CS3235 Part1 Lec2

$_/|/|U_Ch@NgRu$!

December 16, 2021

1 Limitation on one-time pad

1. The key is as long as the message
2. It achieves secrecy only if each key is used to encrypt the message once
   - Two/multiple-time pad is vulnerable to known-plaintext attack
3. Malleability of OTP

1.1 The insecurity of two-time pad

1. k be a random key in $\{0, 1\}^n$
2. Let m_1, m_2 be two message in $\{1, 0\}^n$
3. (a) $c_1 = m_1 \oplus k$

   (b) $c_2 = m_2 \oplus k$
4. $c_1 \oplus c_2 = (m_1 \oplus k) \oplus (m_2 \oplus k) = (m_1 \oplus m_2) \oplus (koplusk) = m_1 \oplus m_2$
5. If both m_1, m_2 are in English(Frequency analysis) or characters of ASCII, with $m_1 \oplus m_2$, many information can be leaked(Trivially broken by a known-plaintext attack)

1.1.1 ”space” method

The general idea is that

- $c_i \oplus c_j = (m_i \oplus k) \oplus (m_j \oplus k) = m_i \oplus j$
- the space in ASCII is 32_{10} or 0x20, when a space in m_i XOR with a char say x in m_j
  - if x is a small letter, it will be transfer into a capital letter(e.g. $a \oplus 0x20 = A$)
  - if if x is a capital letter, it will be transfer into a small letter (e.g. $A \oplus 0x20 = a$)

Under this property, we can use 0x20 to xor with all 11 messages to find out possible space place, then use that space place to find corresponding position plaintext of other ciphertext with $m_i \oplus m_j$

reference :

1. https://github.com/CameronLonsdale/MTP
2. https://www.fatalerrors.org/a/cryptography-experiment-1-many-time-pad-attacks.html

1.2 Malleability of OTP

Malleability is regarding the integrity of messages. When sending a ciphertext $c = k \oplus m$, can the (man in the middle) attacker make arbitrary changes to the content of plaintext message?

- The attacker can XOR the ciphertext with a pattern x
  1. $c \oplus x = (k \oplus m) \oplus x = k \oplus (m \oplus x)$
  2. The decrypted messageis $m \oplus x$,and not m
  3. This remains undetected
  4. The attacker can control the change (it' s predictable): for example, change the bits that are corresponding to the name of the recipient of a $1M transaction

2 Shannon's bad theory

Theorem: If(Gen, Enc, Dec) with message space $\mathcal{M}$ is perfectly secret, then $|\mathcal{K}| \geq |\mathcal{M}|$

2.1 Proof

(proof by contradiction)

1. Given $\varepsilon =$ (E, D) is a cipher defined over$(\mathcal{K}, \mathcal{M}, \mathcal{C})$

   1.1 statement P_1: ε= (E, D) is a perfectly secure cipher

   1.2 statement P_2: $|\mathcal{K}| < |\mathcal{M}|$

2. Define statements P_3 :

   $\exists$ cipher $\varepsilon =$ (E, D) defined over$(\mathcal{K}, \mathcal{M}, \mathcal{C})$ such that $P_1 \land P_2$

3. Assume P_3 is true and we call this cipher $\varepsilon =$ (E, D), by the definition of perfect security,

   3.1 Here we assume that $\forall m_i, m_j \in \mathcal{M} : Pr[m = m_i] = Pr[m = m_j]$ without lose of generality(perfect secrecy should work regardless of distribution of plaintext)

3.2 for message $m_0 \neq m_1 \in \mathcal{M}$, given that Pr[$m_0$]=Pr[$m_1$], then ∀ ciphertext $c_0 \in \mathcal{C}$, $Pr[m = m_0|c = c_0] = Pr[m = m_1|c = c_0]$ (By definition of perfect secrecy)

3.3 So given c_0, by the definition of encryption scheme $\exists m \in \mathcal{M}, k \in \mathcal{K}$ such that $c_0 = E(k, m)$

3.4 define set X:

$$X = \{D(k, c_0)|k \in K\}$$

3.4 by the correctness of cipher, i.e. D(k,E(k,x))=x

$D(k, c_0)$ is a function for k, or $f : \mathcal{K} \to \mathcal{M}$

3.6 Let Y be the range of $D(k, c_0)$, then $|Y| \leq |\mathcal{K}| < |\mathcal{M}|$ (from 3)

3.7 Hence $\exists m \in \mathcal{M} : m \in \mathcal{M} \wedge m \notin Y$, or $\mathcal{M} \backslash Y \neq \emptyset$

3.8 Let $m_1 \in Y \subset \mathcal{M}, m_2 \in \mathcal{M} \backslash Y$, then

3.8.1 $Pr[m = m_1|c = c_0] = Pr[D(k, c_0) = m_1|c = c_0] \neq 0$, because $m_1 \in Y$

3.8.2 $Pr[m = m_2|c = c_0] = Pr[D(k, c_0) = m_2|c = c_0] = 0$, because $m_2 \notin Y$

3.8.3 Then Pr[$m = m_1|c = c_0$]≠ Pr[$m = m_2|c = c_0$] (conflict with 3.1)

4. Therefore P_3 is false, i.e. $\neg P_3$ is true

5. Therefore ∀ cipher ε = (E, D) defined over($\mathcal{K}, \mathcal{M}, \mathcal{C}$),

   ε= (E, D) is a perfectly secure cipher → $|\mathcal{K}| \geq |\mathcal{M}|$

3 Computational secrecy

1. Allowing security fail with chance bounded by negligible function
2. Restricting attention to "efficient" attackers

3.1 Tiny probability of failure

In practice, security failure sith probability of $\leq 2^{-60}$ is considered as tiny probability

4 Stream Cipher

To reduce the size of key, we apply a pseudo-random number generator(PRG(k)) to map a short key k to a much larger bit string

4.1 Unpredictability of PRG(k)

Unpredictability is considered as the highest achievement regarding adversaries' advantage

5 Semantic Security

Would be ok if an encryption scheme leaked information with tiny probability to eavesdroppers with bounded computational resources

We can relax perfect secrecy by

- Allowing security to "fail" with tiny probability
- Restricting attention to "efficient" (computationally bounded) attackers

5.1 Perfect indistinguishability

A different (yet equivalent) definition of perfect secrecy is given, which allows incorporating the probability of failure for computationally bounded attackers later
This definition is also called perfect indistinguishability, which is an alternated definition of perfect secrecy

5.2 Definition of Perfectly indistinguishable

- $\Pi = (Gen, Enc, Dec)$, is defined over ($\mathcal{M}$, $\mathcal{K}$, $\mathcal{C}$)
- We consider a game between a challenger (that runs the encryption scheme) and the adversary

$$PrivK^{\forall context}_{\mathcal{A},\Pi}(n)$$

1. Challenger generate pre-determined key k and a bit b
2. Adversary selects two message m_0, m_1
3. Challenger encrypt message, c=$Enc_{b,k}(m_b)$
4. Adversary decides b' under any possible context(e.g. Attack may have multiple access to a oracle)
5. Adversary success if b=b', i.e. $PrivK^{\forall context}_{\mathcal{A},\Pi}(n) = 1$
6. A encryption scheme Π is said to be Perfectly indistinguishable(semantically secure)if
   - $\forall \mathcal{A}, Pr[PrivK^{\forall context}_{\mathcal{A},\Pi}(n) = 1] \leq \frac{1}{2}$

5.3 Adversary's Semantic Security Advantage

We compute advantage of $\mathcal{A}$ as $SS^*_{adv}[\mathcal{A}, \Pi] = |Pr[win] - \frac{1}{2}|$

5.4 If E is semantically secure, it is perfectly secret

Proof. RTBD

6 Polynomial

A function F: $\mathbb{Z}^+ \to \mathbb{Z}^+$ is polynomial, if $\exists c \in \mathbb{Z}^+$ such that $F(n) < n^c$

7 Negligible

A function F: $\mathbb{Z}^+ \to [0, 1]$ is negligible if $\exists N$, for every polynomial p, it holds that$\forall n \geq N, f(n) < \frac{1}{p(n)}$ for large enough n

- i.e. It dacays faster than any inverse polynomial
- Typical example: f(n)=log(n)*2^{-cn},2^{-n}, $2^{-\sqrt{n}}, n^{-log\ n}$

7.1 Theorem

A function f:$\mathbb{Z}_{\geq 1} \to \mathbb{R}$ is negligible if and only if for all $c > 0$, we have

$$lim_{n \to \infty} f(n) n^c = 0$$

8 Adversary in Computational secrecy

1. Assumption: the adversary is "efficient" but has limited computational power
2. The adversary is arbitrary algorithm in PPT
    - The can execute in polynomial # of steps
    - That has randomized, non-determ.execution
3. Such an adversary is quite powerful
    - Covers all deterministic, efficient algos
    - But less powerful than "perfect secrecy"

9 Poly-bounded function

A function f: $\mathbb{Z}_{\geq 1} \rightarrow \mathbb{R}$ is called poly-bounded if there exists c,d$\in \mathbb{R}_{>0}$ such that for all integers $n \geq 0$, we have

$$|f(n)| \leq n^c + d$$

10 Efficient algorithm

Let A be an algorithm, which takes $\lambda \in \mathbb{Z}_{\geq 1}$ as input, and another input x$\in \{0,1\}^{p(\lambda)}$, where p is a fixed polynomial
We call A an efficient algorithm if $\exists$ a poly-bounded function say t, and a negligible function ϵ, such that $\forall \lambda \in \mathbb{Z}_{\geq 1}$,and $\forall x \in \{0,1\}^{\leq p(\lambda)}$, the probability that the running time of A on input (λ, x) exceeds $t(\lambda)$ is at most $\epsilon(\lambda)$

- The security parameter is λ
- The attacker’ s computation power is bounded by a polynomial with order λ
- The advantage of the efficient attacker is negligible in λ

11 n in secrecy function

- n is a variable chosen by honest parties when they generate key
- n may be known by attacker by Kerckhoffs’s principle
- Function of n is applied for measuring running times of all parties, and the success probability of the adversary success
- Security may fail with probability negligible in n(tiny probability)
- Restrict attention to attackers running in time (at most) polynomial in n