Pseudorandom Generator) A deterministic polynomial time algorithm* $G : \{0, 1\}^s \to \{0, 1\}^{l(s)}$ *is a pseudorandom generator with expansion factor* $l(s)$ *if* $l(s) > s$ *for all* $s \in \mathbb{N}$ *and for every* PPT *distinguisher* D*,we have* $|\Pr_{r \sim \{0,1\}^{l(s)}}[D(r) = 1] - \Pr_{k \sim \{0,1\}^s}[D(G(k)) = 1]| = \text{negl}(s)$*,where the first probability is over* r *chosen uniformly from* $\{0, 1\}^{l(s)}$ *and the second is over* k *chosen uniformly from* $\{0, 1\}^s$.

With PRG,we may construct a computability secure encryption scheme with key length s and message length $l(s)$ and achieve $|\mathscr{K}| < |\mathscr{M}|$