Principle of Morden Cryptography

- Principle 1 – *Formal Definitions*
  - Precise, mathematical model and definition of what security means
- Principle 2 – *Precise Assumptions*
  - Clearly stated and unambiguous
- Principle 3 – *Proofs of Security*
  - Move away from "design-break-tweak"

Importance of definitions

Design

- Definitions are essential for the design, analysis, and usage of crypto
- Developing a precise definition forces the designer to think about what they really want
  - What is essential and (sometimes more important) & what is not
  - Often reveals subtleties of the problem

If you don't understand what you want to achieve, how can you possibly know when (or if) you have achieved it?

Analysis

- Definitions enable meaningful analysis, evaluation, and comparison of schemes
  - Does a scheme satisfy the definition?
  - What definition does it satisfy?

One scheme may be less efficient than another, yet satisfy a stronger security definition.

Usage

- Definitions allow to understand the security guarantees provided by a scheme
- Enable schemes to be used as *components* of a larger system (modularity)
- Enables one scheme to be *substituted* for another if they satisfy the same definition

Assumptions

- With few exceptions, cryptography currently requires *computational assumptions*
  - At least until we prove $P \neq NP$ (even that would not be enough)
- Principle: any such assumptions should be made *explicit*

所有的假设必须显式给出.

Importance of clear assumptions

- Allow researchers to (attempt to) *validate* assumptions by studying them
- Allow meaningful *comparison* between schemes based on different assumptions
  - Useful to understand minimal assumptions needed
- Practical implications if assumptions are wrong
- Enable proofs of security

有最小假设的一般更安全

Proofs of security

- Provide a rigorous *proof* that a construction satisfies a given *definition* under certian specified *assumptions*
- Proofs are crucial in crypto, where there is a malicious attacker trying to "break" the scheme

Limitations

- Crypto remains partly an *art* as well
- Given a proof of security based on certain assumptions, we still need to instantiate the assumption.
  - Validity of various assumptions
- Provably secure schemes can be broken!
  - If the definition does not correspond to the real-world threat model
  - If the assumption is invalid
  - If the implementation is flawed

现实不满足定义
假设不一定对
实现上有漏洞

Defining secure encryption

- Crypto definitions (in general)

  Security guarantee/goal
  - What we want to achieve and/or what we want to prevent the attacker from achieving

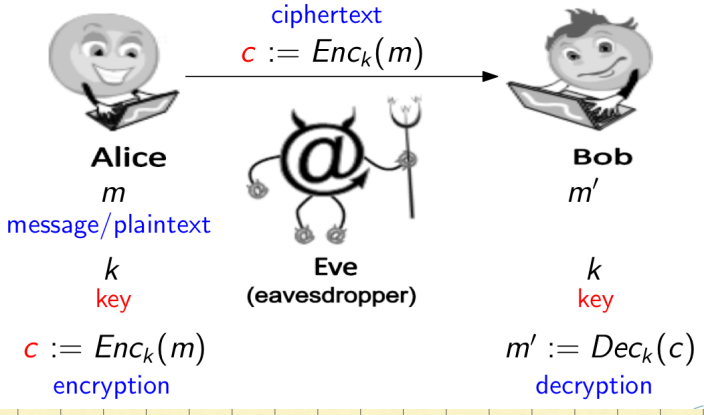
  Threat model
  - What (real-world) capabilities the attacker is assumed to have

安全一般讲的是相对安全

安全目标

威胁模型

Private-key encryption



信道是不安全的，窃听者可以看到密文c.
(此时仍是安全的，因为窃听者不知道m，这是被动安全)

同时窃听者也可以篡改密文c
(若Bob能发现被篡改，是主动安全)

- A *private-key encryption* scheme is defined by a message space $\mathcal{M}$ and algorithms (Gen, Enc, Dec):
  - Gen (*key-generation algorithm*): generates k
  - Enc (*encryption algorithm*): takes key k and message $m \in \mathcal{M}$ as input; outputs ciphertext c: $c \leftarrow Enc_k(m)$
  - Dec (*decryption algorithm*): takes key k and ciphertext c as input; outputs m': $m' := Dec_k(c)$

算法都是公开的，安全取决于密钥k.
一般假设密钥长度小于明文

Theat model for encryption

- *Ciphertext-only attack*

  Known-plaintext attack

  Chosen-plaintext attack → 被动安全

  Chosen-ciphertext attack → 主动安全

stronger

攻击者只知道密文信息

密码设计要求至少是CCA安全的

Probability distribution

- Notation
  - $\mathcal{K}$ (*key space*): set of all possible keys
  - $\mathcal{M}$ (*plaintext space*): set of all possible plaintexts
  - $\mathcal{C}$ (*ciphertext space*): set of all possible ciphertexts

- M: the r.v. denoting the value of the message
  - M ranges over $\mathcal{M}$
  - This reflects the likelihood of different messages being sent by the parties, given the attacker's prior knowledge
  - For example,

    $\Pr[M = \text{“attack today”}] = 0.7$

    $\Pr[M = \text{“don't attack”}] = 0.3$

一般代表了攻击者的先验知识.

- K: the r.v. denoting the key
  - K ranges over $\mathcal{K}$

密钥的取值一般是均匀分布
由Gen决定.

- Fix some encryption scheme (Gen, Enc, Dec)
  - Gen defines a probability distribution for K:

    $\Pr[K = k] = \Pr[\textit{Gen} \text{ outputs key } k]$

- Random variables M and K are independent
  - i.e., the message that a party send does not depend on the key used to encrypt that message

随机变量M与K是独立的

- Fix some encryption scheme (Gen, Enc, Dec), and some distribution for M. Consider the following (randomized) experiment:
  1. Choose a message m, according to the given distribution
  2. Generate a key k using *Gen*
  3. Compute $c \leftarrow \textit{Enc}_k(m)$
- This defines a distribution on the ciphertext. Let C be an r.v. denoting the value of the ciphertext in this experiment

Secure encryption

- "Impossible for the attacker to learn the key"
  - The key is a means to an end, not the end itself
  - Necessary (to some extent) but not sufficient
  - Easy to design an encryption scheme that hides the key completely, but is insecure

Definition 1.1 *Security of encryption* (Ver. 1). An encryption scheme (Gen, Enc, Dec) is *n-secure* if no matter what method Eve employs, the probability that she can recover the true key k from the ciphertext c is at most 2^{-n}.

Definition 1.1 is too weak!

Consider: the secret key k is chosen at random in $\{0,1\}^n$ but our encryption scheme is simply $Enc_k(x) = x$ and $Dec_k(y) = y$.

不限定攻击者的方法，猜中密钥k只能完全随机

当随机出来的k对原文加密后仍是原文，虽满足上述定义，但没意义。

- **Lemma 1.2** Let (Gen, Enc, Dec) be the encryption scheme above. For every function $Eve : \{0,1\}^\ell \to \{0,1\}^n$ and for every $x \in \{0,1\}^\ell$, the probability that $Eve(Enc_k(x))$ is exactly 2^{-n}.

  Proof. This follows because $E_k(x) = x$ and hence $Eve(Enc_k(x)) = Eve(x)$ which is some fixed value $k' \in \{0,1\}^n$ independent of k. Hence the probability that $k = k'$ is 2^{-n}.

  Problem for Ver. 1: Could be hard to learn key, but easy to learn message.

从密文中恢复明文

- **Definition 1.3** *Security of encryption* (Ver. 2). An encryption scheme (Gen, Enc, Dec) is *n-secure* if for every message m no matter what method Eve employs, the probability that she can recover x from the ciphertext c is at most 2^{-n}.

  Problem for Ver. 2: Too strong, for "every message", it is impossible to achieve!

  Example: $Eve(Enc_k(x)) = 0^\ell$ for all x
  $x = 0^\ell$

不能针对特定的m去定义

- **Definition 1.4** *Security of encryption* (Ver. 3). An encryption scheme (Gen, Enc, Dec) is *n-secure* if no matter what method Eve employs, if x is chosen at random from $\{0,1\}^\ell$, the probability that she can recover x from the ciphertext c is at most 2^{-n}.

  Problem for Ver. 3: Still weak!

  Consider an encryption that hides the last $\ell/2$ bits of the message, but completely reveals the first $\ell/2$ bits.The probability of guessing a random message is $2^{-\ell/2}$, and so it would be $\ell/2$-secure.

若加密时保持明文前半部分不变，后半部分加密，那Eve只用猜x的后半部分，此时是$\ell/2$-安全的.

Perfect secrecy

- *Perfect secrecy* (informal)
  - "Regardless of any *prior* information the attacker has about the plaintext, the ciphertext should leak *no additional* information about the plaintext"
  - Attacker's information about the plaintext = attacker-known *distribution* of M
  - Perfect secrecy means that observing the ciphertext should not change the attacker's knowledge about the distribution of M

不管攻击者知道多少关于明文的信息，密文h必须没有泄露明文信息.

- **Definition 1.5** An encryption scheme (Gen, Enc, Dec) with message space $\mathcal{M}$ and ciphertext space $\mathcal{C}$ is *perfectly secure* if for every distribution over $\mathcal{M}$, every $m \in \mathcal{M}$, and every $c \in \mathcal{C}$ with $\Pr[C = c] > 0$, it holds that
$$\Pr[M = m | C = c] = \Pr[M = m]$$
Equivalently, for every set $M \subseteq \{0,1\}^\ell$ of plaintexts, and for every strategy used by Eve, if we choose at random $x \in M$ and a random key $k \in \{0,1\}^n$, then the probability that Eve guesses x after seeing $Enc_k(x)$ is at most $1/|M|$, i.e.,
$$\Pr[Eve(Enc_k(x)) = x] \le 1/|M|$$

Key point: The ciphertext c reveals zero additional information about the plaintext m.
$\Pr[M = m | C = c] = Pr[M = m]$

即M与C是互相独立的，密文未泄露明文的信息

→最大也只是$|M|$，即从密文空间中猜一个.

↓

即Eve从密文m中猜出明文x

- Another two equivalent definitions

Definition 1.6 *Perfect secrecy.* An encryption scheme (Gen, Enc, Dec) with message space $\mathcal{M}$ and ciphertext space $\mathcal{C}$ is *perfectly secure* if and only if for every two distinct plaintexts $\{x_0, x_1\} \in \mathcal{M}$, and for every strategy used by Eve, if we choose at random $b \in \{0,1\}$ and a random key $k \in \{0,1\}^n$, then the probability that Eve guesses x_b after seeing the ciphertext $c = Enc_k(x_b)$ is at most $1/2$.

two-to-many theorem能证明1.5与1.6等价

$\Rightarrow k \leftarrow_R \{0,1\}^n$, $b \leftarrow_R \{0,1\}$, $\Pr[Eve(Enc_k(x_b)) = x_b] \le \frac{1}{2}$

Definition 1.7 *Perfect secrecy.* Two probability distributions X, Y over $\{0,1\}^\ell$ are *identical*, denoted by $X \equiv Y$, if for every $y \in \{0,1\}^\ell$, $\Pr[X = y] = \Pr[Y = y]$. An encryption scheme (Gen, Enc, Dec) is *perfectly secure* if for every pair of plaintexts x, x', we have $Enc_{U_n}(x) \equiv Enc_{U_n}(x')$.

$\Rightarrow \forall m, m' \in M, \forall c \in C, \Pr[Enc_k(m) = c] = \Pr[Enc_k(m') = c]$

即密文的概率分布也是一致的，与密明无关.

$\mathcal{Q}$: Does this mean that for every k, $Enc_k(x) = Enc_k(x')$?

但并不表示对任两个明文加密后的结果一样.

- **Theorem 1.8** (Two-to-Many Theorem) The scheme (Gen, Enc, Dec) is *perfectly secure* if and only if
$$Pr[Eve(Enc_k(x_0)) = x_0] \le 1/2.$$

Proof.

The "only if" part is easy (by definition, this is the special case that $|M| = 2$).

only if：直接取$|M|=2$特例

The "if" part is tricky.

We need to show that if there is some set M and some strategy for Eve to guess a plaintext chosen from M with probability larger than $1/|M|$, then there is also some set M' of size 2 and a strategy Eve' for Eve to guess a plaintext chosen from M' with probability larger than $1/2$.

We fix $x_0 = 0^\ell$ and pick x_1 at random in M. Then it holds that for random key k and message $x_1 \in M$,
$\Pr_{k \leftarrow \{0,1\}^n, x_1 \leftarrow M}[Eve(Enc_k(x_1)) = x_1] > 1/|M|$.
On the other hand, for every choice of k, $x' = Eve(Enc_k(x_0))$ is a fixed string independent on the choice of x_1, and so if we pick x_1 at random in M, then the probability that $x_1 = x'$ is at most $1/|M|$, i.e.,
$\Pr_{k \leftarrow \{0,1\}^n, x_1 \leftarrow M}[Eve(Enc_k(x_0)) = x_1] \le 1/|M|$.
Due to the linearity of expection,there exists some x_1 satisfying
$\Pr[Eve(Enc_k(x_1)) = x_1] > \Pr[Eve(Enc_k(x_0)) = x_1]$. (why?)
We now define a new attacker Eve' as: $Eve'(c) = \begin{cases} x_1, & \text{if } Eve(c) = x_1 \\ x_i, i \in \{0,1\} \text{ at random}, & \text{otherwise} \end{cases}$
This means the probability that $Eve'(Enc_k(x_b)) = x_b$ is larger than $1/2$ (Why?).

反证法命题开始证：

(Gen, Enc, Dec)不是完全安全 $\Rightarrow \Pr[Eve(Enc_k(x_b)) = x_b] > \frac{1}{2}$

即$\exists M$, $\Pr[Eve | Enc_k(x) = x] > \frac{1}{|M|}$ ① $\Rightarrow \Pr[Eve(Enc_k(x_b)) = x_b] > \frac{1}{2}$

取M'，有$|M'| = 2$，此时固定$x_0 = 0^\ell$，$x_1 \leftarrow_R M'$，则对随机$k \in \{0,1\}^n$，有

$\Pr[Eve(Enc_k(x_1)) = x_1] > \frac{1}{|M'|}$ (由于①)

同时$\forall k \in \{0,1\}^n$，$x' = Eve(Enc_k(x_0))$是固定的且与x_1独立.

则 $\Pr[Eve(Enc_k(x_1)) = x_1] \le \frac{1}{|M'|}$ (即随机猜)

所以，$\exists x_1 \in M'$, s.t.,

$\Pr[Eve(Enc_k(x_1)) = x_1] > \Pr[Eve(Enc_k(x_0)) = x_1] \rightarrow P_{1,1} > P_{0,1}$

对Eve的解密有 $Enc_k(x_0) \xrightarrow{P_{0,0}} x_0$, $Enc_k(x_1) \xrightarrow{P_{1,0}} x_0$, $Enc_k(x_1) \xrightarrow{P_{1,1}} x_1$, $Enc_k(x_0) \xrightarrow{P_{0,1}} x_1$

$\Rightarrow P_{0,0} + P_{0,1} = 1$, $P_{0,0} + P_{1,0} = 1$, $P_{1,0} + P_{1,1} = 1$, $P_{0,1} + P_{1,1} = 1$

又$P_{1,1} > P_{0,1}$，故$P_{1,1} > \frac{1}{2}$, $P_{0,1} < \frac{1}{2}$，此时可有Eve'

$Eve'(c) = \begin{cases} x_1, & \text{若 } Eve(c) = x_1 \\ x_i, & i \in \{0,1\} \text{ 随机取，否则} \end{cases}$

此时有$\Pr[Eve'(Enc_k(x_b)) = x_b] > \frac{1}{2}$ ($b \in \{0,1\}$)，因为

$\Pr[Eve'(Enc_k(x_1)) = x_1] \begin{cases} > \frac{1}{2}, & \text{当 } Eve(Enc_k(x_1)) = x_1 \\ = \frac{1}{2}, & \text{当 } Eve(Enc_k(x_1)) = x_0 \end{cases}$

$\Pr[Eve'(Enc_k(x_0)) = x_0] = \frac{1}{2}$

(只有以上三种情况，求期望后大于$\frac{1}{2}$)