现代密码学
Modern Cryptography

张方国

中山大学计算机学院

Office: Room 305, IM School Building

E-mail: isszhfg@mail.sysu.edu.cn

HomePage:https://cse.sysu.edu.cn/content/2460

第六讲 完善保密理论

- 安全分类
- 概率论与信息论基础
- 完善保密的定义和性质
- Shannon定理与一次一密
- 乘积密码

安全分类

- 计算安全
- 可证明安全
- 无条件安全

Computational Security

- 如果使用最好的算法攻破一个密码体制需要至少N次操作，这里的N是一个特定的非常大的数字，我们可以定义这个密码体制是计算安全的。
- 没有一个已知的实际的密码体制在这个定义下可以被证明安全。
- 人们经常经过几种特定的攻击类型来研究计算上的安全性。对一种类型的攻击是安全的，并不表示对其他类型的攻击是安全的。

Provable Security

- 将密码体制的安全性归结为某个经过深入研究的数学难题。
- 这种途径只是说明了安全和另一个问题是相关的，并没有完全证明是安全的。
- 这和证明一个问题是NP完全的有些类似：证明给定的问题和任何其他的NP完全问题的难度是一样的，但没有完全证明这个问题的计算难度。

如何定义安全性

攻击手段

结果或达到的目标

摇晃

缺口

完全摧毁

可证明安全的基本思路（过程）

- **前提：**
  1，密码原语的定义；
  2，密码原语的安全性定义；
  3，密码原语的一个具体构造（方案）Π。
- **假设：**
  某个数学问题P是困难的（可能还有其他辅助假设）。
- **证明：**
  具体构造（方案）Π满足原语的安全性定义（即Π是安全的）

可证明安全的基本思路（过程）

- **方法：反证法+归约**
  假定Π不是(t, ϵ)安全的，则可构造一个算法能够(t', ϵ')求解问题类P。
- **具体过程：**
  1）设A是一个概率多项式时间算法，能够攻破方案Π，即至多运行t时间，并至少以概率ϵ成功；
  2）构造一个叫归约的算法A′，该算法通过调用算法A，试图求解问题P。这里A′只知道A可以攻击Π，但对A如何工作一无所知。指定问题类P的一个实例x，将x嵌入（或利用x构造）密码方案实例Π。

可证明安全的基本思路（过程）

(a) A′调用算法A，但对A来说是和Π交互（不是x）；
(b) 如果A成功攻破了由A′模拟的方案Π，则将允许A′至少以多项式的倒数（即1/p(n)）的成功概率求解出x。

3） 如果ϵ不是可忽略的，则A’成功的概率ϵ/p(n)也是不可忽略的。因为A’是将PPT共识机算法A作为子程序调用，所以A’也是有效的，即存在一个有效的算法求解问题类P。如果一开始假定P是困难的，则得出矛盾。

4）所以，给出关于问题P的一个困难假设的话，就可以证明不存在有效算法A能以不可忽略概率攻破Π。

Unconditional Security

- 对攻击者的计算量没有限制。即使提供了无穷的计算资源，也是无法被攻破的。
- 讨论安全性时，与攻击类型（手段）有关

- 惟密文攻击下无条件安全的密码体制是存在的。

概率论

S: 概率空间 （任意的固定的点集）

x∈*S*： 样点。

E： 事件， 它是*S*的一个子集。

传统概率的定义是由法国数学家拉普拉斯 (Laplace) 提出的：假设一个试验从n个等可能的点中选出一个点而且每个实验必须选出一个点。令m是组成事件E的点的数目。那么m/n我们就称作事件E发生的概率。

英国逻辑学家约翰·维恩 (1834-1923) 和奥地利数学家理查德 (Richard Von Mises 1883-1953) 提出建立在频率理论基础上的统计概率：假设在相同的条件下进行n 次试验， 其中E发生的次数是m， 如果在n充分大时， m/n趋于稳定。那么m/n我们就称作事件E发生的概率。 （极限）

概率的严格定义

设E是随机试验，Ω是它的样本空间。对于E的每一事件A赋于一个实数，记为P(A)，称为事件A的概率。这里P(·)是一个集合函数，P(·)要满足下列条件：

（1）非负性：对于每一个事件A，有$P(A) \geq 0$;

（2）规范性：对于必然事件S，有$P(S)=1$;

（3）可列可加性：设A_1，A_2……是两两互不相容的事件，即对于$i \neq j$，$A_i \cap A_j = \phi$，（$i,j=1,2$……），则有$P(A_1 \cup A_2 \cup \ldots\ldots) = P(A_1) + P(A_2) + \ldots\ldots$

基本性质

1、确定事件：概率空间本身就是一个事件。

2、不可能事件：永远都不可能发生的事件。

3、 Prob [E]$\in$[0,1]

4、事件之间的关系： 包含，和事件，积事件， 差事件，补事件。

基本运算

1、Prob $[E\cup F]$ = Prob $[E]$+ Prob $[F]$ - Prob $[E\cap F]$

2、 $\cup_i E_i=S$， $E_i\cap E_j=\emptyset$，则 $\sum_i$Prob $[E_i]=1$

3、条件概率：在E发生的前提下，F发生的概率

Prob$[F|E]$= Prob $[E\cap F]$/ Prob $[E]$

4、独立事件：如果Prob$[F|E]$= Prob $[F]$，称E、F为独立事件。

5、全概率公式： $\cup_i E_i=S$， $E_i\cap E_j=\emptyset$，那么任何事件A

Prob$[A]$= $\sum_i$ Prob $[A|E_i]$ Prob $[E_i]$

随机变量和概率分布

离散随机变量和其分布函数

1、一个离散随机变量定义在一个离散的样本空间上的函数，它是一个试验的某个数字表示结果。

2、令S为离散空间，X为随机变量。那么X的离散分布函数就是一个由$S \longrightarrow R$的离散值的映射：

$$\text{Prob}\ [X=x_i]=p_i$$

P_i>=0;

\sum_{i=1}^{#S} p_i=1

均匀分布，二元分布

(P,C, K, E, D)是一个特定的密码体制

P和K的概率分布导出了C的概率分布，同样，可以把密文元素**y**看成是随机变量，我们可以使用如下的公示计算出

$$Pr[\mathbf{y}=y]=\sum_{\{K:y\in C(K)\}} Pr[\mathbf{K}=K]Pr[\mathbf{x}=d_K(y)]$$

同样，可以按如下的公式计算条件概率：

$$Pr[\mathbf{y}=y|\mathbf{x}=x]=\sum_{\{K:x=d_K(y)\}} Pr[\mathbf{K}=K]$$

这样，我们可以采用Bayes公式计算：

$$Pr[\mathbf{x}=x|\mathbf{y}=y]=\frac{Pr[\mathbf{x}=x]\times\sum_{\{K:x=d_K(y)\}} Pr[\mathbf{K}=K]}{\sum_{\{K:y\in C(K)\}} Pr[\mathbf{K}=K]Pr[\mathbf{x}=d_K(y)]}$$

Example 2.3 Let $\mathcal{P} = \{a, b\}$ with $\mathbf{Pr}[a] = 1/4, \mathbf{Pr}[b] = 3/4$. Let $\mathcal{K} = \{K_1, K_2, K_3\}$ with $\mathbf{Pr}[K_1] = 1/2, \mathbf{Pr}[K_2] = \mathbf{Pr}[K_3] = 1/4$. Let $\mathcal{C} = \{1, 2, 3, 4\}$, and suppose the encryption functions are defined to be $e_{K_1}(a) = 1, e_{K_1}(b) = 2$; $e_{K_2}(a) = 2, e_{K_2}(b) = 3$; and $e_{K_3}(a) = 3, e_{K_3}(b) = 4$. This cryptosystem can be represented by the following *encryption matrix*:

	a	b
K_1	1	2
K_2	2	3
K_3	3	4

We now compute the probability distribution on $\mathcal{C}$. We obtain the following:

$$\begin{aligned}
\mathbf{Pr}[1] &= \frac{1}{8} \\
\mathbf{Pr}[2] &= \frac{3}{8} + \frac{1}{16} = \frac{7}{16} \\
\mathbf{Pr}[3] &= \frac{3}{16} + \frac{1}{16} = \frac{1}{4} \\
\mathbf{Pr}[4] &= \frac{3}{16}.
\end{aligned}$$

Now we can compute the conditional probability distributions on the plaintext, given that a certain ciphertext has been observed. We have:

$$\begin{aligned}
\mathbf{Pr}[a|1] &= 1 & \mathbf{Pr}[b|1] &= 0 \\
\mathbf{Pr}[a|2] &= \frac{1}{7} & \mathbf{Pr}[b|2] &= \frac{6}{7} \\
\mathbf{Pr}[a|3] &= \frac{1}{4} & \mathbf{Pr}[b|3] &= \frac{3}{4} \\
\mathbf{Pr}[a|4] &= 0 & \mathbf{Pr}[b|4] &= 1.
\end{aligned}$$

完善保密的定义

定义2.3 一个密码体制具有完善保密性，如果对于任意的$x \in \mathrm{P}$和$y \in \mathrm{C}$，都有$Pr[x|y] = Pr[x]$。也就是说，给定密文y，明文x的后验概率等于明文的先验概率。

通俗地说，完善保密性就是攻击者不能通过观察密文获得明文的任何信息。

定理 2.3 假设移位密码的26个密钥都是以相同的概率1/26使用的，则对于任意的明文概率分布，移位密码具有完善保密性。

证明 这里$\mathcal{P}=\mathcal{C}=\mathcal{K}=\mathbb{Z}_{26}$，对于$0\leqslant K\leqslant 25$，加密函数 e_K 定义为$e_K(x)=(x+K) \bmod 26$ $(x\in\mathbb{Z}_{26})$。首先计算$\mathcal{C}$上的概率分布。假设 $y\in\mathbb{Z}_{26}$，则

$$\begin{aligned}\mathbf{Pr}[\mathbf{y}=y] &= \sum_{K\in\mathbb{Z}_{26}}\mathbf{Pr}[\mathbf{K}=K]\mathbf{Pr}[\mathbf{x}=d_K(y)]\\ &= \sum_{K\in\mathbb{Z}_{26}}\frac{1}{26}\mathbf{Pr}[\mathbf{x}=y-K]\\ &= \frac{1}{26}\sum_{K\in\mathbb{Z}_{26}}\mathbf{Pr}[\mathbf{x}=y-K]\end{aligned}$$

现在固定 y，值$(y-K) \bmod 26$ 构成$\mathbb{Z}_{26}$的一个置换。因此有：

$$\sum_{K\in\mathbb{Z}_{26}}\mathbf{Pr}[\mathbf{x}=y-K]=\sum_{K\in\mathbb{Z}_{26}}\mathbf{Pr}[\mathbf{x}=x]=1$$

得到对于任意的 $y\in\mathbb{Z}_{26}$，

$$\mathbf{Pr}[y]=\frac{1}{26}$$

接下来，对于任意的 x, y，我们有：

$$\begin{aligned}\mathbf{Pr}[y \mid x] &= \mathbf{Pr}[\mathbf{K} = (y - x) \bmod 26] \\ &= \frac{1}{26}\end{aligned}$$

（这是因为对于任意的 x, y，满足 $e_K(x) = y$ 的惟一的密钥 $K = (y - x) \bmod 26$。）现在应用 Bayes 定理，很容易计算出：

$$\begin{aligned}\mathbf{Pr}[x \mid y] &= \frac{\mathbf{Pr}[x]\mathbf{Pr}[y \mid x]}{\mathbf{Pr}[y]} \\ &= \frac{\mathbf{Pr}[x]\frac{1}{26}}{\frac{1}{26}} \\ &= \mathbf{Pr}[x]\end{aligned}$$

所以这个密码体制是完善保密的。

Shannon定理

定理 2.4 假设密码体制(P, C, K, E, D)满足|K| = |C| = |P|。该密码体制是完善保密的，当且仅当每个密钥被使用的概率都是1/|K|，并且对于任意的$x \in P$和$y \in C$，存在唯一的密钥K使得$e_K(x) = y$。

“<=”，当这两个条件成立时，类似于上面的定理的证明，可以得出这个密码体制是完善保密的。

“=>”的证明：

证明 假设这个密码体制是完善保密的。由上面可知，对于任意的 $x \in \mathcal{P}$ 和 $y \in \mathcal{C}$，一定至少存在一个密钥 K 满足 $e_K(x)=y$。因此有不等式：

$$\begin{aligned}|\mathcal{C}| &= |\{e_K(x) : K \in \mathcal{K}\}| \\ &\leqslant |\mathcal{K}|\end{aligned}$$

但是我们假设 $|\mathcal{C}|=|\mathcal{K}|$，因此一定有：

$$|\{e_K(x) : K \in \mathcal{K}\}| = |\mathcal{K}|$$

也就是说，不存在两个不同的密钥 K_1 和 K_2 使得 $e_{K_1}(x)=e_{K_2}(x)=y$。因此对于 $x \in \mathcal{P}$ 和 $y \in \mathcal{C}$，刚好存在一个密钥 K 使得 $e_K(x)=y$。

记 $n=|\mathcal{K}|$。设 $\mathcal{P}=\{x_i : 1 \leqslant i \leqslant n\}$ 并且固定一个密文 $y \in \mathcal{C}$。设密钥为 $K_1, K_2, \cdots, K_n$，并且 $e_{K_i}(x_i)=y$，$1 \leqslant i \leqslant n$。使用 Bayes 定理，我们有：

$$\begin{aligned}\mathbf{Pr}[x_i \mid y] &= \frac{\mathbf{Pr}[y \mid x_i]\mathbf{Pr}[x_i]}{\mathbf{Pr}[y]} \\ &= \frac{\mathbf{Pr}[\mathbf{K}=K_i]\mathbf{Pr}[x_i]}{\mathbf{Pr}[y]}\end{aligned}$$

考虑完善保密的条件 $\mathbf{Pr}[x_i \mid y]=\mathbf{Pr}[x_i]$。从这里，我们有 $\mathbf{Pr}[K_i]=\mathbf{Pr}[y]$，$1 \leqslant i \leqslant n$。也就是说，所有的密钥都是等概率使用的（概率为 $\mathbf{Pr}[y]$）。但是密钥的数目为 $\mathcal{K}$，我们得到对任意的 $K \in \mathcal{K}$，$\mathbf{Pr}[K]=1/|\mathcal{K}|$。

一次一密

一个著名的具有完善保密性的密码体制是"一次一密"体制。最早由Vernam在1917年用于报文的自动加密和解密。30年后被Shannon证明了其是完善保密的。

密码体制2.1 一次一密

假设$n \geq 1$是正整数，$\mathrm{P} = \mathrm{C} = \mathrm{K} = (\mathbb{Z}_2)^n$。对于$K \in (\mathbb{Z}_2)^n$，定义$e_K(x)$为$K$和$x$的模2向量和（或者说是两个相关比特串的异或）。因此，如果$x = (x_1, x_2, ..., x_n)$并且$K = (K_1, K_2, ..., K_n)$，则

$$e_K(x) = (x_1 + K_1, ..., x_n + K_n) mod 2$$

解密与加密是一样的。如果$y = (y_1, ..., y_n)$，则

$$d_K(y) = (y_1 + K_1, ..., y_n + K_n) mod 2$$

一次一密

由定理2.4，容易看出“一次一密”提供了完善保密性。但是它有一个不利因素，因为$|K| \geq |P|$意味着秘密使用的密钥量必须至少和明文数量一样多。

密码体制的无条件安全性是基于每个密钥仅用一次这样的一个事实。

在已知明文攻击下是不安全的。

在密码学的发展历史中，人们试图设计一个密钥可以加密相对长的明文的密码体制，并且仍然可以保持一定的计算安全性。

- 使用完善保密的加密方案是不必要的。
- 只要在实践上是不可破译的密码方案就可以实用了：
  如某方案在200年内，使用最快的可用的超级计算机，也不能以大于10^{-30}的概率攻破。

乘积密码

Shannon在1949年介绍了通过“乘积”组合密码体制，这种思想非常重要。

假设$\mathrm{C} = \mathrm{P}$（内嵌式密码体制）。

设$S_1 = (\mathrm{P}, \mathrm{P}, \mathrm{K}_1, \mathrm{E}_1, \mathrm{D}_1)$和$S_2 = (\mathrm{P}, \mathrm{P}, \mathrm{K}_2, \mathrm{E}_2, \mathrm{D}_2)$是两个具有相同明文空间（密文空间）的内嵌式密码体制。那么S_1和S_2的乘积密码体制$S_1 \times S_2$定义为

$$S_1 = (\mathrm{P}, \mathrm{P}, \mathrm{K}_1 \times \mathrm{K}_2, \mathrm{E}, \mathrm{D})$$

乘积密码体制的密钥形式是$K = (K_1, K_2)$，其中$(K_1 \in \mathrm{K}_1)$，$(K_2 \in \mathrm{K}_2)$。

加密规则e_K定义为：

$$e_{(K_1,K_2)}(x) = e_{K_2}(e_{K_1}(x))$$

解密规则d_K定义为：

$$d_{(K_1,K_2)}(y) = d_{K_1}(d_{K_2}(y))$$

密码体制有与密钥空间相关的概率分布，因此，需要定义密钥空间K的概率分布。自然地定义为：

$$Pr[(K_1,K_2)] = Pr[K_1] \times Pr[K_2]$$

乘积密码体制是可交换的，如果$S_1 \times S_2 = S_2 \times S_1$.

乘积密码体制是幂等的，如果$S^2 = S$.

如果密码体制不是幂等的，那么多次迭代有可能提高安全性。

如果S_1和S_2是幂等的,并且是可交换的，则$S_1 \times S_2$是幂等的。

信息论与密码学

- **Claude Shannon** was born on April 30, 1916 in the town of Gaylord, Michigan.
- By the 1980's, Shannon began having problems with his memory and he was later diagnosed with Alzheimer's disease.
- In his final years he was “good-natured as usual” and enjoyed daily visits with his wife, Betsy.Eventually his body failed and he passed away in February 2001.
- **A Mathematical Theory of Communication**.
  Bell Syst. Tech. J.,27:379-423, 1948
- **Communication Theory of Secrecy Systems.**
  Bell Syst.Tech. J., 28:
  656-715,949

信息论与密码学

通信系统与密码系统。消息的加密与破译和信息论密切相关。

通信系统：用信息论观点研究存在随机干扰时通信系统中的信息传输问题[Shannon在1948年发表的“**通信的数学理论**”]。在有扰条件下，发送的消息*m*在噪声干扰下变为*m’*，一般*m’*≠*m*。接收者的任务是从收到的*m* 试图恢复原来的消息。为了使这成为可能，发送者常常要对消息进行**编码**，按一定规则增加一些**多余数字**，以便在出错时使接收者能对其进行**检测**或**纠正**。

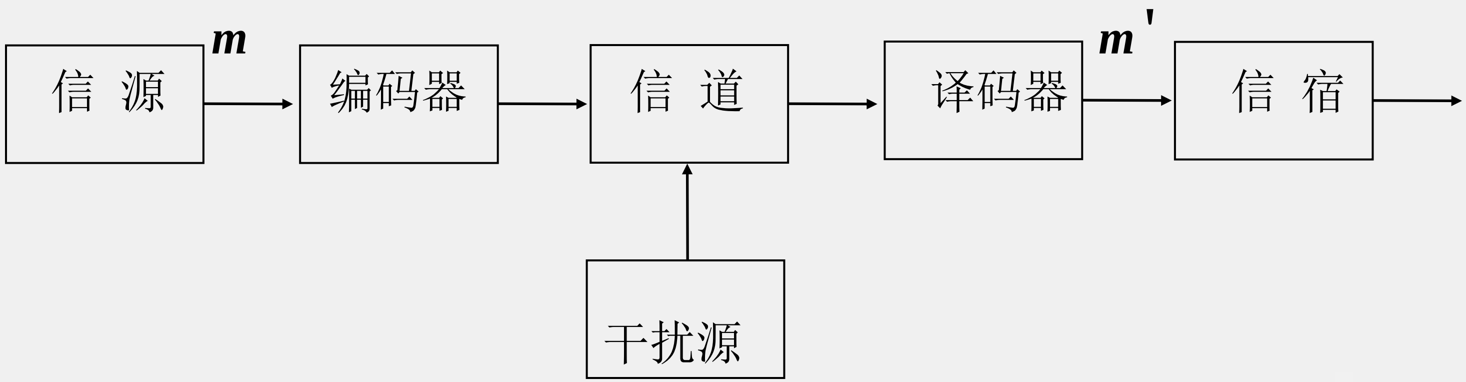

图 通信系统

信息论与密码学

密码系统： 对消息*m*的加密变换的作用类似于向消息注入噪声。密文*c*就相当于经过有扰信道得到的接收消息。密码分析员就相当于有扰信道下原接收者。所不同的是，这种干扰不是信道中的自然干扰，而是发送者有意加进的，目的是使窃听者不能从*c*恢复出原来的消息。[Shannon1949年发表的“**保密系统的通信理论**”]。用信息论的观点对信息保密问题作了全面的阐述。信息论成为研究密码学和密码分析学的一个重要理论基础，宣告了科学的密码学信息理论时代的到来。

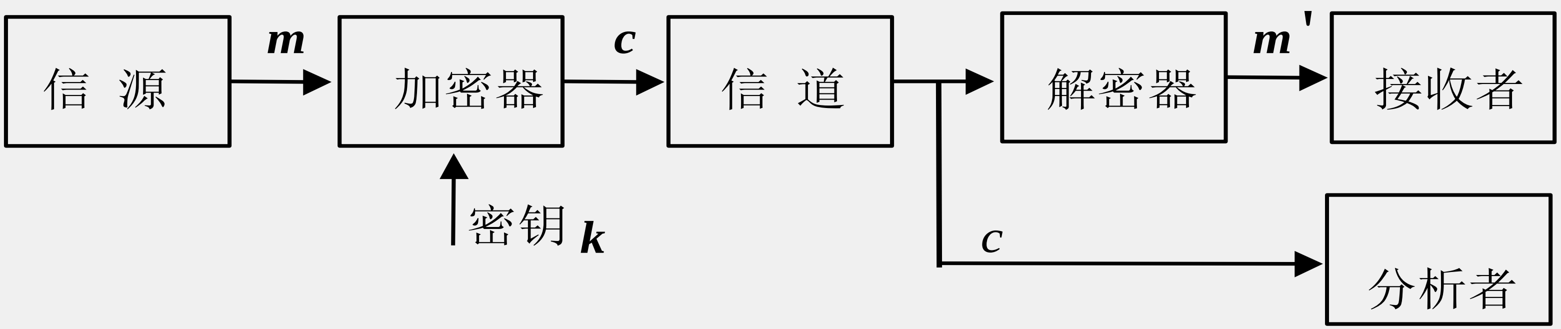

图 保密系统

不确定性和信息量

随机事件的一个特性就是不确定性。比如说Alice成绩优异，则预测她保研的概率是很大的。然而预测密码学考试的第一道题是什么，则非常困难。也就是说，其非确定性要高的多。所以对生活中的事件，如何去度量这个不确定性？

“不确定性”和“信息量”有着重要的关系。比如说，某人在一个晴朗的天气宣布“现在没有下雨”，这多半引不起别人的兴趣（其确定性非常大）。但是，有人宣布发现了“黎曼猜想”的证明，这将引起整个数学界的轰动。

（大家都要关心其真伪，并进行验证）。所以，事情的不确定性越大，其信息量也越大。

信息量不是一成不变的。对不同的对象，在不同的时间都是发生变化的。

信息量和熵

信息量和熵

自信息量：　　　$I(x_i)=-\log_a p_i$

离散集合$X=\{x_i，i=1, \dots, n\}$，$p(x_i)\geq 0$是x_i出现的概率。

事件x_i出现给出的信息量。

事件x_i出现的可能性大小，也是为确定事件x_i的出现所必须付出的信息量。

信息量的单位

$a=2$时为**比特**(bit)。它表示两个等可能事件集中，一个事件出现给出的信息量。

$a=e$为**奈特**(nat)，$a=10$时为**铁特**(Tet)。

1 bit=0.693 nat=0.301 Tet。

平均自信息量： $H(X)=-\sum_i p(x_i)lb\,p(x_i)\geq 0$

集X中事件出现给出的信息的统计平均值，为集X的**熵**(entropy)。*lb:=log_b*

表示X中出现一个事件平均给出的信息量，

集 X中事件的**平均不确定性**(average uncertainty)，

确定集X出现一个符号必须提供的信息量。

熵的性质：

联合熵：

条件熵

Huffman编码与熵

- 1951年， Huffman和他在MIT信息论的同学需要选择是完成学期报告还是期末考试。导师Robert M. Fano给他们的学期报告的题目是，寻找最有效的二进制编码。
- 1952年，David A. Huffman发表了A Method for the Construction of Minimum-Redundancy Codes一文，它一般就叫做Huffman编码。
- “Huffman编码” 又称“熵编码法”，用于数据的无损耗压缩。
- （自学）